Incluye código de acceso a:
MathXL®

david c. lay

álgebra lineal
y sus aplicaciones

4a edición

ALWAYS LEARNING

PEARSON

Código de acceso a MathXL® (en inglés)

Para registrarse en línea, ingrese a **www.mathxl.com**, vaya a la sección *Register*, seleccione la opción *Student* y siga las instrucciones que aparecen en pantalla.

El código de acceso para el estudiante que viene a continuación sólo puede usarse una vez y permite el acceso a esta plataforma durante un año.

Si está adquiriendo un libro nuevo y este código aparece descubierto, es probable que ya haya sido utilizado por alguien más.

IMPORTANTE:
Use una moneda para raspar y descubrir el código de acceso.
No use objetos afilados porque podría dañarlo.
El código de acceso no puede ser reemplazado en caso de daño.

CUARTA EDICIÓN

Álgebra lineal
y sus aplicaciones

David C. Lay
University of Maryland—College Park

Traducción
Ana Elizabeth García Hernández
Traductora especialista en matemáticas

Revisión técnica
Javier Alfaro Pastor
Instituto Tecnológico Autónomo de México

PEARSON

Datos de catalogación bibliográfica

LAY, DAVID C.

Álgebra lineal y sus aplicaciones.
Cuarta edición

PEARSON EDUCACIÓN, México, 2012

ISBN: 978-607-32-1398-1
Área: Matemáticas

Formato: 21 × 27 cm Páginas: 576

ISBN 9780321385178

Edición en español
Dirección Educación Superior: Mario Contreras
Editor sponsor: Gabriela López Ballesteros
e-mail: gabriela.lopezballesteros@pearson.com
Editor de desarrollo: Felipe Hernández Carrasco
Supervisor de Producción: Enrique Trejo Hernández
Gerencia Editorial Latinoamérica: Marisa de Anta

CUARTA EDICIÓN, 2012

D.R. © 2012 por Pearson Educación de México, S.A. de C.V.
Atlacomulco 500-5º piso
Col. Industrial Atoto
53519, Naucalpan de Juárez
Estado de México

Cámara Nacional de la Industria Editorial Mexicana. Reg. núm. 1031.

ISBN 978-607-32-1398-1
ISBN e-book 978-607-32-1399-8
ISBN e-chapter 978-607-32-1400-1

Impreso en México. *Printed in Mexico.*
1 2 3 4 5 6 7 8 9 0 - 15 14 13 12

PEARSON

www.pearsonenespañol.com **ISBN: 978-607-32-1398-1**

A mi esposa, Lillian, y a nuestras hijas
Christina, Deborah y Melissa, cuyo
apoyo, ánimos y devotas
oraciones hicieron posible este libro.

Acerca del autor

David C. Lay tiene una licenciatura de Aurora University (Illinois), y una maestría y un doctorado en la Universidad de California en Los Ángeles. Lay ha sido catedrático e investigador en matemáticas desde 1966, principalmente en la Universidad de Maryland, College Park. También ha trabajado como profesor visitante en la Universidad de Amsterdam, en la Universidad Libre de Amsterdam y en la Universidad de Kaiserslautern, en Alemania. Ha escrito más de 30 artículos de investigación de análisis funcional y álgebra lineal.

Como miembro fundador del Grupo de Estudio del Currículo de Álgebra Lineal patrocinado por la NSF, ha sido líder en el movimiento actual para modernizar el plan de estudios de álgebra lineal. Lay también es coautor de varios libros de matemáticas, entre los que se incluyen, *Introduction to Functional Analysis*, con Angus E. Taylor, *Calculus and its Applications*, con L. J. Goldstein y D. I. Schneider, y *Linear Algebra Gems–Assets for Undergraduate Mathematics*, con D. Carlson, C. R. Johnson y A. D. Porter.

El profesor Lay ha recibido cuatro premios universitarios por excelencia docente, incluido el de Distinguished Scholar Teacher de la Universidad de Maryland en 1996. En 1994 la Mathematical Association of America le otorgó el premio Distinguished College or University Teaching of Mathematics. Ha sido elegido por los estudiantes universitarios como miembro de la Alpha Lambda Delta National Scholastic Honor Society y de la Golden Key National Honor Society. En 1989 Aurora University le concedió el premio Outstanding Alumnus. Lay es miembro de la American Mathematical Society, de la Canadian Mathematical Society, de la International Linear Algebra Society, de la Mathematical Association of America, Sigma Xi, y de la Society for Industrial and Applied Mathematics. Desde 1992 ha formado parte de la junta directiva nacional de la Association of Christians in the Mathematical Sciences.

Contenido

Los capítulos 9 y 10 se encuentran en inglés en el sitio Web del libro.

Apéndices

Prefacio

La respuesta de los estudiantes y profesores a las tres primeras ediciones de *Álgebra lineal y sus aplicaciones* ha sido muy gratificante. Esta cuarta edición brinda un importante apoyo tanto para la enseñanza como para el uso de la tecnología en el curso. Al igual que en las ediciones anteriores, el libro ofrece una introducción elemental actualizada al álgebra lineal y una amplia selección de aplicaciones interesantes. El material es accesible a estudiantes con la madurez que se consigue al finalizar de manera exitosa dos semestres de matemáticas de nivel universitario, por lo general, de cálculo.

El objetivo principal del libro es ayudar a los estudiantes a dominar los conceptos básicos y las habilidades que usarán más adelante en sus carreras. Los temas expuestos siguen las recomendaciones del Grupo de Estudio del Currículo de Álgebra Lineal, las cuales se basan en una cuidadosa investigación de las necesidades reales de los estudiantes y en un consenso entre profesionales de muchas disciplinas que utilizan el álgebra lineal. Esperamos que este curso sea una de las clases de matemáticas más útiles e interesantes para los estudiantes de licenciatura.

LO NUEVO EN ESTA EDICIÓN

El principal objetivo de esta revisión fue actualizar los ejercicios e incluir nuevos contenidos, tanto en el libro como en línea.

1. Más del 25 por ciento de los ejercicios son nuevos o actualizados, en especial los ejercicios computacionales. Los conjuntos de ejercicios son una de las características más importantes de este libro, y estos nuevos ejercicios siguen el mismo estándar elevado de los conjuntos de ejercicios de las tres últimas ediciones. Están diseñados de tal forma que se refieren a los temas importantes de cada una de las secciones anteriores, y permiten que los alumnos desarrollen confianza al motivarlos a practicar y generalizar las nuevas ideas que acaban de estudiar.

2. El 25 por ciento de los ejemplos introductorios de los capítulos son nuevos. Estas introducciones tienen que ver con aplicaciones de álgebra lineal y despiertan el interés en torno al desarrollo del tema que se presenta a continuación. El texto retoma el ejemplo introductorio en una sección al final de cada capítulo.

3. Se incluye un nuevo capítulo, el 8, titulado “Geometría de los espacios vectoriales”, el cual presenta un tema novedoso que mis alumnos han disfrutado estudiar. Las secciones 1, 2 y 3 ofrecen las herramientas geométricas básicas. La sección 6 utiliza estas ideas para estudiar las curvas y superficies de Bézier, las cuales se utilizan en gráficos elaborados con computadora en el campo de la ingeniería y en línea (en Adobe® Illustrator® y Macromedia® FreeHand®). Estas cuatro secciones se pueden cubrir en cuatro o cinco sesiones de clase de 50 minutos.

El segundo curso en las aplicaciones de álgebra lineal suele comenzar con una revisión sustancial de las ideas principales del primer curso. Si una parte del capítulo 8 se encuentra en el primer curso, el segundo podría incluir una breve reseña de las secciones 1 a 3 y, luego, un enfoque de la geometría en las secciones 4 y 5. Eso conduciría, naturalmente, a los capítulos 9 y 10 que se presentan en línea, los cuales se han utilizado junto con el capítulo 8 en varias escuelas en los últimos cinco años.

4. Hay dos nuevos capítulos disponibles en línea en inglés, y se pueden utilizar en un segundo curso:

   Chapter 9. Optimization
   Chapter 10. Finite-State Markov Chains

   Se requiere un código de acceso y está disponible para todos los profesores que adopten el libro. Para más información, visite *www.pearsonhighered.com/irc* o póngase en contacto con su representante de Pearson.
5. Diapositivas de PowerPoint® están disponibles para las 25 secciones principales del texto; también se incluyen más de 75 figuras del texto.

CARACTERÍSTICAS DISTINTIVAS

Introducción temprana a los conceptos clave

Muchas de las ideas fundamentales del álgebra lineal se introducen dentro de las primeras siete lecturas en el contexto concreto de $\mathbb{R}^n$, y después, gradualmente, se examinan desde diferentes puntos de vista. Más adelante, se presentan generalizaciones de estos conceptos como extensiones naturales de ideas familiares, visualizadas a través de la intuición geométrica desarrollada en el capítulo 1. Un logro importante del libro es que el nivel de dificultad es bastante uniforme durante todo el curso.

Una visión moderna de la multiplicación de matrices

Una buena notación es importante, y el libro refleja la manera en que los científicos e ingenieros utilizan el álgebra lineal en la práctica. Las definiciones y demostraciones se centran en las columnas de una matriz antes que en sus entradas. Un tema central es considerar un producto matriz-vector $A\mathbf{x}$ como una combinación lineal de las columnas de A. Este enfoque moderno simplifica muchos argumentos, y vincula las ideas de espacio vectorial con el estudio de sistemas lineales.

Transformaciones lineales

Las transformaciones lineales forman un "hilo" que se entreteje en la trama del libro. Su uso mejora el sentido geométrico del texto. En el capítulo 1, por ejemplo, las transformaciones lineales ofrecen una visión dinámica y gráfica de la multiplicación matriz-vector.

Valores propios y sistemas dinámicos

Los valores propios se presentan muy pronto en el libro, en los capítulos 5 y 7. Como este material se estudia durante varias semanas, los estudiantes tienen más tiempo de lo habitual para aprender y revisar tales conceptos fundamentales. Los valores propios se aplican a sistemas dinámicos discretos y continuos, los cuales se presentan en las secciones 1.10, 4.8 y 4.9, y en las cinco secciones del capítulo 5. Algunos cursos llegan al capítulo 5, después de aproximadamente cinco semanas, cubriendo las secciones 2.8 y 2.9 en vez del capítulo 4. Estas dos secciones opcionales presentan todos los conceptos de espacio vectorial del capítulo 4 necesarios para el capítulo 5.

Ortogonalidad y problemas de mínimos cuadrados

Estos temas reciben un tratamiento más completo que el que se otorga comúnmente en los libros básicos. El Grupo de Estudio del Currículo de Álgebra Lineal ha hecho hincapié en la necesidad de contar con una unidad sustancial de ortogonalidad y problemas de mínimos cuadrados, ya que la ortogonalidad desempeña un importante papel en los cálculos computacionales y en el álgebra lineal numérica, y porque, con frecuencia, en el trabajo práctico surgen sistemas lineales inconsistentes.

CARACTERÍSTICAS PEDAGÓGICAS

Aplicaciones

Una amplia selección de aplicaciones muestra el poder del álgebra lineal para explicar principios fundamentales y simplificar los cálculos en ingeniería, ciencias de la computación, matemáticas, física, biología, economía y estadística. Algunas aplicaciones se presentan en secciones separadas, mientras que otras se explican con ejemplos y ejercicios. Además, cada capítulo se inicia con un ejemplo introductorio que prepara el escenario para algunas aplicaciones del álgebra lineal y sirve de base para el desarrollo de las matemáticas que siguen. Después, el texto considera nuevamente la aplicación en una sección cercana al final del capítulo.

Un fuerte énfasis geométrico

Todos los conceptos importantes en el curso cuentan con una interpretación geométrica, ya que muchos estudiantes aprenden mejor cuando logran visualizar una idea. Aquí se presentan más dibujos de lo habitual, y algunas de las figuras nunca antes se han presentado en un libro de álgebra lineal.

Ejemplos

Este libro dedica una mayor proporción de su material de exposición a ejemplos, en comparación con la mayoría de libros de álgebra lineal. Hay más ejemplos de los que un profesor presenta normalmente en clase. Puesto que los ejemplos se escribieron con sumo cuidado y con detalle, los estudiantes pueden leerlos por su cuenta.

Teoremas y demostraciones

Los resultados importantes se establecen como teoremas. Otros datos útiles se presentan en recuadros, para una fácil localización. La mayoría de los teoremas incluyen demostraciones formales, escritas pensando en el alumno principiante. En algunos casos, los cálculos esenciales de una demostración se muestran en un ejemplo cuidadosamente elegido. Algunas comprobaciones de rutina se dejan para los ejercicios, cuando sea benéfico para los estudiantes.

Problemas de práctica

Antes de cada conjunto de ejercicios se incluyen problemas de práctica seleccionados con gran cuidado. Las soluciones completas se presentan después del conjunto de ejercicios. Estos problemas se centran en los aspectos problemáticos del conjunto de ejercicios o sirven de "calentamiento" para los ejercicios; con frecuencia, las soluciones contienen útiles consejos o advertencias acerca del trabajo que hay que realizar.

Ejercicios

El gran número de ejercicios incluye desde algunos que tienen que ver con cálculos de rutina hasta preguntas conceptuales que requieren de mayor reflexión. Un buen número de preguntas innovadoras destacan las dificultades conceptuales que he encontrado en los documentos de los estudiantes en los últimos años. Cada conjunto de ejercicios está cuidadosamente organizado en el mismo orden general que el libro, de manera que las tareas se pueden encontrar fácilmente cuando solo se ha estudiado una parte de la sección. Una característica notable de los ejercicios es su sencillez numérica. El contenido de los problemas se puede ordenar rápidamente, para que los estudiantes dediquen poco tiempo a los cálculos numéricos. Los ejercicios se concentran en enseñar a razonar antes que en realizar cálculos mecánicos. Los ejercicios de la *cuarta edición* conservan la integridad de los que se incluyeron en la tercera edición, y presentan nuevos problemas para estudiantes y profesores.

Los ejercicios marcados con el símbolo [**M**] están diseñados para trabajarse con la ayuda de un "programa de **M**atrices" (por ejemplo, programas computacionales, como **M**ATLAB®, **M**aple™, **M**athematica®, **M**athCad®, o Derive™, o calculadoras programables con capacidades matriciales, como las que fabrica Texas Instruments).

Preguntas verdadero/falso

Para animar a los estudiantes a leer todo el libro y a pensar críticamente, he desarrollado 300 preguntas sencillas de falso/verdadero que se presentan en 33 secciones del libro, justo después de los problemas computacionales. Estas preguntas se pueden contestar directamente del libro, y preparan al estudiante para los problemas conceptuales que siguen. Los estudiantes aprecian estas preguntas una vez que valoran la importancia de leer con cuidado el libro. Con base en las pruebas de clase y los análisis con los estudiantes, decidí no incluir las respuestas en el libro. Se cuenta con 150 preguntas adicionales de falso/verdadero (casi siempre al final de los capítulos) para comprobar la comprensión del material. El libro presenta solo respuestas con V o F para la mayoría de estas preguntas, pero omite las justificaciones de las respuestas (las cuales, por lo general, requieren de cierto razonamiento).

Ejercicios de escritura

La capacidad de escribir enunciados matemáticos coherentes en español es esencial para todos los estudiantes de álgebra lineal, y no solo para aquellos que cursan un posgrado en matemáticas. El libro incluye muchos ejercicios para los que una justificación por escrito es parte de la respuesta. Los ejercicios conceptuales que requieren una prueba corta, por lo general, incluyen consejos que ayudan a los estudiantes a comenzar. Para todos los ejercicios de escritura de numeración impar, en la parte final del libro, se incluye ya sea una solución o una sugerencia.

Temas computacionales

El libro hace hincapié en los efectos de la computadora tanto en el desarrollo como en la práctica del álgebra lineal en las ciencias y la ingeniería. Las frecuentes notas numéricas llaman la atención en torno a problemas computacionales; además, distinguen entre los conceptos teóricos, como la inversión de matrices, y las implementaciones computacionales, como la factorización LU.

APOYO EN LÍNEA

El sitio Web en *www.pearsonenespañol.com/lay* contiene material de apoyo para el libro de texto. Para los estudiantes, incluye **hojas de repaso** y **exámenes de práctica** (con soluciones) que cubren los temas principales en el libro. Estas secciones provienen directamente de cursos que he impartido en los últimos años. Cada hoja de repaso identifica definiciones clave, así como teoremas y habilidades de una parte específica del libro.

Aplicaciones de los capítulos

El sitio Web también contiene siete estudios de caso, los cuales amplían los temas introducidos al inicio de cada capítulo, al agregar datos del mundo real y la posibilidad de realizar una exploración más profunda. Por otro lado, más de veinte proyectos de aplicación amplían los temas del libro e introducen nuevas aplicaciones, como splines cúbicos, rutas de vuelo de aerolíneas, matrices de dominio en competencias deportivas y códigos de corrección de errores. Algunas aplicaciones matemáticas son técnicas de integración, ubicación de raíces polinomiales, secciones cónicas, superficies cuadráticas y extremos de funciones de dos variables. También se incluyen temas de álgebra lineal numérica, como números de condición, factorizaciones de matrices y el método QR para encontrar valores propios. Entretejidos en cada análisis, se encuentran ejercicios que pueden implicar grandes conjuntos de datos (por lo que requieren de tecnología para su solución).

Introducción a la tecnología

Si el curso incluye un trabajo con MATLAB, Maple, Mathematica o calculadoras TI, se puede leer uno de los proyectos en el sitio Web para tener una introducción a la tecnología.

Archivos de datos

Cientos de archivos contienen datos de 900 ejercicios del texto, estudios de caso y proyectos de aplicación. Los datos están disponibles en *www.pearsonenespañol.com/lay* en una variedad de formatos, para MATLAB, Maple, Mathematica y las calculadoras graficadoras TI-83+/86/89. Al permitir a los alumnos acceder a las matrices y los vectores de un problema particular con solo pulsar unas cuantas teclas, los archivos de datos eliminan los errores de captura de datos y ahorran tiempo en la tarea.

Proyectos MATLAB

Estos proyectos de exploración invitan a los estudiantes a descubrir los aspectos matemáticos y numéricos básicos de álgebra lineal. Escritos por Rick Smith, se han desarrollado para acompañar los cursos de álgebra lineal computacional en la Universidad de Florida, que han utilizado *Álgebra lineal y sus aplicaciones* durante muchos años. Se hace referencia a los proyectos por medio de un icono WEB en puntos adecuados del libro. Alrededor de la mitad de los proyectos exploran conceptos fundamentales, como el espacio columna, la diagonalización y las proyecciones ortogonales; varios proyectos tratan temas numéricos, tales como flops, métodos iterativos y DVS, y algunos más exploran aplicaciones como la interpolación de Lagrange y las cadenas de Markov.

COMPLEMENTOS

Manuales de tecnología para el profesor

Cada manual ofrece una guía detallada para integrar al curso un paquete de software específico o una calculadora gráfica. Los manuales fueron escritos por profesores que ya han utilizado tecnología con este libro. Los siguientes manuales están disponibles para profesores que adopten el libro, a través de Pearson Instructor Resource Center, *www.pearsonhighered.com/irc*: MATLAB (ISBN: 0-321-53365-8), Maple (ISBN: 0-321-75605-3), Mathematica (ISBN: 0-321-38885-2) y TI-83+/86/89 (ISBN: 0-321-38887-9).

AGRADECIMIENTOS

Estoy muy agradecido con muchos grupos de personas que me han ayudado en los últimos años con diversos aspectos de este libro.

Quiero agradecer a Israel Gohberg y Robert Ellis, quienes desde hace más de quince años han colaborado conmigo en la investigación, lo que ha contribuido a formar en gran parte mi punto de vista del álgebra lineal. Para mí, ha sido un privilegio ser un miembro del Grupo de Estudio del Currículo de Álgebra Lineal junto con David Carlson, Charles Johnson y Duane Porter. Sus ideas creativas acerca de la enseñanza del álgebra lineal han influido en este libro de forma significativa.

Agradezco sinceramente a los siguientes revisores por su cuidadoso análisis y sugerencias constructivas:

Rafal Ablamowicz, *Tennessee Technological University*
Brian E. Blank, *Washington University en Saint Louis*
Vahid Dabbaghian-Abdoly, *Simon Fraser University*
James L. Hartman, *The College of Wooster*
Richard P. Kubelka, *San Jose State University*
Martin Nikolov, *University of Connecticut*
Ilya M. Spitkovsky, *College of William & Mary*
John Alongi, *Northwestern University*
Steven Bellenot, *Florida State University*
Herman Gollwitzer, *Drexel University*
David R. Kincaid, *The University of Texas en Austin*
Douglas B. Meade, *University of South Carolina*
Tim Olson, *University of Florida*
Albert L. Vitter III, *Tulane University*

En esta cuarta edición, agradezco a mi hermano, Steven Lay, de Lee University, por su generosa ayuda y aliento, y por su reciente revisión del capítulo 8. Agradezco a Raymond Rosentrater, de Westmont College, por sus útiles consejos y su ayuda con los ejemplos introductorios de los capítulos. Otra talentosa profesora, Judith McDonald, de Washington State University, desarrolló muchos nuevos ejercicios para el libro. Su ayuda y entusiasmo por el libro fue muy refrescante y estimulante.

Agradezco a los expertos en tecnología que trabajaron en los diferentes complementos de la cuarta edición, la preparación de los datos, la redacción de las notas para los profesores, la escritura de notas de tecnología para los estudiantes y por compartir sus proyectos con nosotros: Jeremy Case (MATLAB), Taylor University; Douglas Meade (Maple), University of South Carolina; Michael Miller (calculadora TI), Western Baptist College; y Marie Vanisko (Mathematica), Carroll College.

Agradezco al profesor John Risley y a los estudiantes de posgrado David Aulicino, Sean Burke y Goldberg Hersh por sus conocimientos técnicos para ayudar a desarrollar las tareas en línea que apoyan el libro. Por las pruebas en clase de este apoyo de tareas en línea, estoy muy agradecido con: Agnes Boskovitz, Malcolm Brooks, Elizabeth Ormerod, Alexander Isaev y John Urbas, de la Australian National University; John Scott y Wee Leben, del Montgomery College, Maryland; y Xingru Zhang en SUNY University of Buffalo.

Agradezco la ayuda de Blaise DeSesa, Jean Horn, Roger Lipsett, Paul Lorczak, Thomas Polaski, Sarah Streett y Marie Vanisko, quienes comprobaron la exactitud de los cálculos en el libro.

Por último, agradezco sinceramente al personal de Addison-Wesley por toda su ayuda en el desarrollo y la producción de la cuarta edición: Caroline Celano, editora responsable; Chere Bemelmans, editora de contenido; Tamela Ambush, editora administrativa asociada; Carl Cottrell, productor de medios de comunicación; Jeff Weidenaar, director ejecutivo de marketing; Kendra Bassi, asistente de marketing; y Andrea Nix, diseñadora de texto. Por último, agradezco a tres buenos amigos que han guiado el desarrollo de la obra casi desde el principio con sus sabios consejos y estímulos: Greg Tobin, editor, Laurie Rosatone, editor anterior, y William Hoffman, editor actual. Muchas gracias a todos.

David C. Lay

AGRADECIMIENTOS

Pearson agradece a los profesores usuarios de esta obra y a los centros de estudio por su apoyo y retroalimentación, elementos fundamentales para esta nueva edición de *Álgebra lineal y sus aplicaciones*.

COLOMBIA
Universidad Nacional de Colombia
Departamento de Matemáticas
Gustavo Rubiano

MÉXICO

AGUASCALIENTES
Instituto Tecnológico de Aguascalientes
Ciencias Básicas
Alejandra Espinosa Guzmán
David Ortiz Acosta
Jesús Espino Márquez
José Refugio González López
Judith Mauricio de Anda
Paula Castillo Rosales
Sergio Heraccio Sánchez Calvillo

DISTRITO FEDERAL
Instituto Tecnológico Autónomo de México
Departamento de Matemáticas
Araceli Reyes Guerrero
Marcela González Peláez

Universidad Anáhuac del Sur
Departamento de Matemáticas
José Antonio Bohon Devars

Universidad del Valle de México campus Tlalpan
Departamento de Matemáticas
Juan Andrés Aspiazu Fabián

GUANAJUATO
Instituto Tecnológico de Celaya
Ciencias Básicas
José Carlos Cárdenas Rivera

SAN LUIS POTOSÍ
Universidad Autónoma de San Luis Potosí
Física y Matemáticas
Guadalupe Silva Esparza
J. Socorro Loera Díaz
María del Pilar Yudiche Paz
María Eugenia Noriega Treviño
María Irene Liliana Gallegos García
María Isabel Zermeño Mantante
Miguel Ángel Viramontes Reyna

PUEBLA
Instituto Tecnológico de Estudios Superiores
de Monterrey, Campus Puebla
Departamento Académico de Administración
Escuela de Negocios y Ciencias Sociales
Jorge Alberto González Mendivil
Miguel Guadalupe Díaz Sánchez

Instituto Tecnológico de Puebla
Departamento Ingeniería Industrial
Escuela de Ingeniería
Alfonso Serrano Gálvez

Universidad De Las Américas Puebla
Departamento de Turismo
Escuela de Negocios y Economía
Alfonso Rocha Herrera

Universidad Popular Autónoma del Estado de Puebla
Departamento Administración
Escuela de Negocios
Claudia Malcón Cervera

SINALOA
Instituto Tecnológico de Estudios Superiores
de Monterrey, Campus Sinaloa
Departamento de Ingeniería
Cruz Evelia Sosa Carrillo

Universidad de Occidente, Unidad Culiacán
Departamento de Ingeniería
Raúl Soto Murray

Nota para los estudiantes

Este curso es potencialmente el más interesante y valioso de los cursos de matemáticas de licenciatura. De hecho, algunos estudiantes me han escrito o han hablado conmigo después de la graduación para decirme que aún utilizan este libro de cuando en cuando como una referencia en su carrera en las grandes corporaciones y en las escuelas de posgrado de ingeniería. Los siguientes comentarios ofrecen algunos consejos prácticos e información para ayudarle a dominar el material y disfrutar del curso.

En álgebra lineal, los *conceptos* son tan importantes como los *cálculos*. Los sencillos ejercicios numéricos que se incluyen al principio de cada conjunto de ejercicios solo le ayudarán a comprobar su comprensión de los procedimientos básicos. Más adelante en su carrera, las computadoras harán los cálculos, pero usted tendrá que elegir cuáles son pertinentes, saber interpretar los resultados, y después explicar los resultados a otras personas. Por esta razón, muchos ejercicios en el libro le piden que explique o justifique sus cálculos. Con frecuencia se solicita una explicación por escrito como parte de la respuesta. Para los ejercicios con numeración impar, se incluye ya sea la explicación deseada o, al menos, una buena sugerencia. Debe evitar la tentación de consultar esas respuestas antes de haber tratado de escribir la solución. De lo contrario, es probable que crea que entiende algo cuando en realidad no es así.

Para dominar los conceptos de álgebra lineal, tendrá que leer y releer el texto con cuidado. Los nuevos términos aparecen en negritas, a veces dentro de un recuadro de definición. Al final del libro se incluye un glosario. Algunos hechos importantes se establecen como teoremas o se destacan en recuadros sombreados, para una fácil localización. Le animo a que lea las primeras cinco páginas del prefacio para aprender más acerca de la estructura de este libro. Esto le dará una idea para comprender cómo puede continuar el curso.

En un sentido práctico, el álgebra lineal es un lenguaje. Usted tiene que aprender este lenguaje de la misma manera que un idioma extranjero, esto es, con el trabajo diario. El material que se presenta en una sección no es fácil de entender a menos que haya estudiado a fondo el libro y que haya trabajado los ejercicios de las secciones anteriores. ¡Mantenerse al día con el curso le ahorrará mucho tiempo y angustia!

Notas numéricas

Espero que lea las notas numéricas en el texto, incluso si no está utilizando una computadora o una calculadora gráfica con el libro. En la vida real, la mayoría de las aplicaciones del álgebra lineal implican cálculos numéricos que están sujetos a algún error numérico, aunque quizás este sea muy pequeño. Las notas numéricas le advertirán las posibles dificultades en el uso del álgebra lineal más adelante en su carrera, y si usted estudia las notas ahora, es más probable que las recuerde después.

Si le gusta leer las notas numéricas, es posible que desee tomar un curso más tarde en álgebra lineal numérica. Debido a la gran demanda de mayor capacidad para realizar cálculos, científicos de la computación y matemáticos trabajan en álgebra lineal numérica para desarrollar algoritmos de cálculos más rápidos y más confiables, mientras que los ingenieros eléctricos diseñan computadoras pequeñas y rápidas para ejecutar algoritmos. Este es un campo emocionante, y su primer curso de álgebra lineal le ayudará a prepararse para ello.

1

Ecuaciones lineales en álgebra lineal

EJEMPLO INTRODUCTORIO

Modelos lineales en economía e ingeniería

Al final del verano de 1949, Wassily Leontief, profesor de Harvard, introducía con cuidado la última de sus tarjetas perforadas en la computadora Mark II de la universidad. Las tarjetas contenían información acerca de la economía de Estados Unidos; se trataba de un resumen de más de 250,000 datos generados por la Oficina de Estadística Laboral (U.S. Bureau of Labor) durante dos años de intenso trabajo. Leontief dividió la economía estadounidense en 500 "sectores", que incluían las industrias carbonífera, automotriz, de comunicaciones, etcétera. Para cada sector, escribió una ecuación lineal que describía cómo la industria en cuestión distribuía su producto hacia los otros sectores de la economía. Como la computadora Mark II, una de las más grandes de su época, no podía manejar el sistema resultante de 500 ecuaciones y 500 incógnitas, Leontief redujo el problema a un sistema de 42 ecuaciones y 42 incógnitas.

Programar la Mark II para manejar las 42 ecuaciones de Leontief requirió varios meses de trabajo, y él estaba ansioso por ver cuánto tardaría la computadora en resolver el problema. La máquina emitió zumbidos y sus luces parpadearon durante 56 horas antes de que finalmente arrojara un resultado. En las secciones 1.6 y 2.6 se analizará la naturaleza de esa solución.

Leontief, galardonado en 1973 con el Premio Nobel de Economía, abrió la puerta a una nueva era en la elaboración de modelos matemáticos en economía. Sus esfuerzos en Harvard, en 1949, representaron uno de los primeros usos significativos de las computadoras para analizar lo que, en esa época, era un modelo matemático de gran escala. Desde entonces, investigadores en muchos otros campos han empleado computadoras para analizar modelos matemáticos. Debido a las enormes cantidades de datos implicados, los modelos, por lo regular, son *lineales*; es decir, se describen mediante *sistemas de ecuaciones lineales*.

La importancia del álgebra lineal para diversas aplicaciones ha crecido en proporción directa al incremento de la capacidad de las computadoras, y cada nueva generación de hardware y software dispara la demanda de capacidades aun mayores. Por ello, la ciencia de la computación está fuertemente vinculada con el álgebra lineal a través del explosivo crecimiento de los procesamientos en paralelo y el cálculo a gran escala.

Ahora los científicos e ingenieros trabajan en problemas cada vez más complejos, lo que era impensable hace algunas décadas. Actualmente, ¡el álgebra lineal tiene mayor valor potencial para estudiantes de muchos campos científicos y de negocios que cualquier otra materia de matemáticas! El material que se presenta en este libro ofrece el fundamento para un trabajo posterior en muchas áreas interesantes. A continuación se mencionan unas cuantas posibilidades; otras se describirán más adelante.

- *Exploración petrolera*. Cuando un barco busca depósitos submarinos de petróleo, sus computadoras resuelven *todos los días* miles de sistemas de ecuaciones lineales. Los datos sísmicos de las ecuaciones se obtienen a partir de las ondas de choque submarinas generadas por explosiones de pistolas de aire. Las ondas rebotan en las rocas bajo el agua, y los geófonos conectados a la popa del barco mediante cables de varios kilómetros se encargan de medirlas.

- *Programación lineal.* Actualmente, muchas decisiones empresariales importantes se toman con base en modelos de programación lineal que utilizan cientos de variables. La industria de las aerolíneas, por ejemplo, utiliza la programación lineal para organizar los itinerarios de las tripulaciones de vuelo, monitorizar la ubicación de los aviones o planear la variada agenda de los servicios de apoyo, como las actividades operativas y de mantenimiento en las terminales aéreas.

- *Redes eléctricas.* Los ingenieros utilizan software de simulación para diseñar circuitos eléctricos y microchips, lo que implica millones de transistores. Dicho software se basa en técnicas de álgebra lineal y en sistemas de ecuaciones lineales.

WEB

Los sistemas de ecuaciones lineales constituyen el corazón del álgebra lineal, y este capítulo los utiliza para introducir, de manera sencilla y concreta, algunos de los conceptos centrales del álgebra lineal. Las secciones 1.1 y 1.2 presentan un método sistemático para resolver sistemas de ecuaciones lineales. En este libro se empleará dicho algoritmo para realizar diversos cálculos. Las secciones 1.3 y 1.4 muestran cómo un sistema de ecuaciones lineales es equivalente a una *ecuación vectorial* y a una *ecuación matricial*. Esta equivalencia reducirá problemas que implican combinaciones lineales de vectores a preguntas acerca de sistemas de ecuaciones lineales. Los conceptos fundamentales de generación, independencia lineal y transformaciones lineales, que se estudiarán en la segunda mitad de este capítulo, desempeñarán un papel esencial a lo largo del libro conforme se explore la belleza y el poder del álgebra lineal.

1.1 SISTEMAS DE ECUACIONES LINEALES

Una **ecuación lineal** en las variables $x_1, \dots, x_n$ es una ecuación que puede escribirse en la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1}$$

donde b y los **coeficientes** $a_1, \dots, a_n$ son números reales o complejos, que generalmente se conocen de antemano. El subíndice n puede ser cualquier entero positivo. En los ejemplos y ejercicios del libro, n normalmente está entre 2 y 5. En problemas de la vida real, n podría ser 50 o 5000, o incluso mayor.

Las ecuaciones

$$4x_1 - 5x_2 + 2 = x_1 \quad \text{y} \quad x_2 = 2\left(\sqrt{6} - x_1\right) + x_3$$

son lineales porque se pueden reordenar algebraicamente en la forma de la ecuación (1):

$$3x_1 - 5x_2 = -2 \quad \text{y} \quad 2x_1 + x_2 - x_3 = 2\sqrt{6}$$

Las ecuaciones

$$4x_1 - 5x_2 = x_1x_2 \quad \text{y} \quad x_2 = 2\sqrt{x_1} - 6$$

no son lineales debido a la presencia de x_1x_2 en la primera ecuación y de $\sqrt{x_1}$ en la segunda.

Un **sistema de ecuaciones lineales** (o **sistema lineal**) es una colección de una o más ecuaciones lineales que implican las mismas variables, por ejemplo, $x_1, \dots, x_n$. Un ejemplo es

$$\begin{aligned} 2x_1 - x_2 + 1.5x_3 &= 8 \\ x_1 \qquad\quad - 4x_3 &= -7 \end{aligned} \tag{2}$$

Una **solución** del sistema es una lista de números ($s_1, s_2, \ldots, s_n$) que da validez a cada ecuación cuando se utilizan los valores $s_1, \ldots, s_n$ en lugar de $x_1, \ldots, x_n$, respectivamente. Por ejemplo, (5, 6.5, 3) es una solución del sistema (2) porque al sustituir estos valores en (2) para x_1, x_2, x_3, respectivamente, las ecuaciones se simplifican a $8 = 8$ y $-7 = -7$.

El conjunto de todas las posibles soluciones se llama **conjunto solución** del sistema lineal. Se dice que dos sistemas lineales son **equivalentes** si tienen el mismo conjunto solución. Es decir, cada solución del primer sistema es una solución del segundo sistema, y cada solución del segundo sistema también es una solución del primero.

Es fácil encontrar el conjunto solución de un sistema de dos ecuaciones lineales con dos variables porque equivale a obtener la intersección de dos rectas. Un problema común es

$$\begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$$

Las gráficas de esas ecuaciones son líneas rectas, las cuales se denotan como ℓ_1 y ℓ_2. Un par de números (x_1, x_2) satisface *ambas* ecuaciones del sistema si y solo si el punto (x_1, x_2) está sobre ℓ_1 y ℓ_2. En el sistema anterior, la solución es el único punto (3, 2), lo que puede comprobarse fácilmente. Véase la figura 1.

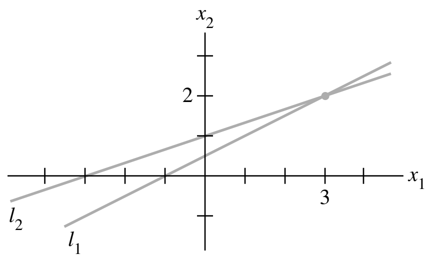


FIGURA 1 Exactamente una solución.

Desde luego, dos rectas no necesitan intersecarse en un solo punto; podrían ser paralelas, o coincidir y, así, "intersecarse" en todos los puntos de la recta. La figura 2 muestra las gráficas que corresponden a los siguientes sistemas:

$$a)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 3 \end{aligned} \qquad\qquad b)\quad \begin{aligned} x_1 - 2x_2 &= -1 \\ -x_1 + 2x_2 &= 1 \end{aligned}$$

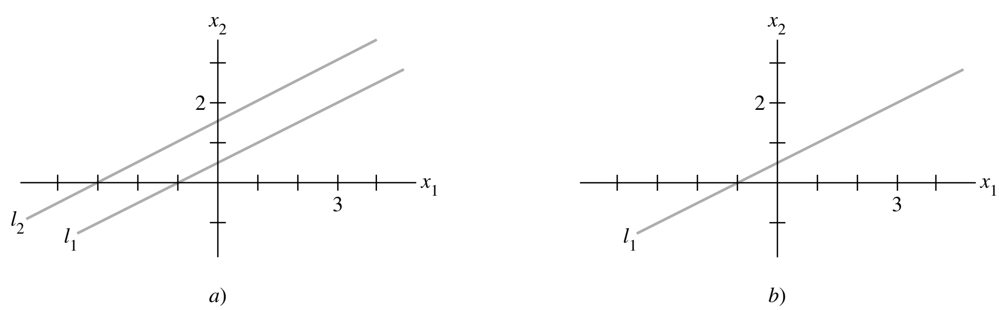


FIGURA 2 *a*) No hay solución. *b*) Número infinito de soluciones.

Las figuras 1 y 2 ilustran el siguiente hecho general acerca de los sistemas lineales, el cual se comprobará en la sección 1.2.

Un sistema de ecuaciones lineales tiene

1. ninguna solución, o
2. exactamente una solución, o
3. un número infinito de soluciones.

Se dice que un sistema de ecuaciones lineales es **consistente** si tiene una solución o un número infinito de soluciones; un sistema es **inconsistente** cuando no tiene ninguna solución.

Notación matricial

La información esencial de un sistema lineal puede registrarse de forma compacta en un arreglo rectangular llamado **matriz**. Dado el sistema

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \tag{3}$$

con los coeficientes de cada variable alineados en columnas, la matriz

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

se llama **matriz coeficiente** (o **matriz de coeficientes**) del sistema (3), y

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \tag{4}$$

se llama **matriz aumentada** del sistema. (Aquí la segunda fila contiene un cero porque la segunda ecuación podría escribirse como $0 \cdot x_1 + 2x_2 - 8x_3 = 8$). La matriz aumentada de un sistema consiste en la matriz de coeficientes con una columna adicional que contiene las constantes de los miembros derechos de las ecuaciones.

El **tamaño** de una matriz indica su número de filas y columnas. La matriz aumentada (4) tiene 3 filas y 4 columnas, por lo que es una matriz de 3×4 (que se lee "3 por 4"). Si m y n son enteros positivos, entonces una **matriz de $\boldsymbol{m \times n}$** es un arreglo rectangular de números con m filas y n columnas. (Siempre va primero el número de filas). La notación matricial simplificará los cálculos en los ejemplos que siguen.

Solución de un sistema lineal

Esta sección y la siguiente describen un algoritmo, o un procedimiento sistemático, para resolver sistemas lineales. La estrategia básica es *remplazar un sistema por otro equivalente (es decir, uno con el mismo conjunto solución) y que sea más fácil resolver*.

En general, use el término x_1 de la primera ecuación de un sistema para eliminar los términos x_1 en las ecuaciones restantes. Después, utilice el término x_2 en la segunda ecuación para eliminar los términos x_2 en las demás ecuaciones, y así sucesivamente, hasta que finalmente obtenga un sistema equivalente de ecuaciones muy sencillo.

Se utilizan tres operaciones básicas para simplificar un sistema lineal: remplazar una ecuación por la suma de esta y un múltiplo de otra ecuación, intercambiar dos ecuaciones, y multiplicar todos los términos de una ecuación por una constante distinta de cero. Después del primer ejemplo, resultará claro por qué esas tres operaciones no alteran el conjunto solución del sistema.

EJEMPLO 1 Resuelva el sistema (3).

SOLUCIÓN Aquí se muestra el procedimiento de eliminación, con y sin notación matricial, y los resultados se colocan uno al lado del otro para facilitar la comparación:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

Mantenga x_1 en la primera ecuación y elimínela en las otras ecuaciones. Para hacerlo, sume la ecuación 1 multiplicada por 4 a la ecuación 3. Después de cierta práctica, estos cálculos se podrán efectuar mentalmente:

$$\begin{array}{r r} 4\cdot[\text{ecuación 1}]: & 4x_1 - 8x_2 + 4x_3 = 0 \\ +\ [\text{ecuación 3}]: & -4x_1 + 5x_2 + 9x_3 = -9 \\ \hline [\text{nueva ecuación 3}]: & -3x_2 + 13x_3 = -9 \end{array}$$

El resultado de este cálculo se escribe en lugar de la tercera ecuación original:

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Ahora, multiplicamos la ecuación 2 por $1/2$ para así obtener 1 como coeficiente de x_2. (Este cálculo simplificará la aritmética en el siguiente paso).

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ -3x_2 + 13x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

Utilice x_2 de la ecuación 2 para eliminar $-3x_2$ en la ecuación 3. El cálculo "mental" es

$$\begin{array}{r r} 3\cdot[\text{ecuación 2}]: & 3x_2 - 12x_3 = 12 \\ +\ [\text{ecuación 3}]: & -3x_2 + 13x_3 = -9 \\ \hline [\text{nueva ecuación 3}]: & x_3 = 3 \end{array}$$

El nuevo sistema tiene forma *triangular*:[1]

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Finalmente, se desea eliminar el término $-2x_2$ de la ecuación 1, pero es más eficiente usar primero x_3 de la ecuación 3 para eliminar los términos $-4x_3$ y $+x_3$ en las ecuaciones 2 y 1. Los dos cálculos "mentales" son

$$\begin{array}{r r} 4\cdot[\text{ec. 3}]: & 4x_3 = 12 \\ +\ [\text{ec. 2}]: & x_2 - 4x_3 = 4 \\ \hline [\text{nueva ec. 2}]: & x_2 = 16 \end{array} \qquad \begin{array}{r r} -1\cdot[\text{ec. 3}]: & -x_3 = -3 \\ +\ [\text{ec. 1}]: & x_1 - 2x_2 + x_3 = 0 \\ \hline [\text{nueva ec. 1}]: & x_1 - 2x_2 = -3 \end{array}$$

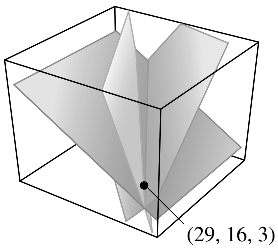
[1] En la siguiente sección se remplazará el término intuitivo *triangular* por uno más preciso.

Es conveniente combinar los resultados de esas dos operaciones:

$$\begin{aligned} x_1 - 2x_2 \qquad &= -3 \\ x_2 \qquad &= 16 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Ahora, una vez que se ha eliminado la columna que está sobre x_3 en la ecuación 3, regrese a x_2 en la ecuación 2 y utilícela para eliminar $-2x_2$ sobre ella. Gracias al trabajo previo con x_3, ahora no hay operaciones que impliquen términos con x_3. Sume dos veces la ecuación 2 a la ecuación 1 para obtener el sistema:

$$\begin{cases} x_1 \qquad\qquad = 29 \\ \qquad x_2 \qquad = 16 \\ \qquad\qquad x_3 = 3 \end{cases} \qquad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$



Cada una de las ecuaciones originales determina un plano en el espacio tridimensional. El punto (29, 16, 3) pertenece a los tres planos.

En esencia, el proceso está terminado. Se observa que la única solución del sistema original es (29, 16, 3). Sin embargo, puesto que son muchos los cálculos realizados, es recomendable comprobar los resultados. Para comprobar que (29, 16, 3) *es* una solución, sustituya esos valores en el lado izquierdo del sistema original, y calcule:

$$\begin{aligned} (29) - 2(16) + (3) &= 29 - 32 + 3 = 0 \\ 2(16) - 8(3) &= 32 - 24 = 8 \\ -4(29) + 5(16) + 9(3) &= -116 + 80 + 27 = -9 \end{aligned}$$

Los resultados concuerdan con el lado derecho del sistema original, de manera que (29, 16, 3) es una solución del sistema. ■

El ejemplo 1 muestra cómo las operaciones con las ecuaciones de un sistema lineal corresponden a las operaciones en las filas adecuadas de la matriz aumentada. Las tres operaciones básicas mencionadas con anterioridad corresponden a las siguientes operaciones en la matriz aumentada.

OPERACIONES ELEMENTALES DE FILA

1. (Remplazo) Sustituir una fila por la suma de sí misma y un múltiplo de otra fila.[2]
2. (Intercambio) Intercambiar dos filas.
3. (Escalamiento) Multiplicar todos los elementos de una fila por una constante diferente de cero.

Las operaciones de fila pueden aplicarse a cualquier matriz, no solo a las matrices aumentadas de un sistema lineal. Dos matrices son **equivalentes por filas** si existe una secuencia de operaciones elementales de fila que transforme una matriz en otra.

Es importante observar que las operaciones de fila son *reversibles*. Si dos filas se intercambian, es posible hacerlas retornar a sus posiciones originales mediante otro intercambio. Si una fila se multiplica por una constante c distinta de cero, entonces al multiplicar la nueva fila por $1/c$ se obtiene la fila original. Por último, considere una operación de remplazo que implica a dos filas —por ejemplo, las filas 1 y 2— y suponga que a la fila 2 se le suma la fila 1 multiplicada por c para producir una nueva fila 2. Para "revertir" esta operación, sume la fila 1 multiplicada por $-c$ a la nueva fila 2 para así obtener la fila 2 original. Véase los ejercicios 29 al 32 al final de esta sección.

[2] Una forma alternativa de expresar la operación de remplazo de filas es: "Sume a una fila un múltiplo de otra fila".

Por el momento, estamos interesados en las operaciones de fila sobre la matriz aumentada de un sistema de ecuaciones lineales. Suponga que un sistema se transforma en otro mediante operaciones de fila. Considerando cada tipo de operación de fila, puede verse que cualquier solución del sistema original continúa siendo una solución del nuevo sistema. A la inversa, puesto que el sistema original se puede obtener mediante operaciones de fila sobre el nuevo sistema, cada solución del nuevo sistema también es solución del sistema original. Este análisis justifica el siguiente enunciado.

Si las matrices aumentadas de dos sistemas lineales son equivalentes por filas, entonces los dos sistemas tienen el mismo conjunto solución.

A pesar de que el ejemplo 1 es largo, después de cierta práctica se desarrollará habilidad para realizar los cálculos con rapidez. En el texto y en los ejercicios de este libro, las operaciones de fila por lo general serán muy fáciles de efectuar, lo que permitirá al lector enfocarse en los conceptos subyacentes. Pero debe aprender a realizar con exactitud las operaciones de fila porque se utilizarán a lo largo del libro.

El resto de esta sección muestra cómo emplear operaciones de fila para determinar el tamaño de un conjunto solución, sin resolver completamente el sistema lineal.

Preguntas de existencia y unicidad

La sección 1.2 mostrará por qué un conjunto solución de un sistema lineal puede no contener ninguna solución, o bien, tener una solución o un número infinito de soluciones. Las respuestas a las siguientes dos preguntas determinarán la naturaleza del conjunto solución de un sistema lineal.

Para determinar qué posibilidad es verdadera para un sistema particular, nos planteamos dos preguntas.

DOS PREGUNTAS FUNDAMENTALES ACERCA DE UN SISTEMA LINEAL

1. ¿El sistema es consistente, es decir, al menos *existe* una solución?
2. Si existe una solución, ¿*solo* hay una, es decir, la solución es *única*?

Estas dos preguntas se presentarán a lo largo del libro, en diversas circunstancias. Esta sección y la siguiente le mostrarán cómo responder a esas preguntas usando operaciones de fila sobre la matriz aumentada.

EJEMPLO 2 Determine si el siguiente sistema es consistente

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

SOLUCIÓN Este es el sistema del ejemplo 1. Suponga que se han efectuado las operaciones de fila necesarias para obtener la forma triangular

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 4x_3 &= 4 \\ x_3 &= 3 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

En este punto, conocemos x_3. Si se sustituyera el valor de x_3 en la ecuación 2, entonces se podría calcular x_2 y, por lo tanto, se podría obtener x_1 de la ecuación 1. Así que existe una solución; el sistema es consistente. (En efecto, x_2 está determinada de manera unívoca por la ecuación 2 ya que x_3 sólo tiene un valor posible, y en consecuencia x_1 está determinada de forma unívoca por la ecuación 1. Por lo tanto, la solución es única). ■

EJEMPLO 3 Determine si el siguiente sistema es consistente

$$\begin{aligned} x_2 - 4x_3 &= 8 \\ 2x_1 - 3x_2 + 2x_3 &= 1 \\ 5x_1 - 8x_2 + 7x_3 &= 1 \end{aligned} \tag{5}$$

SOLUCIÓN La matriz aumentada es

$$\begin{bmatrix} 0 & 1 & -4 & 8 \\ 2 & -3 & 2 & 1 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

Para obtener una x_1 en la primera ecuación, se intercambian las filas 1 y 2:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

Para eliminar el término $5x_1$ en la tercera ecuación, a la fila 3 se suma la fila 1 multiplicada por $-5/2$:

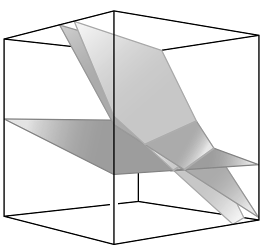
$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & -1/2 & 2 & -3/2 \end{bmatrix} \tag{6}$$

Ahora, use el término x_2 en la segunda ecuación para eliminar el término $-(1/2)x_2$ en la tercera ecuación. A la fila 3, sume la fila 2 multiplicada por $1/2$:

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \tag{7}$$

Ahora la matriz aumentada está en forma triangular. Para interpretarla correctamente, conviene regresar a la notación con ecuaciones:

$$\begin{aligned} 2x_1 - 3x_2 + 2x_3 &= 1 \\ x_2 - 4x_3 &= 8 \\ 0 &= 5/2 \end{aligned} \tag{8}$$



Este sistema es inconsistente porque no existe un punto que pertenezca a los tres planos de manera simultánea.

La ecuación $0 = 5/2$ es una forma abreviada de $0x_1 + 0x_2 + 0x_3 = 5/2$. Este sistema en forma triangular, evidentemente, tiene una contradicción inherente. No existen valores de x_1, x_2, x_3 que satisfagan la ecuación (8) porque la ecuación $0 = 5/2$ nunca es válida. Como (5) y (8) tienen el mismo conjunto solución, entonces el sistema original es inconsistente (es decir, no tiene solución). ■

Preste atención a la matriz aumentada en (7). Su última fila es característica de un sistema inconsistente en forma triangular.

NOTA NUMÉRICA

En problemas del mundo real, los sistemas de ecuaciones lineales se resuelven en computadora. Para una matriz de coeficientes cuadrada, los programas computacionales casi siempre utilizan el algoritmo de eliminación presentado aquí y en la sección 1.2, aunque ligeramente modificado para obtener mayor exactitud.

La gran mayoría de los problemas de álgebra lineal en los negocios y en la industria se resuelven con programas que emplean *aritmética de punto flotante*. Los números se representan como decimales $\pm .d_1 \cdots d_p \times 10^r$, donde r es un entero, y el número p de dígitos a la derecha del punto decimal, por lo general, está entre 8 y 16. La aritmética con dichos números normalmente es inexacta, porque el resultado debe redondearse (o truncarse) al número de dígitos almacenados. Se introduce el "error de redondeo" cuando un número como $1/3$ ingresa a la computadora, ya que su representación decimal debe ser aproximada por una cantidad finita de dígitos. Por fortuna, las inexactitudes en la aritmética de punto flotante rara vez causan problemas. Las notas numéricas en este libro ocasionalmente le advertirán sobre asuntos que deberá considerar más adelante en su carrera profesional.

PROBLEMAS DE PRÁCTICA

A lo largo del libro, es conveniente resolver los problemas de práctica antes de trabajar los ejercicios. Las soluciones se presentan después de cada conjunto de ejercicios.

1. Exprese con palabras la siguiente operación elemental de fila que debe efectuarse en el sistema para resolverlo. [En *a*) es posible más de una respuesta].

a)
$$\begin{aligned} x_1 + 4x_2 - 2x_3 + 8x_4 &= 12 \\ x_2 - 7x_3 + 2x_4 &= -4 \\ 5x_3 - x_4 &= 7 \\ x_3 + 3x_4 &= -5 \end{aligned}$$

b)
$$\begin{aligned} x_1 - 3x_2 + 5x_3 - 2x_4 &= 0 \\ x_2 + 8x_3 \qquad &= -4 \\ 2x_3 \qquad &= 3 \\ x_4 &= 1 \end{aligned}$$

2. La matriz aumentada de un sistema lineal se transformó, mediante operaciones de fila, en la forma que se indica a continuación. Determine si el sistema es consistente.

$$\begin{bmatrix} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

3. ¿$(3, 4, -2)$ es una solución para el siguiente sistema?

$$\begin{aligned} 5x_1 - x_2 + 2x_3 &= 7 \\ -2x_1 + 6x_2 + 9x_3 &= 0 \\ -7x_1 + 5x_2 - 3x_3 &= -7 \end{aligned}$$

4. ¿Para qué valores de h y k es consistente el siguiente sistema?

$$\begin{aligned} 2x_1 - x_2 &= h \\ -6x_1 + 3x_2 &= k \end{aligned}$$

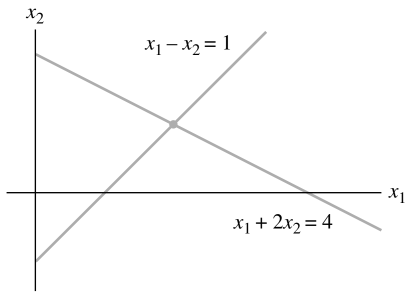
1.1 EJERCICIOS

En los ejercicios 1 a 4 resuelva cada sistema utilizando operaciones elementales de fila sobre las ecuaciones o sobre la matriz aumentada. Siga el procedimiento de eliminación sistemático explicado en esta sección.

1. $\begin{aligned} x_1 + 5x_2 &= 7 \\ -2x_1 - 7x_2 &= -5 \end{aligned}$

2. $\begin{aligned} 3x_1 + 6x_2 &= -3 \\ 5x_1 + 7x_2 &= 10 \end{aligned}$

3. Encuentre el punto (x_1, x_2) que pertenece tanto a la recta $x_1 + 2x_2 = 4$ como a la recta $x_1 - x_2 = 1$. Observe la figura.



4. Obtenga el punto de intersección de las rectas $x_1 + 2x_2 = -13$ y $3x_1 - 2x_2 = 1$.

En los ejercicios 5 y 6 considere que cada matriz es la matriz aumentada de un sistema lineal. Exprese con palabras las siguientes dos operaciones elementales de fila que deben realizarse para resolver el sistema.

5. $\begin{bmatrix} 1 & -4 & -3 & 0 & 7 \\ 0 & 1 & 4 & 0 & 6 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -5 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -6 & 4 & 0 & -1 \\ 0 & 2 & -7 & 0 & 4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 4 & 1 & 2 \end{bmatrix}$

En los ejercicios 7 a 10, la matriz aumentada de un sistema lineal se redujo mediante operaciones de fila a la forma indicada. En cada caso, continúe con las operaciones adecuadas de fila y describa el conjunto solución del sistema original.

7. $\begin{bmatrix} 1 & 7 & 3 & -4 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -5 & 4 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}$

9. $\begin{bmatrix} 1 & -1 & 0 & 0 & -5 \\ 0 & 1 & -2 & 0 & -7 \\ 0 & 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 1 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 3 & 0 & -2 & -7 \\ 0 & 1 & 0 & 3 & 6 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}$

En los ejercicios 11 a 14 resuelva los sistemas.

11. $\begin{aligned} x_2 + 5x_3 &= -4 \\ x_1 + 4x_2 + 3x_3 &= -2 \\ 2x_1 + 7x_2 + x_3 &= -2 \end{aligned}$

12. $\begin{aligned} x_1 - 5x_2 + 4x_3 &= -3 \\ 2x_1 - 7x_2 + 3x_3 &= -2 \\ -2x_1 + x_2 + 7x_3 &= -1 \end{aligned}$

13. $\begin{aligned} x_1 - 3x_3 &= 8 \\ 2x_1 + 2x_2 + 9x_3 &= 7 \\ x_2 + 5x_3 &= -2 \end{aligned}$

14. $\begin{aligned} 2x_1 - 6x_3 &= -8 \\ x_2 + 2x_3 &= 3 \\ 3x_1 + 6x_2 - 2x_3 &= -4 \end{aligned}$

En los ejercicios 15 y 16 determine si los sistemas son consistentes. No resuelva por completo dichos sistemas.

15. $\begin{aligned} x_1 - 6x_2 &= 5 \\ x_2 - 4x_3 + x_4 &= 0 \\ -x_1 + 6x_2 + x_3 + 5x_4 &= 3 \\ -x_2 + 5x_3 + 4x_4 &= 0 \end{aligned}$

16. $\begin{aligned} 2x_1 - 4x_4 &= -10 \\ 3x_2 + 3x_3 &= 0 \\ x_3 + 4x_4 &= -1 \\ -3x_1 + 2x_2 + 3x_3 + x_4 &= 5 \end{aligned}$

17. ¿Las tres rectas $2x_1 + 3x_2 = -1$, $6x_1 + 5x_2 = 0$, y $2x_1 - 5x_2 = 7$ tienen un punto común de intersección? Explique su respuesta.

18. Diga si los tres planos $2x_1 + 4x_2 + 4x_3 = 4$, $x_2 - 2x_3 = -2$, y $2x_1 + 3x_2 = 0$ tienen al menos un punto común de intersección. Explique su respuesta.

En los ejercicios 19 a 22, determine el valor o los valores de h tales que la matriz sea la matriz aumentada de un sistema lineal consistente.

19. $\begin{bmatrix} 1 & h & 4 \\ 3 & 6 & 8 \end{bmatrix}$

20. $\begin{bmatrix} 1 & h & -5 \\ 2 & -8 & 6 \end{bmatrix}$

21. $\begin{bmatrix} 1 & 4 & -2 \\ 3 & h & -6 \end{bmatrix}$

22. $\begin{bmatrix} -4 & 12 & h \\ 2 & -6 & -3 \end{bmatrix}$

En los ejercicios 23 y 24, se citan enunciados clave de esta sección, con ligeras modificaciones (pero manteniendo su validez), o se alteraron de manera que, en algunos casos, son falsos. Marque cada enunciado como verdadero o falso, y *justifique* su respuesta. (Si el enunciado es verdadero, entonces indique la ubicación aproximada donde se presenta un enunciado similar, o haga referencia a una definición o un teorema. Si la afirmación es falsa, indique la ubicación de un enunciado que se haya citado o empleado incorrectamente, o cite un ejemplo que muestre la falsedad del enunciado en todos los casos). Preguntas similares de falso/verdadero se presentarán en muchas secciones del libro.

23. *a*) Todas las operaciones elementales de fila son reversibles.

b) Una matriz de 5×6 tiene seis filas.

c) El conjunto solución de un sistema lineal que incluye a las variables $x_1,\dots, x_n$ es una lista de números $(s_1,\dots, s_n)$ que da validez a cada ecuación del sistema cuando se sustituyen los valores $s_1,\dots, s_n$ por $x_1,\dots, x_n$, respectivamente.

d) Las dos preguntas fundamentales acerca de un sistema lineal incluyen existencia y unicidad.

24. *a*) Dos matrices son equivalentes por filas si tienen el mismo número de filas.

b) En una matriz aumentada, las operaciones elementales de fila no modifican nunca el conjunto solución del sistema lineal asociado.

c) Dos sistemas lineales equivalentes pueden tener diferentes conjuntos solución.

d) Un sistema consistente de ecuaciones lineales tiene una o más soluciones.

25. Encuentre una ecuación que incluya a g, h y k, y que permita que esta matriz aumentada corresponda a un sistema consistente:

$$\begin{bmatrix} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{bmatrix}$$

26. Suponga que el sistema que aparece a continuación es consistente para todos los posibles valores de f y g. ¿Qué puede decirse acerca de los coeficientes c y d? Justifique su respuesta.

$$2x_1 + 4x_2 = f$$
$$cx_1 + dx_2 = g$$

27. Suponga que a, b, c y d son constantes tales que a es diferente de cero y el sistema que aparece a continuación es consistente para todos los posibles valores de f y g. ¿Qué podría decir acerca de los números a, b, c y d? Justifique su respuesta.

$$ax_1 + bx_2 = f$$
$$cx_1 + dx_2 = g$$

28. Construya tres diferentes matrices aumentadas para los sistemas lineales cuyo conjunto solución es $x_1 = 3$, $x_2 = -2$, $x_3 = -1$.

En los ejercicios 29 a 32, encuentre la operación elemental de fila que transforme a la primera matriz en la segunda, y luego encuentre la operación de fila inversa que transforme a la segunda matriz en la primera.

29. $\begin{bmatrix} 0 & -2 & 5 \\ 1 & 3 & -5 \\ 3 & -1 & 6 \end{bmatrix}, \begin{bmatrix} 3 & -1 & 6 \\ 1 & 3 & -5 \\ 0 & -2 & 5 \end{bmatrix}$

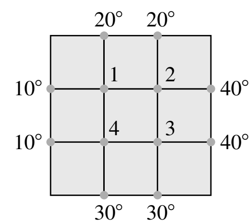
30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & 1 & -2 \end{bmatrix}$

31. $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 4 & -1 & 3 & -6 \end{bmatrix}, \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -2 & 8 \\ 0 & 7 & -1 & -6 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 4 & -12 & 7 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -5 & 0 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 0 & 15 \end{bmatrix}$

Un importante asunto en el estudio de transferencia de calor es determinar la distribución de temperatura de estado estable de una placa delgada cuando se conoce la temperatura en los bordes. Suponga que la placa que se ilustra en la figura representa una sección transversal de una viga de metal, con flujo de calor despreciable en la dirección perpendicular a la placa. Sean $T_1,\dots, T_4$ las temperaturas en los cuatro nodos interiores de la malla en la figura. La temperatura en un nodo es aproximadamente igual al promedio de las temperaturas de los cuatro nodos más cercanos, esto es, a la izquierda, arriba, a la derecha y abajo.[3] Por ejemplo,

$$T_1 = (10 + 20 + T_2 + T_4)/4, \quad \text{o} \quad 4T_1 - T_2 - T_4 = 30$$



33. Escriba un sistema de cuatro ecuaciones cuya solución dé estimaciones de las temperaturas $T_1,\dots, T_4$.

34. Resuelva el sistema de ecuaciones del ejercicio 33. [*Sugerencia:* Para conseguir rapidez en el cálculo, intercambie las filas 1 y 4 antes de iniciar las operaciones de “remplazo”].

[3] Véase Frank M. White, *Heat and Mass Transfer* (Reading, MA: Addison-Wesley Publishing, 1991), pp. 145-149.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

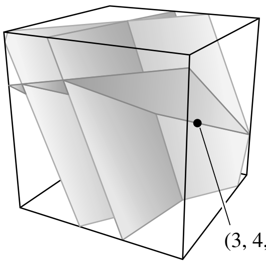
1. *a*) Para efectuar “cálculos a mano”, la mejor elección es intercambiar las ecuaciones 3 y 4. Otra posibilidad es multiplicar la ecuación 3 por $1/5$. O bien, remplazar la ecuación 4 por su suma con la fila 3 multiplicada por $-1/5$. (En cualquier caso, no utilice x_2 en la ecuación 2 para eliminar $4x_2$ en la ecuación 1. Espere hasta que se haya logrado una forma triangular y los términos x_3 y x_4 se hayan eliminado de las primeras dos ecuaciones).

b) El sistema tiene forma triangular. La simplificación ulterior inicia con x_4 en la cuarta ecuación. Utilice x_4 para eliminar todos los términos x_4 sobre ella. Ahora el paso adecuado es sumar la ecuación 4, multiplicada por 2, a la ecuación 1. (Luego, vaya a la ecuación 3 y multiplíquela por $1/2$, y después utilice la ecuación para eliminar los términos x_3 sobre ella).

2. El sistema correspondiente a la matriz aumentada es

$$\begin{aligned} x_1 + 5x_2 + 2x_3 &= -6 \\ 4x_2 - 7x_3 &= 2 \\ 5x_3 &= 0 \end{aligned}$$

La tercera ecuación hace $x_3 = 0$, el cual, desde luego, es un valor permitido para x_3. Después de eliminar los términos x_3 en las ecuaciones 1 y 2, se podría continuar para obtener valores únicos de x_1 y x_2. Así que existe una solución, y es única. Esta situación contrasta con la del ejemplo 3.



Como $(3, 4, -2)$ satisface las dos primeras ecuaciones, está sobre la recta de intersección de los primeros dos planos. Puesto que $(3, 4, -2)$ no satisface las tres ecuaciones, se concluye que no pertenece a los tres planos.

3. Es fácil comprobar si una lista específica de números es una solución. Sean $x_1 = 3$, $x_2 = 4$, y $x_3 = -2$, y encuentre que

$$\begin{aligned} 5(3) - (4) + 2(-2) &= 15 - 4 - 4 = 7 \\ -2(3) + 6(4) + 9(-2) &= -6 + 24 - 18 = 0 \\ -7(3) + 5(4) - 3(-2) &= -21 + 20 + 6 = 5 \end{aligned}$$

Aunque las primeras dos ecuaciones se satisfacen, no sucede lo mismo con la tercera, por lo que $(3, 4, -2)$ no es una solución del sistema. Observe cómo se utilizan los paréntesis cuando se realizan las sustituciones; su uso es muy recomendable para protegerse contra errores aritméticos.

4. Cuando la segunda ecuación se remplaza por su suma con la primera ecuación multiplicada por 3, el sistema se convierte en:

$$\begin{aligned} 2x_1 - x_2 &= h \\ 0 &= k + 3h \end{aligned}$$

Si $k + 3h$ es diferente de cero, el sistema no tiene solución. El sistema es consistente para cualesquiera valores de h y k que produzcan $k + 3h = 0$.

1.2 REDUCCIÓN POR FILAS Y FORMAS ESCALONADAS

En esta sección se perfecciona el método de la sección 1.1 para obtener un nuevo algoritmo de reducción por filas que permitirá analizar cualquier sistema de ecuaciones lineales.[1] Las preguntas fundamentales de existencia y unicidad planteadas en la sección 1.1 podrán responderse utilizando la primera parte del algoritmo.

El algoritmo es aplicable a cualquier matriz, sin importar si esta se considera o no como la matriz aumentada de un sistema lineal. Así, la primera parte de esta sección se ocupa de una matriz rectangular arbitraria y empieza introduciendo dos importantes clases de matrices, que incluyen a las matrices "triangulares" de la sección 1.1. En las definiciones que siguen, una fila o columna *distinta de cero* (o *no nula*) de una matriz será una fila o columna que contenga al menos un elemento diferente de cero; una **entrada principal** de una fila se refiere a la entrada o el elemento diferente de cero que se encuentra más a la izquierda (en una fila distinta de cero).

[1] El algoritmo es una variante de lo que se conoce comúnmente como *eliminación gaussiana*. Un método de eliminación similar para sistemas lineales fue utilizado por matemáticos chinos en el año 250 a. C. El proceso era desconocido en la cultura occidental hasta el siglo XIX, cuando el famoso matemático alemán, Carl Friedrich Gauss, lo descubrió. El ingeniero alemán, Wilhelm Jordan, dio a conocer el algoritmo en un libro sobre geodesia publicado en 1888.

DEFINICIÓN

Una matriz rectangular está en **forma escalonada** (o **forma escalonada por filas**) si tiene las siguientes tres propiedades:

1. Todas las diferentes de cero están arriba de las filas que solo contienen ceros.
2. Cada entrada principal de una fila está en una columna a la derecha de la entrada principal de la fila superior.
3. En una columna todas las entradas debajo de la entrada principal son ceros.

Si una matriz de forma escalonada satisface las siguientes condiciones adicionales, entonces está en **forma escalonada reducida** (o **forma escalonada reducida por filas**):

4. La entrada principal en cada fila diferente de cero es 1.
5. Cada entrada principal 1 es la única entrada distinta de cero en su columna.

Una **matriz escalonada** (o bien, una **matriz escalonada reducida**) está en forma de escalón (o en forma escalonada reducida, respectivamente). La propiedad 2 dice que las entradas principales forman un patrón escalonado (esto es, en forma de *escalera*) que avanza hacia abajo y hacia la derecha de la matriz. La propiedad 3 es una simple consecuencia de la propiedad 2, pero se incluyó para darle mayor énfasis.

Las matrices "triangulares" de la sección 1.1, tales como

$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

están en forma escalonada. De hecho, la segunda matriz está en forma escalonada reducida. A continuación se presentan más ejemplos.

EJEMPLO 1 Las siguientes matrices están en forma escalonada. Las entradas principales (■) pueden tener cualquier valor diferente de cero; las entradas con asterisco ($*$) pueden tener cualquier valor (incluyendo al cero).

$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$$

Las siguientes matrices están en forma escalonada reducida porque las entradas principales son números 1, y hay ceros abajo *y arriba* de cada entrada principal 1.

$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$ ■

Cualquier matriz distinta de cero puede **reducirse por filas** (es decir, transformarse mediante operaciones elementales de fila) para producir más de una matriz en forma escalonada, utilizando diferentes secuencias de operaciones de fila. Sin embargo, la forma escalonada reducida que se obtiene a partir de una matriz es única. El siguiente teorema se demuestra en el apéndice A al final del libro.

TEOREMA 1

Unicidad de la forma escalonada reducida

Cada matriz es equivalente por filas a una, y solo a una, matriz escalonada reducida.

Si una matriz A es equivalente por filas a una matriz escalonada U, entonces U se llama **una forma escalonada** (o una forma escalonada por filas) **de A**; si U está en forma escalonada reducida, entonces **U es la forma escalonada reducida de A**. [La mayoría de los programas de matrices y de las calculadoras con capacidades para trabajar con matrices emplean la abreviatura RREF (por las siglas de *reduced row echelon form*) para referirse a la forma escalonada reducida por filas. Algunos utilizan REF (por las siglas de *row echelon form*) para designar la forma escalonada por filas].

Posiciones pivote

Cuando las operaciones de fila sobre una matriz producen una forma escalonada, las operaciones de fila posteriores para obtener la forma escalonada reducida no cambian la posición de las entradas principales. Como la forma escalonada reducida es única, entonces *las entradas principales siempre están en las mismas posiciones en cualquier forma escalonada obtenida a partir de una matriz dada.* Esas entradas principales corresponden a los números 1 principales de la forma escalonada reducida.

DEFINICIÓN

Una **posición pivote** en una matriz A es una ubicación en A que corresponde a un 1 principal en la forma escalonada reducida de A. Una **columna pivote** es una columna de A que contiene una posición pivote.

En el ejemplo 1, los cuadrados (■) identifican las posiciones pivote. Muchos conceptos fundamentales en los primeros cuatro capítulos estarán relacionados de una u otra manera con las posiciones pivote en una matriz.

EJEMPLO 2 Reduzca por filas la matriz A que se muestra a continuación hasta la forma escalonada, y localice las columnas pivote de A.

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$

SOLUCIÓN Utilice la misma estrategia básica de la sección 1.1. La entrada superior de la columna diferente de cero más a la izquierda de la matriz es la primera posición pivote. En esta posición debe colocarse una entrada diferente de cero, o *pivote*. Una buena opción es intercambiar las filas 1 y 4 (porque los cálculos mentales en el siguiente paso no implicarán fracciones).

$$\begin{array}{c} \text{Pivote} \\ \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \\ \text{Columna pivote} \end{array}$$

Cree ceros debajo del pivote, 1, sumando múltiplos de la primera fila a las filas inferiores, para así obtener la matriz (1) que se muestra a continuación. La posición pivote en la segunda fila debe estar tan a la izquierda como sea posible, es decir, en la segunda columna. Se elige el 2 en esta posición como el siguiente pivote.

$$\begin{array}{c} \text{Pivote} \\ \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \\ \text{Siguiente columna pivote} \end{array} \tag{1}$$

Sume la fila 2 multiplicada por $-5/2$ a la fila 3, y sume la fila 2 multiplicada por $3/2$ a la fila 4.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix} \tag{2}$$

La matriz en (2) es diferente de las que se incluyen en la sección 1.1. ¡No hay manera de crear una entrada principal en la columna 3! (No podemos emplear las filas 1 o 2 porque, al hacerlo, se destruiría el arreglo escalonado de las entradas principales ya obtenidas). Sin embargo, si se intercambian las filas 3 y 4, se puede obtener una entrada principal en la columna 4.

Pivote

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{Forma general:} \quad \begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columnas pivote

La matriz está en forma escalonada y así revela que las columnas 1, 2 y 4 de A son columnas pivote.

Posiciones pivote

$$A = \begin{bmatrix} \boxed{0} & -3 & -6 & 4 & 9 \\ -1 & \boxed{-2} & -1 & 3 & 1 \\ -2 & -3 & 0 & \boxed{3} & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \tag{3}$$

Columnas pivote ■

Un **pivote**, como el que se muestra en el ejemplo 2, es un número distinto de cero en una posición pivote que se utiliza conforme se necesite crear ceros mediante operaciones de fila. Los pivotes en el ejemplo 2 fueron 1, 2 y -5. Observe que esos números no son los mismos que los elementos reales de A en las posiciones pivote indicadas en (3).

Con el ejemplo 2 como guía, es posible describir un procedimiento eficiente para transformar una matriz en una matriz escalonada o escalonada reducida. El estudio cuidadoso y el dominio de este procedimiento rendirán valiosos frutos en este curso.

Algoritmo de reducción por filas

El algoritmo que sigue consta de cuatro pasos y produce una matriz en forma escalonada. Un quinto paso da por resultado una matriz en forma escalonada reducida. Demostraremos este algoritmo con un ejemplo.

EJEMPLO 3 Aplique operaciones elementales de fila para transformar la siguiente matriz a la forma escalonada y, luego, a la forma escalonada reducida:

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

SOLUCIÓN

PASO 1

Se inicia con la columna diferente de cero del extremo izquierdo. Esta es una columna pivote. La posición pivote se ubica en la parte superior.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Columna pivote

PASO 2

Seleccione como pivote una entrada diferente de cero en la columna pivote. Si es necesario, intercambie filas para mover esta entrada a la posición pivote.

Intercambie a las filas 1 y 3. (O bien, también se podrían intercambiar las filas 1 y 2).

Pivote

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

PASO 3

Utilice operaciones de remplazo de filas para crear ceros en todas las posiciones ubicadas debajo del pivote.

Como paso preliminar, se podría dividir la fila superior entre el pivote, 3. Pero con dos números 3 en la columna 1, esto es tan fácil como sumar la fila 1 multiplicada por -1 a la fila 2.

Pivote

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

PASO 4

Cubra (o ignore) la fila que contiene la posición pivote y cubra todas las filas, si las hay, por encima de esta. Aplique los pasos 1 a 3 a la submatriz restante. Repita el proceso hasta que no haya filas diferentes de cero por modificar.

Con la fila 1 cubierta, el paso 1 muestra que la columna 2 es la próxima columna pivote; para el paso 2, seleccione como pivote la entrada "superior" en esa columna.

Pivote

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

Siguiente columna pivote

En el paso 3, se podría insertar un paso adicional de dividir la fila "superior" de la submatriz entre el pivote, 2. En vez de ello, se suma la fila "superior" multiplicada por $-3/2$ a la fila de abajo. Esto produce

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Para el paso 4, cuando se cubre la fila que contiene la segunda posición pivote, se obtiene una nueva submatriz con una sola fila:

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

Pivote

Los pasos 1 a 3 no necesitan aplicarse para esta submatriz, pues ya se ha alcanzado una forma escalonada para la matriz completa. Si se desea la forma escalonada reducida, se efectúa un paso más.

PASO 5

Empezando con la posición pivote del extremo derecho y trabajando hacia arriba y hacia la izquierda, genere ceros arriba de cada pivote. Si un pivote no es 1, conviértalo en 1 mediante una operación de escalamiento.

El pivote del extremo derecho está en la fila 3. Genere ceros sobre él, sumando múltiplos adecuados de la fila 3 a las filas 1 y 2.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \leftarrow \text{Fila 1} + (-6)\cdot \text{fila 3} \\ \leftarrow \text{Fila 2} + (-2)\cdot \text{fila 3} \\ \end{array}$$

El siguiente pivote se encuentra en la fila 2. Se escala esta fila dividiéndola entre el pivote.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \\ \leftarrow \text{Fila escalada por } \frac{1}{2} \\ \end{array}$$

Cree un cero en la columna 2 sumando la fila 2 multiplicada por 9 a la fila 1.

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \leftarrow \text{Fila 1} + (9)\cdot \text{fila 2} \\ \\ \end{array}$$

Finalmente, escale la fila 1 dividiéndola entre el pivote, 3.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{array}{l} \leftarrow \text{Fila escalada por } \frac{1}{3} \\ \\ \end{array}$$

Esta es la forma escalonada reducida de la matriz original. ■

La combinación de los pasos 1 a 4 se conoce como **fase progresiva** del algoritmo de reducción por filas. El paso 5, que produce la única forma escalonada reducida, se conoce como **fase regresiva**.

NOTA NUMÉRICA

En el paso 2 que se describió antes, un programa computacional por lo general selecciona como pivote a la entrada en una columna que tenga el mayor valor absoluto. Esta estrategia, llamada **pivoteo parcial**, se utiliza porque reduce los errores de redondeo en los diversos cálculos.

Soluciones de sistemas lineales

El algoritmo de reducción por filas conduce directamente a una descripción explícita del conjunto solución de un sistema lineal cuando se aplica a la matriz aumentada del sistema.

Suponga, por ejemplo, que la matriz aumentada de un sistema lineal se transformó a la forma escalonada *reducida* equivalente

$$\begin{bmatrix} 1 & 0 & -5 & 1 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Existen tres variables porque la matriz aumentada tiene cuatro columnas. El sistema de ecuaciones asociado es

$$\begin{aligned} x_1 \qquad - 5x_3 &= 1 \\ x_2 + \quad x_3 &= 4 \\ 0 \quad &= 0 \end{aligned} \tag{4}$$

Las variables x_1 y x_2 correspondientes a las columnas pivote se conocen como **variables básicas**[2]. La otra variable, x_3, se denomina **variable libre**.

Siempre que un sistema es consistente, como en (4), el conjunto solución se puede describir explícitamente al despejar en el sistema de ecuaciones *reducido* las variables básicas en términos de las variables libres. Esta operación es posible porque la forma escalonada reducida coloca a cada variable básica en una y solo una ecuación. En (4), despeje x_1 de la primera ecuación y x_2 de la segunda. (Ignore la tercera ecuación, ya que no ofrece restricciones sobre las variables).

$$\begin{cases} x_1 = 1 + 5x_3 \\ x_2 = 4 - x_3 \\ x_3 \text{ es libre} \end{cases} \tag{5}$$

El enunciado "x_3 es libre" significa que existe libertad de elegir cualquier valor para x_3. Una vez hecho esto, las fórmulas en (5) determinan los valores de x_1 y x_2. Por ejemplo, cuando $x_3 = 0$, la solución es (1, 4, 0); cuando $x_3 = 1$, la solución es (6, 3, 1). *Cada asignación diferente de x_3 determina una solución (distinta) del sistema, y cada solución del sistema está determinada por una asignación de x_3.*

EJEMPLO 4 Encuentre la solución general del sistema lineal cuya matriz aumentada se redujo a

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

SOLUCIÓN La matriz está en forma escalonada, pero se desea la forma escalonada reducida antes de despejar las variables. El siguiente paso es completar la reducción por filas. El símbolo $\sim$ antes de una matriz indica que esta es equivalente por filas a la matriz anterior.

$$\begin{bmatrix} 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 2 & -8 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 2 & -8 & 0 & 10 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 6 & 2 & -5 & 0 & 10 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 5 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{bmatrix}$$

---

[2] Algunos libros utilizan el término *variables principales*, ya que corresponden a las columnas que contienen entradas principales.

Existen cinco variables porque la matriz aumentada tiene seis columnas. Ahora el sistema asociado es

$$\begin{aligned} x_1 + 6x_2 \quad\quad + 3x_4 \quad &= 0 \\ x_3 - 4x_4 \quad &= 5 \\ x_5 &= 7 \end{aligned} \tag{6}$$

Las columnas pivote de la matriz son 1, 3 y 5, así que las variables básicas son x_1, x_3 y x_5. Las variables restantes, x_2 y x_4, deben ser libres. Se despejan las variables básicas para obtener la solución general:

$$\begin{cases} x_1 = -6x_2 - 3x_4 \\ x_2 \text{ es libre} \\ x_3 = 5 + 4x_4 \\ x_4 \text{ es libre} \\ x_5 = 7 \end{cases} \tag{7}$$

Observe que el valor de x_5 ya estaba establecido por la tercera ecuación del sistema (6). ■

Descripciones paramétricas de conjuntos solución

Las descripciones en (5) y (7) son *descripciones paramétricas* de conjuntos solución en los cuales las variables libres actúan como parámetros. *Resolver un sistema* significa encontrar una descripción paramétrica del conjunto solución o determinar que el conjunto solución está vacío.

Siempre que un sistema sea consistente y tenga variables libres, el conjunto solución tendrá muchas descripciones paramétricas. Por ejemplo, en el sistema (4), se puede sumar la ecuación 2 multiplicada por 5 a la ecuación 1 para obtener el sistema equivalente

$$\begin{aligned} x_1 + 5x_2 \quad\quad &= 21 \\ x_2 + x_3 &= 4 \end{aligned}$$

Se podría tratar a x_2 como un parámetro y despejar x_1 y x_3 en términos de x_2, y se tendría una descripción exacta del conjunto solución. Sin embargo, para ser consistentes, se establece la convención (arbitraria) de utilizar siempre las variables libres como parámetros para describir un conjunto solución. (La sección de respuestas al final del libro también refleja esta convención).

Cuando un sistema es inconsistente, el conjunto solución es un conjunto vacío, aun cuando el sistema tenga variables libres. En este caso, el conjunto solución *no* tiene representación paramétrica.

Sustitución regresiva

Considere el siguiente sistema, cuya matriz aumentada está en forma escalonada, pero *no* en forma escalonada reducida:

$$\begin{aligned} x_1 - 7x_2 + 2x_3 - 5x_4 + 8x_5 &= 10 \\ x_2 - 3x_3 + 3x_4 + x_5 &= -5 \\ x_4 - x_5 &= 4 \end{aligned}$$

Un programa computacional resolvería este sistema mediante sustitución regresiva, en vez de calcular la forma escalonada reducida. Es decir, el programa despejaría x_4 de la ecuación 3 en términos de x_5, y sustituiría la expresión para x_4 en la ecuación 2; luego, despejaría x_2 de esta última, sustituiría las expresiones para x_2 y x_4 en la ecuación 1 y despejaría x_1.

Nuestro formato matricial para la fase regresiva de reducción por filas, el cual produce la forma escalonada reducida, tiene el mismo número de operaciones aritméticas que la sustitución regresiva. Pero la disciplina del formato matricial reduce de forma sustancial los errores posibles en los cálculos a mano. ¡La mejor estrategia es utilizar solamente la forma escalonada *reducida* para resolver un sistema! La *Guía de estudio* que acompaña a este libro ofrece varias sugerencias útiles para efectuar operaciones de fila de manera exacta y rápida.

NOTA NUMÉRICA

En general, la fase progresiva de reducción por filas es más larga que la fase regresiva. Por lo regular, un algoritmo para resolver un sistema se mide en *flops* (u operaciones de punto flotante). Un **flop** es una operación aritmética ($+$, $-$, $*$, $/$) que se realiza sobre dos números reales de punto flotante.[3] Para una matriz de $n \times (n + 1)$, la reducción a la forma escalonada puede requerir $2n^3/3 + n^2/2 - 7n/6$ flops (que es aproximadamente $2n^3/3$ flops cuando n es moderadamente grande, por ejemplo, $n \geq 30$). En contraste, una reducción adicional a la forma escalonada reducida necesita, a lo sumo, n^2 flops.

Preguntas de existencia y unicidad

Aunque una forma escalonada no reducida es una herramienta poco eficiente para resolver un sistema, esta forma es justamente el medio correcto para responder las dos preguntas fundamentales planteadas en la sección 1.1.

EJEMPLO 5 Determine la existencia y unicidad de las soluciones del sistema

$$\begin{aligned} 3x_2 - 6x_3 + 6x_4 + 4x_5 &= -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 &= 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 &= 15 \end{aligned}$$

SOLUCIÓN La matriz aumentada de este sistema se redujo por filas en el ejemplo 3 a:

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \tag{8}$$

Las variables básicas son x_1, x_2 y x_5; las variables libres son x_3 y x_4. No existe ninguna ecuación del tipo $0 = 1$ que indique la inconsistencia del sistema, así que se podría emplear sustitución regresiva para encontrar una solución. Pero en (8) ya es evidente la *existencia* de una solución. Además, la solución *no es única* porque hay variables libres. Cada diferente asignación de x_3 y x_4 determina una solución distinta. Por lo tanto, el sistema tiene un número infinito de soluciones. ■

Cuando un sistema está en forma escalonada y no contiene ecuaciones del tipo $0 = b$, con b diferente de cero, entonces cada ecuación no nula tiene una variable básica con un coeficiente distinto de cero. Es posible que las variables básicas estén completamente determinadas (sin variables libres) o que al menos una de las variables básicas pueda expresarse en términos de una o más variables libres. En el primer caso, existe una solución única; en el último caso, hay infinidad de soluciones (una para cada asignación de valores a las variables libres).

[3] Tradicionalmente, un *flop* era solo una multiplicación o división, ya que la suma y la resta tomaban mucho menos tiempo y podían ignorarse. Ahora se prefiere la definición de *flop* que aquí se presenta, debido a los avances en la arquitectura computacional. Véase Golub y Van Loan, *Matrix Computations*, 2a. ed. (Baltimore: The Johns Hopkins Press, 1989), pp. 19-20.

Esas observaciones justifican el siguiente teorema.

TEOREMA 2

Teorema de existencia y unicidad

Un sistema lineal es consistente si y solo si la columna más a la derecha de la matriz aumentada *no* es una columna pivote, es decir, si y solo si una forma escalonada de la matriz aumentada *no* tiene filas del tipo

$$[0 \ \cdots \ 0 \quad b] \quad \text{con } b \text{ diferente de cero}$$

Si un sistema lineal es consistente, entonces el conjunto solución contiene: **i.** una única solución, cuando no existen variables libres, o **ii.** una infinidad de soluciones, cuando hay al menos una variable libre.

El siguiente procedimiento indica cómo encontrar y describir todas las soluciones de un sistema lineal.

USO DE LA REDUCCIÓN POR FILAS PARA RESOLVER UN SISTEMA LINEAL

1. Escriba la matriz aumentada del sistema.
2. Emplee el algoritmo de reducción por filas para obtener una matriz aumentada equivalente en forma escalonada. Determine si el sistema es consistente o no. Si no existe solución, deténgase; en caso contrario, continúe con el siguiente paso.
3. Prosiga con la reducción por filas para obtener la forma escalonada reducida.
4. Escriba el sistema de ecuaciones correspondiente a la matriz obtenida en el paso 3.
5. Rescriba cada ecuación no nula del paso 4 de manera que su única variable básica se exprese en términos de cualquiera de las variables libres que aparecen en la ecuación.

PROBLEMAS DE PRÁCTICA

1. Encuentre la solución general del sistema lineal cuya matriz aumentada es

$$\begin{bmatrix} 1 & -3 & -5 & 0 \\ 0 & 1 & 1 & 3 \end{bmatrix}$$

2. Obtenga la solución general del sistema

$$\begin{aligned} x_1 - 2x_2 - x_3 + 3x_4 &= 0 \\ -2x_1 + 4x_2 + 5x_3 - 5x_4 &= 3 \\ 3x_1 - 6x_2 - 6x_3 + 8x_4 &= 2 \end{aligned}$$

1.2 EJERCICIOS

En los ejercicios 1 y 2, determine cuáles matrices están en forma escalonada reducida y cuáles se encuentran solo en forma escalonada.

1. *a*) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$ *b*) $\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ *d*) $\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 4 \end{bmatrix}$

2. *a*) $\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ *b*) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

d) $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 3 y 4 aplique la reducción por filas a las matrices para llevarlas a la forma escalonada reducida. En las matrices original y final encierre en un círculo las posiciones pivote, e indique las columnas pivote.

3. $\begin{bmatrix} 1 & 2 & 4 & 8 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$ **4.** $\begin{bmatrix} 1 & 2 & 4 & 5 \\ 2 & 4 & 5 & 4 \\ 4 & 5 & 4 & 2 \end{bmatrix}$

5. Describa las posibles formas escalonadas de una matriz de 2×2 diferente de cero**.** Utilice los símbolos $\blacksquare$, $*$ y 0, como en la primera parte del ejemplo 1.

6. Repita el ejercicio 5 para una matriz de 3×2 diferente de cero.

Encuentre las soluciones generales de los sistemas cuyas matrices aumentadas se presentan en los ejercicios 7 a 14.

7. $\begin{bmatrix} 1 & 3 & 4 & 7 \\ 3 & 9 & 7 & 6 \end{bmatrix}$ **8.** $\begin{bmatrix} 1 & -3 & 0 & -5 \\ -3 & 7 & 0 & 9 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 1 & -2 & 3 \\ 1 & -3 & 4 & -6 \end{bmatrix}$ **10.** $\begin{bmatrix} 1 & -2 & -1 & 4 \\ -2 & 4 & -5 & 6 \end{bmatrix}$

11. $\begin{bmatrix} 3 & -2 & 4 & 0 \\ 9 & -6 & 12 & 0 \\ 6 & -4 & 8 & 0 \end{bmatrix}$ **12.** $\begin{bmatrix} 1 & 0 & -9 & 0 & 4 \\ 0 & 1 & 3 & 0 & -1 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & -3 & 0 & -1 & 0 & -2 \\ 0 & 1 & 0 & 0 & -4 & 1 \\ 0 & 0 & 0 & 1 & 9 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 0 & -5 & 0 & -8 & 3 \\ 0 & 1 & 4 & -1 & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Los ejercicios 15 y 16 emplean la notación del ejemplo 1 para matrices en forma escalonada. Suponga que cada matriz representa la matriz aumentada para un sistema de ecuaciones lineales. En cada caso, determine si el sistema es consistente. De ser así, determine si la solución es única.

15. *a*) $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 0 & \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & 0 \end{bmatrix}$

16. *a*) $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \end{bmatrix}$

b) $\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare & * \end{bmatrix}$

En los ejercicios 17 y 18, determine el valor o los valores de h tales que la matriz sea la matriz aumentada de un sistema lineal consistente.

17. $\begin{bmatrix} 1 & -1 & 4 \\ -2 & 3 & h \end{bmatrix}$ **18.** $\begin{bmatrix} 1 & -3 & 1 \\ h & 6 & -2 \end{bmatrix}$

En los ejercicios 19 y 20, asigne valores para h y k de manera que el sistema *a*) no tenga solución, *b*) tenga solución única, y *c*) tenga muchas soluciones. Dé respuestas por separado para cada inciso.

19. $\begin{aligned} x_1 + hx_2 &= 2 \\ 4x_1 + 8x_2 &= k \end{aligned}$ **20.** $\begin{aligned} x_1 - 3x_2 &= 1 \\ 2x_1 + hx_2 &= k \end{aligned}$

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique cada respuesta.[4]

21. *a*) En algunos casos, una matriz se puede reducir por filas a más de una matriz en forma escalonada reducida, mediante diferentes secuencias de operaciones de fila.

b) El algoritmo de reducción por filas solamente se aplica a matrices aumentadas para un sistema lineal.

c) En un sistema lineal una variable básica es una variable que corresponde a una columna pivote en la matriz de coeficientes.

d) Encontrar una descripción paramétrica del conjunto solución de un sistema lineal es lo mismo que *resolver* el sistema.

e) Si una fila en una forma escalonada de una matriz aumentada es [0 0 0 5 0], entonces el sistema lineal asociado es inconsistente.

22. *a*) La forma escalonada reducida de una matriz es única.

b) Si cada columna de una matriz aumentada contiene un pivote, entonces el sistema correspondiente es consistente.

c) Las posiciones pivote en una matriz dependen de si se utilizan o no intercambios de filas en el proceso de reducción por filas.

d) Una solución general de un sistema es una descripción explícita de todas las soluciones del sistema.

e) Si un sistema tiene variables libres, entonces el conjunto solución contiene muchas soluciones.

23. Suponga que la matriz coeficiente de un sistema lineal de cuatro ecuaciones con cuatro variables tiene un pivote en cada columna. Explique por qué el sistema tiene solución única.

24. Suponga que un sistema de ecuaciones lineales tiene una matriz *aumentada* de 3×5 cuya quinta columna no es una columna pivote. ¿El sistema es consistente? ¿Por qué?

[4] Preguntas de verdadero/falso de este tipo se presentarán en muchas secciones. Antes de los ejercicios 23 y 24 de la sección 1.1, se describieron algunos métodos para justificar las respuestas.

25. Suponga que la matriz coeficiente de un sistema de ecuaciones lineales tiene una posición pivote en cada fila. Explique por qué el sistema es consistente.

26. Suponga que una matriz de *coeficientes* de 3×5 para un sistema tiene tres columnas pivote. ¿El sistema es consistente? ¿Por qué?

27. Restructure la última frase del teorema 2 empleando el concepto de columnas pivote: "Si un sistema lineal es consistente, entonces la solución es única si y solo si __________".

28. En una matriz aumentada, ¿qué se necesita saber acerca de las columnas pivote para determinar que el sistema lineal es consistente y tiene una solución única?

29. Un sistema de ecuaciones lineales con menos ecuaciones que incógnitas se conoce como *sistema subdeterminado*. ¿Tal sistema puede tener una solución única? Explique su respuesta.

30. Dé un ejemplo de un sistema subdeterminado inconsistente de dos ecuaciones con tres incógnitas.

31. Un sistema de ecuaciones lineales con más ecuaciones que incógnitas se llama *sistema sobredeterminado*. ¿Tal sistema puede ser consistente? Ilustre su respuesta con un sistema específico de tres ecuaciones con dos incógnitas.

32. Considere que una matriz de $n \times (n + 1)$ se simplifica por filas a su forma escalonada reducida. Aproximadamente, ¿qué fracción del número total de operaciones (flops) está implicada en la fase regresiva de la reducción cuando $n = 20$? ¿Y cuando $n = 200$?

Suponga que los datos experimentales están representados por un conjunto de puntos en el plano. Un **polinomio de interpolación** para los datos es un polinomio cuya gráfica pasa por todos los puntos. En el ámbito científico, dicho polinomio se utiliza, por ejemplo, para estimar valores entre los puntos de datos conocidos. Otro uso es en la creación de curvas para imágenes gráficas en el monitor de las computadoras. Un método para construir un polinomio de interpolación consiste en resolver un sistema de ecuaciones lineales.

WEB

33. Encuentre el polinomio de interpolación $p(t) = a_0 + a_1 t + a_2 t^2$ para los datos (1, 6), (2, 15), (3, 28). Es decir, determine a_0, a_1 y a_2 tales que

$$\begin{aligned} a_0 + a_1(1) + a_2(1)^2 &= 6 \\ a_0 + a_1(2) + a_2(2)^2 &= 15 \\ a_0 + a_1(3) + a_2(3)^2 &= 28 \end{aligned}$$

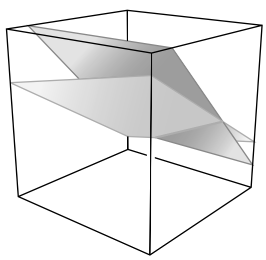
34. [**M**] En un experimento de túnel de viento, la fuerza sobre un proyectil debido a la resistencia del aire se midió a diferentes velocidades:

Velocidad (100 ft/seg)	0	2	4	6	8	10
Fuerza (100 lb)	0	2.90	14.8	39.6	74.3	119

Encuentre un polinomio de interpolación para estos datos y estime la fuerza sobre el proyectil si este viaja a 750 ft/seg. Utilice $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5$. ¿Qué ocurre si usted intenta emplear un polinomio de menor grado que 5? (Por ejemplo, intente un polinomio cúbico).[5]

[5] Los ejercicios marcados con el símbolo [**M**] deben trabajarse con la ayuda de un "programa de matrices" (esto es, un programa computacional, como **M**ATLAB®, **M**aple™, **M**athematica®, **M**athCad® o Derive™, o una calculadora programable con capacidades matriciales, como las fabricadas por Texas Instruments o Hewlett-Packard).

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA



La solución general del sistema de ecuaciones es la recta de intersección de los dos planos.

1. La forma escalonada reducida de la matriz aumentada y el sistema correspondiente son

$$\begin{bmatrix} 1 & 0 & -2 & 9 \\ 0 & 1 & 1 & 3 \end{bmatrix} \quad \text{y} \quad \begin{aligned} x_1 \quad - 2x_3 &= 9 \\ x_2 + \ x_3 &= 3 \end{aligned}$$

Las variables básicas son x_1 y x_2, y la solución general es

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 \text{ es libre} \end{cases}$$

Nota: Es esencial que la solución general describa a cada variable, con cualquier parámetro claramente definido. El siguiente enunciado *no* describe la solución:

$$\begin{cases} x_1 = 9 + 2x_3 \\ x_2 = 3 - x_3 \\ x_3 = 3 - x_2 \quad \text{Solución incorrecta} \end{cases}$$

Esta descripción implica que *tanto* x_2 *como* x_3 son libres, lo que desde luego no es el caso.

2. La matriz aumentada del sistema se reduce por filas:

$$\begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ -2 & 4 & 5 & -5 & 3 \\ 3 & -6 & -6 & 8 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & -3 & -1 & 2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -2 & -1 & 3 & 0 \\ 0 & 0 & 3 & 1 & 3 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Esta matriz escalonada indica que el sistema es *inconsistente*, porque su columna del extremo derecho es una columna pivote; la tercera fila corresponde a la ecuación $0 = 5$. No hay necesidad de realizar más operaciones de fila. Observe que en este problema es irrelevante la presencia de las variables libres porque el sistema es inconsistente.

1.3 ECUACIONES VECTORIALES

Importantes propiedades de sistemas lineales se pueden describir mediante el concepto y la notación de vectores. Esta sección relaciona ecuaciones vectoriales con sistemas ordinarios de ecuaciones. El término *vector* aparece en una variedad de contextos matemáticos y físicos, los cuales se analizarán en el capítulo 4, "Espacios vectoriales". Por ahora, *vector* significará una *lista ordenada de números*. Esta idea sencilla permite realizar, de manera rápida, interesantes e importantes aplicaciones.

Vectores en $\mathbb{R}^2$

Una matriz con una sola columna es un **vector columna**, o simplemente un **vector**. Ejemplos de vectores con dos entradas son

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} .2 \\ .3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

donde w_1 y w_2 son números reales. $\mathbb{R}^2$ (léase "erre dos") denota el conjunto de todos los vectores con dos entradas. La $\mathbb{R}$ representa los números reales que aparecen como entradas en los vectores, y el exponente 2 indica que cada vector contiene dos entradas.[1]

Dos vectores en $\mathbb{R}^2$ son **iguales** si y solo si sus entradas correspondientes son iguales. Así, $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ y $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ *no* son iguales, porque los vectores en $\mathbb{R}^2$ son *pares ordenados* de números reales.

Dados dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^2$, su **suma** es el vector $\mathbf{u} + \mathbf{v}$, que se obtiene al sumar las entradas correspondientes de $\mathbf{u}$ y $\mathbf{v}$. Por ejemplo,

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Considerando un vector $\mathbf{u}$ y un número real c, el **múltiplo escalar** de $\mathbf{u}$ por c es el vector $c\mathbf{u}$, que se obtiene al multiplicar por c cada entrada en $\mathbf{u}$. Por ejemplo,

$$\text{si} \quad \mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{y} \quad c = 5, \qquad \text{entonces} \quad c\mathbf{u} = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$$

[1] La mayor parte del libro se refiere a vectores y matrices que solo tienen entradas reales. Sin embargo, todas las definiciones y los teoremas en los capítulos 1 a 5, y en la mayor parte del resto del libro, conservan su validez si las entradas son números complejos. Vectores y matrices complejos surgen de manera natural en áreas como física e ingeniería eléctrica.

El número c en $c\mathbf{u}$ se llama **escalar**; se escribe en cursivas, y no en negritas, para así distinguirlo del vector **u**.

Es posible combinar las operaciones de multiplicación por un escalar y suma vectorial, como en el siguiente ejemplo.

EJEMPLO 1 A partir de $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, encuentre $4\mathbf{u}$, $(-3)\mathbf{v}$, y $4\mathbf{u} + (-3)\mathbf{v}$.

SOLUCIÓN

$$4\mathbf{u} = \begin{bmatrix} 4 \\ -8 \end{bmatrix}, \qquad (-3)\mathbf{v} = \begin{bmatrix} -6 \\ 15 \end{bmatrix}$$

y

$$4\mathbf{u} + (-3)\mathbf{v} = \begin{bmatrix} 4 \\ -8 \end{bmatrix} + \begin{bmatrix} -6 \\ 15 \end{bmatrix} = \begin{bmatrix} -2 \\ 7 \end{bmatrix}$$ ■

Algunas veces, por conveniencia (y también para ahorrar espacio), en este libro se denota un vector columna como $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ en la forma $(3, -1)$. En este caso, los paréntesis y la coma distinguen al vector $(3, -1)$ de la matriz fila 1×2 $[3 \; -1]$, que se representa con corchetes y sin coma. Así,

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq \begin{bmatrix} 3 & -1 \end{bmatrix}$$

porque las matrices tienen diferentes formas, aunque las entradas sean iguales.

Descripciones geométricas de $\mathbb{R}^2$

Considere un sistema de coordenadas rectangulares en el plano. Como cada punto en el plano está determinado por un par ordenado de números, *es posible identificar un punto geométrico* (a, b) *con el vector columna* $\begin{bmatrix} a \\ b \end{bmatrix}$. Así, puede considerarse a $\mathbb{R}^2$ como el conjunto de todos los puntos del plano. Véase la figura 1.

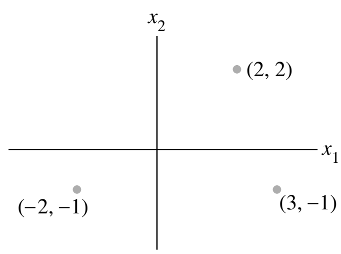


FIGURA 1 Vectores como puntos.

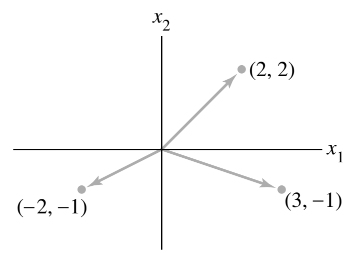


FIGURA 2 Vectores con flechas.

La visualización geométrica de un vector como $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ se facilita al incluir una flecha (un segmento de recta dirigido) desde el origen $(0, 0)$ al punto $(3, -1)$, como en la figura 2. En este caso, los puntos individuales a lo largo de la flecha carecen de significado especial.[2]

La suma de dos vectores tiene una útil representación geométrica. La siguiente regla puede verificarse mediante geometría analítica.

[2] En física, las flechas representan fuerzas y, por lo general, son libres para moverse en el espacio. En la sección 4.1 se analizará esta interpretación de vectores.

Regla del paralelogramo para la adición

Si $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^2$ se representan como puntos en el plano, entonces $\mathbf{u} + \mathbf{v}$ corresponde a un cuarto vértice del paralelogramo cuyos otros vértices son $\mathbf{u}$, $\mathbf{0}$ y $\mathbf{v}$. Véase la figura 3.

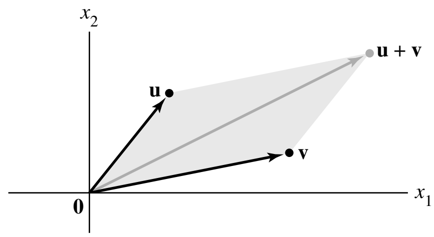


FIGURA 3 Regla del paralelogramo.

EJEMPLO 2 Los vectores $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$ y $\mathbf{u} + \mathbf{v} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ se muestran en la figura 4. ■

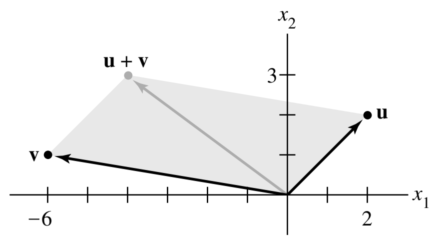
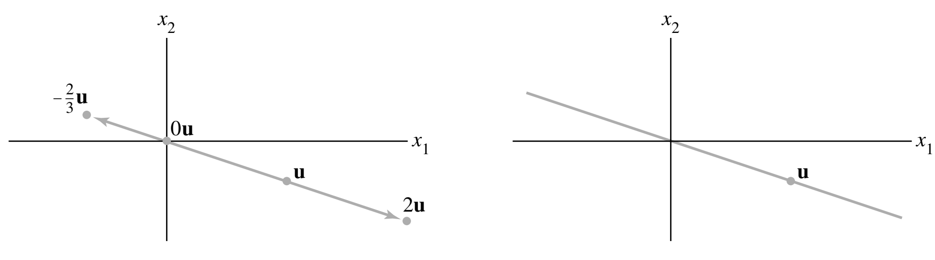


FIGURA 4

El siguiente ejemplo ilustra el hecho de que el conjunto de todos los múltiplos escalares de un vector diferente de cero (también llamado no nulo), fijo, es una recta que pasa por el origen, (0, 0).

EJEMPLO 3 Sea $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. En una gráfica muestre los vectores $\mathbf{u}$, $2\mathbf{u}$ y $-\frac{2}{3}\mathbf{u}$.

SOLUCIÓN Observe la figura 5, donde se indican $\mathbf{u}$, $2\mathbf{u} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$ y $-\frac{2}{3}\mathbf{u} = \begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$. La flecha para $2\mathbf{u}$ es el doble de largo que la flecha para $\mathbf{u}$, y las flechas apuntan en el mismo sentido. La flecha para $-\frac{2}{3}\mathbf{u}$ es dos tercios de la longitud de la flecha para $\mathbf{u}$, y las flechas apuntan en sentidos opuestos. En general, la longitud de la flecha para $c\mathbf{u}$ es $|c|$ veces la



Múltiplos típicos de $\mathbf{u}$

Conjunto de todos los múltiplos de $\mathbf{u}$

FIGURA 5

longitud de la flecha para **u**. [Recuerde que la longitud del segmento de línea de (0, 0) a (a, b) es $\sqrt{a^2 + b^2}$. Esto se analizará en el capítulo 6]. ■

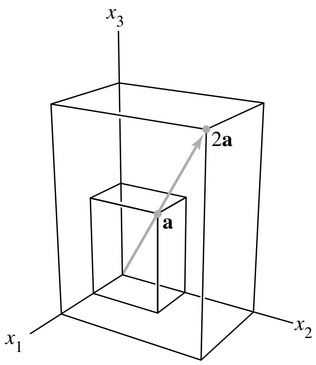


FIGURA 6
Múltiplos escalares.

Vectores en $\mathbb{R}^3$

Los vectores en $\mathbb{R}^3$ son matrices columna de 3×1 con tres entradas. Se representan geométricamente mediante puntos en un espacio coordenado tridimensional; algunas veces se incluyen flechas desde el origen para dar una mayor claridad visual. Los vectores $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ y 2**a** se muestran en la figura 6.

Vectores en $\mathbb{R}^n$

Si n es un entero positivo, $\mathbb{R}^n$ (léase "erre ene") denota la colección de todas las listas (o *n-adas ordenadas*) de n números reales, generalmente escritas como matrices columna de $n \times 1$ del tipo

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

El vector cuyas entradas son todas cero se llama **vector cero** y se denota con **0**. (El número de entradas en **0** será evidente a partir del contexto).

La igualdad de vectores en $\mathbb{R}^n$ y las operaciones de multiplicación escalar y suma vectorial en $\mathbb{R}^n$ se definen entrada por entrada como en $\mathbb{R}^2$. Esas operaciones sobre vectores tienen las siguientes propiedades, las cuales pueden verificarse directamente a partir de las propiedades correspondientes de los números reales. Véase el problema de práctica 1 y los ejercicios 33 y 34 al final de esta sección.

Propiedades algebraicas de $\mathbb{R}^n$

Para todo **u**, **v**, **w** en $\mathbb{R}^n$ y para todos los escalares c y d:

i. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
ii. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
iii. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
iv. $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$,
donde $-\mathbf{u}$ denota $(-1)\mathbf{u}$
v. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
vi. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
vii. $c(d\mathbf{u}) = (cd)(\mathbf{u})$
viii. $1\mathbf{u} = \mathbf{u}$

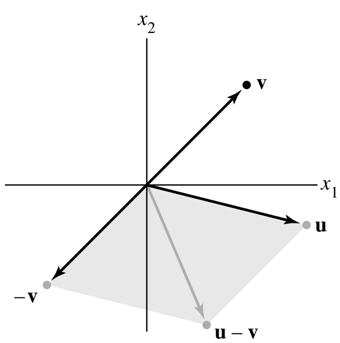


FIGURA 7
Resta vectorial.

Para simplificar la notación, un vector del tipo $\mathbf{u} + (-1)\mathbf{v}$ con frecuencia se escribe como $\mathbf{u} - \mathbf{v}$. La figura 7 muestra a $\mathbf{u} - \mathbf{v}$ como la suma de **u** y $-$**v**.

Combinaciones lineales

Dados los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$ y dados los escalares $c_1, c_2, \ldots, c_p$, el vector **y** definido por

$$\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

se llama **combinación lineal** de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ con **pesos** $c_1, \ldots, c_p$. La propiedad **ii** anterior nos permite omitir los paréntesis al formar la combinación lineal. Los pesos en una combinación lineal pueden ser cualesquiera números reales, incluyendo el cero. Por ejemplo, algunas combinaciones lineales de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ son

$$\sqrt{3}\,\mathbf{v}_1 + \mathbf{v}_2, \quad \tfrac{1}{2}\mathbf{v}_1 \ (= \tfrac{1}{2}\mathbf{v}_1 + 0\mathbf{v}_2) \quad \text{y} \quad \mathbf{0} \ (= 0\mathbf{v}_1 + 0\mathbf{v}_2)$$

EJEMPLO 4 La figura 8 identifica combinaciones lineales seleccionadas de $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. (Observe que los conjuntos de líneas paralelas de la rejilla están trazados mediante múltiplos enteros de $\mathbf{v}_1$ y $\mathbf{v}_2$). Estime las combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_2$ que generan los vectores $\mathbf{u}$ y $\mathbf{w}$.

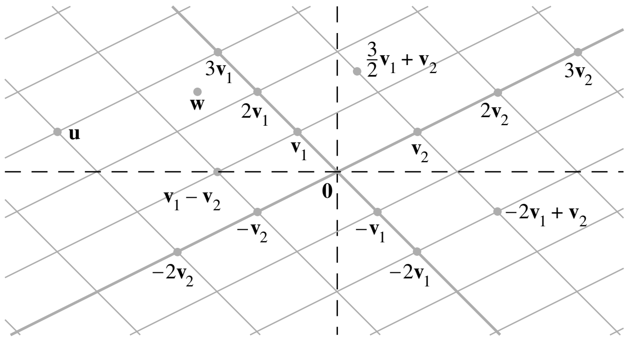


FIGURA 8 Combinaciones lineales de $\mathbf{v}_1$ y $\mathbf{v}_2$.

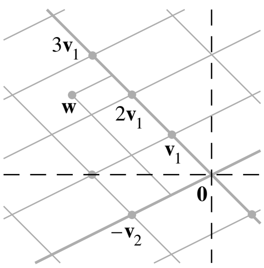


FIGURA 9

SOLUCIÓN La regla del paralelogramo indica que $\mathbf{u}$ es la suma de $3\mathbf{v}_1$ y $-2\mathbf{v}_2$; es decir,

$$\mathbf{u} = 3\mathbf{v}_1 - 2\mathbf{v}_2$$

Esta expresión para $\mathbf{u}$ se puede interpretar como las instrucciones para desplazarse desde el origen a $\mathbf{u}$ por dos trayectorias rectas. Primero, desplácese 3 unidades en la dirección de $\mathbf{v}_1$ hacia $3\mathbf{v}_1$, y luego avance -2 unidades en la dirección de $\mathbf{v}_2$ (paralela a la recta que pasa por $\mathbf{v}_2$ y $\mathbf{0}$). Después, aunque el vector $\mathbf{w}$ no está sobre una línea de la rejilla, $\mathbf{w}$ parece estar a la mitad del camino entre dos pares de rectas de la rejilla, en el vértice de un paralelogramo determinado por $(5/2)\mathbf{v}_1$ y $(-1/2)\mathbf{v}_2$. (Véase la figura 9). Así, una estimación razonable de $\mathbf{w}$ es

$$\mathbf{w} = \tfrac{5}{2}\mathbf{v}_1 - \tfrac{1}{2}\mathbf{v}_2 \qquad \blacksquare$$

El siguiente ejemplo relaciona un problema sobre combinaciones lineales con la pregunta fundamental de existencia que se estudió en las secciones 1.1 y 1.2.

EJEMPLO 5 Sean $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Determine si $\mathbf{b}$ se puede generar (o escribir) como una combinación lineal de $\mathbf{a}_1$ y $\mathbf{a}_2$. Es decir, determine si existen pesos x_1 y x_2 tales que

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b} \qquad (1)$$

Si la ecuación vectorial (1) tiene solución, encuéntrela.

SOLUCIÓN Aplique las definiciones de multiplicación escalar y suma vectorial para rescribir la ecuación vectorial

$$x_1 \underset{\uparrow \atop \mathbf{a}_1}{\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}} + x_2 \underset{\uparrow \atop \mathbf{a}_2}{\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}} = \underset{\uparrow \atop \mathbf{b}}{\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}}$$

que es lo mismo que

$$\begin{bmatrix} x_1 \\ -2x_1 \\ -5x_1 \end{bmatrix} + \begin{bmatrix} 2x_2 \\ 5x_2 \\ 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

y

$$\begin{bmatrix} x_1 + 2x_2 \\ -2x_1 + 5x_2 \\ -5x_1 + 6x_2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix} \tag{2}$$

Los vectores en los miembros izquierdo y derecho de (2) son iguales si y solo si sus entradas correspondientes son iguales. Es decir, x_1 y x_2 hacen válida la ecuación vectorial (1) si y solo si x_1 y x_2 satisfacen el sistema

$$\begin{aligned} x_1 + 2x_2 &= 7 \\ -2x_1 + 5x_2 &= 4 \\ -5x_1 + 6x_2 &= -3 \end{aligned} \tag{3}$$

Para resolver este sistema, se reduce por filas la matriz aumentada del sistema como sigue:[3]

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

La solución de (3) es $x_1 = 3$ y $x_2 = 2$. Así que $\mathbf{b}$ es una combinación lineal de $\mathbf{a}_1$ y $\mathbf{a}_2$, con pesos $x_1 = 3$ y $x_2 = 2$. Es decir,

$$3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$ ■

Observe en el ejemplo 5 que los vectores originales $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}$ son las columnas de la matriz aumentada reducida por filas:

$$\begin{array}{ccc} \left[\begin{array}{ccc} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{array}\right] \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{array} \end{array}$$

Por brevedad, se escribe esta matriz en una forma que identifique sus columnas, a saber,

$$[\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{b}] \tag{4}$$

De la ecuación vectorial (1) es claro cómo escribir esta matriz aumentada, sin realizar los pasos intermedios del ejemplo 5. Tome los vectores en el orden en que aparecen en (1) y colóquelos en las columnas de una matriz como en (4).

El análisis anterior puede modificarse fácilmente para establecer el siguiente hecho fundamental.

Una ecuación vectorial

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

tiene el mismo conjunto solución que el sistema lineal cuya matriz aumentada es

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad \mathbf{b}] \tag{5}$$

En particular, $\mathbf{b}$ se puede generar por una combinación lineal de $\mathbf{a}_1, \ldots, \mathbf{a}_n$ si y solo si existe una solución al sistema lineal correspondiente a la matriz (5).

[3] El símbolo $\sim$ entre matrices denota equivalencia de filas (sección 1.2).

Una de las ideas fundamentales en álgebra lineal es el estudio del conjunto de todos los vectores que se pueden generar o escribir como una combinación lineal de un conjunto fijo $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ de vectores.

DEFINICIÓN

Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en $\mathbb{R}^n$, entonces el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ se denota como Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ y se llama el **subconjunto de $\mathbb{R}^n$ extendido** o **generado por** $\mathbf{v}_1, \ldots, \mathbf{v}_p$. Es decir, Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es el conjunto de todos los vectores que se pueden escribir en la forma

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

con escalares $c_1, \ldots, c_p$.

Preguntar si un vector $\mathbf{b}$ está en Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ equivale a preguntar si la ecuación vectorial

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{b}$$

tiene una solución o, de manera equivalente, si el sistema lineal con la matriz aumentada $[\mathbf{v}_1 \ \cdots \ \mathbf{v}_p \ \mathbf{b}]$ tiene una solución.

Observe que Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ contiene a cada múltiplo escalar de $\mathbf{v}_1$ (por ejemplo), ya que $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. En particular, el vector cero debe estar en Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

Descripción geométrica de Gen {v} y de Gen {u, v}

Sea $\mathbf{v}$ un vector diferente de cero en $\mathbb{R}^3$. Entonces Gen $\{\mathbf{v}\}$ es el conjunto de todos los múltiplos escalares de $\mathbf{v}$, que es el conjunto de puntos sobre la recta en $\mathbb{R}^3$ que pasa por $\mathbf{v}$ y $\mathbf{0}$. Véase la figura 10.

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores diferentes de cero en $\mathbb{R}^3$, y $\mathbf{v}$ no es un múltiplo de $\mathbf{u}$, entonces Gen $\{\mathbf{u}, \mathbf{v}\}$ es el plano en $\mathbb{R}^3$ que contiene a $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{0}$. En particular, Gen $\{\mathbf{u}, \mathbf{v}\}$ contiene la recta en $\mathbb{R}^3$ que pasa por $\mathbf{u}$ y $\mathbf{0}$, y la recta que pasa por $\mathbf{v}$ y $\mathbf{0}$. Véase la figura 11.

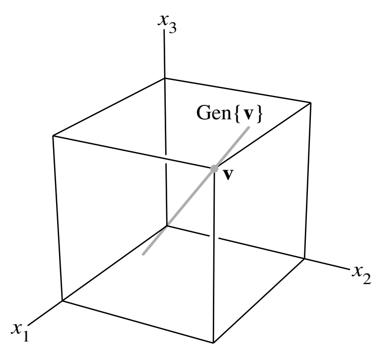


FIGURA 10 Gen $\{\mathbf{v}\}$ como una recta que pasa por el origen.

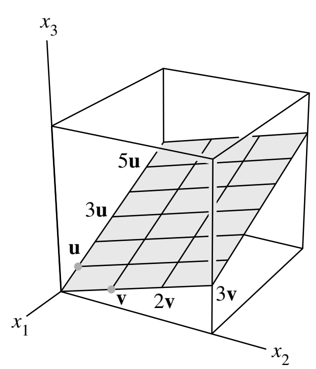


FIGURA 11 Gen $\{\mathbf{u}, \mathbf{v}\}$ como un plano que pasa por el origen.

EJEMPLO 6 Sean $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Entonces Gen $\{\mathbf{a}_1, \mathbf{a}_2\}$ es un plano que pasa por el origen en $\mathbb{R}^3$. ¿Está $\mathbf{b}$ en ese plano?

SOLUCIÓN ¿Tiene solución la ecuación $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$? Para responder a esto, reduzca por filas la matriz aumentada $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}]$:

$$\begin{bmatrix} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{bmatrix}$$

La tercera ecuación es $0 = -2$, la cual muestra que el sistema no tiene solución. La ecuación vectorial $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ no tiene solución, de manera que $\mathbf{b}$ *no* está en Gen $\{\mathbf{a}_1, \mathbf{a}_2\}$. ■

Combinaciones lineales en aplicaciones

El ejemplo final muestra cómo surgen múltiplos escalares y combinaciones lineales cuando una cantidad, como el "costo", se descompone en varias categorías. El principio básico en este ejemplo concierne al costo de fabricar varias unidades de un producto cuando se conoce el costo por unidad:

$$\begin{Bmatrix} \text{número} \\ \text{de unidades} \end{Bmatrix} \cdot \begin{Bmatrix} \text{costo} \\ \text{por unidad} \end{Bmatrix} = \begin{Bmatrix} \text{costo} \\ \text{total} \end{Bmatrix}$$

EJEMPLO 7 Una empresa fabrica dos productos. Para obtener \$1.00 del producto B, la empresa gasta \$0.45 en materiales, \$0.25 en mano de obra y \$0.15 por concepto de costos indirectos. Para obtener \$1.00 del producto C, la empresa gasta \$0.40 en materiales, \$0.30 en mano de obra y \$0.15 en costos indirectos. Sean

$$\mathbf{b} = \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix}$$

Entonces **b** y **c** representan los "costos por dólar de ingreso" para los dos productos.

a) ¿Qué interpretación económica puede darse al vector 100**b**?

b) Suponga que la empresa desea fabricar x_1 dólares del producto B y x_2 dólares del producto C. Dé un vector que describa los diversos costos que tendrá que enfrentar la empresa (por materiales, mano de obra y gastos indirectos).

SOLUCIÓN

a) Calcule

$$100\mathbf{b} = 100\begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} = \begin{bmatrix} 45 \\ 25 \\ 15 \end{bmatrix}$$

El vector 100**b** lista los diversos costos para producir \$100 del producto B, a saber, \$45 por materiales, \$25 por mano de obra y \$15 por gastos indirectos.

b) Los costos de fabricación de x_1 dólares del producto B están dados por el vector $x_1\mathbf{b}$, y los costos para manufacturar x_2 dólares del producto C están dados por $x_2\mathbf{c}$. Así que los costos totales para ambos productos están dados por el vector $x_1\mathbf{b} + x_2\mathbf{c}$. ■

PROBLEMAS DE PRÁCTICA

1. Demuestre que $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ para cualesquiera **u** y **v** en $\mathbb{R}^n$.
2. Obtenga el valor o los valores de h para que **y** esté en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

1.3 EJERCICIOS

En los ejercicios 1 y 2, calcule $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - 2\mathbf{v}$.

1. $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$

2. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

En los ejercicios 3 y 4, muestre los siguientes vectores utilizando flechas en una gráfica xy: $\mathbf{u}$, $\mathbf{v}$, $-\mathbf{v}$, $-2\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, $\mathbf{u} - \mathbf{v}$ y $\mathbf{u} - 2\mathbf{v}$. Observe que $\mathbf{u} - \mathbf{v}$ es el vértice de un paralelogramo cuyos otros vértices son $\mathbf{u}$, $\mathbf{0}$ y $-\mathbf{v}$.

3. $\mathbf{u}$ y $\mathbf{v}$ como en el ejercicio 1

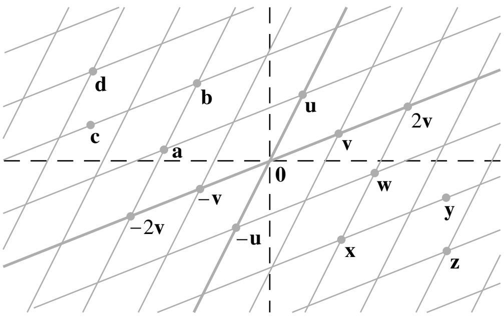
4. $\mathbf{u}$ y $\mathbf{v}$ como en el ejercicio 2

En los ejercicios 5 y 6, escriba un sistema de ecuaciones que sea equivalente a la ecuación vectorial dada.

5. $x_1 \begin{bmatrix} 3 \\ -2 \\ 8 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 0 \\ -9 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$

6. $x_1 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 7 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Utilice la figura adjunta para escribir cada vector listado en los ejercicios 7 y 8 como una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$. ¿Cada vector en $\mathbb{R}^2$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$?



7. Vectores $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ y $\mathbf{d}$

8. Vectores $\mathbf{w}$, $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$

En los ejercicios 9 y 10, escriba una ecuación vectorial que sea equivalente al sistema de ecuaciones dado.

9. $$\begin{aligned} x_2 + 5x_3 &= 0 \\ 4x_1 + 6x_2 - x_3 &= 0 \\ -x_1 + 3x_2 - 8x_3 &= 0 \end{aligned}$$

10. $$\begin{aligned} 3x_1 - 2x_2 + 4x_3 &= 3 \\ -2x_1 - 7x_2 + 5x_3 &= 1 \\ 5x_1 + 4x_2 - 3x_3 &= 2 \end{aligned}$$

En los ejercicios 11 y 12, determine si $\mathbf{b}$ es una combinación lineal de $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

11. $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} 5 \\ -6 \\ 8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$

12. $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} -6 \\ 7 \\ 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 11 \\ -5 \\ 9 \end{bmatrix}$

En los ejercicios 13 y 14, determine si $\mathbf{b}$ es una combinación lineal de los vectores formados a partir de las columnas de la matriz A.

13. $A = \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 5 \\ -2 & 8 & -4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -7 \\ -3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 0 & 5 \\ -2 & 1 & -6 \\ 0 & 2 & 8 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix}$

15. Sean $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} -5 \\ -8 \\ 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 3 \\ -5 \\ h \end{bmatrix}$. ¿Con qué valor (o valores) de h se encuentra $\mathbf{b}$ en el plano generado por $\mathbf{a}_1$ y $\mathbf{a}_2$?

16. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 7 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} h \\ -3 \\ -5 \end{bmatrix}$. ¿Para qué valor (o valores) de h se encuentra $\mathbf{y}$ en el plano generado por $\mathbf{v}_1$ y $\mathbf{v}_2$?

En los ejercicios 17 y 18, liste cinco vectores en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$. Para cada vector, muestre los pesos sobre $\mathbf{v}_1$ y $\mathbf{v}_2$ empleados para generar el vector e indique las tres entradas de este. No realice bosquejos.

17. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$

18. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 0 \end{bmatrix}$

19. Dé una descripción geométrica de Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ para los vectores $\mathbf{v}_1 = \begin{bmatrix} 8 \\ 2 \\ -6 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 12 \\ 3 \\ -9 \end{bmatrix}$

20. Realice una descripción geométrica de Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ para los vectores del ejercicio 18.

21. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Demuestre que $\begin{bmatrix} h \\ k \end{bmatrix}$ está en Gen $\{\mathbf{u}, \mathbf{v}\}$ para todas las h y k.

22. Construya una matriz A de 3×3, con entradas diferentes de cero, y un vector $\mathbf{b}$ en $\mathbb{R}^3$ tal que $\mathbf{b}$ *no* esté en el conjunto generado por las columnas de A.

En los ejercicios 23 y 24, marque cada enunciado como falso o verdadero. Justifique sus respuestas.

23. *a*) Otra notación para el vector $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ es $[-4 \quad 3]$.

b) Los puntos en el plano que corresponden a $\begin{bmatrix} -2 \\ 5 \end{bmatrix}$ y $\begin{bmatrix} -5 \\ 2 \end{bmatrix}$ están sobre una recta que pasa por el origen.

c) Un ejemplo de combinación lineal de los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ es el vector $\frac{1}{2}\mathbf{v}_1$.

d) El conjunto solución del sistema lineal cuya matriz aumentada es $[\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{b}]$ coincide con el conjunto solución de la ecuación $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$.

e) El conjunto Gen $\{\mathbf{u}, \mathbf{v}\}$ siempre se visualiza como un plano que pasa por el origen.

24. *a*) Cuando $\mathbf{u}$ y $\mathbf{v}$ son vectores diferentes de cero, Gen $\{\mathbf{u}, \mathbf{v}\}$ solo contiene la recta que pasa por $\mathbf{u}$ y por el origen, y la recta que pasa por $\mathbf{v}$ y el origen.

b) Cualquier lista de cinco números reales es un vector en $\mathbb{R}^5$.

c) Preguntar si el sistema lineal correspondiente a la matriz aumentada $[\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{b}]$ tiene solución equivale a preguntar si el vector $\mathbf{b}$ está en Gen $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

d) El vector $\mathbf{v}$ resulta cuando un vector $\mathbf{u} - \mathbf{v}$ se suma al vector $\mathbf{v}$.

e) No todos los pesos $c_1, \ldots, c_p$ en una combinación lineal $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ pueden ser cero.

25. Sean $A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -4 \end{bmatrix}$. Denote las columnas de A por $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, y sea $W =$ Gen $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

a) ¿Está $\mathbf{b}$ en $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$? ¿Cuántos vectores hay en $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$?

b) ¿Está $\mathbf{b}$ en W? ¿Cuántos vectores hay en W?

c) Demuestre que $\mathbf{a}_1$ está en W. [*Sugerencia:* Las operaciones de fila son innecesarias].

26. Sean $A = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 8 & 5 \\ 1 & -2 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 10 \\ 3 \\ 7 \end{bmatrix}$ y sea W el conjunto de todas las combinaciones lineales de las columnas de A.

a) ¿Está $\mathbf{b}$ en W?

b) Demuestre que la segunda columna de A está en W.

27. Una compañía minera posee dos minas. En un día, la mina #1 produce mineral que contiene 30 toneladas métricas de cobre y 600 kilogramos de plata, mientras que, también en un día, la mina #2 produce mineral que contiene 40 toneladas métricas de cobre y 380 kilogramos de plata. Sean $\mathbf{v}_1 = \begin{bmatrix} 30 \\ 600 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 40 \\ 380 \end{bmatrix}$. Así, $\mathbf{v}_1$ y $\mathbf{v}_2$ representan la "producción diaria" de la mina #1 y la mina #2, respectivamente.

a) ¿Qué interpretación física puede darse al vector $5\mathbf{v}_1$?

b) Suponga que la compañía opera la mina #1 durante x_1 días y la mina #2 por x_2 días. Escriba una ecuación vectorial cuya solución dé el número de días que cada mina debería operar para producir 240 toneladas de cobre y 2824 kilogramos de plata. No resuelva la ecuación.

c) [**M**] Resuelva la ecuación en *b*).

28. Una planta eléctrica de vapor quema dos tipos de carbón: antracita (A) y bituminoso (B). Por cada tonelada de A que se quema, la planta produce 27.6 millones de Btu de calor, 3100 gramos (g) de dióxido de sulfuro, y 250 g de contaminantes sólidos (partículas). Por cada tonelada de B que se quema, la planta produce 30.2 millones de Btu, 6400 g de dióxido de sulfuro, y 360 g de contaminantes sólidos (partículas).

a) ¿Cuánto calor produce la planta cuando quema x_1 toneladas de A y x_2 toneladas de B?

b) Suponga que la producción de la planta de vapor está descrita por un vector que lista las cantidades de calor, dióxido de sulfuro y contaminantes sólidos. Exprese esta producción como una combinación lineal de dos vectores, suponiendo que la planta quema x_1 toneladas de A y x_2 toneladas de B.

c) [**M**] Durante cierto tiempo, la planta de vapor produjo 162 millones de Btu de calor, 23,610 g de dióxido de sulfuro y 1623 g de contaminantes sólidos. Determine cuántas toneladas de cada tipo debe haber quemado la planta. Como parte de la solución, incluya una ecuación vectorial.

29. Sean $\mathbf{v}_1, \ldots, \mathbf{v}_k$ puntos en $\mathbb{R}^3$ y suponga que para $j = 1, \ldots, k$ un objeto con masa m_j está localizado en el punto $\mathbf{v}_j$. Los físicos llaman *masas puntuales* a esos objetos. La masa total del sistema de masas puntuales es

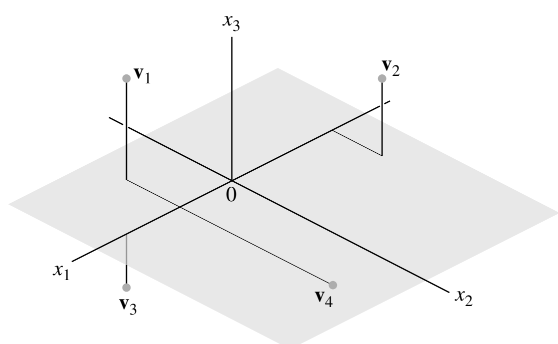
$$m = m_1 + \cdots + m_k$$

El *centro de gravedad* (o *centro de masa*) del sistema es

$$\overline{\mathbf{v}} = \frac{1}{m}[m_1\mathbf{v}_1 + \cdots + m_k\mathbf{v}_k]$$

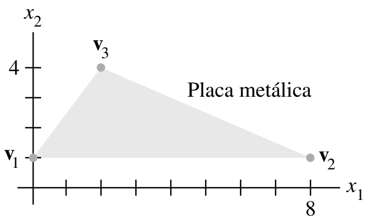
Calcule el centro de gravedad del sistema que consiste en las siguientes masas puntuales (véase la figura):

Punto	Masa
$\mathbf{v}_1 = (2, -2, 4)$	4 g
$\mathbf{v}_2 = (-4, 2, 3)$	2 g
$\mathbf{v}_3 = (4, 0, -2)$	3 g
$\mathbf{v}_4 = (1, -6, 0)$	5 g



30. Sea $\mathbf{v}$ el centro de masa de un sistema de masas puntuales localizadas en $\mathbf{v}_1, \ldots, \mathbf{v}_k$ como en el ejercicio 29. ¿Está $\mathbf{v}$ en Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$? Explique su respuesta.

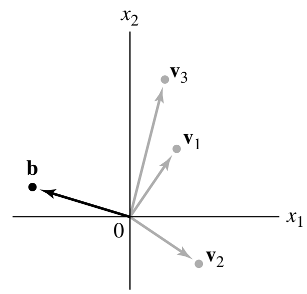
31. Una delgada placa triangular de densidad y grosor uniformes tiene vértices en $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$ y $\mathbf{v}_3 = (2, 4)$, como en la figura que aparece a continuación; la masa de la placa es de 3 g.



a) Encuentre las coordenadas (x, y) del centro de masa de la placa. Este "punto de equilibrio" de la placa coincide con el centro de masa de un sistema que consta de tres masas puntuales de 1 g colocadas en los vértices de la placa.

b) Determine cómo distribuir una masa adicional de 6 g en los tres vértices de la placa para así mover su punto de equilibrio a (2, 2). [*Sugerencia:* Sean w_1, w_2 y w_3 las masas agregadas a los tres vértices, de manera que $w_1 + w_2 + w_3 = 6$].

32. Considere los vectores $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{b}$ en $\mathbb{R}^2$, que se muestran en la figura. ¿Tiene solución la ecuación $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$? ¿Es única la solución? Utilice la figura para explicar sus respuestas.



33. Con los vectores $\mathbf{u} = (u_1,\ldots, u_n)$, $\mathbf{v} = (v_1,\ldots, v_n)$ y $\mathbf{w} = (w_1,\ldots, w_n)$, verifique las siguientes propiedades algebraicas de $\mathbb{R}^n$.

a) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

b) $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ para cada escalar c

34. Utilice el vector $\mathbf{u} = (u_1,\ldots, u_n)$ para verificar las siguientes propiedades algebraicas de $\mathbb{R}^n$.

a) $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$

b) $c(d\mathbf{u}) = (cd)\mathbf{u}$ para todos los escalares c y d

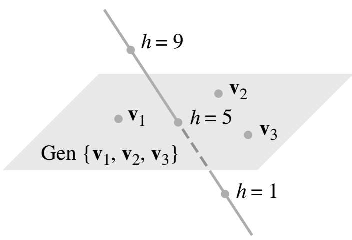
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Tome los vectores arbitrarios $\mathbf{u} = (u_1,\ldots, u_n)$ y $\mathbf{v} = (v_1,\ldots, v_n)$ en $\mathbb{R}^n$, y calcule

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1 + v_1, \ldots, u_n + v_n) && \text{Definición de suma vectorial} \\ &= (v_1 + u_1, \ldots, v_n + u_n) && \text{Conmutatividad de la adición en } \mathbb{R} \\ &= \mathbf{v} + \mathbf{u} && \text{Definición de suma vectorial} \end{aligned}$$

2. El vector $\mathbf{y}$ pertenece a Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si y solo si existen escalares x_1, x_2, x_3 tales que

$$x_1\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + x_2\begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$



Los puntos $\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$ están sobre la recta que se interseca con el plano cuando $h = 5$.

Esta ecuación vectorial es equivalente a un sistema de tres ecuaciones lineales con tres incógnitas. Si la matriz aumentada de este sistema se reduce por filas, se encuentra que

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}$$

Este sistema es consistente si y solo si no existe pivote en la cuarta columna. Es decir, $h - 5$ debe ser 0. Así, $\mathbf{y}$ está en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si y solo si $h = 5$.

Recuerde: La presencia de una variable libre en un sistema no garantiza que este sea consistente.

1.4 ECUACIÓN MATRICIAL $A\mathbf{x} = \mathbf{b}$

Una idea fundamental en álgebra lineal consiste en ver una combinación lineal de vectores como el producto de una matriz y un vector. La siguiente definición permite reformular algunos de los conceptos de la sección 1.3 desde nuevos puntos de vista.

DEFINICIÓN

> Si A es una matriz de $m \times n$, con columnas $\mathbf{a}_1, \ldots, \mathbf{a}_n$, y si $\mathbf{x}$ está en $\mathbb{R}^n$, entonces el **producto de** A **y** $\mathbf{x}$, denotado como $A\mathbf{x}$, es **la combinación lineal de las columnas de** A **utilizando como pesos las entradas correspondientes en x;** es decir,
>
> $$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Observe que $A\mathbf{x}$ está definido solamente si el número de columnas de A es igual al número de entradas en $\mathbf{x}$.

EJEMPLO 1

a) $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$
$$= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b) $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7\begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$ ■

EJEMPLO 2 Para $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ en $\mathbb{R}^m$, escriba la combinación lineal $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$ como una matriz por un vector.

SOLUCIÓN Coloque $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ en las columnas de una matriz A y coloque los pesos 3, -5 y 7 en un vector $\mathbf{x}$. Es decir,

$$3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3 = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = A\mathbf{x}$$ ■

La sección 1.3 mostró cómo escribir un sistema de ecuaciones lineales como una ecuación vectorial que implica una combinación lineal de vectores. Por ejemplo, el sistema

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 4 \\ -5x_2 + 3x_3 &= 1 \end{aligned} \tag{1}$$

es equivalente a

$$x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 2 \\ -5 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{2}$$

Como en el ejemplo 2, la combinación lineal en el lado izquierdo es una matriz por un vector, de manera que (2) se convierte en

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \tag{3}$$

La ecuación (3) tiene la forma $A\mathbf{x} = \mathbf{b}$. Tal ecuación se llama **ecuación matricial**, para distinguirla de una ecuación vectorial como la que se muestra en (2).

Observe cómo la matriz en (3) es justamente la matriz de coeficientes del sistema (1). Cálculos similares indican que cualquier sistema de ecuaciones lineales, o cualquier ecuación vectorial como (2), se puede escribir como una ecuación matricial equivalente en la forma $A\mathbf{x} = \mathbf{b}$. Esta sencilla observación se utilizará de manera recurrente a lo largo del texto.

A continuación se presenta el resultado formal.

TEOREMA 3

Si A es una matriz de $m \times n$, con columnas $\mathbf{a}_1,\ldots,\mathbf{a}_n$, y si $\mathbf{b}$ está en $\mathbb{R}^m$, la ecuación matricial

$$A\mathbf{x} = \mathbf{b} \tag{4}$$

tiene el mismo conjunto solución que la ecuación vectorial

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{5}$$

la cual, a la vez, tiene el mismo conjunto solución que el sistema de ecuaciones lineales cuya matriz aumentada es

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad \mathbf{b}] \tag{6}$$

El teorema 3 constituye una poderosa herramienta para comprender problemas de álgebra lineal, porque ahora un sistema de ecuaciones lineales puede verse en tres formas diferentes, pero equivalentes: como una ecuación matricial, como una ecuación vectorial o como un sistema de ecuaciones lineales. Siempre que usted construya un modelo matemático de un problema de la vida real, tendrá libertad para elegir qué punto de vista es más natural. Además, será posible pasar de una formulación del problema a otra, según sea conveniente. En cualquier caso, la ecuación matricial (4), la ecuación vectorial (5) y el sistema de ecuaciones se resuelven de la misma manera: por reducción de filas de la matriz aumentada (6). Más adelante se analizarán otros métodos de solución.

Existencia de soluciones

La definición de $A\mathbf{x}$ conduce directamente al siguiente hecho que resulta útil.

La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución si y solo si $\mathbf{b}$ es una combinación lineal de las columnas de A.

En la sección 1.3 se consideró la pregunta de existencia: "¿Está $\mathbf{b}$ en Gen $\{\mathbf{a}_1,\ldots,\mathbf{a}_n\}$?". De manera equivalente: "¿Es consistente $A\mathbf{x} = \mathbf{b}$?". Un problema de existencia más difícil consiste en determinar si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente *para toda* $\mathbf{b}$ posible.

EJEMPLO 3 Sean $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. ¿La ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para todas las posibles b_1, b_2, b_3?

SOLUCIÓN Se reduce por filas la matriz aumentada para $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{bmatrix}$$

La tercera entrada en la columna 4 es igual a $b_1 - \frac{1}{2}b_2 + b_3$. La ecuación $A\mathbf{x} = \mathbf{b}$ *no* es consistente para toda $\mathbf{b}$ porque algunas asignaciones de $\mathbf{b}$ pueden hacer que $b_1 - \frac{1}{2}b_2 + b_3$ sea diferente de cero. ■

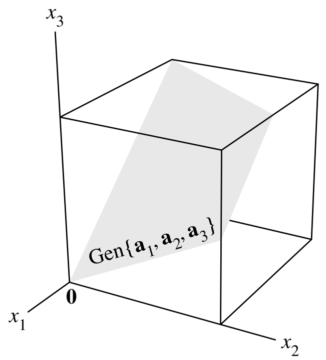


FIGURA 1 Las columnas de $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ generan un plano a través de $\mathbf{0}$.

La matriz reducida del ejemplo 3 da una descripción de todas las $\mathbf{b}$ para las cuales la ecuación $A\mathbf{x} = \mathbf{b}$ *es* consistente: las entradas en $\mathbf{b}$ deben satisfacer

$$b_1 - \tfrac{1}{2}b_2 + b_3 = 0$$

Esta es la ecuación de un plano que pasa por el origen en $\mathbb{R}^3$. El plano es el conjunto de todas las combinaciones lineales de las tres columnas de A. Véase la figura 1.

La ecuación $A\mathbf{x} = \mathbf{b}$ del ejemplo 3 no es consistente para todas las $\mathbf{b}$ porque la forma escalonada de A tiene una fila de ceros. Si A tuviera un pivote en las tres filas, no habría que preocuparse por los cálculos en la columna aumentada, ya que, en este caso, una forma escalonada de la matriz aumentada no tendría una fila como $[0 \quad 0 \quad 0 \quad 1]$.

En el siguiente teorema, la frase "las columnas de A generan a $\mathbb{R}^m$" significa que *cada* $\mathbf{b}$ en $\mathbb{R}^m$ es una combinación lineal de las columnas de A. En general, un conjunto de vectores $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^m$ **genera** a $\mathbb{R}^m$ si cada vector en $\mathbb{R}^m$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_p$; es decir, si Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} = \mathbb{R}^m$.

TEOREMA 4

Sea A una matriz de $m \times n$. Entonces, los siguientes enunciados son lógicamente equivalentes. Es decir, para una A particular, todos los enunciados son verdaderos o todos son falsos.

a) Para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución.

b) Cada $\mathbf{b}$ en $\mathbb{R}^m$ es una combinación lineal de las columnas de A.

c) Las columnas de A generan $\mathbb{R}^m$.

d) A tiene una posición pivote en cada fila.

El teorema 4 es uno de los teoremas más útiles en este capítulo. Los enunciados *a*), *b*) y *c*) son equivalentes a causa de la definición de $A\mathbf{x}$ y lo que significa para un conjunto de vectores generar $\mathbb{R}^m$. El análisis posterior al ejemplo 3 revela por qué *a*) y *d*) son equivalentes; al final de la sección se presenta una prueba de ello. Los ejercicios aportan ejemplos de cómo emplear el teorema 4.

Advertencia: El teorema 4 se refiere a una *matriz* de *coeficientes*, no a una matriz aumentada. Si una matriz aumentada $[A \quad \mathbf{b}]$ tiene una posición pivote en cada fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ puede o no ser consistente.

Cálculo de $A\mathbf{x}$

Los cálculos en el ejemplo 1 se apoyaron en la definición del producto de una matriz A y un vector $\mathbf{x}$. El siguiente ejemplo sencillo conducirá a un método más eficiente para calcular las entradas en $A\mathbf{x}$ cuando los problemas se resuelvan a mano.

EJEMPLO 4 Calcule $A\mathbf{x}$, donde $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

SOLUCIÓN A partir de la definición,

$$\begin{aligned}
\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} 2 \\ -1 \\ 6 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ -3 \\ 8 \end{bmatrix} \\
&= \begin{bmatrix} 2x_1 \\ -x_1 \\ 6x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ 5x_2 \\ -2x_2 \end{bmatrix} + \begin{bmatrix} 4x_3 \\ -3x_3 \\ 8x_3 \end{bmatrix} \qquad (7) \\
&= \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{bmatrix}
\end{aligned}$$

La primera entrada en el producto $A\mathbf{x}$ es una suma de productos (algunas veces se llama *producto punto*), empleando la primera fila de A y las entradas en $\mathbf{x}$. Es decir,

$$\begin{bmatrix} 2 & 3 & 4 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ \\ \\ \end{bmatrix}$$

Esta matriz muestra cómo calcular de forma directa la primera entrada en $A\mathbf{x}$, sin escribir todos los cálculos indicados en (7). De manera similar, la segunda entrada en $A\mathbf{x}$ puede calcularse multiplicando las entradas en la segunda fila de A por las entradas correspondientes en $\mathbf{x}$, y después sumando los productos resultantes:

$$\begin{bmatrix} & & \\ -1 & 5 & -3 \\ & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \\ -x_1 + 5x_2 - 3x_3 \\ \\ \end{bmatrix}$$

De manera semejante, la tercera entrada en $A\mathbf{x}$ puede determinarse con la tercera fila de A y las entradas en $\mathbf{x}$. ■

Regla fila-vector para calcular $A\mathbf{x}$

Si el producto $A\mathbf{x}$ está definido, entonces la i-ésima entrada en $A\mathbf{x}$ es la suma de los productos de las entradas correspondientes de la fila i de A y del vector $\mathbf{x}$.

EJEMPLO 5

a) $$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 + 2 \cdot 3 + (-1) \cdot 7 \\ 0 \cdot 4 + (-5) \cdot 3 + 3 \cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

b) $$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + (-3) \cdot 7 \\ 8 \cdot 4 + 0 \cdot 7 \\ (-5) \cdot 4 + 2 \cdot 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$$

c) $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1 \cdot r + 0 \cdot s + 0 \cdot t \\ 0 \cdot r + 1 \cdot s + 0 \cdot t \\ 0 \cdot r + 0 \cdot s + 1 \cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$$ ■

Por definición, la matriz del ejemplo 5*c*) con números 1 en la diagonal y ceros en las demás posiciones se llama **matriz identidad** y se denota con I. Los cálculos en el inciso *c*) indican que $I\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^3$. Existe una análoga matriz identidad de $n \times n$, algunas veces denotada como I_n. Al igual que en *c*), $I_n\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

Propiedades del producto matriz-vector $A\mathbf{x}$

Es importante el contenido del próximo teorema y se utilizará a lo largo del libro. La demostración se apoya en la definición de $A\mathbf{x}$ y en las propiedades algebraicas de $\mathbb{R}^n$.

TEOREMA 5

> Si A es una matriz de $m \times n$, $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$, y c es un escalar, entonces:
>
> *a*) $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;
>
> *b*) $A(c\mathbf{u}) = c(A\mathbf{u})$.

DEMOSTRACIÓN En aras de la sencillez, se toma $n = 3$, $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$, y $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^3$. (La demostración para el caso general es similar). Para $i = 1, 2, 3$, sean u_i y v_i las i-ésimas entradas en $\mathbf{u}$ y $\mathbf{v}$, respectivamente. Para probar el enunciado *a*), calcule $A(\mathbf{u} + \mathbf{v})$ como una combinación lineal de las columnas de A empleando como pesos las entradas en $\mathbf{u} + \mathbf{v}$.

$$A(\mathbf{u}+\mathbf{v}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{bmatrix}$$

Entradas en $\mathbf{u} + \mathbf{v}$

$$= (u_1 + v_1)\mathbf{a}_1 + (u_2 + v_2)\mathbf{a}_2 + (u_3 + v_3)\mathbf{a}_3$$

Columnas de A

$$\begin{aligned} &= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3) \\ &= A\mathbf{u} + A\mathbf{v} \end{aligned}$$

Para demostrar el enunciado *b*), calcule $A(c\mathbf{u})$ como una combinación lineal de las columnas de A utilizando como pesos las entradas en $c\mathbf{u}$.

$$\begin{aligned} A(c\mathbf{u}) = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} cu_1 \\ cu_2 \\ cu_3 \end{bmatrix} &= (cu_1)\mathbf{a}_1 + (cu_2)\mathbf{a}_2 + (cu_3)\mathbf{a}_3 \\ &= c(u_1\mathbf{a}_1) + c(u_2\mathbf{a}_2) + c(u_3\mathbf{a}_3) \\ &= c(u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) \\ &= c(A\mathbf{u}) \end{aligned}$$

■

NOTA NUMÉRICA

> Con la finalidad de optimizar un algoritmo computacional para calcular $A\mathbf{x}$, la secuencia de operaciones implicaría datos almacenados en ubicaciones adyacentes de memoria. Los algoritmos profesionales más ampliamente utilizados para cálculos matriciales están escritos en Fortran, un lenguaje que almacena una matriz como un conjunto de columnas. Tales algoritmos calculan $A\mathbf{x}$ como una combinación lineal de las columnas de A. En contraste, si un programa está escrito en el conocido lenguaje C, el cual almacena matrices por filas, entonces $A\mathbf{x}$ debería calcularse mediante la regla alternativa que usa las filas de A.

DEMOSTRACIÓN DEL TEOREMA 4 Como se indicó después del teorema 4, los enunciados *a*), *b*) y *c*) son lógicamente equivalentes. De manera que es suficiente demostrar (para una matriz arbitraria A) que *a*) y *d*) son ambos verdaderos o ambos falsos. Eso vinculará a los cuatro enunciados.

Sea U una forma escalonada de A. Dada $\mathbf{b}$ en $\mathbb{R}^m$, la matriz aumentada $[A \quad \mathbf{b}]$ se puede reducir por filas a una matriz aumentada $[U \quad \mathbf{d}]$ para alguna $\mathbf{d}$ en $\mathbb{R}^m$:

$$[A \quad \mathbf{b}] \sim \cdots \sim [U \quad \mathbf{d}]$$

Si el enunciado *d*) es verdadero, entonces cada fila de U contiene una posición pivote y tal vez no exista un pivote en la columna aumentada. De manera que $A\mathbf{x} = \mathbf{b}$ tiene una solución para cualquier $\mathbf{b}$, y *a*) es verdad. Si *d*) es falso, la última fila de U solo tiene ceros. Sea $\mathbf{d}$ cualquier vector con un 1 en su última entrada. En tal caso, $[U \quad \mathbf{d}]$ representa un sistema *inconsistente*. Como las operaciones de fila son reversibles, $[U \quad \mathbf{d}]$ se puede transformar a la forma $[A \quad \mathbf{b}]$. El nuevo sistema $A\mathbf{x} = \mathbf{b}$ también es inconsistente, y *a*) es falso. ■

PROBLEMAS DE PRÁCTICA

1. Sean $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$. Es posible demostrar que $\mathbf{p}$ es una solución de $A\mathbf{x} = \mathbf{b}$. Con base en este hecho, presente a $\mathbf{b}$ como una combinación lineal específica de las columnas de A.

2. Sean $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Compruebe el teorema 5*a*) en este caso mediante el cálculo de $A(\mathbf{u} + \mathbf{v})$ y $A\mathbf{u} + A\mathbf{v}$.

1.4 EJERCICIOS

En los ejercicios 1 a 4 calcule los productos utilizando: *a*) la definición, como en el ejemplo 1, y *b*) la regla fila-vector para calcular $A\mathbf{x}$. Si un producto está indefinido, explique por qué.

1. $\begin{bmatrix} -4 & 2 \\ 1 & 6 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}\begin{bmatrix} 5 \\ -1 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 6 \end{bmatrix}\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 & -4 \\ 3 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

En los ejercicios 5 a 8, aplique la definición de $A\mathbf{x}$ para escribir la ecuación matricial como una ecuación vectorial, o viceversa.

5. $\begin{bmatrix} 1 & 2 & -3 & 1 \\ -2 & -3 & 1 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 2 & -3 \\ 3 & 2 \\ 8 & -5 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -21 \\ 1 \\ -49 \\ 11 \end{bmatrix}$

7. $x_1\begin{bmatrix} 4 \\ -1 \\ 7 \\ -4 \end{bmatrix} + x_2\begin{bmatrix} -5 \\ 3 \\ -5 \\ 1 \end{bmatrix} + x_3\begin{bmatrix} 7 \\ -8 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 0 \\ -7 \end{bmatrix}$

8. $z_1\begin{bmatrix} 2 \\ -4 \end{bmatrix} + z_2\begin{bmatrix} -1 \\ 5 \end{bmatrix} + z_3\begin{bmatrix} -4 \\ 3 \end{bmatrix} + z_4\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 12 \end{bmatrix}$

En los ejercicios 9 y 10, primero escriba el sistema como una ecuación vectorial y después como una ecuación matricial.

9. $\begin{aligned} 5x_1 + x_2 - 3x_3 &= 8 \\ 2x_2 + 4x_3 &= 0 \end{aligned}$

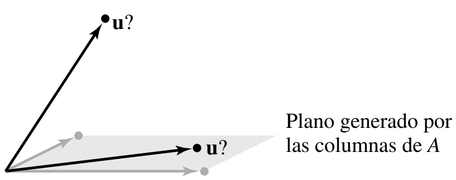
10. $\begin{aligned} 4x_1 - x_2 &= 8 \\ 5x_1 + 3x_2 &= 2 \\ 3x_1 - x_2 &= 1 \end{aligned}$

Considerando A y $\mathbf{b}$ en los ejercicios 11 y 12, escriba la matriz aumentada para el sistema lineal que corresponde a la ecuación matricial $A\mathbf{x} = \mathbf{b}$. Después, resuelva el sistema y escriba la solución como un vector:

11. $A = \begin{bmatrix} 1 & 3 & -4 \\ 1 & 5 & 2 \\ -3 & -7 & 6 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 4 \\ 12 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -4 & 2 \\ 5 & 2 & 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$

13. Sean $\mathbf{u} = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$ y $A = \begin{bmatrix} 3 & -5 \\ -2 & 6 \\ 1 & 1 \end{bmatrix}$. ¿Está $\mathbf{u}$ en el plano en $\mathbb{R}^3$ generado por las columnas de A? (Véase la figura). ¿Por qué?



¿Dónde está $\mathbf{u}$?

14. Sean $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \\ 4 \end{bmatrix}$ y $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{bmatrix}$. ¿Está $\mathbf{u}$ en el subconjunto de $\mathbb{R}^3$ generado por las columnas de A? ¿Por qué?

15. Sean $A = \begin{bmatrix} 3 & -1 \\ -9 & 3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Demuestre que la ecuación $A\mathbf{x} = \mathbf{b}$ no tiene solución para todas las posibles $\mathbf{b}$, y describa el conjunto de todas las $\mathbf{b}$ para las cuales $A\mathbf{x} = \mathbf{b}$ *sí* tiene solución.

16. Repita el ejercicio 15 considerando

$$A = \begin{bmatrix} 1 & -2 & -1 \\ -2 & 2 & 0 \\ 4 & -1 & 3 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Los ejercicios 17 a 20 se refieren a las matrices A y B que se presentan a continuación. Realice los cálculos pertinentes para justificar sus respuestas y mencione un teorema adecuado.

$$A = \begin{bmatrix} 1 & 3 & 0 & 3 \\ -1 & -1 & -1 & 1 \\ 0 & -4 & 2 & -8 \\ 2 & 0 & 3 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 4 & 1 & 2 \\ 0 & 1 & 3 & -4 \\ 0 & 2 & 6 & 7 \\ 2 & 9 & 5 & -7 \end{bmatrix}$$

17. ¿Cuántas filas de A contienen una posición pivote? ¿La ecuación $A\mathbf{x} = \mathbf{b}$ tiene solución para cada $\mathbf{b}$ en $\mathbb{R}^4$?

18. ¿Cada vector en $\mathbb{R}^4$ se puede escribir como una combinación lineal de las columnas de la matriz B anterior? ¿Las columnas de B generan a $\mathbb{R}^3$?

19. ¿Todo vector en $\mathbb{R}^4$ se puede escribir como una combinación lineal de las columnas de la matriz A anterior? ¿Las columnas de A generan a $\mathbb{R}^4$?

20. ¿Las columnas de B generan a $\mathbb{R}^4$? ¿La ecuación $B\mathbf{x} = \mathbf{y}$ tiene solución para cada $\mathbf{y}$ en $\mathbb{R}^4$?

21. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$. ¿$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ genera a $\mathbb{R}^4$? ¿Por qué?

22. Sean $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 9 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ -2 \\ -6 \end{bmatrix}$. ¿$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ genera a $\mathbb{R}^3$? ¿Por qué?

En los ejercicios 23 y 24, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) La ecuación $A\mathbf{x} = \mathbf{b}$ se reconoce como una *ecuación vectorial*.

b) Un vector $\mathbf{b}$ es una combinación lineal de las columnas de una matriz A si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución.

c) La ecuación $A\mathbf{x} = \mathbf{b}$ es consistente si la matriz aumentada $[A \quad \mathbf{b}]$ tiene una posición pivote en cada fila.

d) La primera entrada en el producto $A\mathbf{x}$ es una suma de productos.

e) Si las columnas de una matriz A de $m \times n$ generan a $\mathbb{R}^m$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para cada $\mathbf{b}$ en $\mathbb{R}^m$.

f) Si A es una matriz de $m \times n$ y si la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente para alguna $\mathbf{b}$ en $\mathbb{R}^m$, entonces A no puede tener una posición pivote en cada fila.

24. *a*) Cada ecuación matricial $A\mathbf{x} = \mathbf{b}$ corresponde a una ecuación vectorial con el mismo conjunto solución.

b) Si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces $\mathbf{b}$ está en el conjunto generado por las columnas de A.

c) Cualquier combinación lineal de vectores siempre se puede escribir en la forma $A\mathbf{x}$ para una matriz A y un vector $\mathbf{x}$ adecuados.

d) Si la matriz coeficiente A tiene una posición pivote en cada fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente.

e) El conjunto solución de un sistema lineal cuya matriz aumentada es $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}]$ coincide con el conjunto solución de $A\mathbf{x} = \mathbf{b}$, si $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$.

f) Si A es una matriz de $m \times n$ cuyas columnas no generan a $\mathbb{R}^m$, entonces la ecuación $\mathrm{A}\mathbf{x} = \mathbf{b}$ es consistente para toda $\mathbf{b}$ en $\mathbb{R}^m$.

25. Observe que $\begin{bmatrix} 4 & -3 & 1 \\ 5 & -2 & 5 \\ -6 & 2 & -3 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix}$. Con base en este hecho (sin realizar operaciones de fila), encuentre escalares c_1, c_2, c_3 tales que $\begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix} = c_1 \begin{bmatrix} 4 \\ 5 \\ -6 \end{bmatrix} + c_2 \begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$.

26. Sean $\mathbf{u} = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$. Es posible demostrar que $2\mathbf{u} - 3\mathbf{v} - \mathbf{w} = \mathbf{0}$. Apóyese en este hecho (sin realizar operaciones de fila) para obtener x_1 y x_2 que satisfagan la ecuación $\begin{bmatrix} 7 & 3 \\ 2 & 1 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$.

27. Rescriba la siguiente ecuación matricial (numérica) en forma simbólica como una ecuación vectorial, utilizando los símbolos $\mathbf{v}_1, \mathbf{v}_2, \ldots$ para los vectores y $c_1, c_2, \ldots$ para los escalares. Defina qué representa cada símbolo, utilizando los datos presentados en la ecuación matricial.

$$\begin{bmatrix} -3 & 5 & -4 & 9 & 7 \\ 5 & 8 & 1 & -2 & -4 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \\ 2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 \\ -11 \end{bmatrix}$$

28. Considere que $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3$ y $\mathbf{v}$ son vectores en $\mathbb{R}^5$, mientras que x_1, x_2 y x_3 denotan escalares. Escriba la siguiente ecuación vectorial como una ecuación matricial. Identifique cualquier símbolo que utilice.

$$x_1\mathbf{q}_1 + x_2\mathbf{q}_2 + x_3\mathbf{q}_3 = \mathbf{v}$$

29. Construya una matriz de 3×3, no en forma escalonada, cuyas columnas generen a $\mathbb{R}^3$. Demuestre que la matriz que construyó tiene la propiedad deseada.

30. Construya una matriz de 3×3, no en forma escalonada, cuyas columnas *no* generen a $\mathbb{R}^3$. Demuestre que la matriz que construyó tiene la característica deseada.

31. Sea A una matriz de 3×2. Explique por qué la ecuación $A\mathbf{x} = \mathbf{b}$ no puede ser consistente para toda $\mathbf{b}$ en $\mathbb{R}^3$. Generalice su argumento al caso de una A arbitraria con más filas que columnas.

32. ¿Un conjunto de tres vectores en $\mathbb{R}^4$ podría generar a $\mathbb{R}^4$? Explique su respuesta. ¿Y qué hay respecto de n vectores en $\mathbb{R}^m$ cuando n es menor que m?

33. Suponga que A es una matriz de 4×3, y $\mathbf{b}$ es un vector en $\mathbb{R}^4$ con la propiedad de que $A\mathbf{x} = \mathbf{b}$ tiene una solución única. ¿Qué podría decir sobre la forma escalonada reducida de A? Justifique su respuesta.

34. Considere que A es una matriz de 3×4, $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores en $\mathbb{R}^3$, y que $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$. Suponga que $\mathbf{v}_1 = A\mathbf{u}_1$ y $\mathbf{v}_2 = A\mathbf{u}_2$ para algunos vectores $\mathbf{u}_1$ y $\mathbf{u}_2$ en $\mathbb{R}^4$. ¿Qué hecho permite concluir que el sistema $A\mathbf{x} = \mathbf{w}$ es consistente? (*Nota:* $\mathbf{u}_1$ y $\mathbf{u}_2$ denotan vectores, y no entradas escalares de vectores).

35. Sean A una matriz de 5×3, $\mathbf{y}$ un vector en $\mathbb{R}^3$, y $\mathbf{z}$ un vector en $\mathbb{R}^5$. Suponga que $A\mathbf{y} = \mathbf{z}$. ¿Qué hecho permite concluir que el sistema $A\mathbf{x} = 5\mathbf{z}$ es consistente?

36. Suponga que A es una matriz de 4×4 y $\mathbf{b}$ un vector en $\mathbb{R}^4$ tal que $A\mathbf{x} = \mathbf{b}$ tiene una solución única. Explique por qué las columnas de A deben generar a $\mathbb{R}^4$.

[**M**] En los ejercicios 37 a 40, determine si las columnas de la matriz generan a $\mathbb{R}^4$.

37. $\begin{bmatrix} 7 & 2 & -5 & 8 \\ -5 & -3 & 4 & -9 \\ 6 & 10 & -2 & 7 \\ -7 & 9 & 2 & 15 \end{bmatrix}$ **38.** $\begin{bmatrix} 4 & -5 & -1 & 8 \\ 3 & -7 & -4 & 2 \\ 5 & -6 & -1 & 4 \\ 9 & 1 & 10 & 7 \end{bmatrix}$

39. $\begin{bmatrix} 10 & -7 & 1 & 4 & 6 \\ -8 & 4 & -6 & -10 & -3 \\ -7 & 11 & -5 & -1 & -8 \\ 3 & -1 & 10 & 12 & 12 \end{bmatrix}$

40. $\begin{bmatrix} 5 & 11 & -6 & -7 & 12 \\ -7 & -3 & -4 & 6 & -9 \\ 11 & 5 & 6 & -9 & -3 \\ -3 & 4 & -7 & 2 & 7 \end{bmatrix}$

41. [**M**] Encuentre una columna de la matriz del ejercicio 39 que se pueda eliminar y, aun así, las restantes columnas de la matriz sigan generando a $\mathbb{R}^4$.

42. [**M**] Encuentre una columna de la matriz del ejercicio 40 que se pueda eliminar y, aun así, las restantes columnas sigan generando a $\mathbb{R}^4$. ¿Se puede eliminar más de una columna?

WEB

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. La ecuación matricial

$$\begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

es equivalente a la ecuación vectorial

$$3\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} - 2\begin{bmatrix} 5 \\ 1 \\ -8 \end{bmatrix} + 0\begin{bmatrix} -2 \\ 9 \\ -1 \end{bmatrix} - 4\begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$$

que expresa a $\mathbf{b}$ como una combinación lineal de las columnas de A.

2.

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} \\ A(\mathbf{u} + \mathbf{v}) &= \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 + 20 \\ 3 + 4 \end{bmatrix} = \begin{bmatrix} 22 \\ 7 \end{bmatrix} \\ A\mathbf{u} + A\mathbf{v} &= \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 11 \end{bmatrix} + \begin{bmatrix} 19 \\ -4 \end{bmatrix} = \begin{bmatrix} 22 \\ 7 \end{bmatrix} \end{aligned}$$

1.5 CONJUNTOS SOLUCIÓN DE SISTEMAS LINEALES

Los conjuntos solución de sistemas lineales son importantes objetos de estudio en álgebra lineal. Más adelante, se presentarán en diversos contextos. Esta sección utiliza notación vectorial para dar descripciones explícitas y geométricas de tales conjuntos solución.

Sistemas lineales homogéneos

Se dice que un sistema de ecuaciones lineales es **homogéneo** si se puede escribir en la forma $A\mathbf{x} = \mathbf{0}$, donde A es una matriz de $m \times n$, y $\mathbf{0}$ es el vector cero en $\mathbb{R}^m$. Tal sistema $A\mathbf{x} = \mathbf{0}$ *siempre* tiene al menos una solución, a saber, $\mathbf{x} = \mathbf{0}$ (el vector cero en $\mathbb{R}^n$). Esta solución cero generalmente se conoce como **solución trivial**. Para una ecuación dada $A\mathbf{x} = \mathbf{0}$, la pregunta importante es si existe una **solución no trivial**, es decir, un vector $\mathbf{x}$ diferente de cero que satisfaga $A\mathbf{x} = \mathbf{0}$. El teorema de existencia y unicidad de la sección 1.2 (teorema 2) conduce de inmediato al siguiente resultado.

La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial si y solo si la ecuación tiene al menos una variable libre.

EJEMPLO 1 Determine si el siguiente sistema homogéneo tiene una solución no trivial. Luego, describa el conjunto solución.

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

SOLUCIÓN Sea A la matriz de coeficientes del sistema; la matriz aumentada $[A \quad \mathbf{0}]$ se reduce por filas a una forma escalonada:

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Como x_3 es una variable libre, entonces $A\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales (una para cada asignación de x_3). Para describir el conjunto solución, continúe la reducción por filas de $[A \quad \mathbf{0}]$ hasta su forma escalonada *reducida*:

$$\begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_1 \quad - \tfrac{4}{3}x_3 &= 0 \\ x_2 \quad &= 0 \\ 0 &= 0 \end{aligned}$$

Al despejar las variables básicas x_1 y x_2 se obtiene $x_1 = \frac{4}{3}x_3$, $x_2 = 0$, con x_3 libre. Como un vector, la solución general de $A\mathbf{x} = \mathbf{0}$ tiene la forma

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} = x_3\mathbf{v}, \quad \text{donde } \mathbf{v} = \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

Aquí se factorizó x_3 de la expresión para la solución general vectorial. Esto muestra que cada solución de $A\mathbf{x} = \mathbf{0}$ es, en este caso, un múltiplo escalar de $\mathbf{v}$. La solución se obtiene al considerar $x_3 = 0$. Geométricamente, el conjunto solución es una recta que pasa por $\mathbf{0}$ en $\mathbb{R}^3$. Véase la figura 1. ■

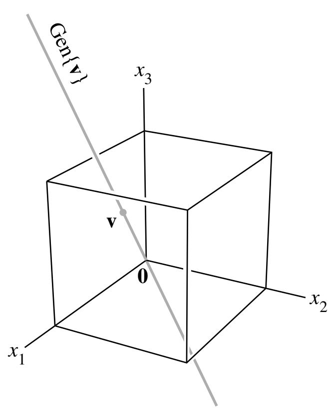


FIGURA 1

Observe que una solución no trivial $\mathbf{x}$ puede tener algunas entradas cero, siempre y cuando no todas ellas sean cero.

EJEMPLO 2 Una sola ecuación lineal puede tratarse como un sencillo sistema de ecuaciones. Describa todas las soluciones del “sistema” homogéneo

$$10x_1 - 3x_2 - 2x_3 = 0 \tag{1}$$

SOLUCIÓN No hay necesidad de una notación matricial. Despeje la variable básica x_1 en términos de las variables libres. La solución general es $x_1 = .3x_2 + .2x_3$, con x_2 y x_3 libres. Como un vector, la solución general es

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} .2x_3 \\ 0 \\ x_3 \end{bmatrix}$$

$$= x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix}} + x_3 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix}} \quad (\text{con } x_2, x_3 \text{ libres}) \tag{2}$$

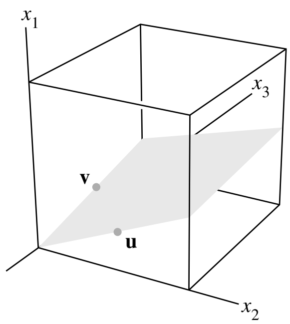


FIGURA 2

Este cálculo indica que cada solución de (1) es una combinación lineal de los vectores $\mathbf{u}$ y $\mathbf{v}$, que se muestran en (2). Es decir, el conjunto solución es Gen $\{\mathbf{u}, \mathbf{v}\}$. Puesto que $\mathbf{u}$ y $\mathbf{v}$ no son múltiplos entre sí, el conjunto solución es un plano que pasa por el origen. Véase la figura 2. ■

Los ejemplos 1 y 2, junto con los ejercicios, ilustran el hecho de que el conjunto solución de una ecuación homogénea $A\mathbf{x} = \mathbf{0}$ siempre se puede expresar de manera explícita como Gen $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ para vectores adecuados $\mathbf{v}_1, \dots, \mathbf{v}_p$. Si la única solución es el vector cero, entonces el conjunto solución es Gen $\{\mathbf{0}\}$. Si la ecuación $A\mathbf{x} = \mathbf{0}$ solo tiene una variable libre, el conjunto solución es una recta que pasa por el origen, como en la figura 1. Un plano que pasa por el origen, como en la figura 2, da una buena imagen mental para el conjunto solución de $A\mathbf{x} = \mathbf{0}$ cuando existen dos o más variables libres. Observe, sin embargo, que se puede emplear una figura similar para visualizar Gen $\{\mathbf{u}, \mathbf{v}\}$ aun cuando $\mathbf{u}$ y $\mathbf{v}$ no estén relacionados con soluciones de $A\mathbf{x} = \mathbf{0}$. Véase la figura 11 de la sección 1.3.

Forma vectorial paramétrica

La ecuación original (1) para el plano del ejemplo 2 es una descripción *implícita* del plano. Al resolver esta ecuación se obtiene una descripción *explícita* del plano como el conjunto generado por $\mathbf{u}$ y $\mathbf{v}$. La ecuación (2) se llama **ecuación vectorial paramétrica** del plano. Algunas veces dicha ecuación se escribe como

$$\mathbf{x} = s\mathbf{u} + t\mathbf{v} \quad (s, t \text{ en } \mathbb{R})$$

para poner de relieve que los parámetros varían sobre todos los números reales. En el ejemplo 1, la ecuación $\mathbf{x} = x_3\mathbf{v}$ (con x_3 libre), o $\mathbf{x} = t\mathbf{v}$ (con t en $\mathbb{R}$), es la ecuación vectorial paramétrica de una recta. Siempre que un conjunto solución se describa explícitamente con vectores como en los ejemplos 1 y 2, se dirá que la solución está en **forma vectorial paramétrica**.

Soluciones de sistemas no homogéneos

Cuando un sistema lineal no homogéneo tiene muchas soluciones, la solución general se puede escribir en forma vectorial paramétrica como un vector más una combinación lineal arbitraria de vectores que satisfagan el sistema homogéneo correspondiente.

EJEMPLO 3 Describa todas las soluciones de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$$

SOLUCIÓN Aquí A es la matriz de coeficientes del ejemplo 1. Al efectuar operaciones de fila sobre $[A \quad \mathbf{b}]$ se obtiene

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 \quad - \tfrac{4}{3}x_3 &= -1 \\ x_2 \qquad &= 2 \\ 0 &= 0 \end{aligned}$$

Por lo tanto, $x_1 = -1 + \frac{4}{3}x_3,\ x_2 = 2$, $x_2 = 2$ y x_3 es libre. Como un vector, la solución general de $A\mathbf{x} = \mathbf{b}$ tiene la forma

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 + \frac{4}{3}x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} = \underset{\substack{\uparrow \\ \mathbf{p}}}{\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}} + x_3 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}}$$

La ecuación $\mathbf{x} = \mathbf{p} + x_3\mathbf{v}$, o, escribiendo t como un parámetro general,

$$\mathbf{x} = \mathbf{p} + t\mathbf{v} \quad (t \text{ en } \mathbb{R}) \tag{3}$$

describe el conjunto solución de $A\mathbf{x} = \mathbf{b}$ en forma vectorial paramétrica. Recuerde del ejemplo 1 que el conjunto solución de $A\mathbf{x} = \mathbf{0}$ tiene la ecuación vectorial paramétrica

$$\mathbf{x} = t\mathbf{v} \quad (t \text{ en } \mathbb{R}) \tag{4}$$

[con la misma $\mathbf{v}$ que aparece en (3)]. Así, las soluciones de $A\mathbf{x} = \mathbf{b}$ se obtienen sumando el vector $\mathbf{p}$ a las soluciones de $A\mathbf{x} = \mathbf{0}$. El propio vector $\mathbf{p}$ es justamente una solución particular de $A\mathbf{x} = \mathbf{b}$ [correspondiente a $t = 0$ en (3)]. ■

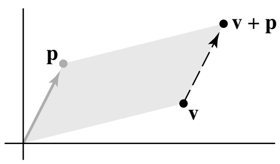


FIGURA 3
La suma de $\mathbf{p}$ a $\mathbf{v}$ traslada $\mathbf{v}$ a $\mathbf{v} + \mathbf{p}$.

Para describir el conjunto solución de $A\mathbf{x} = \mathbf{b}$ en forma geométrica, podemos pensar que la suma vectorial es una *traslación*. Dados $\mathbf{v}$ y $\mathbf{p}$ en $\mathbb{R}^2$ o $\mathbb{R}^3$, el efecto de sumar $\mathbf{p}$ a $\mathbf{v}$ es *mover* a $\mathbf{v}$ en una dirección paralela a la recta que pasa por $\mathbf{p}$ y $\mathbf{0}$. Se dice que **p traslada a v** hacia $\mathbf{v} + \mathbf{p}$. Véase la figura 3. Si cada punto sobre la recta L en $\mathbb{R}^2$ o $\mathbb{R}^3$ es trasladado por un vector $\mathbf{p}$, el resultado es una recta paralela a L. Véase la figura 4.

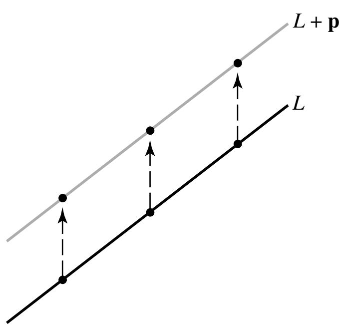


FIGURA 4
Recta trasladada.

Suponga que L es la recta que pasa por $\mathbf{0}$ y $\mathbf{v}$, descrita mediante la ecuación (4). Sumar $\mathbf{p}$ a cada punto sobre L produce la recta trasladada que describe la ecuación (3). Observe que $\mathbf{p}$ está sobre la recta en la ecuación (3). Así, (3) es **la ecuación de la recta que pasa por p paralela a v.** Entonces *el conjunto solución de* $A\mathbf{x} = \mathbf{b}$ *es una recta que pasa por* $\mathbf{p}$ *paralela al conjunto solución de* $A\mathbf{x} = \mathbf{0}$. La figura 5 ilustra este caso.

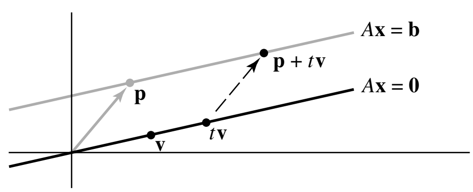


FIGURA 5 Conjuntos solución paralelos de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$.

La relación entre los conjuntos solución de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$, que se ilustra en la figura 5, se generaliza a cualquier ecuación *consistente* $A\mathbf{x} = \mathbf{b}$, aunque el conjunto solución será más grande que una recta cuando existan muchas variables libres. El siguiente teorema da el enunciado preciso. Para una demostración, véase el ejercicio 25.

TEOREMA 6

Suponga que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para alguna $\mathbf{b}$ dada, y sea $\mathbf{p}$ una solución. El conjunto solución de $A\mathbf{x} = \mathbf{b}$ es el conjunto de todos los vectores de la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{v}_h$ es cualquier solución de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

El teorema 6 dice que si $A\mathbf{x} = \mathbf{b}$ tiene una solución, entonces el conjunto solución se obtiene al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{0}$, empleando cualquier solución particular $\mathbf{p}$ de $A\mathbf{x} = \mathbf{b}$ para dicha traslación. La figura 6 ilustra el caso en que existen dos variables libres. Aun cuando $n > 3$, nuestra imagen mental del conjunto solución de un sistema consistente $A\mathbf{x} = \mathbf{b}$ (con $\mathbf{b} \neq \mathbf{0}$) es un solo punto diferente de cero, o una recta o un plano que no pasan por el origen.

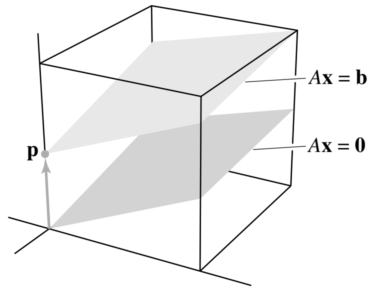


FIGURA 6 Conjuntos solución paralelos de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$.

Advertencia: El teorema 6 y la figura 6 solamente se aplican a una ecuación $A\mathbf{x} = \mathbf{b}$ que tiene al menos una solución $\mathbf{p}$ diferente de cero. Cuando $A\mathbf{x} = \mathbf{b}$ no tiene solución, entonces el conjunto solución es vacío.

El siguiente algoritmo resume los cálculos que se muestran en los ejemplos 1, 2 y 3.

ESCRITURA DE UN CONJUNTO SOLUCIÓN (DE UN SISTEMA CONSISTENTE) EN FORMA VECTORIAL PARAMÉTRICA

1. Reduzca por filas la matriz aumentada a su forma escalonada reducida.
2. Exprese cada variable básica en términos de cualquiera de las variables libres presentes en una ecuación.
3. Escriba una solución típica $\mathbf{x}$ como un vector cuyas entradas dependen de las variables libres (si las hay).
4. Descomponga $\mathbf{x}$ en una combinación lineal de vectores (con entradas numéricas) empleando las variables libres como parámetros.

PROBLEMAS DE PRÁCTICA

1. Cada una de las siguientes ecuaciones determina un plano en $\mathbb{R}^3$. ¿Se intersecan los dos planos? Si es así, describa su intersección

$$\begin{aligned} x_1 + 4x_2 - 5x_3 &= 0 \\ 2x_1 - x_2 + 8x_3 &= 9 \end{aligned}$$

2. Escriba la solución general de $10x_1 - 3x_2 - 2x_3 = 7$ en forma vectorial paramétrica, y relacione el conjunto solución con el obtenido en el ejemplo 2.

1.5 EJERCICIOS

En los ejercicios 1 a 4, determine si el sistema tiene una solución no trivial. Intente usar lo menos posible las operaciones de fila.

1. $$\begin{aligned} 2x_1 - 5x_2 + 8x_3 &= 0 \\ -2x_1 - 7x_2 + x_3 &= 0 \\ 4x_1 + 2x_2 + 7x_3 &= 0 \end{aligned}$$

2. $$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ -2x_1 - 3x_2 - 4x_3 &= 0 \\ 2x_1 - 4x_2 + 9x_3 &= 0 \end{aligned}$$

3. $$\begin{aligned} -3x_1 + 4x_2 - 8x_3 &= 0 \\ -2x_1 + 5x_2 + 4x_3 &= 0 \end{aligned}$$

4. $$\begin{aligned} 5x_1 - 3x_2 + 2x_3 &= 0 \\ -3x_1 - 4x_2 + 2x_3 &= 0 \end{aligned}$$

En los ejercicios 5 y 6, siga el método de los ejemplos 1 y 2 para escribir el conjunto solución del sistema homogéneo en forma vectorial paramétrica.

5. $$\begin{aligned} 2x_1 + 2x_2 + 4x_3 &= 0 \\ -4x_1 - 4x_2 - 8x_3 &= 0 \\ -3x_2 - 3x_3 &= 0 \end{aligned}$$

6. $$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 0 \\ 2x_1 + x_2 - 3x_3 &= 0 \\ -1x_1 + x_2 &= 0 \end{aligned}$$

En los ejercicios 7 a 12, describa todas las soluciones de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica, donde A es equivalente por filas a la matriz dada.

7. $\begin{bmatrix} 1 & 3 & -3 & 7 \\ 0 & 1 & -4 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -3 & -8 & 5 \\ 0 & 1 & 2 & -4 \end{bmatrix}$

9. $\begin{bmatrix} 3 & -6 & 6 \\ -2 & 4 & -2 \end{bmatrix}$

10. $\begin{bmatrix} -1 & -4 & 0 & -4 \\ 2 & -8 & 0 & 8 \end{bmatrix}$

11. $\begin{bmatrix} 1 & -4 & -2 & 0 & 3 & -5 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 1 & -2 & 3 & -6 & 5 & 0 \\ 0 & 0 & 0 & 1 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

13. Suponga que el conjunto solución de un cierto sistema de ecuaciones lineales se describe como $x_1 = 5 + 4x_3$, $x_2 = -2 - 7x_3$, con x_3 libre. Use vectores para describir este conjunto como una recta en $\mathbb{R}^3$.

14. Suponga que el conjunto solución de un cierto sistema de ecuaciones lineales se describe como $x_1 = 5x_4$, $x_2 = 3 - 2x_4$, $x_3 = 2 + 5x_4$, con x_4 libre. Utilice vectores para describir este conjunto como una "recta" en $\mathbb{R}^4$.

15. Describa y compare los conjuntos solución de $x_1 + 5x_2 - 3x_3 = 0$ y $x_1 + 5x_2 - 3x_3 = -2$.

16. Describa y compare los conjuntos solución de $x_1 - 2x_2 + 3x_3 = 0$ y $x_1 - 2x_2 + 3x_3 = 4$.

17. Siga el método del ejemplo 3 para describir las soluciones del siguiente sistema en forma vectorial paramétrica. También, dé una descripción geométrica del conjunto solución y compárelo con el que obtuvo en el ejercicio 5.

$$\begin{aligned} 2x_1 + 2x_2 + 4x_3 &= 8 \\ -4x_1 - 4x_2 - 8x_3 &= -16 \\ -3x_2 - 3x_3 &= 12 \end{aligned}$$

18. Como en el ejercicio 17, describa las soluciones del siguiente sistema en forma vectorial paramétrica, y realice una comparación geométrica con el conjunto solución del ejercicio 6.

$$\begin{aligned} x_1 + 2x_2 - 3x_3 &= 5 \\ 2x_1 + x_2 - 3x_3 &= 13 \\ -x_1 + x_2 &= -8 \end{aligned}$$

En los ejercicios 19 y 20, encuentre la ecuación paramétrica de la recta que pasa por **a** y es paralela a **b**.

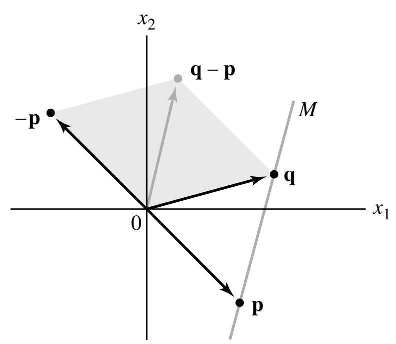
19. $\mathbf{a} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

20. $\mathbf{a} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -7 \\ 6 \end{bmatrix}$

En los ejercicios 21 y 22, obtenga una ecuación paramétrica de la recta M que pasa a través de **p** y **q**. [*Sugerencia:* M es paralela al vector $\mathbf{q} - \mathbf{p}$. Véase la figura que aparece más abajo].

21. $\mathbf{p} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

22. $\mathbf{p} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$



En los ejercicios 23 y 24, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) Una ecuación homogénea siempre es consistente.

b) La ecuación $A\mathbf{x} = \mathbf{0}$ da una descripción explícita de su conjunto solución.

c) La ecuación homogénea $A\mathbf{x} = \mathbf{0}$ tiene la solución trivial si y solo si la ecuación tiene al menos una variable libre.

d) La ecuación $\mathbf{x} = \mathbf{p} + t\mathbf{v}$ describe una recta que pasa por **v** y es paralela a **p**.

e) El conjunto solución de $A\mathbf{x} = \mathbf{b}$ es el conjunto de todos los vectores de la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{v}_h$ es cualquier solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

24. *a*) Un sistema homogéneo de ecuaciones puede ser inconsistente.

b) Si **x** es una solución no trivial de $A\mathbf{x} = \mathbf{0}$, entonces cada entrada en **x** es distinta de cero.

c) El efecto de sumar **p** a un vector es mover a dicho vector en una dirección paralela a **p**.

d) La ecuación $A\mathbf{x} = \mathbf{b}$ es homogénea si el vector cero es una solución.

e) Si $A\mathbf{x} = \mathbf{b}$ es consistente, entonces el conjunto solución de $A\mathbf{x} = \mathbf{b}$ se obtiene por traslación del conjunto solución de $A\mathbf{x} = \mathbf{0}$.

25. Demuestre el teorema 6:

a) Suponga que $\mathbf{p}$ es una solución de $A\mathbf{x} = \mathbf{b}$, de manera que $A\mathbf{p} = \mathbf{b}$. Sea $\mathbf{v}_h$ cualquier solución de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$, y sea $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$. Demuestre que $\mathbf{w}$ es una solución de $A\mathbf{x} = \mathbf{b}$.

b) Sea $\mathbf{w}$ cualquier solución de $A\mathbf{x} = \mathbf{b}$, y defina $\mathbf{v}_h = \mathbf{w} - \mathbf{p}$. Demuestre que $\mathbf{v}_h$ es una solución de $A\mathbf{x} = \mathbf{0}$. Esto demuestra que cada solución de $A\mathbf{x} = \mathbf{b}$ tiene la forma $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, donde $\mathbf{p}$ es una solución particular de $A\mathbf{x} = \mathbf{b}$ y $\mathbf{v}_h$ una solución de $A\mathbf{x} = \mathbf{0}$.

26. Suponga que A es la matriz *cero* de 3×3 (con cero en todas las entradas). Describa el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

27. Suponga que $A\mathbf{x} = \mathbf{b}$ tiene una solución. Explique por qué la solución es única precisamente cuando $A\mathbf{x} = \mathbf{0}$ tiene solo la solución trivial.

En los ejercicios 28 a 31, *a*) ¿la ecuación $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial? *b*) ¿La ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda posible $\mathbf{b}$?

28. A es una matriz de 3×3 con tres posiciones pivote.

29. A es una matriz de 4×4 con tres posiciones pivote.

30. A es una matriz de 2×5 con dos posiciones pivote.

31. A es una matriz de 3×2 con dos posiciones pivote.

32. Si $\mathbf{b} \neq \mathbf{0}$, ¿el conjunto solución de $A\mathbf{x} = \mathbf{b}$ puede ser un plano que pasa por el origen? Explique su respuesta.

33. Construya una matriz A, diferente de cero, de 3×3 tal que el vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ sea una solución de $A\mathbf{x} = \mathbf{0}$.

34. Construya una matriz A, diferente de cero, de 3×3 tal que el vector $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ sea una solución de $A\mathbf{x} = \mathbf{0}$.

35. A partir de $A = \begin{bmatrix} -1 & -3 \\ 7 & 21 \\ -2 & -6 \end{bmatrix}$, encuentre por inspección una solución no trivial de $A\mathbf{x} = \mathbf{0}$. [*Sugerencia:* Piense en la ecuación $A\mathbf{x} = \mathbf{0}$ escrita como una ecuación vectorial].

36. A partir de $A = \begin{bmatrix} 3 & -2 \\ -6 & 4 \\ 12 & -8 \end{bmatrix}$, encuentre por inspección una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

37. Construya una matriz A de 2×2 tal que el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$ sea la recta en $\mathbb{R}^2$ que pasa a través de (4, 1) y el origen. Luego, encuentre un vector $\mathbf{b}$ en $\mathbb{R}^2$ tal que el conjunto solución de $A\mathbf{x} = \mathbf{b}$ *no* sea una recta en $\mathbb{R}^2$ paralela al conjunto solución de $A\mathbf{x} = \mathbf{0}$. ¿Por qué esto *no* contradice al teorema 6?

38. Sean A una matriz de $m \times n$, y $\mathbf{w}$ un vector en $\mathbb{R}^n$ que satisface la ecuación $A\mathbf{x} = \mathbf{0}$. Demuestre que para cualquier escalar c, el vector $c\mathbf{w}$ también satisface $A\mathbf{x} = \mathbf{0}$. [Es decir, demuestre que $A(c\mathbf{w}) = \mathbf{0}$].

39. Suponga que A es una matriz de $m \times n$, y que $\mathbf{v}$ y $\mathbf{w}$ son vectores en $\mathbb{R}^n$ tales que $A\mathbf{v} = \mathbf{0}$ y $A\mathbf{w} = \mathbf{0}$. Explique por qué $A(\mathbf{v} + \mathbf{w})$ debe ser el vector cero. Luego, explique por qué $A(c\mathbf{v} + d\mathbf{w}) = \mathbf{0}$ para cada par de escalares c y d.

40. Suponga que A es una matriz de 3×3 y $\mathbf{b}$ es un vector en $\mathbb{R}^3$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ *no* tiene solución. ¿Existe un vector $\mathbf{y}$ en $\mathbb{R}^3$ tal que la ecuación $A\mathbf{x} = \mathbf{y}$ tiene una solución única? Justifique su respuesta.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & 4 & -5 & 0 \\ 2 & -1 & 8 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & -5 & 0 \\ 0 & -9 & 18 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -2 & -1 \end{bmatrix}$$

$$\begin{aligned} x_1 \quad + 3x_3 &= 4 \\ x_2 - 2x_3 &= -1 \end{aligned}$$

Por lo tanto, $x_1 = 4 - 3x_3$, $x_2 = -1 + 2x_3$, con x_3 libre. La solución general en forma vectorial paramétrica es

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 - 3x_3 \\ -1 + 2x_3 \\ x_3 \end{bmatrix} = \underset{\uparrow \atop \mathbf{p}}{\begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix}} + x_3 \underset{\uparrow \atop \mathbf{v}}{\begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}}$$

La intersección de los dos planos es la recta que pasa por $\mathbf{p}$ en la dirección de $\mathbf{v}$.

2. La matriz aumentada $[10 \quad -3 \quad -2 \quad 7]$ es equivalente por filas a $[1 \quad -.3 \quad -.2 \quad .7]$, y la solución general es $x_1 = .7 + .3x_2 + .2x_3$, con x_2 y x_3 libres. Es decir,

$$\begin{aligned}\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} .7 + .3x_2 + .2x_3 \\ x_2 \\ x_3 \end{bmatrix} &= \begin{bmatrix} .7 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} .3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} .2 \\ 0 \\ 1 \end{bmatrix} \\ &= \mathbf{p} + x_2\mathbf{u} + x_3\mathbf{v}\end{aligned}$$

El conjunto solución de la ecuación no homogénea $A\mathbf{x} = \mathbf{b}$ es el plano trasladado $\mathbf{p} + \text{Gen}\,\{\mathbf{u}, \mathbf{v}\}$, que pasa por $\mathbf{p}$ y es paralelo al conjunto solución de la ecuación homogénea del ejemplo 2.

1.6 APLICACIONES DE SISTEMAS LINEALES

Tal vez usted espere que un problema de la vida real que implica álgebra lineal tenga solo una solución, o quizá ninguna solución**.** La finalidad de esta sección es mostrar cómo, de manera natural, surgen sistemas lineales con muchas soluciones. Aquí las aplicaciones provienen de áreas como economía, química y flujo en redes.

Un sistema homogéneo en economía

WEB

El sistema de 500 ecuaciones con 500 variables, que se mencionó en la introducción de este capítulo, ahora se conoce como un modelo de “entrada-salida” (o “producción”) de Leontief.[1] En la sección 2.6 se examinará este modelo con más detalle, cuando se disponga de más bases teóricas y de una mejor notación. Por ahora, se considerará un “modelo de intercambio” más sencillo, que también se debe a Leontief.

Suponga que la economía de una nación se divide en muchos sectores, como manufactura, comunicaciones, entretenimiento y servicios. Considere que se conoce la producción total anual de cada sector y se sabe exactamente cómo esta producción se divide o se “intercambia” entre los otros sectores de la economía. Al valor total en dólares de la producción de un sector se le llama **precio** de esa producción. Leontief probó el siguiente resultado.

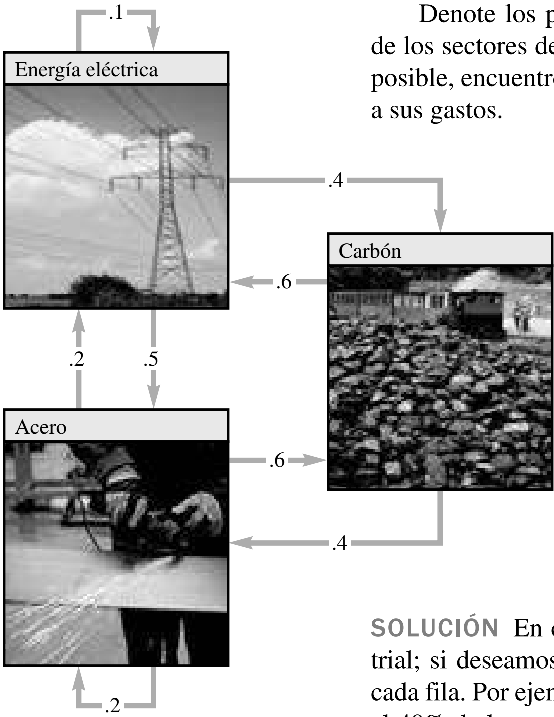
> Existen *precios de equilibrio* que se pueden asignar a las producciones totales de varios sectores, de tal forma que el ingreso de cada sector equilibra exactamente sus gastos.

El siguiente ejemplo ilustra cómo encontrar los precios de equilibrio.

EJEMPLO 1 Suponga que una economía comprende las industrias carbonífera, eléctrica y del acero, y que la producción de cada sector se distribuye entre los diversos sectores como se muestra en la tabla 1, página 50: las entradas en una columna representan las partes fraccionales de la producción total de un sector industrial.

La segunda columna de la tabla 1, por ejemplo, dice que la producción total del sector eléctrico se divide como sigue: 40% a la industria del carbón, 50% a la del acero, y el restante 10% a la industria eléctrica. (El sector eléctrico trata a este 10% como un gasto en el que se incurre con la finalidad de operar su negocio). Como se deben considerar todas las producciones, las fracciones decimales en cada columna deben sumar 1.

[1] Véase Wassily W. Leontief, “Input-Output Economics”, *Scientific American*, octubre de 1951, pp. 15-21.



Denote los precios (es decir, valores en dólares) del total de las producciones anuales de los sectores del carbón, eléctrico y del acero mediante p_C, p_E y p_S, respectivamente. Si es posible, encuentre los precios de equilibrio que hacen que los ingresos de cada sector igualen a sus gastos.

TABLA 1 Una economía básica

Distribución de la producción por sectores:			
Del carbón	**Eléctrico**	**Del acero**	**Comprada por:**
.0	.4	.6	S. del carbón
.6	.1	.2	S. eléctrico
.4	.5	.2	S. del acero

SOLUCIÓN En cada columna se indica hacia dónde va la producción de cada sector industrial; si deseamos saber qué necesita cada sector como insumos, hay que leer a lo largo de cada fila. Por ejemplo, la primera fila de la tabla 1 dice que el sector del carbón recibe (y paga) el 40% de la producción del sector eléctrico y el 60% del producto de la industria del acero. Puesto que los respectivos valores de las producciones totales son p_E y p_S, la industria del carbón debe gastar $.4p_E$ dólares por compartir el producto del sector eléctrico y $.6p_S$ por la producción de la industria del acero. Así, los gastos totales del sector carbonífero son $.4p_E + .6p_S$. Para hacer que el ingreso de la industria del carbón, p_C, sea igual a sus gastos, escribimos

$$p_C = .4p_E + .6p_S \tag{1}$$

La segunda fila de la tabla de intercambio indica que el sector eléctrico gasta $0.6p_C$ en la industria del carbón, $0.1p_E$ en energía eléctrica, y $0.2p_S$ en la industria del acero. Así que el requerimiento de ingreso/gasto para el sector eléctrico es

$$p_E = .6p_C + .1p_E + .2p_S \tag{2}$$

Finalmente, la tercera fila de la tabla de intercambio conduce al último requerimiento:

$$p_S = .4p_C + .5p_E + .2p_S \tag{3}$$

Para resolver el sistema de ecuaciones (1), (2) y (3), se transfieren todas las incógnitas a los lados izquierdos de las ecuaciones y se combinan términos semejantes. [Por ejemplo, en el lado izquierdo de (2), se escribe $p_E - .1p_E$ como $.9p_E$.]

$$\begin{aligned} p_C - .4p_E - .6p_S &= 0 \\ -.6p_C + .9p_E - .2p_S &= 0 \\ -.4p_C - .5p_E + .8p_S &= 0 \end{aligned}$$

El siguiente paso es reducir por filas. Aquí, para simplificar, los decimales se redondean a dos cifras.

$$\begin{bmatrix} 1 & -.4 & -.6 & 0 \\ -.6 & .9 & -.2 & 0 \\ -.4 & -.5 & .8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & -.66 & .56 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -.94 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general es $p_C = .94p_S$, $p_E = .85p_S$, y p_S es libre. El vector de precios de equilibrio para la economía tiene la forma

$$\mathbf{p} = \begin{bmatrix} p_C \\ p_E \\ p_S \end{bmatrix} = \begin{bmatrix} .94p_S \\ .85p_S \\ p_S \end{bmatrix} = p_S \begin{bmatrix} .94 \\ .85 \\ 1 \end{bmatrix}$$

Cualquier asignación (no negativa) para p_S da por resultado una asignación de precios de equilibrio. Por ejemplo, si se toma p_S como 100 (o \$100 millones), entonces $p_C = 94$ y $p_E = 85$. Los ingresos y gastos de cada sector serán iguales si la producción de carbón se cotiza en \$94 millones, la de energía eléctrica en \$85 millones, y la de acero en \$100 millones. ■

Balanceo de ecuaciones químicas

Las ecuaciones químicas describen las cantidades de sustancias que se consumen y producen en reacciones químicas. Por ejemplo, cuando el gas propano se quema, el propano (C_3H_8) se combina con oxígeno (O_2) para formar dióxido de carbono (CO_2) y agua (H_2O), de acuerdo con una ecuación de la forma

$$(x_1)C_3H_8 + (x_2)O_2 \rightarrow (x_3)CO_2 + (x_4)H_2O \tag{4}$$

Para "balancear" esta ecuación, un químico debe encontrar números $x_1, \ldots, x_4$ tales que los números totales de átomos de carbón (C), hidrógeno(H) y oxígeno (O) en el lado izquierdo concuerden con los números de átomos correspondientes en el lado derecho (porque, en la reacción, los átomos no se crean ni se destruyen).

Un método sistemático para balancear ecuaciones químicas es colocar una ecuación vectorial que describa el número de átomos de cada tipo presentes en una reacción. Como la ecuación (4) implica a tres tipos de átomos (carbón, hidrógeno y oxígeno), construya un vector en $\mathbb{R}^3$ para cada reactante y producto en (4) que liste los números de "átomos por molécula", como sigue:

$$C_3H_8\colon \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix},\ O_2\colon \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix},\ CO_2\colon \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix},\ H_2O\colon \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \begin{array}{l} \leftarrow \text{Carbón} \\ \leftarrow \text{Hidrógeno} \\ \leftarrow \text{Oxígeno} \end{array}$$

Para balancear la ecuación (4), los coeficientes $x_1, \ldots, x_4$ deben satisfacer

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

Para resolver, mueva todos los términos a la izquierda (cambiando los signos en los vectores tercero y cuarto):

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

La reducción por filas de la matriz aumentada para esta ecuación conduce a la solución general

$$x_1 = \tfrac{1}{4}x_4,\ x_2 = \tfrac{5}{4}x_4,\ x_3 = \tfrac{3}{4}x_4, \text{ con } x_4 \text{ libre}$$

Como los coeficientes en una ecuación química deben ser enteros, tome $x_4 = 4$; en tal caso, $x_1 = 1$, $x_2 = 5$ y $x_3 = 3$. La ecuación balanceada es

$$C_3H_8 + 5O_2 \rightarrow 3CO_2 + 4H_2O$$

La ecuación también estaría balanceada si, por ejemplo, cada coeficiente se duplicara. Sin embargo, para la mayoría de los propósitos, los químicos prefieren utilizar una ecuación balanceada cuyos coeficientes sean los números más pequeños posibles.

Flujo de redes

WEB

Los sistemas de ecuaciones lineales se originan naturalmente cuando científicos, ingenieros o economistas estudian el flujo de alguna cantidad a través de una red. Por ejemplo, los planificadores urbanos y los ingenieros de tráfico monitorizan el patrón de flujo de tránsito en una rejilla de las calles de la ciudad. Los ingenieros eléctricos calculan el flujo de corriente a través de circuitos eléctricos. Y los economistas analizan la distribución de productos del fabricante al consumidor a través de una red de mayoristas y minoristas. Para muchas redes, los sistemas de ecuaciones implican cientos o incluso miles de variables y ecuaciones.

Una *red* consiste en un conjunto de puntos llamados *uniones*, o *nodos*, con líneas o arcos llamados *ramas*, que conectan algunos o todos los nodos. Se indica la dirección y el sentido de flujo en cada rama, y la cantidad de flujo (o tasa) se denota con una variable.

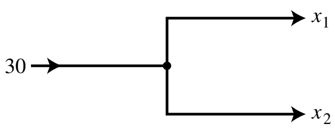


FIGURA 1
Una unión, también llamada nodo.

La suposición básica en el flujo de red es que el flujo total en la red es igual al flujo total de salida de la red, y que el flujo total en un nodo es igual al flujo total de salida en dicho nodo. Por ejemplo, la figura 1 muestra 30 unidades que fluyen por una rama hacia una unión; x_1 y x_2 denotan los flujos de salida del nodo a través de otras ramas. Como el flujo se "conserva" en cada unión, entonces $x_1 + x_2 = 30$. En forma similar, el flujo en cada nodo se describe mediante una ecuación lineal. El problema de análisis de redes es determinar el flujo en cada rama cuando se tiene información parcial (como el flujo de entrada y salida de la red).

EJEMPLO 2 La red de la figura 2 representa el flujo del tránsito (en vehículos por hora) en varias calles de un solo sentido en el centro de Baltimore en un día común, poco después del mediodía. Determine el patrón de flujo general para la red.

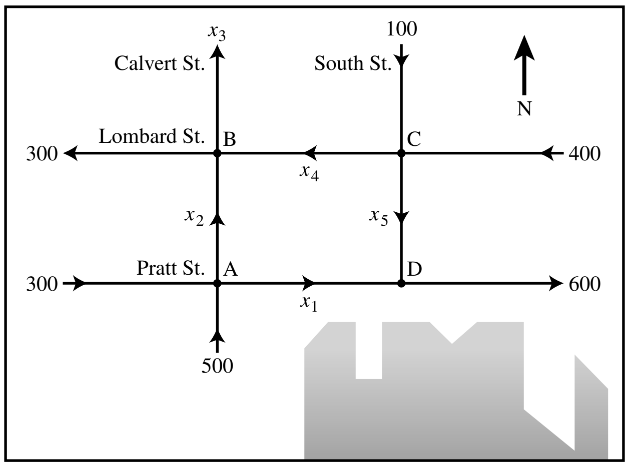


FIGURA 2 Calles de Baltimore.

SOLUCIÓN Escriba las ecuaciones que describen el flujo, y después encuentre la solución general del sistema. Marque las intersecciones de las calles (nodos) y los flujos desconocidos en las ramas, como se indica en la figura 2. En cada intersección, iguale el flujo de entrada al de salida.

Intersección	Flujo de entrada		Flujo de salida
A	$300 + 500$	$=$	$x_1 + x_2$
B	$x_2 + x_4$	$=$	$300 + x_3$
C	$100 + 400$	$=$	$x_4 + x_5$
D	$x_1 + x_5$	$=$	600

También, el flujo total de entrada en la red (500 + 300 + 100 + 400) es igual al flujo total de salida (300 + x_3 + 600), que se simplifica a $x_3 = 400$. Combine esta ecuación con un reordenamiento de las primeras cuatro ecuaciones para obtener el siguiente sistema de ecuaciones:

$$\begin{array}{rcrcrcrcrcl} x_1 & + & x_2 & & & & & & & = & 800 \\ & & x_2 & - & x_3 & + & x_4 & & & = & 300 \\ & & & & & & x_4 & + & x_5 & = & 500 \\ x_1 & & & & & & & + & x_5 & = & 600 \\ & & & & x_3 & & & & & = & 400 \end{array}$$

La reducción por filas de la matriz aumentada conduce a

$$\begin{array}{rcrcrcrcrcl} x_1 & & & & & & & + & x_5 & = & 600 \\ & & x_2 & & & & & - & x_5 & = & 200 \\ & & & & x_3 & & & & & = & 400 \\ & & & & & & x_4 & + & x_5 & = & 500 \end{array}$$

El patrón de flujo general para la red está descrito por

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ es libre} \end{cases}$$

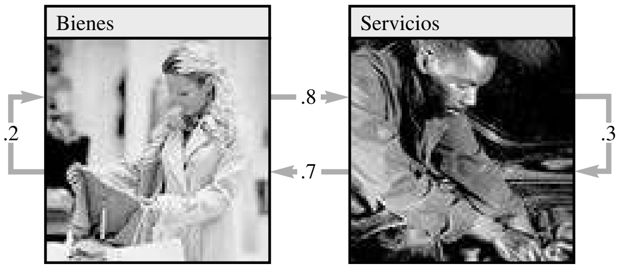
Un flujo negativo en una rama de la red corresponde a un flujo en el sentido opuesto al indicado en el modelo. Como las calles en este problema son de un solo sentido, ninguna de las variables puede ser negativa. Este hecho conduce a ciertas limitaciones sobre los posibles valores de las variables. Por ejemplo, $x_5 \leq 500$ porque x_4 no puede ser negativa. En el problema de práctica 2 se consideran otras restricciones sobre las variables. ■

PROBLEMAS DE PRÁCTICA

1. Suponga que una economía tiene tres sectores: agricultura, minería y manufactura. Agricultura vende el 5% de su producción a minería y el 30% a manufactura, y retiene el resto. Minería vende el 20% de su producto a agricultura y el 70% a manufactura, conservando el resto. Manufactura vende el 20% de su producción a agricultura y el 30% a minería, y retiene lo restante. Determine la tabla de intercambio para esta economía; utilice las columnas para describir cómo la producción de cada sector se intercambia entre los tres sectores.
2. Considere el flujo de red que se analizó en el ejemplo 2. Determine el posible rango de valores de x_1 y x_2. [*Sugerencia:* El ejemplo mencionó que $x_5 \leq 500$. ¿Qué implicaciones tiene esto sobre x_1 y x_2? Además, considere el hecho de que $x_5 \geq 0$].

1.6 EJERCICIOS

1. Suponga que una economía solo tiene dos sectores: bienes y servicios. Cada año, el sector de bienes vende el 80% de su producción al de servicios y retiene el resto, mientras que el sector de servicios vende el 70% de su producción al sector de bienes y conserva lo restante. Encuentre los precios de equilibrio para las producciones anuales de los sectores de bienes y servicios que permiten igualar el ingreso con el gasto de cada sector.



2. Encuentre otro conjunto de precios de equilibrio para la economía del ejemplo 1. Suponga la misma economía, pero que ahora se utiliza el yen japonés en vez del dólar estadounidense para medir los valores de las diversas producciones de los sectores. ¿Esto modificaría el problema en algún sentido? Analícelo.

3. Considere una economía con tres sectores: combustibles y energía, manufactura y servicios. Combustibles y energía vende el 80% de su producción a manufactura, el 10% a servicios, y retiene el resto. Manufactura vende el 10% de su producción a combustibles y energía, el 80% a servicios, y conserva lo restante. Servicios vende el 20% a combustible y energía, el 40% a manufactura, y retiene el resto.

   a) Construya la tabla de intercambio para esta economía.

   b) Desarrolle un sistema de ecuaciones que permita determinar los precios con los cuales se igualen los ingresos y gastos de cada sector. Después, escriba la matriz aumentada que puede reducirse por filas para obtener esos precios.

   c) [**M**] Encuentre un conjunto de precios de equilibrio cuando el precio para la producción de servicios es de 100 unidades.

4. Suponga que una economía tiene cuatro sectores: minero, maderero, de energía y del transporte. Minería vende el 10% de su producción al sector maderero, el 60% a energía, y conserva el resto. El sector maderero vende el 15% de su producción a minería, el 50% a energía, el 20% al sector del transporte, y conserva lo restante. Energía vende el 20% de su producción a minería, el 15% al sector maderero, el 20% al sector del transporte, y retiene el resto. El sector del transporte vende el 20% de su salida de producción a minería, el 10% al sector maderero, el 50% a energía, y conserva lo restante.

   a) Construya la tabla de intercambio para esta economía.

   b) [**M**] Encuentre un conjunto de precios de equilibrio para esta economía.

5. Una economía tiene cuatro sectores: agricultura, manufactura, servicios y transporte. El sector agrícola vende el 20% de su producción a manufactura, el 30% a servicios, el 30% al sector del transporte, y conserva el resto. Manufactura vende el 35% de su producción al sector agrícola, el 35% a servicios, el 20% al sector del transporte, y conserva el resto. Servicios vende el 10% de su salida de producción a agricultura, el 20% a manufactura, el 20% al sector del transporte, y conserva el resto. El sector del transporte vende el 20% de su salida de producción a agricultura, el 30% a manufactura, el 20% a servicios, y conserva el remanente.

   a) Construya la tabla de intercambio para esta economía.

   b) [**M**] Encuentre un conjunto de precios de equilibrio para la economía si el valor del producto del sector del transporte es de \$10.00 por unidad.

   c) El sector de servicios lanza una exitosa campaña para "comer productos frescos de granja", e incrementa su participación con el sector agrícola al 40%, mientras que su participación con el sector manufacturero cae al 10%. Construya la tabla de intercambio para esta nueva economía.

   d) [**M**] Encuentre un conjunto de precios de equilibrio para esta nueva economía si el valor del sector del transporte continúa a \$10.00 por unidad. ¿Qué efectos tuvo la campaña de "comer productos frescos de granja" sobre los precios de equilibrio para los sectores en esta economía?

En los ejercicios 6 a 11, balancee las ecuaciones químicas utilizando el enfoque de ecuación vectorial analizado en esta sección.

6. El óxido de aluminio y el carbón reaccionan para crear el elemento aluminio y dióxido de carbono:

   $$Al_2O_3 + C \rightarrow Al + CO_2$$

   [Para cada compuesto, construya un vector que liste los números de átomos de aluminio, oxígeno y carbón].

7. El Alka-Seltzer contiene bicarbonato de sodio ($NaHCO_3$) y ácido cítrico ($H_3C_6H_5O_7$). Cuando se disuelve una tableta en agua, la siguiente reacción produce citrato de sodio, agua y dióxido de carbono (gas):

   $$NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + H_2O + CO_2$$

8. La piedra caliza, $CaCO_3$, neutraliza el ácido, H_3O, en la lluvia ácida, mediante la siguiente ecuación sin balancear:

   $$H_3O + CaCO_3 \rightarrow H_2O + Ca + CO_2$$

9. El sulfhídrico de boro reacciona violentamente con agua para formar ácido bórico y gas sulfhídrico de hidrógeno (que expide olor a huevo podrido). La ecuación sin balancear es

   $$B_2S_3 + H_2O \rightarrow H_3BO_3 + H_2S$$

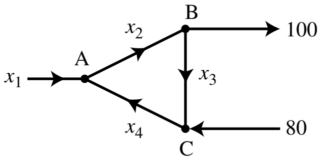
10. [**M**] Si es posible, utilice aritmética exacta o un formato racional para los cálculos al balancear la siguiente reacción química:

    $$PbN_6 + CrMn_2O_8 \rightarrow Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$$

11. [**M**] La reacción química que aparece a continuación se puede utilizar en algunos procesos industriales, como la producción de arsénico (AsH_3). Utilice aritmética exacta o un formato racional para los cálculos al balancear esta ecuación.

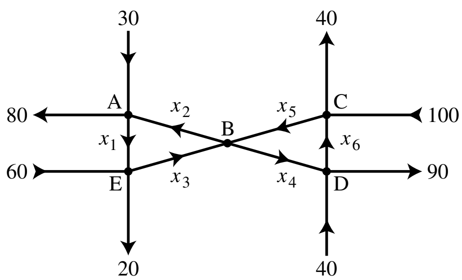
    $$MnS + As_2Cr_{10}O_{35} + H_2SO_4$$
    $$\rightarrow HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

12. Encuentre el patrón de flujo general de la red que se ilustra en la figura. Suponiendo que todos los flujos son no negativos, ¿cuál es el valor más pequeño posible para x_4?



13. *a*) Encuentre el patrón de flujo general de la red que se ilustra en la figura.

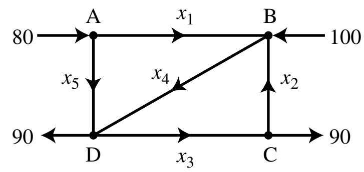
b) Suponiendo que los flujos deben ser en los sentidos indicados, encuentre los flujos mínimos en las ramas denotadas como x_2, x_3, x_4 y x_5.



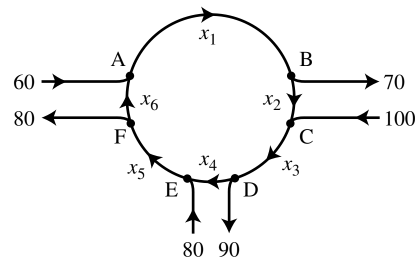
14. *a*) Encuentre el patrón de tráfico general de la red de autopistas que se representa en la figura. (Las tasas de flujo están en vehículos/minuto).

b) Describa el patrón de tráfico general cuando está cerrada la ruta cuyo flujo es x_5.

c) Cuando $x_5 = 0$, ¿cuál es el valor mínimo de x_4?



15. En Inglaterra las intersecciones con frecuencia se construyen como circuitos en forma de glorieta de un solo sentido, como el que se ilustra en la figura. Suponga que el tráfico debe circular en los sentidos indicados. Encuentre la solución general del flujo de red. Determine el valor más pequeño posible para x_6.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Escriba los porcentajes como decimales. Como todas las salidas deben tomarse en cuenta, cada columna debe sumar 1. Este hecho ayuda a llenar cualquier entrada faltante.

Distribución de la producción por sectores:			
Agricultura	**Minería**	**Manufactura**	**Comprada por:**
.65	.20	.20	Agricultura
.05	.10	.30	Minería
.30	.70	.50	Manufactura

2. Como $x_5 \le 500$, las ecuaciones D y A para x_1 y x_2 implican que $x_1 \ge 100$ y $x_2 \le 700$. El hecho de que $x_5 \ge 0$ implica que $x_1 \le 600$ y $x_2 \ge 200$. Así, $100 \le x_1 \le 600$, y $200 \le x_2 \le 700$.

1.7 INDEPENDENCIA LINEAL

Las ecuaciones homogéneas de la sección 1.5 se pueden estudiar desde una perspectiva diferente si las escribimos como ecuaciones vectoriales. De esta manera, la atención se transfiere de las soluciones desconocidas de $A\mathbf{x} = \mathbf{0}$ a los vectores que aparecen en las ecuaciones vectoriales.

Por ejemplo, considere la ecuación

$$x_1\begin{bmatrix}1\\2\\3\end{bmatrix}+x_2\begin{bmatrix}4\\5\\6\end{bmatrix}+x_3\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\tag{1}$$

Esta ecuación tiene una solución trivial, desde luego, donde $x_1 = x_2 = x_3 = 0$. Al igual que en la sección 1.5, el asunto principal es si la solución trivial es la *única*.

DEFINICIÓN

Se dice que un conjunto indexado de vectores $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ en $\mathbb{R}^n$ es **linealmente independiente** si la ecuación vectorial

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

solo tiene la solución trivial. Se dice que el conjunto $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ es **linealmente dependiente** si existen pesos $c_1,\ldots,c_p$, no todos cero, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}\tag{2}$$

La ecuación (2) se llama **relación de dependencia lineal** entre $\mathbf{v}_1,\ldots,\mathbf{v}_p$ cuando no todos los pesos son cero. Un conjunto indexado es linealmente dependiente si y solo si no es linealmente independiente. Por brevedad, puede decirse que $\mathbf{v}_1,\ldots,\mathbf{v}_p$ son linealmente dependientes cuando queremos decir que $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ es un conjunto linealmente dependiente. Se utiliza una terminología semejante para los conjuntos linealmente independientes.

EJEMPLO 1 Sean $\mathbf{v}_1=\begin{bmatrix}1\\2\\3\end{bmatrix}$, $\mathbf{v}_2=\begin{bmatrix}4\\5\\6\end{bmatrix}$ y $\mathbf{v}_3=\begin{bmatrix}2\\1\\0\end{bmatrix}$.

a) Determine si el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

b) Si es posible, encuentre una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

SOLUCIÓN

a) Se debe determinar si existe una solución no trivial de la ecuación (1) que aparece más arriba. Las operaciones de fila sobre la matriz aumentada asociada indican que

$$\begin{bmatrix}1&4&2&0\\2&5&1&0\\3&6&0&0\end{bmatrix}\sim\begin{bmatrix}1&4&2&0\\0&-3&-3&0\\0&0&0&0\end{bmatrix}$$

Como es evidente, x_1 y x_2 son variables básicas, y x_3 es libre. Cada valor de x_3 distinto de cero determina una solución no trivial de (1). Así que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son linealmente dependientes (y, por tanto, no son linealmente independientes).

b) Para encontrar una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, se reduce por filas completamente a la matriz aumentada y se escribe el nuevo sistema:

$$\begin{bmatrix}1&0&-2&0\\0&1&1&0\\0&0&0&0\end{bmatrix}\qquad\begin{aligned}x_1\qquad\; - 2x_3 &= 0\\ x_2 + \;\; x_3 &= 0\\ 0 &= 0\end{aligned}$$

Así, $x_1 = 2x_3$, $x_2 = -x_3$, y x_3 es libre. Seleccione cualquier valor distinto de cero para x_3; por ejemplo, $x_3 = 5$. De esta manera, $x_1 = 10$ y $x_2 = -5$. Sustituya estos valores en la ecuación (1) para obtener

$$10\mathbf{v}_1 - 5\mathbf{v}_2 + 5\mathbf{v}_3 = \mathbf{0}$$

Esta es una (entre una infinidad) de las posibles relaciones de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. ■

Independencia lineal de las columnas de una matriz

Suponga que, en vez de utilizar un conjunto de vectores, se inicia con una matriz $A = [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n]$. En tal caso, la ecuación matricial $A\mathbf{x} = \mathbf{0}$ se puede escribir como

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

Cada relación de dependencia lineal entre las columnas de A corresponde a una solución no trivial de A$\mathbf{x} = \mathbf{0}$. Así, tenemos el siguiente resultado importante.

> Las columnas de una matriz A son linealmente independientes si y solo si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene *solo* la solución trivial. (3)

EJEMPLO 2 Determine si las columnas de la matriz $A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$ son linealmente independientes.

SOLUCIÓN Para estudiar $A\mathbf{x} = \mathbf{0}$, se reduce por filas la matriz aumentada:

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix}$$

En este punto, es claro que hay tres variables básicas y ninguna variable libre. Así, la ecuación $A\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial, y las columnas de A son linealmente independientes. ■

Conjuntos de uno o dos vectores

Un conjunto que solo tiene un vector —por ejemplo, $\mathbf{v}$— es linealmente independiente si y solo si $\mathbf{v}$ no es el vector cero. Esto se debe a que la ecuación vectorial $x_1\mathbf{v} = \mathbf{0}$ solo tiene la solución trivial cuando $\mathbf{v} \neq \mathbf{0}$. El vector cero es linealmente dependiente porque $x_1\mathbf{0} = \mathbf{0}$ tiene muchas soluciones no triviales.

El siguiente ejemplo explicará la naturaleza de un conjunto linealmente dependiente de dos vectores.

EJEMPLO 3 Determine si los siguientes conjuntos de vectores son linealmente independientes.

a) $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$ *b*) $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

SOLUCIÓN

a) Observe que $\mathbf{v}_2$ es un múltiplo de $\mathbf{v}_1$, a saber, $\mathbf{v}_2 = 2\mathbf{v}_1$. Así que, $-2\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$, lo que muestra que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente.

b) Desde luego, los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ *no* son múltiplos entre sí. ¿Podrían ser linealmente dependientes? Suponga que c y d satisfacen

$$c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}$$

Si $c \neq 0$, entonces es posible despejar $\mathbf{v}_1$ en términos de $\mathbf{v}_2$, a saber, $\mathbf{v}_1 = (-d/c)\mathbf{v}_2$. Este resultado es imposible porque $\mathbf{v}_1$ *no* es múltiplo de $\mathbf{v}_2$. Así que c debe ser cero. De manera similar, d también debe ser cero. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto linealmente independiente. ■

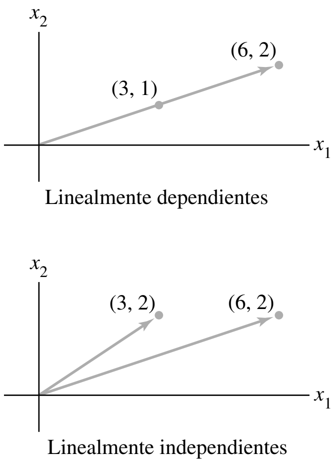


FIGURA 1

Los argumentos del ejemplo 3 indican que siempre es posible determinar por *inspección* cuándo un conjunto de dos vectores es linealmente dependiente. Las operaciones de fila son innecesarias. Basta con comprobar si al menos uno de los vectores es un escalar multiplicado por el otro. (La prueba solo se aplica a conjuntos de *dos* vectores).

Un conjunto de dos vectores $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente dependiente si al menos uno de los vectores es un múltiplo del otro. El conjunto es linealmente independiente si y solo si ninguno de los vectores es un múltiplo del otro.

En términos geométricos, dos vectores son linealmente dependientes si y solo si ambos están sobre la misma recta que pasa por el origen. La figura 1 muestra los vectores del ejemplo 3.

Conjuntos de dos o más vectores

La prueba del siguiente teorema es similar a la solución del ejemplo 3. Los detalles se presentan al final de esta sección.

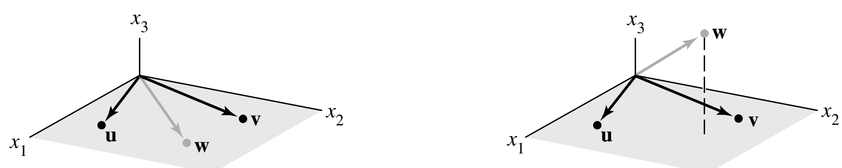
TEOREMA 7 **Caracterización de conjuntos linealmente dependientes**

Un conjunto indexado $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ de dos o más vectores es linealmente dependiente si y solo si al menos uno de los vectores en S es una combinación lineal de los otros. De hecho, si S es linealmente dependiente y $\mathbf{v}_1 \neq \mathbf{0}$, entonces alguna $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores precedentes, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

Advertencia: El teorema 7 *no* dice que *cada* vector en un conjunto linealmente dependiente es una combinación lineal de los vectores precedentes. Un vector en un conjunto linealmente dependiente puede no ser combinación lineal de los otros vectores. Véase el problema de práctica 3.

EJEMPLO 4 Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describa el conjunto generado por $\mathbf{u}$ y $\mathbf{v}$, y explique por qué un vector $\mathbf{w}$ está en Gen $\{\mathbf{u}, \mathbf{v}\}$ si y solo si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente.

SOLUCIÓN Los vectores $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes porque ninguno de ellos es múltiplo del otro, de manera que generan un plano en $\mathbb{R}^3$. (Véase la sección 1.3). De hecho, Gen $\{\mathbf{u}, \mathbf{v}\}$ es el plano x_1x_2 (con $x_3 = 0$). Si $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$, entonces $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente, de acuerdo con el teorema 7. A la inversa, suponga que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente. Por el teorema 7, algún vector en $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es una combinación lineal de los vectores precedentes (porque $\mathbf{u} \neq \mathbf{0}$). Ese vector debe ser $\mathbf{w}$, porque $\mathbf{v}$ no es múltiplo de $\mathbf{u}$. Así, $\mathbf{w}$ está en Gen $\{\mathbf{u}, \mathbf{v}\}$. Véase la figura 2. ■



Linealmente dependientes, $\mathbf{w}$ en Gen $\{\mathbf{u}, \mathbf{v}\}$

Linealmente independientes, $\mathbf{w}$ *no* está en Gen $\{\mathbf{u}, \mathbf{v}\}$

FIGURA 2 Dependencia lineal en $\mathbb{R}^3$.

El ejemplo 4 se generaliza a cualquier conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ en $\mathbb{R}^3$ con $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes. El conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ será linealmente dependiente si y solo si $\mathbf{w}$ está en el plano generado por $\mathbf{u}$ y $\mathbf{v}$.

Los siguientes dos teoremas describen casos especiales en los cuales la dependencia lineal de un conjunto es automática. Aún más, el teorema 8 será un resultado clave para trabajar en capítulos posteriores.

TEOREMA 8

> Si un conjunto contiene más vectores que entradas en cada vector, entonces el conjunto es linealmente dependiente. Es decir, cualquier conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es linealmente dependiente si $p > n$.

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

FIGURA 3
Si $p > n$, las columnas son linealmente dependientes.

DEMOSTRACIÓN Sea $A = [\mathbf{v}_1 \;\; \cdots \;\; \mathbf{v}_p]$. Entonces, A es $n \times p$, y la ecuación $A\mathbf{x} = \mathbf{0}$ corresponde a un sistema de n ecuaciones con p incógnitas. Si $p > n$, hay más variables que ecuaciones, por lo que debe existir una variable libre. Así, $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, y las columnas de A son linealmente dependientes. Véase la figura 3 para una versión matricial de este teorema. ■

Advertencia: El teorema 8 no dice nada acerca del caso en que el número de vectores en el conjunto *no* excede el número de entradas en cada vector.

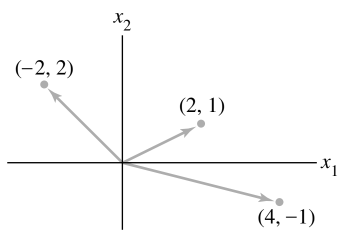


FIGURA 4
Un conjunto linealmente dependiente en $\mathbb{R}^2$.

EJEMPLO 5 Los vectores $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ son linealmente dependientes de acuerdo con el teorema 8, ya que hay tres vectores en el conjunto y solamente existen dos entradas en cada vector. Sin embargo, observe que ninguno de los vectores es un múltiplo de alguno de los otros. Véase la figura 4. ■

TEOREMA 9

> Si un conjunto $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ contiene al vector cero, entonces el conjunto es linealmente dependiente.

DEMOSTRACIÓN Al renumerar los vectores, se puede suponer que $\mathbf{v}_1 = \mathbf{0}$. Así, la ecuación $1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p = \mathbf{0}$ indica que S es linealmente dependiente. ■

EJEMPLO 6 Por inspección, determine si el conjunto dado es linealmente dependiente.

a) $\begin{bmatrix} 1 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 8 \end{bmatrix}$ *b*) $\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \end{bmatrix}$ *c*) $\begin{bmatrix} -2 \\ 4 \\ 6 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ -9 \\ 15 \end{bmatrix}$

SOLUCIÓN

a) El conjunto contiene cuatro vectores, cada uno de los cuales tiene solamente tres entradas. Así que, de acuerdo con el teorema 8, el conjunto es linealmente dependiente.

b) El teorema 8 no se aplica aquí porque el número de vectores no excede al número de entradas en cada vector. Como el vector cero se encuentra en el conjunto, este último es linealmente dependiente, de acuerdo con el teorema 9.

c) Compare las entradas correspondientes de los dos vectores. El segundo vector parece ser el primer vector multiplicado por $-3/2$. La relación es válida para los primeros tres pares de entradas, pero falla para el cuarto par. Así, ninguno de los vectores es múltiplo de otro y, por lo tanto, son linealmente independientes. ■

En general, la sección se debería leer *varias* veces para asimilar plenamente un concepto tan importante como el de independencia lineal. Las notas en la *Guía de estudio* para esta sección ayudarán a construir imágenes mentales de las principales ideas de álgebra lineal. Por ejemplo, es importante leer con cuidado la siguiente demostración, ya que ilustra cómo *se emplea* la definición de independencia lineal.

DEMOSTRACIÓN DEL TEOREMA 7 (Caracterización de conjuntos linealmente dependientes) Si alguna $\mathbf{v}_j$ en S es una combinación lineal de los demás vectores, entonces $\mathbf{v}_j$ se puede restar en ambos lados de la ecuación, produciendo así una relación de dependencia lineal con un peso diferente de cero (-1) sobre $\mathbf{v}_j$. [Por ejemplo, si $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, entonces $\mathbf{0} = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \cdots + 0\mathbf{v}_p$.]. Por lo tanto, S es linealmente dependiente.

A la inversa, suponga que S es linealmente dependiente. Si $\mathbf{v}_1$ es cero, entonces es una combinación lineal (trivial) de los demás vectores en S. De otra forma, $\mathbf{v}_1 \neq \mathbf{0}$, y existen pesos $c_1, \ldots, c_p$, sin que todos sean cero, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = 0$$

Sea j el subíndice más grande para el cual $c_j \neq 0$. Si $j = 1$, entonces $c_1\mathbf{v}_1 = \mathbf{0}$, lo que es imposible porque $\mathbf{v}_1 \neq \mathbf{0}$. De manera que $j > 1$, y

$$\begin{aligned} c_1\mathbf{v}_1 + \cdots + c_j\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_p &= \mathbf{0} \\ c_j\mathbf{v}_j &= -c_1\mathbf{v}_1 - \cdots - c_{j-1}\mathbf{v}_{j-1} \\ \mathbf{v}_j &= \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1} \end{aligned}$$

■

PROBLEMAS DE PRÁCTICA

Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \\ 7 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$ y $\mathbf{z} = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

1. Determine si los siguientes conjuntos son linealmente independientes y explique por qué: $\{\mathbf{u}, \mathbf{v}\}, \{\mathbf{u}, \mathbf{w}\}, \{\mathbf{u}, \mathbf{z}\}, \{\mathbf{v}, \mathbf{w}\}, \{\mathbf{v}, \mathbf{z}\}$ y $\{\mathbf{w}, \mathbf{z}\}$.
2. ¿Las respuestas al problema 1 implican que $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente independiente?
3. Para determinar si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente dependiente, ¿es aconsejable comprobar si, por ejemplo, $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{z}$?
4. ¿El conjunto $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ es linealmente dependiente?

1.7 EJERCICIOS

En los ejercicios 1 a 4, determine si los vectores son linealmente independientes. Justifique cada respuesta.

1. $\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -8 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -8 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix}$

4. $\begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -9 \end{bmatrix}$

En los ejercicios 5 a 8, determine si las columnas de la matriz forman un conjunto linealmente independiente. Justifique sus respuestas.

5. $\begin{bmatrix} 0 & -3 & 9 \\ 2 & 1 & -7 \\ -1 & 4 & -5 \\ 1 & -4 & -2 \end{bmatrix}$

6. $\begin{bmatrix} -4 & -3 & 0 \\ 0 & -1 & 5 \\ 1 & 1 & -5 \\ 2 & 1 & -10 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 4 & -3 & 0 \\ -2 & -7 & 5 & 1 \\ -4 & -5 & 7 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & -2 & 3 & 2 \\ -2 & 4 & -6 & 2 \\ 0 & 1 & -1 & 3 \end{bmatrix}$

En los ejercicios 9 y 10, *a*) ¿para qué valores de h está $\mathbf{v}_3$ en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$, y *b*) ¿para qué valores de h el conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente *dependiente*? Justifique sus respuestas.

9. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ -6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 5 \\ -7 \\ h \end{bmatrix}$

10. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ 15 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -5 \\ h \end{bmatrix}$

En los ejercicios 11 a 14, encuentre el valor o valores de h para los cuales los vectores son linealmente *dependientes*. Justifique sus respuestas.

11. $\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ h \end{bmatrix}$

12. $\begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 9 \\ h \\ 3 \end{bmatrix}$

13. $\begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ -9 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ h \\ -9 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ h \end{bmatrix}$

Por inspección, determine si los vectores en los ejercicios 15 a 20 son linealmente *independientes*. Justifique sus respuestas.

15. $\begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 7 \end{bmatrix}$

16. $\begin{bmatrix} 2 \\ -4 \\ 8 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \\ -12 \end{bmatrix}$

17. $\begin{bmatrix} 5 \\ -3 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 2 \\ 4 \end{bmatrix}$

18. $\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \end{bmatrix}$

19. $\begin{bmatrix} -8 \\ 12 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ 4 \\ -7 \end{bmatrix}, \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique cada respuesta con base en una cuidadosa lectura del texto.

21. *a*) Las columnas de una matriz A son linealmente independientes si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene la solución trivial.
 b) Si S es un conjunto linealmente dependiente, entonces cada vector es una combinación lineal de los otros vectores en S.
 c) Las columnas de cualquier matriz de 4×5 son linealmente dependientes.
 d) Si $\mathbf{x}$ y $\mathbf{y}$ son linealmente independientes, y si $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ es linealmente dependiente, entonces $\mathbf{z}$ está en Gen $\{\mathbf{x}, \mathbf{y}\}$.

22. *a*) Si $\mathbf{u}$ y $\mathbf{v}$ son linealmente independientes, y si $\mathbf{w}$ está en Gen $\{\mathbf{u}, \mathbf{v}\}$, entonces $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente dependiente.
 b) Si tres vectores en $\mathbb{R}^3$ están en el mismo plano en $\mathbb{R}^3$, entonces son linealmente dependientes.
 c) Si un conjunto contiene menos vectores que entradas en los vectores, entonces el conjunto es linealmente independiente.
 d) Si un conjunto en $\mathbb{R}^n$ es linealmente dependiente, entonces el conjunto contiene más de n vectores.

En los ejercicios 23 a 26, describa las posibles formas escalonadas de la matriz. Utilice la notación del ejemplo 1 de la sección 1.2.

23. A es una matriz de 2×2 con columnas linealmente dependientes.

24. A es una matriz de 3×3 con columnas linealmente independientes.

25. A es una matriz de 4×2, $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, y $\mathbf{a}_2$ no es múltiplo de $\mathbf{a}_1$.

26. A es una matriz de 4×3, $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$, tal que $\{\mathbf{a}_1, \mathbf{a}_2\}$ es linealmente independiente y $\mathbf{a}_3$ no está en Gen $\{\mathbf{a}_1, \mathbf{a}_2\}$.

27. ¿Cuántas columnas pivote debe tener una matriz de 6×4 si sus columnas son linealmente independientes? ¿Por qué?

28. ¿Cuántas columnas pivote debe tener una matriz de 4×6 si sus columnas generan a $\mathbb{R}^4$? ¿Por qué?

29. Construya matrices A y B de 3×2 tales que $A\mathbf{x} = \mathbf{0}$ tenga una solución no trivial, pero $B\mathbf{x} = \mathbf{0}$ tenga solamente la solución trivial.

30. *a*) Llene el espacio del siguiente enunciado: "Si A es una matriz de $m \times n$, entonces las columnas de A son linealmente independientes si y solo si A tiene ______ columnas pivote".
 b) Explique por qué es verdadero el enunciado en *a*).

Los ejercicios 31 y 32 deberían resolverse *sin efectuar operaciones de fila*. [*Sugerencia:* Escriba $A\mathbf{x} = \mathbf{0}$ como una ecuación vectorial].

31. A partir de $A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$, observe que la tercera columna es la suma de las dos primeras. Encuentre una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

32. A partir de $A = \begin{bmatrix} 4 & 3 & -5 \\ -2 & -2 & 4 \\ -2 & -3 & 7 \end{bmatrix}$, observe que la primera columna menos la segunda multiplicada por 3 es igual a la tercera columna. Determine una solución no trivial de $A\mathbf{x} = \mathbf{0}$.

En los ejercicios 33 a 38 cada enunciado es verdadero (en todos los casos) o falso (para al menos un ejemplo). Si es falso, idee un ejemplo específico para demostrar que el enunciado no siempre es válido. Tal ejemplo se llama *contraejemplo* del enunciado. Si un enunciado es verdadero, dé una justificación. (Un ejemplo específico no puede explicar por qué un enunciado siempre es verdadero. Aquí es necesario realizar más trabajo que en los ejercicios 21 y 22).

33. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$, y $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente dependiente.

34. Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $\mathbb{R}^4$ y $\mathbf{v}_2$ no es un múltiplo escalar de $\mathbf{v}_1$, entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es linealmente independiente.

35. Si $\mathbf{v}_1, \ldots, \mathbf{v}_5$ están en $\mathbb{R}^5$ y $\mathbf{v}_3 = \mathbf{0}$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ es linealmente dependiente.

36. Si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ están en $\mathbb{R}^3$ y $\mathbf{v}_3$ *no* es combinación lineal de $\mathbf{v}_1, \mathbf{v}_2$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

37. Si $\mathbf{v}_1, \ldots, \mathbf{v}_4$ están en $\mathbb{R}^4$, y $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ también es linealmente dependiente.

38. Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es un conjunto linealmente independiente de vectores en $\mathbb{R}^4$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ también es linealmente independiente. [*Sugerencia:* Piense en $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + 0 \cdot \mathbf{v}_4 = \mathbf{0}$].

39. Suponga que A es una matriz de $m \times n$ tal que para toda $\mathbf{b}$ en $\mathbb{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene, a lo sumo, una solución.

Use la definición de independencia lineal para explicar por qué las columnas de A deben ser linealmente independientes.

40. Suponga que una matriz A de $m \times n$ tiene n columnas pivote. Explique por qué para cada $\mathbf{b}$ en $\mathbb{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ tiene, a lo sumo, una solución. [*Sugerencia:* Explique por qué $A\mathbf{x} = \mathbf{b}$ no puede tener un número infinito de soluciones].

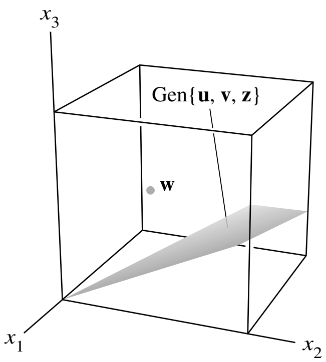
[**M**] En los ejercicios 41 y 42, utilice tantas columnas de A como sea posible para construir una matriz B tal que la ecuación $B\mathbf{x} = \mathbf{0}$ sólo tenga la solución trivial. Para comprobar su trabajo, resuelva $B\mathbf{x} = \mathbf{0}$.

41. $A = \begin{bmatrix} 3 & -4 & 10 & 7 & -4 \\ -5 & -3 & -7 & -11 & 15 \\ 4 & 3 & 5 & 2 & 1 \\ 8 & -7 & 23 & 4 & 15 \end{bmatrix}$

42. $A = \begin{bmatrix} 12 & 10 & -6 & 8 & 4 & -14 \\ -7 & -6 & 4 & -5 & -7 & 9 \\ 9 & 9 & -9 & 9 & 9 & -18 \\ -4 & -3 & -1 & 0 & -8 & 1 \\ 8 & 7 & -5 & 6 & 1 & -11 \end{bmatrix}$

43. [**M**] Con A y B del ejercicio 41, seleccione una columna $\mathbf{v}$ de A que no se haya utilizado en la construcción de B y determine si $\mathbf{v}$ está en el conjunto generado por las columnas de B. (Describa sus cálculos).

44. [**M**] Repita el ejercicio 43 con las matrices A y B del ejercicio 42. Luego, dé una explicación del resultado, suponiendo que B se construyó de acuerdo con las especificaciones.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sí. En cada caso, ningún vector es múltiplo del otro. Por lo tanto, cada conjunto es linealmente independiente.
2. No. La observación en el problema de práctica 1, por sí misma, no dice nada sobre la independencia lineal de $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$.
3. No. Cuando se comprueba la independencia lineal, por lo general es poco útil la idea de comprobar si un vector seleccionado es una combinación lineal de los demás. Puede ocurrir que el vector seleccionado no sea una combinación lineal de los otros y, aun así, el conjunto entero de vectores sea linealmente dependiente. En este problema de práctica, $\mathbf{w}$ no es una combinación lineal de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{z}$.
4. Sí, de acuerdo con el teorema 8. Existen más vectores (cuatro) que entradas (tres) en ellos.

1.8 INTRODUCCIÓN A LAS TRANSFORMACIONES LINEALES

La diferencia entre una ecuación matricial $A\mathbf{x} = \mathbf{b}$ y la ecuación vectorial asociada $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ es tan solo cuestión de notación. Sin embargo, es posible encontrar una ecuación matricial $A\mathbf{x} = \mathbf{b}$ en álgebra lineal (y en aplicaciones como gráficos generados por computadora y procesamiento de señales) que no esté directamente relacionada con combinaciones lineales de vectores. Esto sucede cuando se piensa en la matriz A como un objeto que "actúa" sobre un vector $\mathbf{x}$ multiplicándolo para producir un nuevo vector $A\mathbf{x}$.

Por ejemplo, en las ecuaciones

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}_{\mathbf{b}} \quad \text{y} \quad \underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{0}}$$

se observa que la multiplicación por A transforma a $\mathbf{x}$ en $\mathbf{b}$, y transforma a $\mathbf{u}$ en el vector cero. Véase la figura 1.

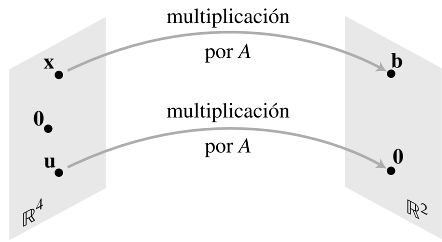


FIGURA 1 Transformación de vectores por medio de multiplicación matricial.

Desde este nuevo punto de vista, resolver la ecuación $A\mathbf{x} = \mathbf{b}$ equivale a encontrar todos los vectores $\mathbf{x}$ en $\mathbb{R}^4$ que se transforman en el vector $\mathbf{b}$ en $\mathbb{R}^2$ como resultado de la "acción" de la multiplicación por A.

La correspondencia de $\mathbf{x}$ a $A\mathbf{x}$ es una *función* de un conjunto de vectores a otro. Este concepto generaliza la noción común de una función como una regla que transforma un número real en otro.

Una **transformación** (o **función** o **mapeo**) T de $\mathbb{R}^n$ a $\mathbb{R}^m$ es una regla que asigna a cada vector $\mathbf{x}$ en $\mathbb{R}^n$ un vector $T(\mathbf{x})$ en $\mathbb{R}^m$. El conjunto $\mathbb{R}^n$ se llama el **dominio** de T, y $\mathbb{R}^m$ se llama el **codominio** de T. La notación $T : \mathbb{R}^n \to \mathbb{R}^m$ indica que el dominio de T es $\mathbb{R}^n$ y que el codominio es $\mathbb{R}^m$. Para $\mathbf{x}$ en $\mathbb{R}^n$, el vector $T(\mathbf{x})$ en $\mathbb{R}^m$ es la **imagen** de $\mathbf{x}$ (bajo la acción de T). El conjunto de todas las imágenes $T(\mathbf{x})$ es el **rango** de T. Véase la figura 2.

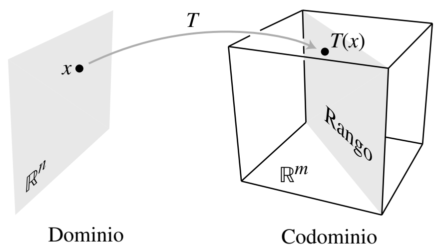


FIGURA 2 Dominio, codominio y rango de $T : \mathbb{R}^n \to \mathbb{R}^m$.

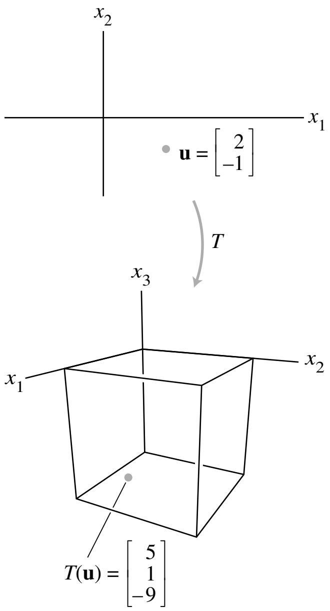
Es muy importante la nueva terminología de esta sección porque un enfoque dinámico del producto matriz-vector es la clave para entender diversas ideas en álgebra lineal, y para construir modelos matemáticos de sistemas físicos que evolucionan en el tiempo. Estos *sistemas dinámicos* se analizarán en las secciones 1.10, 4.8 y 4.9, así como en el capítulo 5.

Transformaciones matriciales

Lo que resta de esta sección se centra en mapeos asociados con la multiplicación matricial. Para cada $\mathbf{x}$ en $\mathbb{R}^n$, $T(\mathbf{x})$ se calcula como $A\mathbf{x}$, donde A es una matriz de $m \times n$. Para simplificar, algunas veces esta *transformación matricial* se denota como $\mathbf{x} \mapsto A\mathbf{x}$. Observe que el dominio de T es $\mathbb{R}^n$ cuando A tiene n columnas, y el codominio de T es $\mathbb{R}^m$ cuando cada columna de A tiene m entradas. El rango de T es el conjunto de todas las combinaciones lineales de las columnas de A, porque cada imagen $T(\mathbf{x})$ es de la forma $A\mathbf{x}$.

EJEMPLO 1 Sean $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, y defina una transformación $T : \mathbb{R}^2 \times \mathbb{R}^3$ por $T(\mathbf{x}) = A\mathbf{x}$, de manera que

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$



a) Encuentre $T(\mathbf{u})$, la imagen de $\mathbf{u}$ bajo la transformación T.

b) Encuentre una $\mathbf{x}$ en $\mathbb{R}^2$ cuya imagen bajo T sea $\mathbf{b}$.

c) ¿Hay más de una $\mathbf{x}$ cuya imagen bajo T sea $\mathbf{b}$?

d) Determine si $\mathbf{c}$ está en el rango de la transformación T.

SOLUCIÓN

a) Calcule

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

b) Resuelva $T(\mathbf{x}) = \mathbf{b}$ para $\mathbf{x}$. Es decir, resuelva $A\mathbf{x} = \mathbf{b}$, o

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \tag{1}$$

Empleando el método analizado en la sección 1.4, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \tag{2}$$

Así que $x_1 = 1.5$, $x_2 = -.5$ y $\mathbf{x} = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$. La imagen de esta $\mathbf{x}$ bajo T es el vector $\mathbf{b}$ dado.

c) Cualquier $\mathbf{x}$ cuya imagen bajo T es $\mathbf{b}$ debe satisfacer la ecuación (1). A partir de (2), es evidente que la ecuación (1) tiene una solución única, por lo que hay exactamente una $\mathbf{x}$ cuya imagen es $\mathbf{b}$.

d) El vector $\mathbf{c}$ está en el rango de T si $\mathbf{c}$ es la imagen de alguna $\mathbf{x}$ en $\mathbb{R}^2$, es decir, si $\mathbf{c} = T(\mathbf{x})$ para alguna $\mathbf{x}$. Esto es justamente otra manera de preguntar si el sistema $A\mathbf{x} = \mathbf{c}$ es consistente. Para encontrar la respuesta, reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

La tercera ecuación, $0 = -35$, indica que el sistema es inconsistente. De manera que $\mathbf{c}$ *no* está en el rango de T. ■

La pregunta del ejemplo 1*c*) es un problema de *unicidad* para un sistema de ecuaciones lineales, traducido aquí al lenguaje de transformaciones lineales: ¿Es $\mathbf{b}$ la imagen de una *única* $\mathbf{x}$ en $\mathbb{R}^n$? De manera similar, el ejemplo 1*d*) es un problema de *existencia*: *¿Existe* una $\mathbf{x}$ cuya imagen sea $\mathbf{c}$?

Las siguientes dos transformaciones matriciales se pueden visualizar geométricamente. Ambas refuerzan el enfoque dinámico de una matriz como algo que transforma vectores en otros vectores. La sección 2.7 incluye otros interesantes ejemplos relacionados con los gráficos generados por computadora.

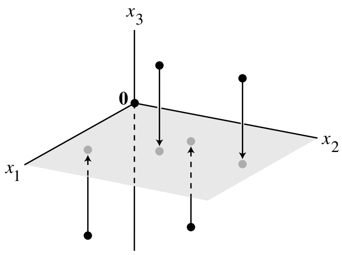


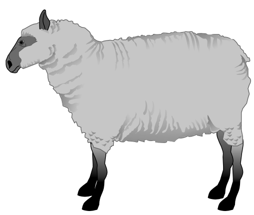
FIGURA 3
Una transformación proyección.

EJEMPLO 2 Si $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ *proyecta* puntos de $\mathbb{R}^3$ sobre el plano x_1x_2 porque

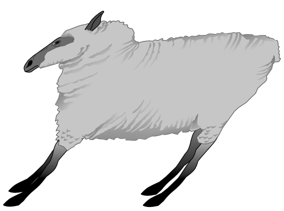
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

Véase la figura 3. ■

EJEMPLO 3 Sea $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. La transformación $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(\mathbf{x}) = A\mathbf{x}$ se llama una **transformación de trasquilado**. Se puede demostrar que si T actúa sobre cada punto del cuadrado 2×2 que se ilustra en la figura 4, entonces el conjunto de imágenes forma el paralelogramo sombreado. La idea fundamental es demostrar que T mapea segmentos de recta sobre segmentos de recta (como se muestra en el ejercicio 27) y después comprobar que los vértices del cuadrado se mapean sobre los vértices del paralelogramo. Por ejemplo, la imagen del punto $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ es $T(\mathbf{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, y la imagen de $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ es $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$. T deforma el cuadrado como si la parte superior de este se empujara hacia la derecha manteniendo fija la base. La transformación de trasquilado se presenta en física, geología y cristalografía. ■



borrego



borrego deformado por una transformación de trasquilado

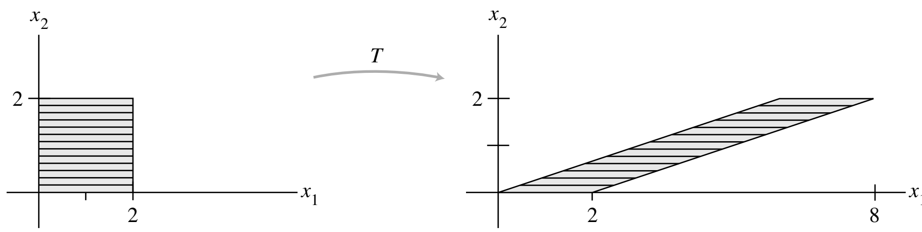


FIGURA 4 Una transformación de trasquilado.

Transformaciones lineales

El teorema 5 de la sección 1.4 establece que si A es de $m \times n$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ tiene las propiedades

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \quad \text{y} \quad A(c\mathbf{u}) = cA\mathbf{u}$$

para toda $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^n$ y todos los escalares c. Estas propiedades, expresadas en notación funcional, identifican a la más importante clase de transformaciones en álgebra lineal.

DEFINICIÓN

Una transformación (o mapeo) T es **lineal** si:

i. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todas las $\mathbf{u}$, $\mathbf{v}$ en el dominio de T;
ii. $T(c\mathbf{u}) = cT(\mathbf{u})$ para todos los escalares c y para todas las $\mathbf{u}$ en el dominio de T.

Cada transformación matricial es una transformación lineal. En los capítulos 4 y 5 se analizarán importantes ejemplos de transformaciones lineales que no son transformaciones matriciales.

Las transformaciones lineales *preservan las operaciones de suma vectorial y multiplicación escalar*. La propiedad **i** dice que el resultado $T(\mathbf{u} + \mathbf{v})$ de primero sumar **u** y **v** en $\mathbb{R}^n$ y después aplicar T es lo mismo que primero aplicar T a **u** y **v**, y luego sumar $T(\mathbf{u})$ y $T(\mathbf{v})$ en $\mathbb{R}^m$. Esas dos propiedades conducen fácilmente a los siguientes útiles resultados.

Si T es una transformación lineal, entonces

$$T(\mathbf{0}) = \mathbf{0} \tag{3}$$

y

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{4}$$

para todos los vectores **u**, **v** en el dominio de T y para todos los escalares c, d.

La propiedad (3) se deriva de la condición **ii** en la definición, porque $T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$. La propiedad (4) requiere tanto de **i** como de **ii**:

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Observe que *si una transformación satisface la ecuación (4) para cualesquiera* **u**, **v** *y* c, d, *entonces debe ser lineal*. (Establezca $c = d = 1$ para preservación de la suma, y $d = 0$ para preservación de la multiplicación escalar). La aplicación repetida de (4) produce una útil generalización:

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p) \tag{5}$$

En física e ingeniería, la ecuación (5) se conoce como *principio de superposición*. Piense que $\mathbf{v}_1,\ldots, \mathbf{v}_p$ son señales que entran a un sistema y $T(\mathbf{v}_1),\ldots, T(\mathbf{v}_p)$ son las respuestas de ese sistema a las señales. El sistema satisface el principio de superposición si para cualquier entrada expresada como una combinación lineal de tales señales, la respuesta del sistema es *la misma* combinación lineal de las respuestas a las señales individuales. En el capítulo 4 retomaremos esta idea.

EJEMPLO 4 Dado un escalar r, defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ por $T(\mathbf{x}) = r\mathbf{x}$. T se llama una **contracción** cuando $0 \leq r \leq 1$, y una **dilatación** cuando $r > 1$. Sea $r = 3$, y demuestre que T es una transformación lineal.

SOLUCIÓN Sean **u**, **v** en $\mathbb{R}^2$ y sean c, d escalares. De esta forma,

$$\begin{aligned} T(c\mathbf{u} + d\mathbf{v}) &= 3(c\mathbf{u} + d\mathbf{v}) && \text{Definición de } T \\ &= 3c\mathbf{u} + 3d\mathbf{v} && \Big\} \text{ Aritmética vectorial} \\ &= c(3\mathbf{u}) + d(3\mathbf{v}) && \\ &= cT(\mathbf{u}) + dT(\mathbf{v}) \end{aligned}$$

Así, T es una transformación lineal porque satisface (4). Véase la figura 5. ■

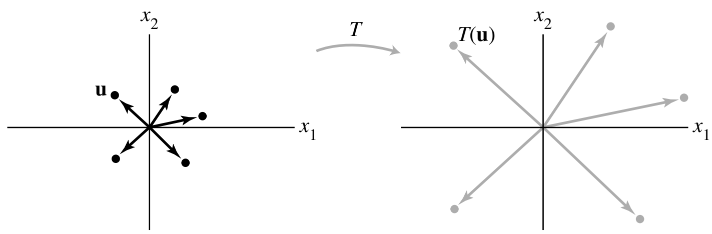


FIGURA 5 Una transformación de dilatación.

EJEMPLO 5 Defina una transformación $T : \mathbb{R}^2 \to \mathbb{R}^2$ mediante

$$T(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Encuentre las imágenes bajo T de $\mathbf{u} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

SOLUCIÓN

$$T(\mathbf{u}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \qquad T(\mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

$$T(\mathbf{u} + \mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

Observe que $T(\mathbf{u} + \mathbf{v})$, evidentemente, es igual a $T(\mathbf{u}) + T(\mathbf{v})$. En la figura 6 es claro que T hace girar a $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} + \mathbf{v}$ en el sentido antihorario, en torno al origen, en un ángulo de 90°. De hecho, T transforma el paralelogramo entero determinado por $\mathbf{u}$ y $\mathbf{v}$ en el determinado por $T(\mathbf{u})$y $T(\mathbf{v})$. (Véase el ejercicio 28). ■

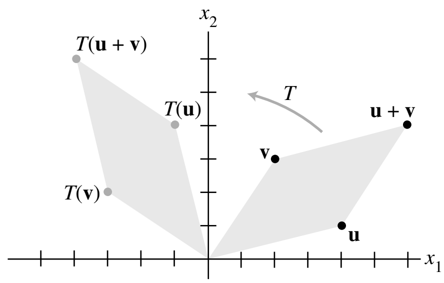


FIGURA 6 Una transformación de rotación.

El ejemplo final no es geométrico, sino que muestra cómo un mapeo lineal puede transformar un tipo de datos en otro.

EJEMPLO 6 Una compañía fabrica dos productos, B y C. Utilizando los datos del ejemplo 7 de la sección 1.3, construya una matriz de "costo unitario", $U = [\mathbf{b} \quad \mathbf{c}]$, cuyas columnas describan los "costos por dólar de producción" para ambos bienes:

$$\begin{array}{cc} & \text{Producto} \\ & \begin{array}{cc} \text{B} & \text{C} \end{array} \\ U = & \begin{bmatrix} .45 & .40 \\ .25 & .35 \\ .15 & .15 \end{bmatrix} \end{array} \begin{array}{l} \text{Materiales} \\ \text{Mano de obra} \\ \text{Gastos indirectos} \end{array}$$

Sea $\mathbf{x} = (x_1, x_2)$ el vector de "producción", correspondiente a x_1 dólares del producto B y a x_2 dólares del producto C, y defina $T : \mathbb{R}^2 \to \mathbb{R}^3$ mediante

$$T(\mathbf{x}) = U\mathbf{x} = x_1 \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} + x_2 \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix} = \begin{bmatrix} \text{Costo total de materiales} \\ \text{Costo total de mano de obra} \\ \text{Costo total de gastos indirectos} \end{bmatrix}$$

El mapeo T transforma una lista de cantidades de producción (medidas en dólares) en una lista de costos totales. La linealidad de este mapeo se refleja de dos maneras:

1. Si la producción se incrementa en un factor de, por ejemplo, 4, de $\mathbf{x}$ a $4\mathbf{x}$, entonces los costos se incrementarán por el mismo factor, de $T(\mathbf{x})$ a $4T(\mathbf{x})$.

2. Si $\mathbf{x}$ y $\mathbf{y}$ son vectores de producción, entonces el vector de costo total asociado con la producción combinada $\mathbf{x} + \mathbf{y}$ es precisamente la suma de los vectores de costo $T(\mathbf{x})$ y $T(\mathbf{y})$. ■

PROBLEMAS DE PRÁCTICA

1. Suponga que $T : \mathbb{R}^5 \to \mathbb{R}^2$ y $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A y para cada $\mathbf{x}$ en $\mathbb{R}^5$. ¿Cuántas filas y columnas tiene A?
2. Sea $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Dé una descripción geométrica de la transformación $\mathbf{x} \mapsto A\mathbf{x}$.
3. El segmento de recta de $\mathbf{0}$ a un vector $\mathbf{u}$ es el conjunto de puntos de la forma $t\mathbf{u}$, donde $0 \le t \le 1$. Demuestre que una transformación lineal T mapea este segmento en el segmento entre $\mathbf{0}$ y $T(\mathbf{u})$.

1.8 EJERCICIOS

1. Sea $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, y defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ por $T(\mathbf{x}) = A\mathbf{x}$.

   Encuentre las imágenes bajo T de $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

2. Sea $A = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 3 \\ 6 \\ -9 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

   Defina $T : \mathbb{R}^3 \to \mathbb{R}^3$ por $T(\mathbf{x}) = A\mathbf{x}$. Encuentre $T(\mathbf{u})$ y $T(\mathrm{v})$.

En los ejercicios 3 a 6, con T definida por $T(\mathbf{x}) = A\mathbf{x}$, encuentre un vector $\mathbf{x}$ cuya imagen bajo T sea $\mathbf{b}$, y analice si $\mathbf{x}$ es único.

3. $A = \begin{bmatrix} 1 & 0 & -3 \\ -3 & 1 & 6 \\ 2 & -2 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -3 \\ 2 & -5 & 6 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -6 \\ -4 \\ -5 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -5 & -7 \\ -3 & 7 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & -3 & 2 \\ 3 & -8 & 8 \\ 0 & 1 & 2 \\ 1 & 0 & 8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 3 \\ 10 \end{bmatrix}$

7. Sea A una matriz de 6×5. ¿Qué valores de a y b permiten definir $T : \mathbb{R}^a \to \mathbb{R}^b$ mediante $T(\mathbf{x}) = A\mathbf{x}$?

8. ¿Cuántas filas y columnas debe tener una matriz A para poder definir un mapeo de $\mathbb{R}^5$ a $\mathbb{R}^7$ con la regla $T(\mathbf{x}) = A\mathbf{x}$?

Para los ejercicios 9 y 10, encuentre todos los vectores $\mathbf{x}$ en $\mathbb{R}^4$ que son mapeados en el vector cero por la transformación $\mathbf{x} \mapsto A\mathbf{x}$ para la matriz A dada.

9. $A = \begin{bmatrix} 1 & -3 & 5 & -5 \\ 0 & 1 & -3 & 5 \\ 2 & -4 & 4 & -4 \end{bmatrix}$

10. $A = \begin{bmatrix} 3 & 2 & 10 & -6 \\ 1 & 0 & 2 & -4 \\ 0 & 1 & 2 & 3 \\ 1 & 4 & 10 & 8 \end{bmatrix}$

11. Sean $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, y A la matriz del ejercicio 9. ¿$\mathbf{b}$ está en el rango de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$? ¿Por qué?

12. Sean $\mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 4 \end{bmatrix}$, y A la matriz del ejercicio 10. ¿$\mathbf{b}$ está en el rango de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$? ¿Por qué?

En los ejercicios 13 a 16, utilice un sistema de coordenadas rectangulares para graficar $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, y sus imágenes bajo la transformación T dada. (Elabore un esquema grande y separado para cada ejercicio). Describa geométricamente lo que hace T a cada vector $\mathbf{x}$ en $\mathbb{R}^2$.

13. $T(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

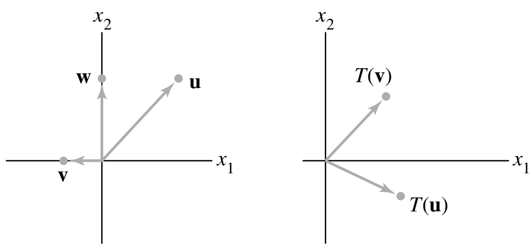
14. $T(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

15. $T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

16. $T(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

17. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ en $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ y mapea $\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ en $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$. Considerando el hecho de que T es lineal, encuentre las imágenes bajo T de $2\mathbf{u}$, $3\mathbf{v}$ y $2\mathbf{u} + 3\mathbf{v}$.

18. La figura muestra los vectores **u**, **v** y **w**, junto con las imágenes $T(\mathbf{u})$ y $T(\mathbf{v})$ bajo la acción de una transformación lineal $T : \mathbb{R}^2 \to \mathbb{R}^2$. Copie esta figura cuidadosamente, y dibuje la imagen de $T(\mathbf{w})$ con la mayor exactitud posible. [*Sugerencia:* Primero, escriba **w** como una combinación lineal de **u** y **v**.]



19. Sean $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{y}_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$ y $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$ y sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea $\mathbf{e}_1$ en $\mathbf{y}_1$, y $\mathbf{e}_2$ en $\mathbf{y}_2$. Encuentre las imágenes de $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ y $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

20. Sean $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$ y sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal que mapea **x** en $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$. Encuentre una matriz A tal que $T(\mathbf{x})$ sea $A\mathbf{x}$ para cada **x**.

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Una transformación lineal es un tipo especial de función.

b) Si A es una matriz de 3×5, y T es una transformación definida por $T(\mathbf{x}) = A\mathbf{x}$, entonces el dominio de T es $\mathbb{R}^3$.

c) Si A es una matriz de $m \times n$, entonces el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es $\mathbb{R}^m$.

d) Cada transformación lineal es una transformación matricial.

e) Una transformación T es lineal si y solo si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

para cualesquiera $\mathbf{v}_1$ y $\mathbf{v}_2$ en el dominio de T y para todos los escalares c_1 y c_2.

22. *a*) El rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es el conjunto de todas las combinaciones lineales de las columnas de A.

b) Cada transformación matricial es una transformación lineal.

c) Si $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal y si **c** está en $\mathbb{R}^m$, entonces una pregunta de unicidad es: "¿Está **c** en el rango de T?".

d) Una transformación lineal preserva las operaciones de suma vectorial y multiplicación escalar.

e) Una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ siempre mapea el origen de $\mathbb{R}^n$ al origen de $\mathbb{R}^m$.

23. Defina $f : \mathbb{R} \to \mathbb{R}$ por $f(x) = mx + b$.

a) Demuestre que f es una transformación lineal cuando $b = 0$.

b) Encuentre una propiedad de una transformación lineal que se viole cuando $b \neq 0$.

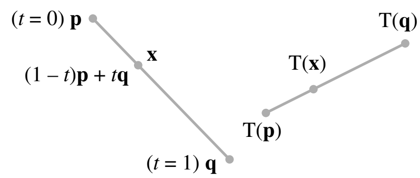
c) ¿Por qué f se llama una función lineal?

24. Una *transformación afín* $T : \mathbb{R}^n \to \mathbb{R}^m$ tiene la forma $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, donde A es una matriz de $m \times n$ y **b** está en $\mathbb{R}^m$. Demuestre que T *no* es una transformación lineal cuando $\mathbf{b} \neq \mathbf{0}$. (Las transformaciones afines son importantes en los gráficos generados por computadora).

25. Dados $\mathbf{v} \neq \mathbf{0}$ y **p** en $\mathbb{R}^n$, la recta que pasa por **p** en la dirección de **v** tiene la ecuación paramétrica $\mathbf{x} = \mathbf{p} + t\mathbf{v}$. Demuestre que una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ mapea esta recta sobre otra recta o sobre un solo punto (una *recta degenerada*).

26. *a*) Demuestre que la recta que pasa por los vectores **p** y **q** en $\mathbb{R}^n$ se puede escribir en la forma paramétrica $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$. (Consulte la figura de los ejercicios 21 y 22 de la sección 1.5).

b) El segmento de recta de **p** a **q** es el conjunto de puntos de la forma $(1 - t)\mathbf{p} + t\mathbf{q}$ para $0 \leq t \leq 1$ (como que se muestra en la figura de abajo). Demuestre que una transformación lineal T mapea este segmento de recta sobre un segmento de recta o sobre un solo punto.



27. Sean **u** y **v** vectores linealmente independientes en $\mathbb{R}^3$, y sea P el plano que pasa por **u**, **v** y **0**. La ecuación paramétrica de P es $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (con s, t en $\mathbb{R}$). Demuestre que una transformación lineal $T : \mathbb{R}^3 \to \mathbb{R}^3$ mapea P sobre un plano a través de **0**, o sobre una recta que pasa por **0**, o justo sobre el origen en $\mathbb{R}^3$. ¿Qué se puede decir acerca de $T(\mathbf{u})$ y $T(\mathbf{v})$ para que la imagen del plano P sea un plano?

28. Sean **u** y **v** vectores en $\mathbb{R}^n$. Es posible demostrar que el conjunto P de todos los puntos en el paralelogramo determinado por **u** y **v** tiene la forma $a\mathbf{u} + b\mathbf{v}$, para $0 \leq a \leq 1$, $0 \leq b \leq 1$. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Explique por qué la imagen, bajo la transformación T, de un punto en P está en el paralelogramo determinado por $T(\mathbf{u})$ y $T(\mathbf{v})$.

29. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal que refleja cada punto a través del eje x_2. Elabore dos esquemas semejantes a la figura 6, que ilustren las propiedades **i** y **ii** de una transformación lineal.

30. Suponga que los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_p$ generan a $\mathbb{R}^n$, y que $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación lineal. Considere que $T(\mathbf{v}_i) = \mathbf{0}$ para $i = 1, \ldots, p$. Demuestre que T es la transformación cero. Es decir, demuestre que si **x** es cualquier vector en $\mathbb{R}^n$, entonces $T(\mathbf{x}) = \mathbf{0}$.

31. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, y sea $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ un conjunto linealmente dependiente en $\mathbb{R}^n$. Explique por qué el conjunto $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ es linealmente dependiente.

En los ejercicios 32 a 36, los vectores columna están escritos como filas, como $\mathbf{x} = (x_1, x_2)$, y $T(\mathbf{x})$ se escribe como $T(x_1, x_2)$.

32. Demuestre que la transformación T definida por $T(x_1, x_2) = (x_1 - 2|x_2|, x_1 - 4x_2)$ no es lineal.

33. Demuestre que la transformación T definida por $T(x_1, x_2) = (x_1 - 2x_2, x_1 - 3, 2x_1 - 5x_2)$ no es lineal.

34. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación que refleja a cada vector $\mathbf{x} = (x_1, x_2, x_3)$ a través del plano $x_3 = 0$ sobre $T(\mathbf{x}) = (x_1, x_2, -x_3)$. Demuestre que T es una transformación lineal. [Para algunas ideas, véase el ejemplo 4].

35. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación que proyecta a cada vector $\mathbf{x} = (x_1, x_2, x_3)$ sobre el plano $x_2 = 0$, de manera que $T(\mathbf{x}) = (x_1, 0, x_3)$. Demuestre que T es una transformación lineal.

36. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Suponga que $\{\mathbf{u}, \mathbf{v}\}$ es un conjunto linealmente independiente, pero $\{T(\mathbf{u}), T(\mathbf{v})\}$ es un conjunto linealmente dependiente. Demuestre que $T(\mathbf{x}) = \mathbf{0}$ tiene una solución no trivial. [*Sugerencia:* Considere el hecho de que $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$ para algunos pesos c_1 y c_2, sin que ambos sean iguales a cero].

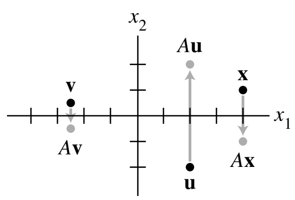
[**M**] En los ejercicios 37 y 38, las matrices determinan una transformación lineal T. Encuentre todas las $\mathbf{x}$ tales que $T(\mathbf{x}) = \mathbf{0}$.

37. $\begin{bmatrix} 2 & 3 & 5 & -5 \\ -7 & 7 & 0 & 0 \\ -3 & 4 & 1 & 3 \\ -9 & 3 & -6 & -4 \end{bmatrix}$ **38.** $\begin{bmatrix} 3 & 4 & -7 & 0 \\ 5 & -8 & 7 & 4 \\ 6 & -8 & 6 & 4 \\ 9 & -7 & -2 & 0 \end{bmatrix}$

39. [**M**] Sea $\mathbf{b} = \begin{bmatrix} 8 \\ 7 \\ 5 \\ -3 \end{bmatrix}$ y sea A la matriz del ejercicio 37. ¿Está $\mathbf{b}$ en el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$? Si es así, obtenga una $\mathbf{x}$ cuya imagen bajo la transformación sea $\mathbf{b}$.

40. [**M**] Sea $\mathbf{b} = \begin{bmatrix} 4 \\ -4 \\ -4 \\ -7 \end{bmatrix}$ y sea A la matriz del ejercicio 38. ¿Está $\mathbf{b}$ en el rango de la transformación $\mathbf{x} \mapsto A\mathbf{x}$? Si así es, encuentre una $\mathbf{x}$ cuya imagen bajo la transformación sea $\mathbf{b}$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

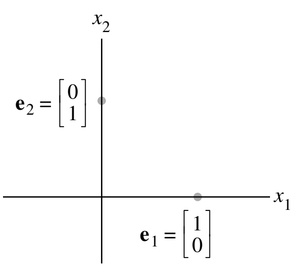


La transformación $\mathbf{x} \mapsto A\mathbf{x}$.

1. A debe tener cinco columnas para que $A\mathbf{x}$ esté definido. A debe tener dos filas para que el codominio de T esté en $\mathbb{R}^2$.
2. Dibuje algunos puntos aleatorios (vectores) sobre papel milimétrico para ver qué ocurre. Un punto como (4, 1) se mapea en (4, -1). La transformación $\mathbf{x} \mapsto A\mathbf{x}$ refleja los puntos a través del eje x (o eje x_1).
3. Sea $\mathbf{x} = t\mathbf{u}$ para alguna t tal que $0 \leq t \leq 1$. Como T es lineal, entonces $T(t\mathbf{u}) = tT(\mathbf{u})$, que es un punto sobre el segmento de línea entre $\mathbf{0}$ y $T(\mathbf{u})$.

1.9 MATRIZ DE UNA TRANSFORMACIÓN LINEAL

Siempre que una transformación lineal T se origina geométricamente o se describe con palabras, surge el deseo de tener una "fórmula" para $T(\mathbf{x})$. El análisis que sigue muestra que cada transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ en realidad es una transformación matricial $\mathbf{x} \mapsto A\mathbf{x}$, y que importantes propiedades de T están íntimamente relacionadas con propiedades de A. La clave para encontrar A es observar que T está plenamente determinada por su acción sobre las columnas de la matriz identidad de $n \times n$, I_n.



EJEMPLO 1 Las columnas de $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ son $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suponga que T es una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^3$ tal que

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

Sin otra información adicional, encuentre una fórmula para la imagen de una $\mathbf{x}$ arbitraria en $\mathbb{R}^2$.

SOLUCIÓN Escriba

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{1}$$

Como T es una transformación *lineal*,

$$\begin{aligned} T(\mathbf{x}) &= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) \qquad (2) \\ &= x_1 \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix} \end{aligned}$$

■

El paso de la ecuación (1) a la ecuación (2) explica por qué el conocimiento de $T(\mathbf{e}_1)$ y $T(\mathbf{e}_2)$ es suficiente para determinar $T(\mathbf{x})$ para cualquier $\mathbf{x}$. Además, ya que (2) expresa a $T(\mathbf{x})$ como una combinación lineal de vectores, podemos colocar esos vectores en las columnas de una matriz A y así escribir (2) como

$$T(\mathbf{x}) = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\mathbf{x}$$

TEOREMA 10

Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Así, existe una única matriz A tal que

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^n$$

De hecho, A es la matriz de $m \times n$ cuya j-ésima columna es el vector $T(\mathbf{e}_j)$, donde $\mathbf{e}_j$ es la j-ésima columna de la matriz identidad en $\mathbb{R}^n$:

$$A = [T(\mathbf{e}_1) \quad \cdots \quad T(\mathbf{e}_n)] \tag{3}$$

DEMOSTRACIÓN Escriba $\mathbf{x} = I_n\mathbf{x} = [\mathbf{e}_1 \quad \cdots \quad \mathbf{e}_n]\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, y utilice la linealidad de T para calcular

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1 T(\mathbf{e}_1) + \cdots + x_n T(\mathbf{e}_n) \\ &= \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x} \end{aligned}$$

En el ejercicio 33 nos ocuparemos de la unicidad de A. ■

La matriz A en (3) se llama **matriz estándar para la transformación lineal T**.

Ahora se sabe que cada transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ puede verse como una transformación matricial, y viceversa. El término *transformación lineal* se enfoca sobre una propiedad de un mapeo, mientras que la *transformación matricial* describe cómo se implementa tal mapeo, lo que se ilustra en los ejemplos 2 y 3.

EJEMPLO 2 Encuentre la matriz estándar A para la transformación de dilatación $T(\mathbf{x}) = 3\mathbf{x}$, para $\mathbf{x}$ en $\mathbb{R}^2$.

SOLUCIÓN Escriba

$$T(\mathbf{e}_1) = 3\mathbf{e}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad \text{y} \quad T(\mathbf{e}_2) = 3\mathbf{e}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \qquad \blacksquare$$

EJEMPLO 3 Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que hace girar a cada punto de $\mathbb{R}^2$ alrededor del origen un ángulo φ, en sentido antihorario, si φ es positivo. Geométricamente, se podría demostrar que esta transformación es lineal. (Véase la figura 6 de la sección 1.8). Encuentre la matriz estándar A de esta transformación.

SOLUCIÓN $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gira a $\begin{bmatrix} \cos\varphi \\ \operatorname{sen}\varphi \end{bmatrix}$, y $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ gira a $\begin{bmatrix} -\operatorname{sen}\varphi \\ \cos\varphi \end{bmatrix}$. Véase la figura 1. Por el teorema 10,

$$A = \begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

El ejemplo 5 de la sección 1.8 es un caso especial de esta transformación, con $\varphi = \pi/2$. $\blacksquare$

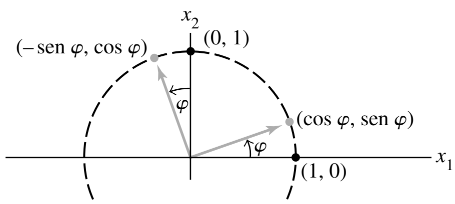


FIGURA 1 Una transformación de rotación.

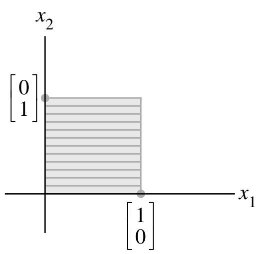


FIGURA 2 El cuadrado unitario.

Transformaciones lineales geométricas de $\mathbb{R}^2$

Los ejemplos 2 y 3 ilustran transformaciones lineales que se describen geométricamente. Las tablas 1 a 4 muestran otras transformaciones lineales geométricas comunes del plano. Como las transformaciones son lineales, estas quedan completamente determinadas por su acción sobre las columnas de I_2. En vez de solo mostrar las imágenes de $\mathbf{e}_1$ y $\mathbf{e}_2$, las tablas indican cómo una transformación afecta a un cuadrado unitario (figura 2).

Es posible construir otras transformaciones diferentes a partir de las listadas en las tablas 1 a 4; basta aplicar una transformación tras otra. Por ejemplo, un trasquilado horizontal podría ir seguido de una reflexión en el eje x_2. La sección 2.1 mostrará que tal *composición* de transformaciones lineales es lineal. (También, véase el ejercicio 34).

Preguntas de existencia y unicidad

El concepto de transformación lineal ofrece una nueva manera de entender las preguntas de existencia y unicidad que se plantearon antes. Las dos definiciones que aparecen después de las tablas 1 a 4 aportan una terminología adecuada para las transformaciones.

TABLA 1 Reflexiones

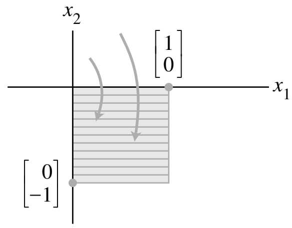
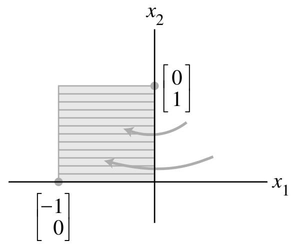
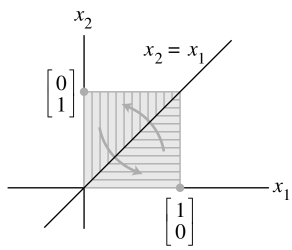
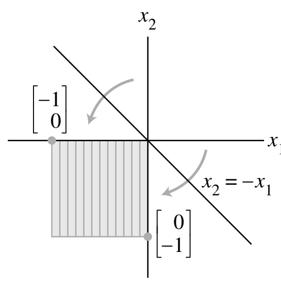
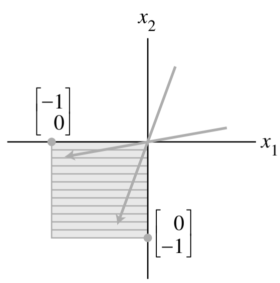
Transformación	**Imagen del cuadrado unitario**	**Matriz estándar**
Reflexión a través del eje x_1		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflexión a través del eje x_2		$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflexión a través de la recta $x_2 = x_1$		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Reflexión a través de la recta $x_2 = -x_1$		$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$
Reflexión a través del origen		$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

TABLA 2 Contracciones y expansiones

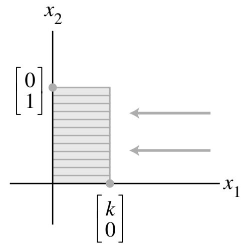
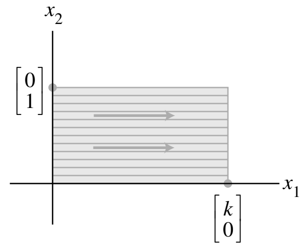
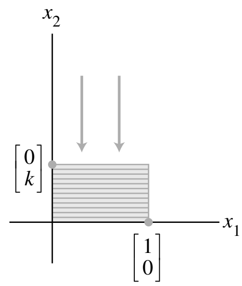
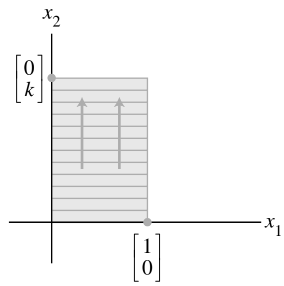
Transformación	Imagen del cuadrado unitario		Matriz estándar
Contracción y expansión horizontal	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}k & 0\\0 & 1\end{bmatrix}$
Contracción y expansión vertical	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}1 & 0\\0 & k\end{bmatrix}$

TABLA 3 Trasquilados

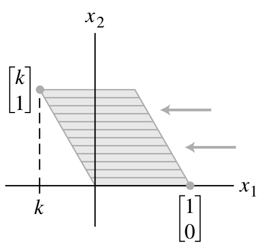
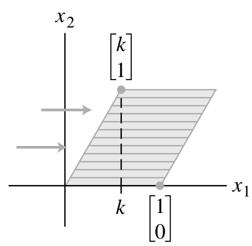
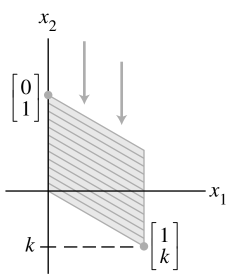
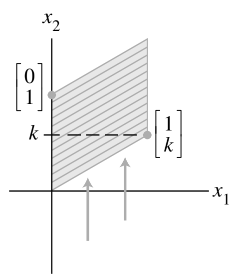
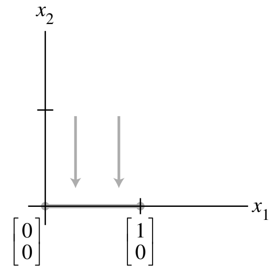
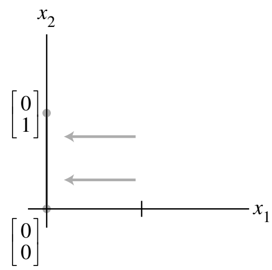
Transformación	Imagen del cuadrado unitario		Matriz estándar
Trasquilado horizontal	 $k < 0$	 $k > 0$	$\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$
Trasquilado vertical	 $k < 0$	 $k > 0$	$\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$

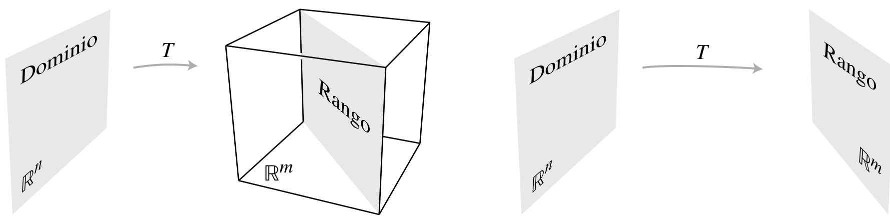
TABLA 4 Proyecciones

Transformación	Imagen del cuadrado unitario	Matriz estándar
Proyección sobre el eje x_1		$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Proyección sobre el eje x_2		$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

DEFINICIÓN

Se dice que un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ es **sobre** $\mathbb{R}^m$ si cada $\mathbf{b}$ en $\mathbb{R}^m$ es la imagen de *al menos una* $\mathbf{x}$ en $\mathbb{R}^n$.

De manera equivalente, T es sobre $\mathbb{R}^m$ cuando todo el rango de T es codominio $\mathbb{R}^m$. Es decir, T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si, para cada $\mathbf{b}$ en el codominio $\mathbb{R}^m$, existe al menos una solución de $T(\mathbf{x}) = \mathbf{b}$. "¿$T$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$?" es una pregunta de existencia. El mapeo T *no* es sobre cuando existe alguna $\mathbf{b}$ en $\mathbb{R}^m$ para la cual la ecuación $T(\mathbf{x}) = \mathbf{b}$ no tiene solución. Véase la figura 3.



T *no* es sobre $\mathbb{R}^m$ $\qquad$ T *es* sobre $\mathbb{R}^m$

FIGURA 3 ¿El rango de T es todo $\mathbb{R}^m$?

DEFINICIÓN

Se dice que un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ es **uno a uno** si cada $\mathbf{b}$ en $\mathbb{R}^m$ es la imagen de *a lo sumo una* $\mathbf{x}$ en $\mathbb{R}^n$.

De manera equivalente, T es uno a uno si, para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $T(\mathbf{x}) = \mathbf{b}$ tiene una única solución o ninguna solución. "¿T es uno a uno?" es una pregunta de unicidad. El mapeo T *no* es uno a uno cuando algún $\mathbf{b}$ en $\mathbb{R}^m$ es la imagen de más de un vector en $\mathbb{R}^n$. Si no existe tal $\mathbf{b}$, entonces T es uno a uno. Véase la figura 4.

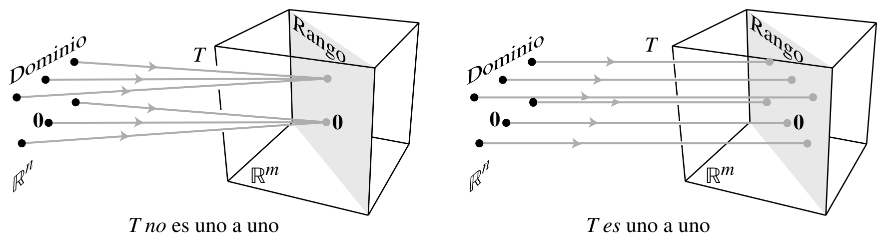


FIGURA 4 ¿Cada **b** es la imagen de a lo sumo un vector?

Las transformaciones de proyección que se ilustran en la tabla 4 *no son* uno a uno y *no* mapean $\mathbb{R}^2$ sobre $\mathbb{R}^2$. Las transformaciones en las tablas 1, 2 y 3 son uno a uno *y* sí mapean $\mathbb{R}^2$ sobre $\mathbb{R}^2$. Otras posibilidades se muestran en los dos ejemplos siguientes.

El ejemplo 4 y los teoremas que siguen muestran cómo las propiedades funcionales de ser un mapeo sobre y uno a uno están relacionadas con importantes conceptos estudiados antes en este capítulo.

EJEMPLO 4 Sea T la transformación lineal cuya matriz estándar es

$$A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

¿T mapea $\mathbb{R}^4$ sobre $\mathbb{R}^3$? ¿T es un mapeo uno a uno?

SOLUCIÓN Como A está en forma escalonada, podemos ver a la vez que A tiene una posición pivote en cada fila. De acuerdo con el teorema 4 de la sección 1.4, para cada $\mathbf{b}$ en $\mathbb{R}^3$, la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente. En otras palabras, la transformación lineal T mapea $\mathbb{R}^4$ (su dominio) sobre $\mathbb{R}^3$. Sin embargo, ya que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una variable libre (porque hay cuatro variables y solamente tres variables básicas), cada $\mathbf{b}$ es la imagen de más de una $\mathbf{x}$. Es decir, T *no es* uno a uno. ■

TEOREMA 11

Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces T es uno a uno si y solo si la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene únicamente la solución trivial.

DEMOSTRACIÓN Como T es lineal, $T(\mathbf{0}) = \mathbf{0}$. Si T es uno a uno, entonces la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene a lo sumo una solución y, por lo tanto, solo la solución trivial. Si T no es uno a uno, entonces existe una $\mathbf{b}$ que es la imagen de al menos dos diferentes vectores en $\mathbb{R}^n$, por ejemplo, $\mathbf{u}$ y $\mathbf{v}$. Es decir, $T(\mathbf{u}) = \mathbf{b}$ y $T(\mathbf{v}) = \mathbf{b}$. Pero, como T es lineal,

$$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

El vector $\mathbf{u} - \mathbf{v}$ no es cero porque $\mathbf{u} \neq \mathbf{v}$. En consecuencia, la ecuación $T(\mathbf{x}) = \mathbf{0}$ tiene más de una solución. Así, las dos condiciones del teorema son ambas verdaderas o ambas son falsas. ■

TEOREMA 12

Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal y sea A su matriz estándar. De esta forma,

a) T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si y solo si las columnas de A generan a $\mathbb{R}^m$;

b) T es uno a uno si y solo si las columnas de A son linealmente independientes.

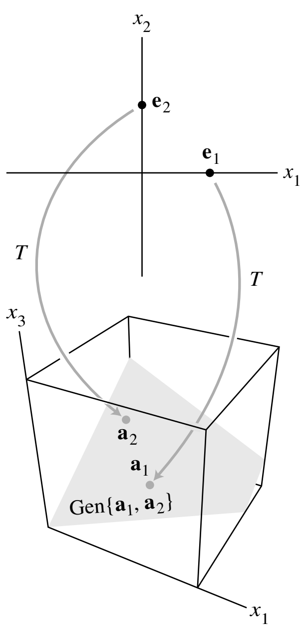
DEMOSTRACIÓN

a) De acuerdo con el teorema 4 de la sección 1.4, las columnas de A generan a $\mathbb{R}^m$ si y solo si para cada $\mathbf{b}$ en $\mathbb{R}^m$ la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente; en otras palabras, si y solo si para cada $\mathbf{b}$, la ecuación $T(\mathbf{x}) = \mathbf{b}$ tiene al menos una solución. Esto es cierto si y solo si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$.

b) Las ecuaciones $T(\mathbf{x}) = \mathbf{0}$ y $A\mathbf{x} = \mathbf{0}$ son iguales excepto por la notación. Así, de acuerdo con el teorema 11, T es uno a uno si y solo si $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Esto ocurre si y solo si las columnas de A son linealmente independientes, como ya se indicó en el enunciado (3) del recuadro en la sección 1.7. ■

El enunciado *a*) del teorema 12 es equivalente al enunciado "T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si y solo si cada vector en $\mathbb{R}^m$ es una combinación lineal de las columnas de A". Véase el teorema 4 de la sección 1.4

En el siguiente ejemplo y en algunos ejercicios posteriores, los vectores columna están escritos en filas, como $\mathbf{x} = (x_1, x_2)$, mientras que $T(\mathbf{x})$ se escribe como $T(x_1, x_2)$ en vez de emplear la manera más formal $T((x_1, x_2))$.



La transformación T no es sobre $\mathbb{R}^3$.

EJEMPLO 5 Sea $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Demuestre que T es una transformación lineal uno a uno. ¿T mapea $\mathbb{R}^2$ sobre $\mathbb{R}^3$?

SOLUCIÓN Cuando $\mathbf{x}$ y $T(\mathbf{x})$ se escriben como vectores columna, es posible determinar por inspección la matriz estándar de T, visualizando el cálculo fila-vector de cada entrada en $A\mathbf{x}$.

$$T(\mathbf{x}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4}$$

Así, T es claramente una transformación lineal; su matriz estándar A se muestra en (4). Las columnas de A son linealmente independientes porque no son múltiplos entre sí. De acuerdo con el teorema 12*b*), T es uno a uno. Para determinar si T es sobre $\mathbb{R}^3$, examine el espacio generado por las columnas de A. Como A es de 3×2, las columnas de A generan a $\mathbb{R}^3$ si y solo si A tiene 3 posiciones pivote, de acuerdo con el teorema 4. Esto es imposible, ya que A solo tiene 2 columnas. Así, las columnas de A no generan a $\mathbb{R}^3$, y la transformación lineal asociada no es sobre $\mathbb{R}^3$. ■

PROBLEMA DE PRÁCTICA

Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que primero efectúa un trasquilado horizontal que mapea $\mathbf{e}_2$ en $\mathbf{e}_2 - .5\mathbf{e}_1$ (pero deja inalterado a $\mathbf{e}_1$), y después refleja el resultado a través del eje x_2. Suponiendo que T es lineal, encuentre su matriz estándar. [*Sugerencia:* Determine la ubicación final de las imágenes de $\mathbf{e}_1$ y $\mathbf{e}_2$].

1.9 EJERCICIOS

En los ejercicios 1 a 10, suponga que T es una transformación lineal. Encuentre la matriz estándar de T.

1. $T:\mathbb{R}^2 \to \mathbb{R}^4$, $T(\mathbf{e}_1) = (3, 1, 3, 1)$ y $T(\mathbf{e}_2) = (-5, 2, 0, 0)$, donde $\mathbf{e}_1 = (1, 0)$ y $\mathbf{e}_2 = (0, 1)$.

2. $T:\mathbb{R}^3 \to \mathbb{R}^2$, $T(\mathbf{e}_1) = (1, 4)$, $T(\mathbf{e}_2) = (-2, 9)$ y $T(\mathbf{e}_3) = (3, -8)$, donde $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son las columnas de la matriz identidad de 3×3.

3. $T:\mathbb{R}^2 \to \mathbb{R}^2$ es una transformación de trasquilado vertical que mapea $\mathbf{e}_1$ en $\mathbf{e}_1 - 3\mathbf{e}_2$, pero deja inalterado a $\mathbf{e}_2$.

4. $T:\mathbb{R}^2 \to \mathbb{R}^2$ es una transformación de trasquilado horizontal que no altera a $\mathbf{e}_1$ y mapea $\mathbf{e}_2$ en $\mathbf{e}_2 + 2\mathbf{e}_1$.

5. $T:\mathbb{R}^2 \to \mathbb{R}^2$ hace girar a los puntos (en torno al origen) a través de un ángulo de $\pi/2$ radianes (en sentido antihorario).

6. $T:\mathbb{R}^2 \to \mathbb{R}^2$ hace girar a los puntos (en torno al origen) a través de un ángulo de $-3\pi/2$ radianes (en el sentido horario).

7. $T:\mathbb{R}^2 \to \mathbb{R}^2$ primero hace girar puntos a través de $-3\pi/4$ radianes (en el sentido horario) y después los refleja a través del eje horizontal x_1. [*Sugerencia:* Considere que $T(\mathbf{e}_1) = (-1/\sqrt{2}, 1/\sqrt{2})$].

8. $T:\mathbb{R}^2 \to \mathbb{R}^2$ primero realiza una transformación de trasquilado horizontal que transforma a $\mathbf{e}_2$ en $\mathbf{e}_2 + 2\mathbf{e}_1$ (dejando inalterado a $\mathbf{e}_1$) y después refleja los puntos a través de la recta $x_2 = -x_1$.

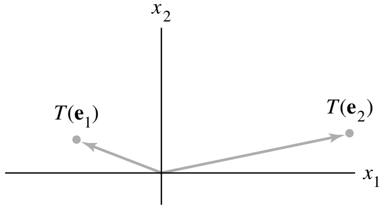
9. $T:\mathbb{R}^2 \to \mathbb{R}^2$ primero refleja los puntos a través del eje horizontal x_1 y luego los hace girar $-\pi/2$ radianes.

10. $T:\mathbb{R}^2 \to \mathbb{R}^2$ primero refleja los puntos a través del eje horizontal x_1 y luego los refleja a través de la recta $x_2 = x_1$.

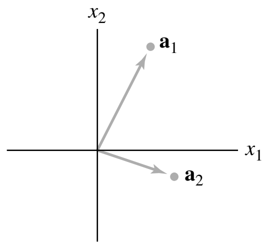
11. Una transformación lineal $T:\mathbb{R}^2 \to \mathbb{R}^2$ primero refleja los puntos a través del eje x_1 y luego los refleja a través del eje x_2. Demuestre que T también se puede describir como una transformación lineal que hace girar los puntos en torno al origen. ¿Cuál es el ángulo de esa rotación?

12. Demuestre que la transformación del ejercicio 10 es meramente una rotación en torno al origen. ¿Cuál es el ángulo de rotación?

13. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal tal que $T(\mathbf{e}_1)$ y $T(\mathbf{e}_2)$ son los vectores que se muestran en la figura. Con base en la figura, dibuje el vector $T(2, 1)$.



14. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal con matriz estándar $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, donde $\mathbf{a}_1$ y $\mathbf{a}_2$ se muestran en la figura, en la parte superior de la columna 2. Utilizando la figura, dibuje la imagen de $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ bajo la transformación T.



En los ejercicios 15 y 16, llene las entradas faltantes de la matriz, suponiendo que la ecuación es válida para todos los valores de las variables.

15. $\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 - 4x_2 \\ x_1 - x_3 \\ -x_2 + 3x_3 \end{bmatrix}$

16. $\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 - 2x_2 \\ x_1 + 4x_2 \\ x_2 \end{bmatrix}$

En los ejercicios 17 a 20, demuestre que T es una transformación lineal encontrando una matriz que implemente el mapeo. Observe que $x_1, x_2, \ldots$ no son vectores, sino entradas en vectores.

17. $T(x_1, x_2, x_3, x_4) = (x_1 + 2x_2, 0, 2x_2 + x_4, x_2 - x_4)$

18. $T(x_1, x_2) = (x_1 + 4x_2, 0, x_1 - 3x_2, x_1)$

19. $T(x_1, x_2, x_3) = (x_1 - 5x_2 + 4x_3, x_2 - 6x_3)$

20. $T(x_1, x_2, x_3, x_4) = 3x_1 + 4x_3 - 2x_4$ (Observe que: $T:\mathbb{R}^4 \to \mathbb{R}$)

21. Sea $T:\mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal tal que $T(x_1, x_2) = (x_1 + x_2, 4x_1 + 5x_2)$. Encuentre $\mathbf{x}$ tal que $T(\mathbf{x}) = (3, 8)$.

22. Sea $T:\mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal con $T(x_1, x_2) = (2x_1 - x_2, -3x_1 + x_2, 2x_1 - 3x_2)$. Encuentre $\mathbf{x}$ tal que $T(\mathbf{x}) = (0, -1, -4)$.

En los ejercicios 23 y 24, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) Una transformación lineal $T:\mathbb{R}^n \to \mathbb{R}^m$ está completamente determinada por sus efectos sobre las columnas de la matriz identidad de $n \times n$.

b) Si $T:\mathbb{R}^2 \to \mathbb{R}^2$ hace girar los vectores un ángulo φ en torno al origen, entonces T es una transformación lineal.

c) Cuando dos transformaciones lineales se realizan una tras otra, el efecto combinado no siempre es una transformación lineal.

d) Un mapeo $T:\mathbb{R}^n \to \mathbb{R}^m$ es sobre $\mathbb{R}^m$ si cada vector $\mathbf{x}$ en $\mathbb{R}^n$ se mapea sobre algún vector en $\mathbb{R}^m$.

e) Si A es una matriz de 3×2, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no puede ser uno a uno.

24. *a*) Si A es una matriz de 4×3, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^3$ sobre $\mathbb{R}^4$.

b) Cada transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ es una transformación matricial.

c) Las columnas de la matriz estándar para una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$ son las imágenes de las columnas de la matriz identidad de $n \times n$ bajo T.

d) Un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ es uno a uno si cada vector en $\mathbb{R}^n$ se mapea sobre un único vector en $\mathbb{R}^m$.

e) La matriz estándar de una transformación de trasquilado horizontal de $\mathbb{R}^2$ a $\mathbb{R}^2$ tiene la forma $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$, donde a y b son ± 1.

En los ejercicios 25 a 28, determine si la transformación lineal especificada es *a*) uno a uno o *b*) sobre. Justifique cada respuesta.

25. La transformación en el ejercicio 17.

26. La transformación en el ejercicio 2.

27. La transformación en el ejercicio 19.

28. La transformación en el ejercicio 14.

En los ejercicios 29 y 30, describa las posibles formas escalonadas de la matriz estándar para una transformación lineal T. Utilice la notación del ejemplo 1 de la sección 1.2

29. $T : \mathbb{R}^3 \to \mathbb{R}^4$ es uno a uno. **30.** $T : \mathbb{R}^4 \to \mathbb{R}^3$ es sobre.

31. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, y sea A su matriz estándar. Complete el siguiente enunciado para hacerlo verdadero: "T es uno a uno si y solo si A tiene _____ columnas pivote". Explique por qué el enunciado es verdadero. [*Sugerencia:* Consulte los ejercicios de la sección 1.7].

32. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, y sea A su matriz estándar. Complete el siguiente enunciado para hacerlo verdadero: "T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$ si y solo si A tiene _______ columnas pivote". Encuentre algunos teoremas que expliquen por qué el enunciado es verdadero.

33. Verifique la unicidad de A en el teorema 10. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal tal que $T(\mathbf{x}) = B\mathbf{x}$ para alguna matriz B de $m \times n$. Demuestre que si A es la matriz estándar de T, entonces $A = B$. [*Sugerencia:* Demuestre que A y B tienen las mismas columnas].

34. Sean $S : \mathbb{R}^p \to \mathbb{R}^n$ y $T : \mathbb{R}^n \to \mathbb{R}^m$ transformaciones lineales. Demuestre que el mapeo $\mathbf{x} \mapsto T(S(\mathbf{x}))$ es una transformación lineal (de $\mathbb{R}^p$ a $\mathbb{R}^m$). [*Sugerencia:* Calcule $T(S(c\mathbf{u} + d\mathbf{v}))$ para $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^p$ y escalares c y d. Justifique cada paso del cálculo, y explique por qué este proceso conduce a la conclusión deseada].

35. Si una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ mapea $\mathbb{R}^n$ *sobre* $\mathbb{R}^m$, ¿es posible encontrar una relación entre m y n? Si T es uno a uno, ¿qué se puede decir acerca de m y n?

36. ¿Por qué la pregunta "¿La transformación lineal T es sobre?" es una pregunta de existencia?

[**M**] En los ejercicios 37 a 40, sea T la transformación lineal cuya matriz estándar se presenta. En los ejercicios 37 y 38, determine si T es un mapeo uno a uno. En los ejercicios 39 y 40, determine si T mapea $\mathbb{R}^5$ sobre $\mathbb{R}^5$. Justifique sus respuestas.

37. $\begin{bmatrix} -5 & 6 & -5 & -6 \\ 8 & 3 & -3 & 8 \\ 2 & 9 & 5 & -12 \\ -3 & 2 & 7 & -12 \end{bmatrix}$ **38.** $\begin{bmatrix} 7 & 5 & 9 & -9 \\ 5 & 6 & 4 & -4 \\ 4 & 8 & 0 & 7 \\ -6 & -6 & 6 & 5 \end{bmatrix}$

39. $\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$

40. $\begin{bmatrix} 9 & 43 & 5 & 6 & -1 \\ 14 & 15 & -7 & -5 & 4 \\ -8 & -6 & 12 & -5 & -9 \\ -5 & -6 & -4 & 9 & 8 \\ 13 & 14 & 15 & 3 & 11 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

WEB

Observe lo que ocurre a $\mathbf{e}_1$ y $\mathbf{e}_2$. Véase la figura 5. Primero, $\mathbf{e}_1$ no sufre alteraciones por el trasquilado y después se refleja en $-\mathbf{e}_1$. Así, $T(\mathbf{e}_1) = -\mathbf{e}_1$. Segundo, $\mathbf{e}_2$ va a $\mathbf{e}_2 - .5\mathbf{e}_1$ por

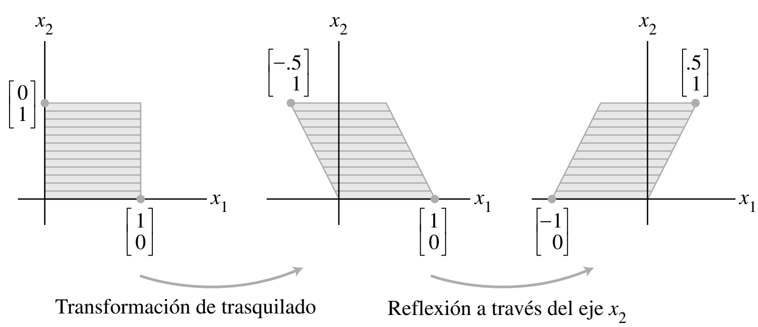


FIGURA 5 Composición de dos transformaciones.

la transformación de trasquilado. Como la reflexión a través del eje x_2 cambia a $\mathbf{e}_1$ en $-\mathbf{e}_1$ y deja inalterado a $\mathbf{e}_2$, el vector $\mathbf{e}_2 - .5\mathbf{e}_1$ va a $\mathbf{e}_2 + .5\mathbf{e}_1$. Así, $T(\mathbf{e}_2) = \mathbf{e}_2 + .5\mathbf{e}_1$. Por lo tanto, la matriz estándar de T es

$$\begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} -\mathbf{e}_1 & \mathbf{e}_2 + .5\mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} -1 & .5 \\ 0 & 1 \end{bmatrix}$$

1.10 MODELOS LINEALES EN LOS NEGOCIOS, CIENCIA E INGENIERÍA

En esta sección todos los modelos matemáticos son *lineales*; es decir, cada uno de ellos describe un problema mediante una ecuación lineal, por lo general en forma vectorial o matricial. El primer modelo concierne al campo de la nutrición, pero en realidad es representativo de una técnica general en problemas de programación lineal. El segundo modelo pertenece al campo de la ingeniería eléctrica. El tercer modelo introduce el concepto de una *ecuación lineal en diferencias*, una poderosa herramienta matemática para estudiar procesos dinámicos en una amplia variedad de campos, como ingeniería, ecología, economía, telecomunicaciones y ciencias administrativas. Los modelos lineales son importantes porque los fenómenos naturales con frecuencia son lineales o casi lineales cuando las variables implicadas se mantienen dentro de límites razonables. Además, los modelos lineales se adaptan más fácilmente al cálculo computacional que los complejos modelos no lineales.

Conforme estudie cada modelo, preste atención en cómo su linealidad refleja alguna propiedad del sistema que se modela.

Elaboración de una dieta nutritiva para bajar de peso

WEB

La fórmula para la dieta Cambridge, una conocida dieta de la década de 1980, se basó en años de investigación. Un equipo de científicos, encabezados por el doctor Alan H. Howard, desarrollaron esta dieta en la Universidad de Cambridge después de más de ocho años de trabajo clínico con pacientes obesos.[1] La fórmula de la dieta baja en calorías combina un balance preciso de carbohidratos, proteínas de alta calidad y grasa, junto con vitaminas, minerales, oligoelementos y electrolitos. Millones de personas han seguido la dieta para lograr una rápida y sustancial pérdida de peso.

Para alcanzar las cantidades y proporciones deseadas de nutrientes, Howard tuvo que incorporar una gran variedad de alimentos en la dieta. Cada alimento suministraba varios de los ingredientes requeridos, pero no en las proporciones correctas. Por ejemplo, la leche sin grasa (descremada) fue la principal fuente de proteína, pero contenía mucho calcio. De manera que se utilizó harina de soya para aportar proteína porque contiene poco calcio. Sin embargo, proporcionalmente la harina de soya contiene mucha grasa, así que se adicionó suero de leche porque contiene menos grasa en relación con el calcio. Por desgracia, el suero de leche contiene muchos carbohidratos…

El siguiente ejemplo ilustra el problema a pequeña escala. En la tabla 1 se listan tres de los ingredientes en la dieta, junto con las cantidades de ciertos nutrientes aportados por 100 gramos (g) de cada ingrediente.[2]

EJEMPLO 1 Si es posible, encuentre alguna combinación de leche descremada, harina de soya y suero de leche que aporte las cantidades exactas de proteínas, grasas y carbohidratos suministrados por la dieta diaria (tabla 1).

[1] El primer anuncio de este rápido régimen para bajar de peso se publicó en el *International Journal of Obesity* (1978) **2**, 321-332.

[2] Ingredientes en la dieta desde 1984; datos de nutrientes en ingredientes adaptados del USDA Agricultural Handbooks, núms. 8-1 y 8-6, 1976.

TABLA 1

Cantidades (en g) suministradas por 100 g de ingrediente				Cantidades (en g) suministradas por la dieta Cambridge en un día
Nutriente	**Leche descremada**	**Harina de soya**	**Suero de leche**	
Proteínas	36	51	13	33
Carbohidratos	52	34	74	45
Grasa	0	7	1.1	3

SOLUCIÓN Dejemos que x_1, x_2 y x_3, respectivamente, denoten el número de unidades (100 g) de esos alimentos. Un enfoque al problema consiste en deducir ecuaciones por separado para cada nutriente. Por ejemplo, el producto

$$\left\{\begin{matrix} x_1 \text{ unidades de} \\ \text{leche descremada} \end{matrix}\right\} \cdot \left\{\begin{matrix} \text{proteínas por unidad} \\ \text{de leche descremada} \end{matrix}\right\}$$

da la cantidad de proteína suministrada por x_1 unidades de leche descremada. A esta cantidad, le agregaríamos productos similares de harina de soya y suero de leche, igualando la suma resultante con la cantidad de proteína que se requiere. Se tendrían que realizar cálculos similares para cada nutriente.

Un método más eficiente, y más sencillo conceptualmente, consiste en considerar un "vector nutriente" para cada alimento y elaborar solo una ecuación vectorial. La cantidad de nutrientes aportados por x_1 unidades de leche descremada es el múltiplo escalar

$$\overset{\text{Escalar}}{\left\{\begin{matrix} x_1 \text{ unidades de} \\ \text{leche descremada} \end{matrix}\right\}} \cdot \overset{\text{Vector}}{\left\{\begin{matrix} \text{nutrientes por unidad} \\ \text{de leche descremada} \end{matrix}\right\}} = x_1\mathbf{a}_1 \tag{1}$$

donde $\mathbf{a}_1$ es la primera columna de la tabla 1. Sean $\mathbf{a}_2$ y $\mathbf{a}_3$ los vectores correspondientes para harina de soya y suero de leche, respectivamente, y sea $\mathbf{b}$ el vector que lista el total de nutrientes requerido (la última columna de la tabla). Luego, $x_2\mathbf{a}_2$ y $x_3\mathbf{a}_3$ dan los nutrientes suministrados por x_2 unidades de harina de soya y x_3 unidades de suero de leche, respectivamente. De esta forma, la ecuación relevante es

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b} \tag{2}$$

La reducción por filas de la matriz aumentada del sistema de ecuaciones correspondiente indica que

$$\begin{bmatrix} 36 & 51 & 13 & 33 \\ 52 & 34 & 74 & 45 \\ 0 & 7 & 1.1 & 3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & .277 \\ 0 & 1 & 0 & .392 \\ 0 & 0 & 1 & .233 \end{bmatrix}$$

A tres dígitos significativos, la dieta requiere .277 unidades de leche descremada, .392 unidades de harina de soya y .233 unidades de suero de leche, con el objetivo de aportar las cantidades deseadas de proteínas, carbohidratos y grasa. ■

Es importante que los valores encontrados para x_1, x_2 y x_3 no sean negativos. Esto es necesario para que la solución sea físicamente realizable. (Por ejemplo, ¿cómo se podrían emplear $-.233$ unidades de suero de leche?) Para un gran número de requerimientos nutricionales, será necesario utilizar un mayor número de alimentos para así generar un sistema de ecuaciones con una solución "no negativa". Por consiguiente, se necesitaría examinar muchísimas combinaciones diferentes de alimentos para encontrar un sistema de ecuaciones con tal solución. De hecho, el diseñador de la dieta Cambridge fue capaz de proporcionar 31 nutrientes en cantidades precisas empleando solo 33 ingredientes.

El problema de la elaboración de la dieta conduce al sistema *lineal* (2) porque la cantidad de nutrientes aportada por cada alimento se puede escribir como un múltiplo escalar de un vector, como en (1). Es decir, los nutrientes suministrados por un alimento son *proporcionales* a la cantidad del alimento agregado a la mezcla de la dieta. Además, cada nutriente en la mezcla es la *suma* de las cantidades de los diversos alimentos.

Los problemas de formulación de dietas especializadas para humanos y ganado ocurren con frecuencia. Por lo regular, dichos problemas se tratan mediante técnicas de programación lineal. Nuestro método de construir ecuaciones vectoriales a menudo simplifica la tarea de formular tales problemas.

Ecuaciones lineales y redes eléctricas

WEB

El flujo de corriente en una red eléctrica sencilla se puede describir por un sistema de ecuaciones lineales. Una fuente de voltaje, como una batería, genera una corriente de electrones que fluye a través de la red. Cuando la corriente pasa por un resistor (ya sea una bombilla o un motor), parte del voltaje se "consume"; de acuerdo con la ley de Ohm, esta "caída de voltaje" en el resistor está dada por

$$V = RI$$

donde el voltaje V se mide en *volts*, la resistencia R en *ohms* (denotados con el símbolo Ω), y el flujo de corriente I en *amperes* (A).

La red de la figura 1 contiene tres circuitos cerrados. Las corrientes que fluyen en los circuitos 1, 2 y 3 se denotan con I_1, I_2 e I_3, respectivamente. Las direcciones asignadas a las *corrientes de circuito* son arbitrarias. Si una corriente resulta negativa, entonces la dirección real del flujo de corriente es opuesta a la que se indica en la figura. Si la dirección de corriente que se muestra se aleja del lado positivo de una batería (-|⊢), alrededor del lado negativo, entonces el voltaje es positivo; de otra forma, el voltaje es negativo.

El flujo de corriente en un circuito está regido por la siguiente ley.

LEY DE KIRCHHOFF SOBRE EL VOLTAJE

La suma algebraica de las caídas de voltaje RI en una dirección alrededor de un circuito es igual a la suma algebraica de las fuentes de voltaje en la misma dirección alrededor del circuito.

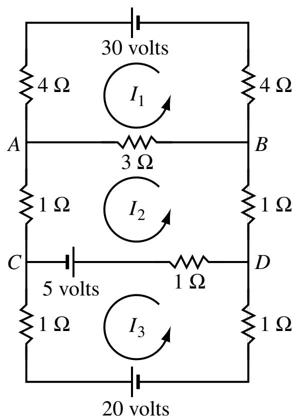


FIGURA 1

EJEMPLO 2 Determine las corrientes de circuito en la red de la figura 1.

SOLUCIÓN Para el circuito 1 la corriente I_1 fluye a través de tres resistores, y la suma de las caídas de voltaje RI es

$$4I_1 + 4I_1 + 3I_1 = (4 + 4 + 3)I_1 = 11I_1$$

La corriente en el circuito 2 también fluye en parte del circuito 1, por la pequeña *rama* entre A y B. Ahí la caída RI asociada es $3I_2$ volts. Sin embargo, la dirección de corriente para la rama AB en el circuito 1 es opuesta a la elegida para el flujo en el circuito 2, de manera que la suma algebraica de todas las caídas RI para el circuito 1 es $11I_1 - 3I_2$. Como el voltaje en el circuito 1 es $+30$ volts, la ley de voltaje de Kirchhoff implica que

$$11I_1 - 3I_2 = 30$$

La ecuación para el circuito 2 es

$$-3I_1 + 6I_2 - I_3 = 5$$

El término $-3I_1$ viene del flujo de la corriente del circuito 1 por la rama AB (con una caída de voltaje negativa porque ahí el flujo de corriente es opuesto al flujo en el circuito 2). El término $6I_2$ es la suma de todas las resistencias en el circuito 2, multiplicada por la corriente del circuito. El término $-I_3 = -1 \cdot I_3$ proviene de la corriente del circuito 3 que fluye por el resistor de 1 ohm en la rama CD, en la dirección opuesta al flujo en el circuito 2. La ecuación para el circuito 3 es

$$-I_2 + 3I_3 = -25$$

Observe que la batería de 5 volts en la rama *CD* se cuenta como parte de los circuitos 2 y 3, pero se considera de -5 volts para el circuito 3 debido a la dirección elegida para la corriente en el circuito 3. La batería de 20 volts es negativa por la misma razón.

Las corrientes de circuito se encuentran resolviendo el sistema

$$\begin{aligned} 11I_1 - 3I_2 \phantom{{}- I_3} &= 30 \\ -3I_1 + 6I_2 - I_3 &= 5 \\ - I_2 + 3I_3 &= -25 \end{aligned} \tag{3}$$

Las operaciones de fila sobre la matriz aumentada conducen a la solución: $I_1 = 3$ A, $I_2 = 1$ A e $I_3 = -8$ A. El valor negativo de I_3 indica que la corriente real en el circuito 3 fluye en la dirección opuesta a la que se muestra en la figura 1. ■

Es conveniente observar el sistema (3) como una ecuación vectorial:

$$\underset{\uparrow\atop \mathbf{r}_1}{I_1\begin{bmatrix} 11 \\ -3 \\ 0 \end{bmatrix}} + \underset{\uparrow\atop \mathbf{r}_2}{I_2\begin{bmatrix} -3 \\ 6 \\ -1 \end{bmatrix}} + \underset{\uparrow\atop \mathbf{r}_3}{I_3\begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}} = \underset{\uparrow\atop \mathbf{v}}{\begin{bmatrix} 30 \\ 5 \\ -25 \end{bmatrix}} \tag{4}$$

La primera entrada de cada vector concierne al primer circuito, y de manera similar para la segunda y tercera entradas. El primer vector resistor $\mathbf{r}_1$ lista la resistencia en los diversos circuitos por donde fluye la corriente I_1. Una resistencia se registra como negativa cuando I_1 fluye contra la dirección de flujo en otro circuito. Examine la figura 1 y vea cómo calcular las entradas en $\mathbf{r}_1$; luego, haga lo mismo con $\mathbf{r}_2$ y $\mathbf{r}_3$. La forma matricial de la ecuación (4),

$$R\mathbf{i} = \mathbf{v}, \quad \text{donde,} \quad R = [\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3\,] \quad \text{e} \quad \mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}$$

da una versión matricial de la ley de Ohm. Si todas las corrientes de circuito se seleccionan en la misma dirección (por ejemplo, en sentido antihorario), entonces todas las entradas fuera de la diagonal principal de R serán negativas.

De una mirada, la ecuación matricial $R\mathbf{i} = \mathbf{v}$ permite distinguir fácilmente la linealidad de este modelo. Por ejemplo, si se duplica el vector voltaje, entonces lo mismo ocurre con el vector corriente. Además, es válido el *principio de superposición*. Es decir, la solución de la ecuación (4) es la suma de las soluciones de las ecuaciones

$$R\mathbf{i} = \begin{bmatrix} 30 \\ 0 \\ 0 \end{bmatrix}, \qquad R\mathbf{i} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix} \quad \text{y} \quad R\mathbf{i} = \begin{bmatrix} 0 \\ 0 \\ -25 \end{bmatrix}$$

Cada ecuación aquí corresponde al circuito con una sola fuente de voltaje (las otras fuentes se remplazan con los alambres que cierran cada circuito). El modelo para el flujo de corriente es *lineal* porque precisamente las leyes de Ohm y de Kirchhoff son lineales. La caída de voltaje en un resistor es *proporcional* a la corriente que fluye por él (Ohm), y la *suma* de las caídas de voltaje en un circuito iguala a la suma de las fuentes de voltaje en el circuito (Kirchhoff).

Las corrientes de circuito en una red se pueden emplear para determinar la corriente en cualquier rama de la red. Si solo una corriente de circuito pasa por una rama, como de *B* a *D* en la figura 1, la corriente de la rama es igual a la corriente de circuito. Si más de una corriente de circuito pasan por una rama, como de *A* a *B*, la corriente de rama es la suma algebraica de las corrientes de circuito en la rama (*ley de Kirchhoff sobre la corriente*). Por ejemplo, la corriente en la rama *AB* es $I_1 - I_2 = 3 - 1 = 2$ A, en la dirección de I_1. La corriente en la rama *CD* es $I_2 - I_3 = 9$ A.

Ecuaciones en diferencias

En muchos campos, como ecología, economía e ingeniería, surge la necesidad de modelar matemáticamente un sistema dinámico que cambia en el tiempo. Diversos aspectos del sistema se miden en intervalos de tiempo discretos, produciendo así una secuencia de vectores $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$,… Las entradas en $\mathbf{x}_k$ brindan información sobre el *estado* del sistema en el momento de la k-ésima medición.

Si existe una matriz A tal que $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$ y, en general,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para} \quad k = 0, 1, 2, \ldots \tag{5}$$

entonces (5) se llama **ecuación lineal en diferencias** (o **relación de recurrencia**). Con esta ecuación, se pueden calcular $\mathbf{x}_1$, $\mathbf{x}_2$, y así sucesivamente, si $\mathbf{x}_0$ se conoce. Las secciones 4.8 y 4.9, así como algunas otras del capítulo 5, desarrollarán fórmulas para $\mathbf{x}_k$ y describirán qué ocurre a $\mathbf{x}_k$ conforme k se incrementa indefinidamente. El análisis que se presenta a continuación ilustra cómo puede originarse una ecuación en diferencias.

Un tema de interés para los demógrafos es el movimiento de poblaciones o grupos de gente de una región a otra. El modelo sencillo que se incluye aquí considera los cambios en la población de una cierta ciudad y sus suburbios durante un periodo de años.

Fije un año inicial —por ejemplo, 2000— y denote las poblaciones de la ciudad y los suburbios de ese año mediante r_0 y s_0, respectivamente. Sea $\mathbf{x}_0$ el vector de población

$$\mathbf{x}_0 = \begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \quad \begin{array}{l} \text{Población de la ciudad, 2000} \\ \text{Población suburbana, 2000} \end{array}$$

Para 2001 y los años subsiguientes, denote las poblaciones de la ciudad y de los suburbios mediante los vectores

$$\mathbf{x}_1 = \begin{bmatrix} r_1 \\ s_1 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} r_2 \\ s_2 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}, \ldots$$

Nuestro objetivo es describir matemáticamente cómo podrían estar relacionados esos vectores.

Suponga que estudios demográficos revelan que, cada año, cerca del 5% de la población de la ciudad se muda a los suburbios (lo que significa que el 95% permanece en la ciudad), mientras que el 3% de la población suburbana cambia su residencia a la ciudad (en tanto que el 97% permanece en los suburbios). Véase la figura 2.

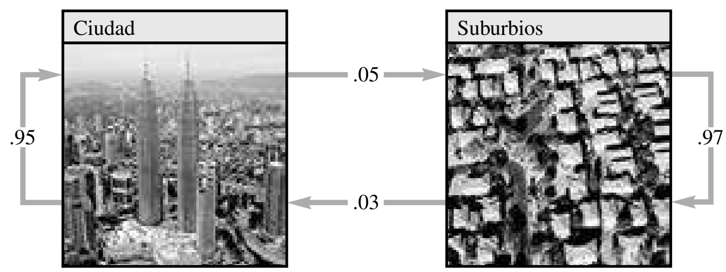


FIGURA 2 Porcentaje anual de migración entre la ciudad y los suburbios.

Después de un año, los habitantes originales de la ciudad, r_0, están ahora distribuidos entre la ciudad y los suburbios como

$$\begin{bmatrix} .95r_0 \\ .05r_0 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} \quad \begin{array}{l} \text{Permanecen en la ciudad} \\ \text{Se mudan a los suburbios} \end{array} \tag{6}$$

Los habitantes de los suburbios en 2000, s_0, están distribuidos un año después como

$$s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} \quad \begin{array}{l} \text{Se mudan a la ciudad} \\ \text{Permanecen en los suburbios} \end{array} \tag{7}$$

Los vectores en (6) y (7) explican cómo se distribuye la población en 2001.[3] Así,

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} + s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

Es decir,

$$\mathbf{x}_1 = M\mathbf{x}_0 \tag{8}$$

donde M es la **matriz de migración** determinada por la siguiente tabla:

$$\begin{array}{ccl} \multicolumn{2}{c}{\text{De:}} & \\ \text{Ciudad} & \text{Suburbios} & \text{A:} \\ \left[\begin{matrix} .95 \\ .05 \end{matrix}\right. & \left.\begin{matrix} .03 \\ .97 \end{matrix}\right] & \begin{matrix} \text{Ciudad} \\ \text{Suburbios} \end{matrix} \end{array}$$

La ecuación (8) describe cómo cambió la población de 2000 a 2001. Si los porcentajes de migración permanecen constantes, entonces el cambio de 2001 a 2002 está dado por

$$\mathbf{x}_2 = M\mathbf{x}_1$$

y, de manera similar, de 2002 a 2003 y en los años subsiguientes. En general,

$$\mathbf{x}_{k+1} = M\mathbf{x}_k \quad \text{para} \quad k = 0, 1, 2, \ldots \tag{9}$$

La secuencia de vectores $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ describe las poblaciones de la ciudad y los suburbios durante un periodo de años.

EJEMPLO 3 Calcule la población de la región que se acaba de describir para los años 2001 y 2002, considerando que la población en el año 2000 era de 600,000 habitantes en la ciudad y de 400,000 en los suburbios.

SOLUCIÓN La población inicial en 2000 es $\mathbf{x}_0 = \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$. Para 2001,

$$\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

Para 2002,

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix} = \begin{bmatrix} 565{,}440 \\ 434{,}560 \end{bmatrix} \qquad \blacksquare$$

El modelo para el movimiento poblacional en (9) es *lineal* porque la correspondencia $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ es una transformación lineal. La linealidad depende de dos hechos: el número de personas que eligen cambiar su residencia de una área a otra es *proporcional* al número de personas en esa área, como se muestra en (6) y (7), y el efecto acumulativo de esas elecciones se encuentra *sumando* el movimiento de las personas de las diferentes áreas.

PROBLEMA DE PRÁCTICA

Encuentre una matriz A y vectores $\mathbf{x}$ y $\mathbf{b}$ tales que el problema del ejemplo 1 signifique resolver la ecuación $A\mathbf{x} = \mathbf{b}$.

[3] Para simplificar, se ignoran otros factores que influyen en la población, como nacimientos, muertes y movimientos migratorios hacia la región que comprende la ciudad y los suburbios, así como hacia fuera de ella.

1.10 EJERCICIOS

1. Por lo regular, el empaque de un cereal indica el número de calorías y las cantidades de proteínas, carbohidratos y grasa contenidas en una porción del producto. A continuación se dan las cantidades para dos cereales comunes. Suponga que se preparará una mezcla de esos dos cereales, de manera que contenga exactamente 295 calorías, 9 g de proteínas, 48 g de carbohidratos y 8 g de grasa.

   a) Establezca una ecuación vectorial para este problema. Incluya un enunciado que explique el significado de cada variable empleada en la ecuación.

   b) Escriba una ecuación matricial equivalente, y determine si es posible preparar la mezcla deseada de los dos cereales.

Información nutricional por cada porción

Nutriente	General Mills Cheerios®	Quaker® 100% Natural Cereal
Calorías	110	130
Proteínas (g)	4	3
Carbohidratos (g)	20	18
Grasa (g)	2	5

2. Una porción de Shredded Wheat aporta 160 calorías, 5 g de proteína, 6 g de fibra y 1 g de grasa. Una porción de Crispix® aporta 110 calorías, 2 g de proteína, .1 g de fibra y .4 g de grasa.

   a) Establezca una matriz B y un vector $\mathbf{u}$ tal que $B\mathbf{u}$ dé las cantidades de calorías, proteínas, fibra y grasa contenidas en una mezcla de tres porciones de Shredded Wheat y dos porciones de Crispix.

   b) [**M**] Suponga que se desea un cereal con más fibra que Crispix, pero con menos calorías que Shredded Wheat. ¿Es posible que una mezcla de ambos cereales proporcione 130 calorías, 3.20 g de proteína, 2.46 g de fibra y .64 g de grasa? Si es posible, ¿cuál es la mezcla?

3. Después de tomar una clase sobre nutrición, una consumidora asidua de los productos de Annie's®, a quien le gusta el producto Mac and Cheese (macarrones con queso), decide mejorar los niveles de proteína y fibra en su almuerzo favorito agregando brócoli y pollo enlatado. La información nutricional de los alimentos mencionados en este ejercicio se indica en la siguiente tabla.

Información nutricional por cada porción

Nutriente	Mac and Cheese	Brócoli	Pollo	Shells
Calorías	270	51	70	260
Proteína (g)	10	5.4	15	9
Fibra (g)	2	5.2	0	5

   a) [**M**] Si ella quiere limitar su almuerzo a 400 calorías, pero desea obtener 30 g de proteína y 10 g de fibra, ¿qué proporciones de Mac and Cheese, brócoli y pollo debería utilizar?

   b) [**M**] Ella encontró que no había mucho brócoli en las proporciones del inciso *a*), así que decidió cambiar del clásico Mac and Cheese a Annie's Whole Wheat Shells and White Cheddar (pasta integral y queso cheddar). ¿Qué proporciones de cada alimento debería emplear para lograr los mismos objetivos que en el inciso *a*)?

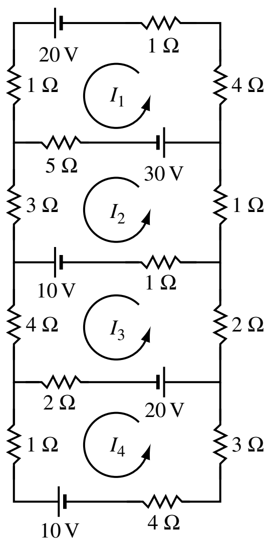
4. La dieta Cambridge aporta .8 g de calcio por día, además de los nutrientes listados en la tabla 1 del ejemplo 1. Las cantidades de calcio por unidad (100 g) que aportan los tres ingredientes en la dieta Cambridge son las siguientes: 1.26 g de leche descremada, .19 g de harina de soya y .8 g de suero de leche. Otro ingrediente en la mezcla de la dieta es proteína aislada de soya, que aporta los siguientes nutrientes en cada unidad: 80 g de proteínas, 0 g de carbohidratos, 3.4 g de grasa y .18 g de calcio.

   a) Establezca una ecuación matricial cuya solución determine las cantidades de leche descremada, harina de soya, suero de leche y proteína aislada de soya necesarias para suministrar las cantidades precisas de proteínas, carbohidratos, grasa y calcio en la dieta Cambridge. Indique qué representan las variables de la ecuación.

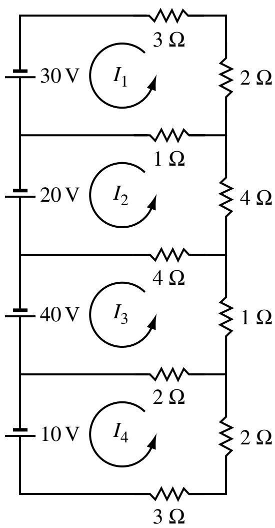
   b) [**M**] Resuelva la ecuación en *a*) y analice la respuesta.

En los ejercicios 5 a 8, escriba una ecuación matricial que determine las corrientes del circuito. [**M**] Si dispone de MATLAB o de algún otro programa de matrices, resuelva el sistema para las corrientes del circuito.

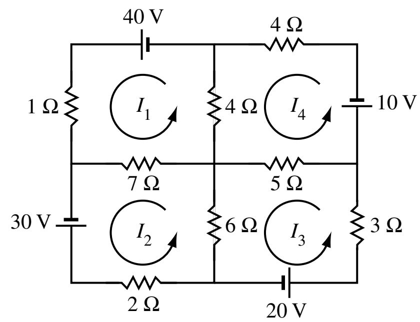
5.



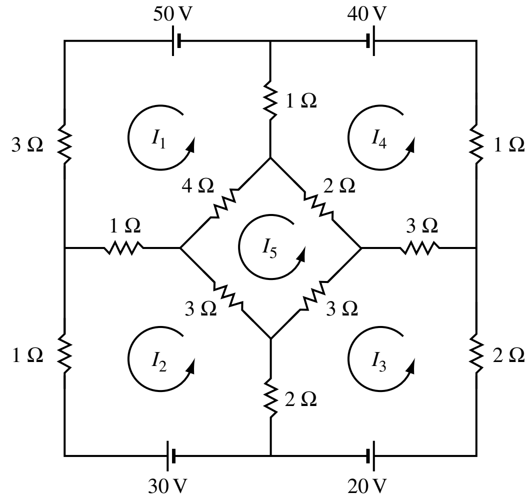
6.



7.



8.



9. En cierta región, cada año, cerca del 7% de la población de una ciudad se muda a los suburbios, y alrededor del 5% de la población suburbana cambia su residencia a la ciudad. En 2010, había 800,000 residentes en la ciudad y 500,000 en los suburbios. Establezca una ecuación en diferencias que describa esta situación, donde $\mathbf{x}_0$ sea la población inicial en 2010. Luego, estime las poblaciones en la ciudad y en los suburbios dos años después, es decir, en 2012. (Ignore otros factores que podrían influir en el tamaño de las poblaciones).

10. Cada año, en cierta región, cerca del 6% de la población de una ciudad se muda a los suburbios, y alrededor del 4% de la población suburbana se muda a la ciudad. En 2010, había 10,000,000 de residentes en la ciudad y 800,000 en los suburbios. Establezca una ecuación en diferencias que describa esta situación, donde $\mathbf{x}_0$ sea la población inicial en 2010. Luego, estime las poblaciones en la ciudad y en los suburbios dos años después, es decir, en 2012.

11. En 1994, la población de California era de 31,524,000 habitantes, y la población que vivía en Estados Unidos, pero *fuera* de California, era de 228,680,000 habitantes. Durante el año, se estimó que 516,100 personas se mudaron de California a otro lugar en Estados Unidos, mientras que 381,262 personas se mudaron a California desde diversos lugares del país.[4]

a) Establezca la matriz de migración para esta situación, utilizando cinco lugares decimales para las tasas de migración entrante y saliente de California. Explique cómo obtuvo la matriz de migración.

b) [**M**] Calcule las poblaciones proyectadas para el año 2000 en California y en otras partes de Estados Unidos, suponiendo que las tasas de migración no cambian durante el periodo de 6 años. (Esos cálculos no toman en cuenta nacimientos, muertes o la migración sustancial de personas a California y a otros lugares de Estados Unidos provenientes de otros países).

[4] Datos de migración suministrados por la Unidad de Investigación Demográfica del Departamento de Finanzas del estado de California.

12. [**M**] Budget® Rent A Car en Wichita, Kansas, tiene una flotilla de 500 vehículos, distribuidos en tres sucursales. Un auto rentado en una sucursal puede devolverse en cualquiera de las tres. Las diversas fracciones de autos devueltos a las tres sucursales se muestran en la matriz que aparece a continuación. Suponga que un lunes hay 295 autos en el aeropuerto (o que se rentan ahí), 55 autos en la sucursal de la zona este, y 150 automóviles en el establecimiento de la zona oeste. ¿Cuál será la distribución aproximada de autos para el miércoles?

$$\begin{array}{cl} \text{Autos rentados en:} & \\ \begin{array}{ccc}\text{Aeropuerto} & \text{Este} & \text{Oeste}\end{array} & \text{Devueltos en:} \\ \begin{bmatrix} .97 & .05 & .10 \\ .00 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix} & \begin{array}{l}\text{Aeropuerto}\\ \text{Este}\\ \text{Oeste}\end{array} \end{array}$$

13. [**M**] Sean M y $\mathbf{x}_0$ como en el ejemplo 3.

a) Calcule los vectores poblacionales $\mathbf{x}_k$ para $k = 1,\ldots, 20$. Analice sus hallazgos.

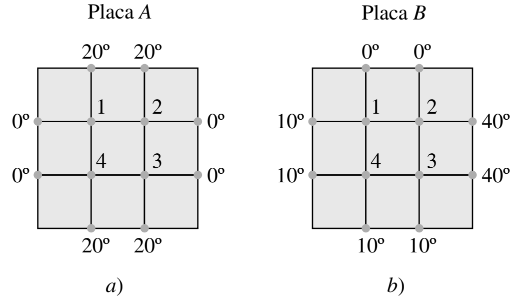
b) Repita el inciso *a*) considerando una población inicial de 350,000 en la ciudad y 650,000 en los suburbios. ¿Qué encontró?

14. [**M**] Estudie cómo los cambios en las temperaturas de los bordes de una placa de acero afectan a las temperaturas en los puntos interiores de la placa.

a) Comience por estimar las temperaturas T_1, T_2, T_3, T_4 en cada uno de los conjuntos de cuatro puntos de la placa que se señalan en la figura. En cada caso, el valor de T_k se aproxima mediante el promedio de las temperaturas en los cuatro puntos más cercanos. Consulte los ejercicios 33 y 34 de la sección 1.1, donde los valores (en grados) son (20, 27.5, 30, 22.5). ¿Cómo se relaciona esta lista de valores con los resultados obtenidos para los puntos en los conjuntos *a*) y *b*)?

b) Sin efectuar cálculos, ¿qué ocurre con las temperaturas interiores en *a*) cuando todas las temperaturas en los bordes se multiplican por 3? Compruebe su suposición.

c) Finalmente, haga una conjetura general sobre la correspondencia de la lista de ocho temperaturas en los bordes con la lista de las cuatro temperaturas interiores.



SOLUCIÓN AL PROBLEMA DE PRÁCTICA

$$A = \begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

CAPÍTULO 1 EJERCICIOS COMPLEMENTARIOS

1. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. (Si el enunciado es válido, cite hechos o teoremas pertinentes. Si es falso, explique por qué o dé un contraejemplo que muestre por qué el enunciado no es verdadero en cada caso).

a) Cada matriz es equivalente por filas a una única matriz en forma escalonada.

b) Cualquier sistema de n ecuaciones lineales con n variables tiene a lo sumo n soluciones.

c) Si un sistema de ecuaciones lineales tiene dos soluciones diferentes, debe tener un número infinito de soluciones.

d) Si un sistema de ecuaciones lineales no tiene variables libres, entonces tiene una solución única.

e) Si una matriz aumentada $[A \quad \mathbf{b}]$ se transforma en $[C \quad \mathbf{d}]$ mediante operaciones elementales de fila, entonces las ecuaciones $A\mathbf{x} = \mathbf{b}$ y $C\mathbf{x} = \mathbf{d}$ tienen exactamente los mismos conjuntos solución.

f) Si un sistema $A\mathbf{x} = \mathbf{b}$ tiene más de una solución, entonces lo mismo sucede con el sistema $A\mathbf{x} = \mathbf{0}$.

g) Si A es una matriz de $m \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para alguna $\mathbf{b}$, entonces las columnas de A generan a $\mathbb{R}^m$.

h) Si una matriz aumentada $[A \quad \mathbf{b}]$ se puede transformar en una forma escalonada mediante operaciones elementales de fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente.

i) Si las matrices A y B son equivalentes por filas, tienen la misma forma escalonada reducida.

j) La ecuación $A\mathbf{x} = \mathbf{0}$ tiene la solución trivial si y solo si no hay variables libres.

k) Si A es una matriz de $m \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para cada $\mathbf{b}$ en $\mathbb{R}^m$, entonces A tiene m columnas pivote.

l) Si una matriz A de $m \times n$ tiene una posición pivote en cada fila, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución única para cada $\mathbf{b}$ en $\mathbb{R}^m$.

m) Si una matriz A de $n \times n$ tiene n posiciones pivote, entonces la forma escalonada reducida de A es la matriz identidad de $n \times n$.

n) Si A y B son matrices de 3×3 con tres posiciones pivote cada una, entonces A se puede transformar en B mediante operaciones elementales de fila.

o) Si A es una matriz de $m \times n$, si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos dos soluciones diferentes, y si la ecuación $A\mathbf{x} = \mathbf{c}$ es consistente, entonces la ecuación $A\mathbf{x} = \mathbf{c}$ tiene muchas soluciones.

p) Si A y B son matrices de $m \times n$ equivalentes por filas y si las columnas de A generan a $\mathbb{R}^m$, entonces también lo hacen las columnas de B.

q) Si ninguno de los vectores en el conjunto $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ en $\mathbb{R}^3$ es un múltiplo de los otros vectores, entonces S es linealmente independiente.

r) Si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es linealmente independiente, entonces $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ no están en $\mathbb{R}^2$.

s) En algunos casos, es posible que cuatro vectores generen a $\mathbb{R}^5$.

t) Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^m$, entonces $-\mathbf{u}$ está en Gen $\{\mathbf{u}, \mathbf{v}\}$.

u) Si $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son vectores diferentes de cero en $\mathbb{R}^2$, entonces $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$.

v) Si $\mathbf{w}$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, entonces $\mathbf{u}$ es una combinación lineal de $\mathbf{v}$ y $\mathbf{w}$.

w) Supongamos que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ están en $\mathbb{R}^5$, $\mathbf{v}_2$ no es múltiplo de $\mathbf{v}_1$, y $\mathbf{v}_3$ no es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente.

x) Una transformación lineal es una función.

y) Si A es una matriz de 6×5, la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ no puede mapear $\mathbb{R}^5$ sobre $\mathbb{R}^6$.

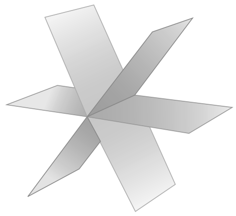
z) Si A es una matriz de $m \times n$ con m columnas pivote, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es un mapeo uno a uno.

2. Sean a y b números reales. Describa los posibles conjuntos solución de la ecuación (lineal) $ax = b$. [*Sugerencia:* El número de soluciones depende de a y b].

3. Las soluciones (x, y, z) de una sola ecuación lineal

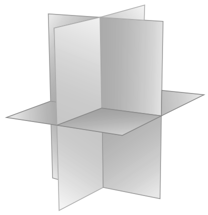
$$ax + by + cz = d$$

forman un plano en $\mathbb{R}^3$ cuando a, b y c no son todos cero. Construya conjuntos de tres ecuaciones lineales cuyas gráficas *a*) se intersecan en una sola recta, *b*) se intersecan en un solo punto, y *c*) no tienen puntos en común. Las siguientes figuras muestran gráficas características.



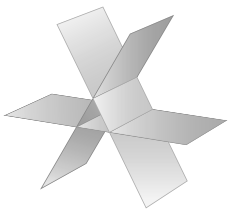
Tres planos se intersecan en una recta

a)



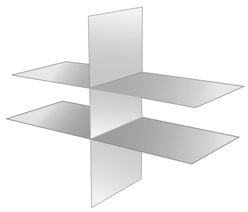
Tres planos se intersecan en un punto

b)



Tres planos sin intersección

c)



Tres planos sin intersección

c')

4. Suponga que la matriz coeficiente de un sistema lineal de tres ecuaciones con tres variables tiene una posición pivote en cada columna. Explique por qué el sistema tiene una solución única.

5. Determine h y k tales que el conjunto solución del sistema: **i.** es un conjunto vacío, **ii.** contiene una solución única, y **iii.** contiene un número infinito de soluciones.

a) $\begin{aligned} x_1 + 3x_2 &= k \\ 4x_1 + hx_2 &= 8 \end{aligned}$ *b*) $\begin{aligned} -2x_1 + hx_2 &= 1 \\ 6x_1 + kx_2 &= -2 \end{aligned}$

6. Considere el problema de determinar si el siguiente sistema de ecuaciones es consistente:

$$\begin{aligned} 4x_1 - 2x_2 + 7x_3 &= -5 \\ 8x_1 - 3x_2 + 10x_3 &= -3 \end{aligned}$$

a) Defina vectores adecuados, y replantee el problema en términos de combinaciones lineales. Luego, resuélvalo.

b) Defina una matriz adecuada, y replantee el problema utilizando la frase "columnas de A".

c) Defina una transformación lineal T adecuada empleando la matriz en *b*), y exponga el problema en términos de T.

7. Considere el problema de determinar si el siguiente sistema de ecuaciones es consistente para toda b_1, b_2, b_3:

$$\begin{aligned} 2x_1 - 4x_2 - 2x_3 &= b_1 \\ -5x_1 + x_2 + x_3 &= b_2 \\ 7x_1 - 5x_2 - 3x_3 &= b_3 \end{aligned}$$

a) Defina vectores adecuados y replantee el problema en términos de Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Luego, resuélvalo.

b) Defina una matriz adecuada y replantee el problema utilizando la frase "columnas de A".

c) Defina una transformación lineal T adecuada utilizando la matriz en *b*), y replantee el problema en términos de T.

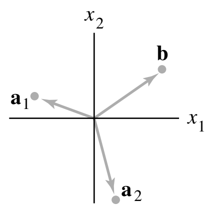
8. Describa las posibles formas escalonadas de la matriz A. Use la notación del ejemplo 1 de la sección 1.2.

a) A es una matriz de 2×3 cuyas columnas generan a $\mathbb{R}^2$.

b) A es una matriz de 3×3 cuyas columnas generan a $\mathbb{R}^3$.

9. Escriba el vector $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ como la suma de dos vectores, uno sobre la recta $\{(x, y): y = 2x\}$ y el otro sobre la recta $\{(x, y): y = x/2\}$.

10. Sean $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}$ los vectores en $\mathbb{R}^2$ que se muestran en la figura, y sea $A = [\mathbf{a}_1 \;\; \mathbf{a}_2]$. ¿Tiene solución la ecuación $A\mathbf{x} = \mathbf{b}$? Si es así, ¿es única la solución? Explique su respuesta.



11. Construya una matriz A de 2×3, que no esté en forma escalonada, de manera que la solución de $A\mathbf{x} = \mathbf{0}$ sea una recta en $\mathbb{R}^3$.

12. Construya una matriz A de 2×3, que no esté en forma escalonada, de manera que la solución de $A\mathbf{x} = \mathbf{0}$ sea un plano en $\mathbb{R}^3$.

13. Escriba la forma escalonada *reducida* de una matriz A de 3×3 de manera que las primeras dos columnas de A sean columnas pivote y

$$A\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

14. Determine el valor o valores de a de manera que $\left\{\begin{bmatrix} 1 \\ a \end{bmatrix}, \begin{bmatrix} a \\ a+2 \end{bmatrix}\right\}$ sea linealmente independiente.

15. En *a*) y *b*), suponga que los vectores son linealmente independientes. ¿Qué puede decir acerca de los números $a, \ldots, f$? Justifique sus respuestas. [*Sugerencia:* Utilice un teorema para *b*)].

a) $\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \end{bmatrix}$ *b*) $\begin{bmatrix} a \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \\ 1 \end{bmatrix}$

16. Con base en el teorema 7 de la sección 1.7, explique por qué las columnas de la matriz A son linealmente independientes.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$$

17. Explique por qué un conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en $\mathbb{R}^5$ debe ser linealmente independiente cuando $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente y $\mathbf{v}_4$ *no* está en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

18. Suponga que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto linealmente independiente en $\mathbb{R}^n$. Demuestre que $\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2\}$ también es linealmente independiente.

19. Suponga que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son distintos puntos sobre una línea en $\mathbb{R}^3$. La recta no necesita pasar por el origen. Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente.

20. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, y suponga que $T(\mathbf{u}) = \mathbf{v}$. Demuestre que $T(-\mathbf{u}) = -\mathbf{v}$.

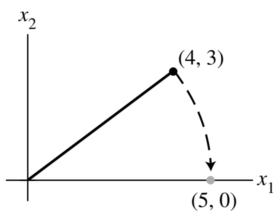
21. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ una transformación lineal que refleja cada vector a través del plano $x_2 = 0$. Es decir, $T(x_1, x_2, x_3) = (x_1, -x_2, x_3)$. Encuentre la matriz estándar de T.

22. Sea A una matriz de 3×3 con la propiedad de que la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^3$ sobre $\mathbb{R}^3$. Explique por qué la transformación debe ser uno a uno.

23. Una *rotación de Givens* es una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^n$ empleada en programas de cómputo para crear una entrada cero en un vector (por lo general, una columna de una matriz). La matriz estándar de una rotación de Givens en $\mathbb{R}^2$ tiene la forma

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \qquad a^2 + b^2 = 1$$

Encuentre a y b tales que $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ gire a $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$.



Una rotación de Givens en $\mathbb{R}^2$.

WEB

24. La siguiente ecuación describe una rotación de Givens en $\mathbb{R}^3$. Encuentre a y b.

$$\begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2\sqrt{5} \\ 3 \\ 0 \end{bmatrix}, \qquad a^2 + b^2 = 1$$

25. Un gran edificio de apartamentos se construirá empleando técnicas de construcción modular. El arreglo de los apartamentos, en cualquier piso en particular, se selecciona a partir de tres planes básicos. El plan A tiene 18 apartamentos en un piso, incluyendo 3 unidades de 3 recámaras, 7 unidades de 2 recámaras, y 8 unidades de 1 recámara. Cada piso del plan B incluye 4 unidades de 3 recámaras, 4 unidades de 2 recámaras, y 8 unidades de 1 recámara. Cada piso del plan C tiene 5 unidades de 3 recámaras, 3 unidades de 2 recámaras, y 9 unidades de 1 recámara. Suponga que el edificio tiene un total de x_1 pisos del plan A, x_2 pisos del plan B, y x_3 pisos del plan C.

a) ¿Qué interpretación puede darse al vector $x_1 \begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix}$?

b) Escriba una combinación lineal formal de vectores que exprese el número total de apartamentos de una, dos y tres recámaras del edificio.

c) [**M**] ¿Es posible diseñar el edificio con exactamente 66 unidades de tres recámaras, 74 unidades de dos recámaras, y 136 unidades de una recámara? Si es así, ¿existe más de una forma de hacerlo? Explique su respuesta.

2 Álgebra de matrices

EJEMPLO INTRODUCTORIO

Modelos de computadora en el diseño de aeronaves

Para diseñar la siguiente generación de aeronaves comerciales y militares, los ingenieros de Boeing's Phantom Works usan el modelado tridimensional (3D) y la dinámica de fluidos computacional (DFC). De esta manera, estudian el flujo de aire alrededor de un avión virtual para responder importantes preguntas de diseño antes de crear los modelos físicos. Este método ha reducido en forma drástica los tiempos y costos del ciclo de diseño; y el álgebra lineal desempeña un papel de suma importancia en el proceso.

La aeronave virtual comienza como un modelo matemático "de alambre" que existe solo en la memoria de la computadora y en las terminales de presentación gráfica. (Se muestra el modelo de un Boeing 777). Este modelo matemático organiza e influye en cada paso del diseño y la manufactura de la aeronave, tanto en el exterior como en el interior. El análisis de DFC concierne a la superficie exterior.

Aunque tal vez el acabado exterior de la aeronave parezca liso, la geometría de la superficie es complicada. Además de alas y fuselaje, un avión tiene cabinas, estabilizadores, dispositivos de sustentación, aletas y alerones. La forma en que el aire fluye alrededor de estas estructuras determina cómo se desplaza la aeronave por el cielo. Las ecuaciones que describen el flujo del aire son complicadas y deben tomar en cuenta la admisión de los motores, los gases expelidos y las estelas que dejan las alas de la aeronave. Para estudiar el flujo del aire, los ingenieros necesitan de una descripción sumamente detallada de la superficie de la aeronave.

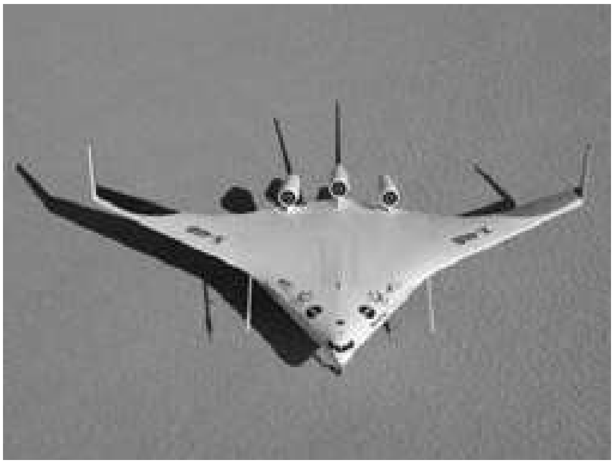
Una computadora crea un modelo de la superficie al superponer, primero, una malla tridimensional de "cuadros" sobre el modelo de alambre original. En esta malla, los cuadros caen completamente dentro o completamente fuera de la aeronave, o se intersecan con la superficie de esta. La computadora selecciona los cuadros que se intersecan con la superficie y los subdivide, reteniendo solo los más pequeños que aún se intersecan con la superficie. El proceso de subdivisión se repite hasta que la malla se vuelve extremadamente fina. Una malla típica puede incluir más de 400,000 cuadros.

El proceso para encontrar el flujo de aire alrededor de la aeronave implica la solución repetida de un sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$ que puede implicar hasta dos millones de ecuaciones y variables. El vector $\mathbf{b}$ cambia a cada momento, con base en datos que provienen de la malla y de las soluciones de ecuaciones previas. Usando las computadoras más rápidas disponibles comercialmente, un equipo de Phantom Works puede tardar desde unas cuantas horas hasta varios días para configurar y resolver un solo problema de flujo de aire. Después de que el equipo analiza la solución, podrá hacer pequeños cambios en la superficie de la aeronave, y comenzar de nuevo todo el proceso. Es posible que se necesiten miles de corridas de DFC.

En este capítulo se presentan dos conceptos importantes que ayudan en la solución de los enormes sistemas de ecuaciones de este tipo:

- *Matrices particionadas:* Un sistema típico de ecuaciones de DFC tiene una matriz de coeficientes "dispersa" con la mayoría de entradas iguales a cero. El agrupar las variables en forma correcta conduce a una matriz particionada con muchos bloques de ceros. En la sección 2.4 se presenta este tipo de matrices y se describen algunas de sus aplicaciones.

- *Factorizaciones de matrices:* El sistema de ecuaciones es complicado, incluso cuando está escrito con matrices particionadas. Para simplificar aún más los cálculos, el programa computacional DFC aplicado en el Boeing utiliza lo que se conoce como factorización LU de la matriz de coeficientes. En la sección 2.5 se analiza la factorización LU y otras factorizaciones matriciales útiles. Más adelante, en varias partes de este libro, aparecen más detalles respecto de las factorizaciones.



La moderna DFC ha revolucionado el diseño de aeronaves. El Boeing Blended Wing Body se encuentra en diseño y entrará en funcionamiento en 2020 o tal vez antes.

Para poder analizar una solución de un sistema de flujo de aire, los ingenieros tienen que visualizar el flujo de aire sobre la superficie de la aeronave; para ello, utilizan gráficos generados por computadora, y el álgebra lineal proporciona las herramientas para trazarlas. El modelo de marco de alambre de la superficie de la aeronave se almacena como datos en muchas matrices. Una vez que se presenta la imagen en una pantalla de computadora, los ingenieros pueden modificarla a escala, hacer acercamientos y alejamientos de zonas pequeñas, y hacerla girar para ver partes que quedan ocultas en determinado ángulo. Cada una de estas operaciones se realiza con una multiplicación adecuada de matrices. En la sección 2.7 se explican las ideas básicas.

WEB

Nuestra capacidad para analizar y resolver ecuaciones aumentará considerablemente al adquirir la habilidad de realizar operaciones algebraicas con matrices. Más aún, las definiciones y los teoremas de este capítulo ofrecen algunas herramientas básicas para manejar las múltiples aplicaciones del álgebra lineal que implican dos o más matrices. Para matrices cuadradas, el teorema de la matriz invertible de la sección 2.3 reúne la mayoría de los conceptos tratados anteriormente en el libro. En las secciones 2.4 y 2.5 se examinan matrices particionadas y factorizaciones de matrices, las cuales aparecen en la mayor parte de los usos modernos del álgebra lineal. En las secciones 2.6 y 2.7 se describen dos aplicaciones interesantes del álgebra de matrices a la economía y a los gráficos generados por computadora.

2.1 OPERACIONES DE MATRICES

Si A es una matriz de $m \times n$, es decir, una matriz con m filas y n columnas, entonces la entrada escalar en la i-ésima fila y la j-ésima columna de A se denota mediante a_{ij} y se llama entrada (i, j) de A. Véase la figura 1. Por ejemplo, la entrada (3, 2) es el número a_{32} en la tercera fila, segunda columna. Cada columna de A es una lista de m números reales, que identifica un vector en $\mathbb{R}^m$. Con frecuencia, estas columnas se denotan mediante $\mathbf{a}_1, \ldots, \mathbf{a}_n$, y la matriz A se escribe como

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

Observe que el número a_{ij} es la i-ésima entrada (desde arriba) del j-ésimo vector columna $\mathbf{a}_j$.

Las **entradas diagonales** en una matriz $A = [a_{ij}]$ de $m \times n$ son $a_{11}, a_{22}, a_{33}, \ldots$, y forman la **diagonal principal** de A. Una **matriz diagonal** es una matriz cuadrada de $n \times n$ cuyas entradas no diagonales son cero. Un ejemplo es la matriz identidad de $n \times n$, I_n. Una matriz de $m \times n$ cuyas entradas son todas cero es una **matriz cero** o **matriz nula** y se escribe como 0. El tamaño de una matriz cero, por lo general, es evidente a partir del contexto.

$$\begin{array}{r c}
 & \begin{array}{ccccc} & & \text{Columna } j & & \end{array} \\
\text{Fila } i & \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} = A \\
 & \begin{array}{ccccc} \uparrow & & \uparrow & & \uparrow \\ \mathbf{a}_1 & & \mathbf{a}_j & & \mathbf{a}_n \end{array}
\end{array}$$

FIGURA 1 Notación matricial.

Suma y múltiplos escalares

La aritmética para vectores que se describió anteriormente tiene una extensión natural hacia las matrices. Se dice que dos matrices son **iguales** si tienen el mismo tamaño (es decir, el mismo número de filas y de columnas) y si sus columnas correspondientes son iguales, lo que equivale a decir que sus entradas correspondientes son iguales. Si A y B son matrices de $m \times n$, entonces la **suma** $A + B$ es la matriz de $m \times n$ cuyas columnas son las sumas de las columnas correspondientes en A y B. Puesto que la suma vectorial de las columnas se realiza por entradas, cada entrada en $A + B$ es la suma de las entradas correspondientes de A y B. La suma $A + B$ está definida solo cuando A y B son del mismo tamaño.

EJEMPLO 1 Sean

$$A = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}$$

Luego,

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

pero $A + C$ *no* está definida porque A y C tienen diferentes tamaños. ■

Si r es un escalar y A es una matriz, entonces el **múltiplo escalar** rA es la matriz cuyas columnas son r veces las columnas correspondientes de A. Al igual que sucede con los vectores, $-A$ significa $(-1)A$, y $A - B$ es igual que $A + (-1)B$.

EJEMPLO 2 Si A y B son las matrices del ejemplo 1, entonces

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$ ■

En el ejemplo 2 no fue necesario calcular $A - 2B$ como $A + (-1)2B$ porque las reglas usuales del álgebra se aplican a las sumas y los múltiplos escalares de matrices, como se verá en el siguiente teorema.

TEOREMA 1

Sean A, B y C matrices del mismo tamaño, y sean r y s escalares.

a) $A + B = B + A$
b) $(A + B) + C = A + (B + C)$
c) $A + 0 = A$
d) $r(A + B) = rA + rB$
e) $(r + s)A = rA + sA$
f) $r(sA) = (rs)A$

Cada igualdad del teorema 1 se comprueba al mostrar que la matriz del lado izquierdo tiene el mismo tamaño que la del lado derecho y que las columnas correspondientes son iguales. El tamaño no es problema porque A, B y C son del mismo tamaño. La igualdad de columnas es consecuencia inmediata de las propiedades análogas de los vectores. Por ejemplo, si las columnas j-ésimas de A, B y C son $\mathbf{a}_j$, $\mathbf{b}_j$ y $\mathbf{c}_j$, respectivamente, entonces las columnas j-ésimas de $(A + B) + C$ y de $A + (B + C)$ son

$$(\mathbf{a}_j + \mathbf{b}_j) + c_j \quad \text{y} \quad \mathbf{a}_j + (\mathbf{b}_j + \mathbf{c}_j),$$

respectivamente. Ya que estas dos sumas vectoriales son iguales para cada j, la propiedad b) queda comprobada.

Debido a la propiedad asociativa de la suma, se puede escribir simplemente $A + B + C$ para la suma, lo cual se calcula como $(A + B) + C$, o bien, como $A + (B + C)$. Lo mismo se aplica a sumas de cuatro o más matrices.

Multiplicación de matrices

Cuando una matriz B multiplica a un vector $\mathbf{x}$, transforma a $\mathbf{x}$ en el vector $B\mathbf{x}$. Si después este vector se multiplica, a la vez, por una matriz A, el vector resultante es $A(B\mathbf{x})$. Véase la figura 2.

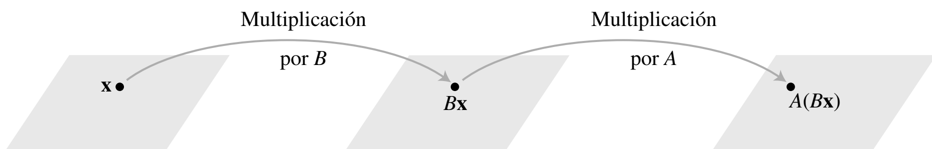


FIGURA 2 Multiplicación por B y luego por A.

Así, $A(B\mathbf{x})$ se produce a partir de $\mathbf{x}$ por una *composición* de mapeos, las transformaciones lineales estudiadas en la sección 1.8. Nuestro objetivo es representar este mapeo compuesto como la multiplicación por una sola matriz, que se denota con AB, de manera que

$$A(B\mathbf{x}) = (AB)\mathbf{x} \tag{1}$$

Véase la figura 3.

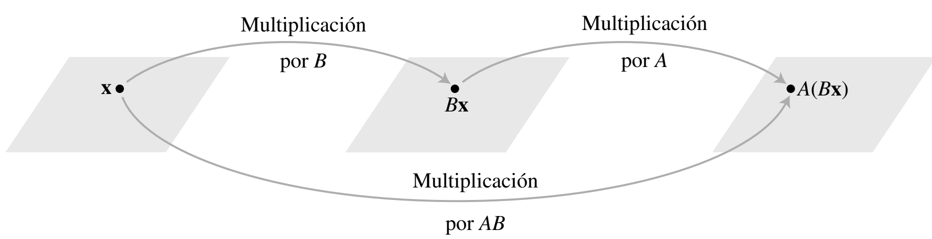


FIGURA 3 Multiplicación por AB.

Si A es de $m \times n$, B es de $n \times p$ y $\mathbf{x}$ está en $\mathbb{R}^p$, las columnas de B se denotan como $\mathbf{b}_j, \ldots, \mathbf{b}_p$, y las entradas de $\mathbf{x}$ como $x_1, \ldots, x_p$. Por consiguiente,

$$B\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_p\mathbf{b}_p$$

Por la linealidad de la multiplicación por A,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1) + \cdots + A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1 + \cdots + x_pA\mathbf{b}_p \end{aligned}$$

El vector $A(B\mathbf{x})$ es una combinación lineal de los vectores $A\mathbf{b}_1,\ldots, A\mathbf{b}_p$, usando las entradas de $\mathbf{x}$ como pesos. En notación matricial, esta combinación lineal se escribe como

$$A(B\mathbf{x}) = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]\mathbf{x}$$

Así, la multiplicación por $[A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$ transforma a $\mathbf{x}$ en $A(B\mathbf{x})$. ¡Ya encontramos la matriz buscada!

DEFINICIÓN

Si A es una matriz de $m \times n$, y si B es una matriz de $n \times p$ con columnas $\mathbf{b}_1,\ldots, \mathbf{b}_p$ entonces el producto AB es la matriz de $m \times p$ cuyas columnas son $A\mathbf{b}_1,\ldots, A\mathbf{b}_p$. Es decir,

$$AB = A[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p] = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p]$$

Esta definición hace que la ecuación (1) sea verdadera para toda $\mathbf{x}$ en $\mathbb{R}^p$. La ecuación (1) demuestra que el mapeo compuesto de la figura 3 es una transformación lineal y que su matriz estándar es AB. *La multiplicación de matrices corresponde a la composición de transformaciones lineales*.

EJEMPLO 3 Calcule AB, donde $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

SOLUCIÓN Escriba $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$, y calcule:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ -1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 13 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix} = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Luego,

$$AB = A[\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{array}$$

■

Observe que como la primera columna de AB es $A\mathbf{b}_1$, esta columna es una combinación lineal de las columnas de A usando como pesos las entradas de $\mathbf{b}_1$. Un enunciado similar es verdadero para cada columna de AB.

Cada columna de AB es una combinación lineal de las columnas de A usando pesos de la columna correspondiente de B.

Evidentemente, el número de columnas de A debe corresponder al número de filas en B para que una combinación lineal como $A\mathbf{b}_1$ esté definida. Además, la definición de AB muestra que *AB tiene el mismo número de filas que A y el mismo número de columnas que B*.

EJEMPLO 4 Si A es una matriz de 3×5 y B una matriz de 5×2, ¿cuáles son los tamaños de AB y de BA, si tales productos están definidos?

SOLUCIÓN Como A tiene 5 columnas y B tiene 5 filas, el producto AB está definido y es una matriz de 3×2:

$$\underset{3\times 5}{\overset{A}{\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}}} \underset{5\times 2}{\overset{B}{\begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix}}} = \underset{3\times 2}{\overset{AB}{\begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix}}}$$

Hay correspondencia

Tamaño de AB

El producto BA *no* está definido, porque las dos columnas de B no se corresponden con las tres filas de A. ■

La definición de AB es importante para el trabajo teórico y las aplicaciones, pero la siguiente regla ofrece un método más eficiente para calcular cada una de las entradas de AB cuando se resuelven a mano problemas sencillos.

REGLA FILA-COLUMNA PARA CALCULAR *AB*

Si el producto AB está definido, entonces la entrada en la fila i y la columna j de AB es la suma de los productos de las entradas correspondientes de la fila i de A y la columna j de B. Si $(AB)_{ij}$ denota la entrada (i, j) en AB, y si A es una matriz de $m \times n$, entonces

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Para comprobar esta regla, sea $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$. La columna j de AB es $A\mathbf{b}_j$, y podemos calcular $A\mathbf{b}_j$ usando la regla fila-vector para calcular $A\mathbf{x}$, como vimos en la sección 1.4. La i-ésima entrada de $A\mathbf{b}_j$ es la suma de los productos de entradas correspondientes de la fila i de A y del vector $\mathbf{b}_j$, que es precisamente el cálculo descrito en la regla para calcular la entrada (i, j) de AB.

EJEMPLO 5 Use la regla fila-columna para calcular dos de las entradas de AB para las matrices del ejemplo 3. Una inspección de los números implicados aclarará cómo los dos métodos para calcular AB producen la misma matriz.

SOLUCIÓN Para encontrar la entrada de la fila 1 y la columna 3 de AB, considere la fila 1 de A y la columna 3 de B. Multiplique las entradas correspondientes y sume los resultados, como se muestra a continuación:

$$AB = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

Para la entrada en la fila 2 y la columna 2 de AB, use la fila 2 de A y la columna 2 de B:

$$\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

■

EJEMPLO 6 Encuentre las entradas de la segunda fila de AB, donde

$$A = \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix}, \qquad B = \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

SOLUCIÓN De acuerdo con la regla fila-columna, las entradas de la segunda fila de AB provienen de la fila 2 de A (y de las columnas de B):

$$\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix}$$ ■

Observe que puesto que el ejemplo 6 pedía solamente la segunda fila de AB, se podría haber escrito solamente la segunda fila de A a la izquierda de B para calcular

$$\begin{bmatrix} -1 & 3 & -4 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 1 \end{bmatrix}$$

Esta observación acerca de las filas de AB es cierta en general, y es consecuencia de la regla fila-columna. Dejemos que $\text{fila}_i\,(A)$ denote la i-ésima fila de una matriz A. De esta forma,

$$\text{fila}_i\,(AB) = \text{fila}_i\,(A) \cdot B \tag{2}$$

Propiedades de la multiplicación de matrices

El siguiente teorema lista las propiedades estándar de la multiplicación de matrices. Recuerde que I_m representa la matriz identidad de $m \times m$, y que $I_m\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^m$.

TEOREMA 2 Sea A una matriz de $m \times n$, y sean B y C matrices con tamaños para los que las sumas y los productos indicados están definidos.

a) $A(BC) = (AB)C$ (ley asociativa de la multiplicación)

b) $A(B + C) = AB + AC$ (ley distributiva izquierda)

c) $(B + C)A = BA + CA$ (ley distributiva derecha)

d) $r(AB) = (rA)B = A(rB)$
para cualquier escalar r

e) $I_mA = A = AI_n$ (identidad para la multiplicación de matrices)

DEMOSTRACIÓN Las propiedades *b*) a *e*) se consideran en los ejercicios. La propiedad *a*) es consecuencia de que la multiplicación de matrices corresponde a la composición de transformaciones lineales (que son funciones) y se sabe (o es fácil comprobar) que la composición de funciones es asociativa. A continuación se presenta otra demostración de *a*) que se basa en la "definición de columna" del producto de dos matrices. Sea

$$C = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_p]$$

Por la definición de multiplicación de matrices,

$$BC = [\, B\mathbf{c}_1 \quad \cdots \quad B\mathbf{c}_p \,]$$
$$A(BC) = [\, A(B\mathbf{c}_1) \quad \cdots \quad A(B\mathbf{c}_p) \,]$$

Recuerde de la ecuación (1) que la definición de AB hace que $A(B\mathbf{x}) = (AB)\mathbf{x}$ para toda $\mathbf{x}$, de manera que

$$A(BC) = [(AB)\mathbf{c}_1 \ \cdots \ (AB)\mathbf{c}_p] = (AB)C \qquad \blacksquare$$

Las leyes asociativa y distributiva de los teoremas 1 y 2 expresan, en esencia, que es posible agregar o eliminar parejas de paréntesis en expresiones matriciales de la misma manera que sucede en el álgebra de números reales. En particular, se puede escribir el producto como ABC y calcularlo ya sea como $A(BC)$ o $(AB)C$.[1] De manera similar, un producto de cuatro matrices $ABCD$ se puede calcular como $A(BCD)$ o $(ABC)D$ o $A(BC)D$, y así sucesivamente. No importa cómo se agrupen las matrices al realizar el cálculo de un producto, siempre y cuando se conserve el orden de izquierda a derecha de las matrices.

El orden de izquierda a derecha en productos es importante porque, en general, AB y BA, *n*o son iguales. Esto no es sorprendente, ya que las columnas de AB son combinaciones lineales de las columnas de A, mientras que las columnas de BA se construyen a partir de las columnas de B. La posición de los factores en el producto AB se enfatiza al decir que A se *multiplica por la derecha* por B o que B *se multiplica por la izquierda* por A. Si $AB = BA$, se dice que A y B **conmutan** una con la otra.

EJEMPLO 7 Sea $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Muestre que estas matrices no conmutan. Es decir, compruebe que $AB \neq BA$.

SOLUCIÓN

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$
$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix} \qquad \blacksquare$$

El ejemplo 7 ilustra la primera de la siguiente lista de diferencias importantes entre el álgebra de matrices y el álgebra de números reales. Para ver ejemplos de estas diferencias, consulte los ejercicios 9 a 12.

ADVERTENCIAS:

1. En general, $AB \neq BA$.
2. Las leyes de la cancelación *no* se aplican en la multiplicación de matrices. Es decir, si $AB = AC$, en general *no* es cierto que $B = C$. (Véase el ejercicio 10).
3. Si un producto AB es la matriz cero, en general, *no* se puede concluir que $A = 0$ o $B = 0$. (Véase el ejercicio 12).

Potencias de una matriz

WEB

Si A es una matriz de $n \times n$ y k es un entero positivo, entonces A^k denota el producto de k

[1] Cuando B es cuadrada y C tiene menos columnas que las filas de A, es más eficiente calcular $A(BC)$ en vez de $(AB)C$.

copias de A:

$$A^k = \underbrace{A \cdots A}_{k}$$

Si A es diferente de cero y $\mathbf{x}$ está en $\mathbb{R}^n$, entonces $A^k\mathbf{x}$ es el resultado de multiplicar repetidamente k veces a $\mathbf{x}$ por la izquierda por A. Si $k = 0$, entonces $A^0\mathbf{x}$ debería ser la misma $\mathbf{x}$. Por consiguiente, A^0 se interpreta como la matriz identidad. Las potencias de matrices son útiles tanto en la teoría como en las aplicaciones (secciones 2.6, 4.9, y más adelante en el libro).

La transpuesta de una matriz

Dada una matriz A de $m \times n$, la **transpuesta** de A es la matriz de $n \times m$, que se denota con A^T, cuyas columnas se forman a partir de las filas correspondientes de A.

EJEMPLO 8 Sean

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

Por lo tanto,

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$

■

TEOREMA 3

Sean A y B matrices cuyos tamaños son adecuados para las siguientes sumas y productos.

a) $(A^T)^T = A$

b) $(A + B)^T = A^T + B^T$

c) Para cualquier escalar r, $(rA)^T = rA^T$

d) $(AB)^T = B^T A^T$

Las demostraciones de *a*) a *c*) son directas y se omiten. Para *d*), véase el ejercicio 33. Por lo regular, $(AB)^T$ no es igual a $A^T B^T$, aun cuando A y B tengan tamaños tales que el producto $A^T B^T$ esté definido.

La generalización del teorema 3*d*) a productos de más de dos factores puede expresarse con palabras de la siguiente forma:

La transpuesta de un producto de matrices es igual al producto de sus transpuestas en orden *inverso*.

Los ejercicios incluyen ejemplos numéricos que ilustran las propiedades de las transpuestas.

NOTAS NUMÉRICAS

1. La manera más rápida de obtener AB en una computadora depende de la forma en que la computadora guarde las matrices en su memoria. Los algoritmos estándar de gran eficiencia, tales como los de LAPACK, calculan AB por columnas, como en nuestra definición del producto. (Una versión de LAPACK escrita en C++ calcula AB por filas).
2. La definición de AB se presta en sí misma al procesamiento paralelo en una computadora. Las columnas de B se asignan individualmente o en grupos a diferentes procesadores que, de manera independiente, y por lo tanto simultánea, calculan las columnas correspondientes de AB.

PROBLEMAS DE PRÁCTICA

1. Puesto que los vectores en $\mathbb{R}^n$ se pueden considerar como matrices de $n \times 1$, las propiedades de las transpuestas del teorema 3 también se aplican a vectores. Sean

$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

Calcule $(A\mathbf{x})^T$, $\mathbf{x}^TA^T$, $\mathbf{x}\mathbf{x}^T$ y $\mathbf{x}^T\mathbf{x}$. ¿Está definida $A^T\mathbf{x}^T$?

2. Sean A una matriz de 4×4 y sea $\mathbf{x}$ un vector en $\mathbb{R}^4$. ¿Cuál es la forma más rápida de calcular $A^2\mathbf{x}$? Cuente las multiplicaciones.

2.1 EJERCICIOS

En los ejercicios 1 y 2, calcule cada suma o producto si la matriz está definida. Si alguna expresión no está definida, explique por qué. Sean

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

1. $-2A$, $B - 2A$, AC, CD
2. $A + 3B$, $2C - 3E$, DB, EC

En lo que resta de este conjunto de ejercicios y los que siguen, suponga que todas las expresiones matriciales están definidas. Es decir, los tamaños de las matrices y los vectores implicados "ajustan" de manera correcta.

3. Sea $A = \begin{bmatrix} 2 & -5 \\ 3 & -2 \end{bmatrix}$, calcule $3I_2 - A$ y $(3I_2)A$.

4. Calcule $A - 5I_3$ y $(5I_3)A$, donde

$$A = \begin{bmatrix} 5 & -1 & 3 \\ -4 & 3 & -6 \\ -3 & 1 & 2 \end{bmatrix}.$$

En los ejercicios 5 y 6 calcule el producto AB de dos maneras: *a*) con la definición, donde $A\mathbf{b}_1$ y $A\mathbf{b}_2$ se calculan por separado, y *b*) por la regla de la fila-columna para obtener AB.

5. $A = \begin{bmatrix} -1 & 3 \\ 2 & 4 \\ 5 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix}$

6. $A = \begin{bmatrix} 4 & -3 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 \\ 3 & -2 \end{bmatrix}$

7. Si A es una matriz de 5×3 y el producto AB es 5×7, ¿cuál es el tamaño de B?

8. ¿Cuántas filas tiene B si BC es una matriz de 5×4?

9. Sean $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 9 \\ -3 & k \end{bmatrix}$. ¿Qué valor(es) de k, si los hay, harán que $AB = BA$?

10. Sea $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$ y $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$. Compruebe que $AB = AC$ y, sin embargo, $B \neq C$.

11. Sean $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ y $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Calcule AD y DA. Explique cómo cambian las columnas o filas de A cuando se multiplica por D por la derecha o por la izquierda. Encuentre una matriz B de 3×3, que no sea la matriz identidad o la matriz cero, tal que $AB = BA$.

12. Sea $A = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}$. Construya una matriz B de 2×2 tal que AB sea igual a la matriz cero. Utilice para B dos diferentes columnas no nulas (distintas de cero).

13. Sean $\mathbf{r}_1, \ldots, \mathbf{r}_p$ vectores en $\mathbb{R}^n$, y sea Q una matriz de $m \times n$. Escriba la matriz $[Q\mathbf{r}_1 \ \cdots \ Q\mathbf{r}_p]$ como un *producto* de dos matrices (ninguna de ellas igual a la matriz identidad).

14. Sea U la matriz de 3×2 de costos descrita en el ejemplo 6 de la sección 1.8. La primera columna de U lista los costos por dólar de producción para elaborar el producto B, y la segunda columna lista los costos por dólar de producción para el artículo C. (Los costos están por categorías de materiales, mano de obra y gastos indirectos). Sea $\mathbf{q}_1$ un vector en $\mathbb{R}^2$ que liste la producción (medida en dólares) de los productos B y C manufacturados durante el primer trimestre del año, y sean $\mathbf{q}_2$, $\mathbf{q}_3$ y $\mathbf{q}_4$ los vectores análogos que listan las cantidades de productos B y C manufacturados en el segundo, tercero y cuarto trimestres, respectivamente. Dé una descripción económica de los datos en la matriz UQ, donde $Q = \{\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3 \quad \mathbf{q}_4\}$.

Los ejercicios 15 y 16 tratan de matrices arbitrarias A, B y C para las cuales las sumas y los productos indicados están definidos. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

15. *a*) Si A y B son matrices de 2×2 con columnas $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{b}_1$, $\mathbf{b}_2$, respectivamente, entonces $AB = [\mathbf{a}_1\mathbf{b}_1 \quad \mathbf{a}_2\mathbf{b}_2]$.

b) Cada columna de AB es una combinación lineal de las columnas de B usando los pesos de la columna correspondiente de A.

c) $AB + AC = A(B + C)$

d) $A^T + B^T = (A + B)^T$

e) La transpuesta de un producto de matrices es igual al producto de sus transpuestas en el mismo orden.

16. *a*) La primera fila de AB es la primera fila de A multiplicada por B por la derecha.

b) Si A y B son matrices de 3×3 y $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$, entonces $AB = [A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3]$.

c) Si A es una matriz de $n \times n$, entonces $(A^2)^T = (A^T)^2$

d) $(ABC)^T = C^TA^TB^T$

e) La transpuesta de una suma de matrices es igual a la suma de sus transpuestas.

17. Si $A = \begin{bmatrix} 1 & -3 \\ -3 & 5 \end{bmatrix}$ y $AB = \begin{bmatrix} -3 & -11 \\ 1 & 17 \end{bmatrix}$, determine la primera y la segunda columnas de B.

18. Suponga que la tercera columna de B está conformada en su totalidad por ceros. ¿Qué se puede decir acerca de la tercera columna de AB?

19. Suponga que la tercera columna de B es la suma de las dos primeras columnas. ¿Qué se puede decir acerca de la tercera columna de AB? ¿Por qué?

20. Suponga que las dos primeras columnas de B, $\mathbf{b}_1$ y $\mathbf{b}_2$, son guales. ¿Qué se puede decir acerca de las columnas de AB? ¿Por qué?

21. Suponga que la última columna de AB está conformada en su totalidad por ceros, pero B, por sí sola, no tiene columnas de ceros. ¿Qué se puede decir acerca de las columnas de A?

22. Demuestre que si las columnas de B son linealmente dependientes, entonces también lo son las columnas de AB.

23. Suponga que $CA = I_n$ (la matriz identidad de $n \times n$). Demuestre que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Explique por qué A no puede tener más columnas que filas.

24. Suponga que A es una matriz de $3 \times n$ cuyas columnas generan a $\mathbb{R}^3$. Explique cómo construir una matriz D de $n \times 3$ tal que $AD = I_3$.

25. Suponga que A es una matriz de $m \times n$, y que existan las matrices C y D de $n \times m$, tales que $CA = I_n$ y $AD = I_m$. Demuestre que $m = n$ y $C = D$. [*Sugerencia:* Piense en el producto CAD].

26. Suponga que $AD = I_m$ (la matriz identidad de $m \times m$). Demuestre que para toda $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución. [*Sugerencia:* Piense en la ecuación $AD\mathbf{b} = \mathbf{b}$]. Explique por qué A no puede tener más filas que columnas.

En los ejercicios 27 y 28, considere los vectores en $\mathbb{R}^n$ como matrices de $n \times 1$. Para $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, el producto de matrices $\mathbf{u}^T\mathbf{v}$ es una matriz de 1×1, que se llama **producto escalar**, o **producto interno**, de $\mathbf{u}$ y $\mathbf{v}$. Por lo general, se escribe como un único número real sin corchetes. El producto de matrices $\mathbf{u}\mathbf{v}^T$ es una matriz de $n \times n$, que se llama **producto exterior** de $\mathbf{u}$ y $\mathbf{v}$. Los productos $\mathbf{u}^T\mathbf{v}$ y $\mathbf{u}\mathbf{v}^T$ se presentarán más adelante en el libro.

27. Sea $\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ -5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Calcule $\mathbf{u}^T\mathbf{v}$, $\mathbf{v}^T\mathbf{u}$, $\mathbf{u}\mathbf{v}^T$ y $\mathbf{v}\mathbf{u}^T$.

28. Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^n$, ¿cómo se relacionan $\mathbf{u}^T\mathbf{v}$ y $\mathbf{v}^T\mathbf{u}$? ¿Y cómo se relacionan $\mathbf{u}\mathbf{v}^T$ y $\mathbf{v}\mathbf{u}^T$?

29. Compruebe el teorema 2*b*) y 2*c*). Use la regla fila-columna. La entrada (i, j) de $A(B + C)$ se puede escribir como

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

o

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

30. Compruebe el teorema 2*d*). [*Sugerencia:* La entrada (i, j) en $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$].

31. Demuestre que $I_mA = A$, donde A es una matriz de $m \times n$. Suponga que $I_m\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^m$.

32. Demuestre que $AI = A$ cuando A es una matriz de $m \times n$. [*Sugerencia:* Use la definición (de columna) de AI_n].

33. Demuestre el teorema 3*d*). [*Sugerencia:* Considere la j-ésima fila de $(AB)^T$].

34. Dé una fórmula para $(AB\mathbf{x})^T$, donde $\mathbf{x}$ es un vector, y A y B son matrices con los tamaños adecuados.

35. **[M]** Lea la documentación de su programa de matrices y escriba los comandos que producirían las siguientes matrices (sin introducir cada entrada de la matriz).

a) Una matriz de 4×5 de ceros.

b) Una matriz de 5×3 de unos.

c) La matriz identidad de 5×5.

d) Una matriz diagonal de 4×4, con entradas diagonales 3, 4, 2, 5.

Una forma útil de someter a prueba ideas nuevas o de hacer conjeturas en álgebra de matrices es realizar cálculos con matrices seleccionadas en forma aleatoria. La comprobación de una propiedad para unas cuantas matrices no demuestra que la propiedad sea válida en general, pero permite que la propiedad sea más creíble. Además, es posible descubrir si una propiedad es falsa realizando unos cuantos cálculos.

36. **[M]** Escriba el comando o los comandos necesarios para crear una matriz de 5×6 con entradas aleatorias. ¿Dentro de qué rango de números se encuentran las entradas? Diga cómo crear aleatoriamente una matriz de 4×4 con entradas enteras entre -9 y 9. [*Sugerencia:* Si x es un número aleatorio tal que $0 < x < 1$, entonces $-9.5 < 19(x - .5) < 9.5$].

37. **[M]** Construya matrices aleatorias A y B de 4×4, y compruebe si $AB = BA$. La mejor manera de hacer esto es calcular $AB - BA$ y comprobar si esta diferencia es la matriz cero. Después compruebe $AB - BA$ para tres pares más de matrices aleatorias de 4×4. Escriba un informe de sus conclusiones.

38. **[M]** Construya una matriz aleatoria A de 5×5 y compruebe si $(A + I)(A - I) = A^2 - I$. La mejor manera de hacer esto es calcular $(A + I)(A - I) - (A^2 - I)$, y verificar que esta diferencia sea la matriz cero. Realícelo para tres matrices al azar. Luego, someta a prueba $(A + B)(A - B) = A^2 - B^2$ de la misma forma para tres pares de matrices aleatorias de 4×4. Escriba un informe de sus conclusiones.

39. **[M]** Use al menos tres pares de matrices aleatorias A y B de 4×4 para someter a prueba las igualdades $(A + B)^T = A^T + B^T$ y $(AB)^T = B^TA^T$, así como $(AB)^T = A^TB^T$ (Véase el ejercicio 37). Escriba un informe de sus conclusiones. [*Nota:* La mayoría de los programas de matrices usan A' para representar A^T].

40. **[M]** Sea

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Calcule S^k para $k = 2, \dots, 6$.

41. **[M]** Describa con palabras qué ocurre cuando se calcula A^5, A^{10}, A^{20} y A^{30} para

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \\ 1/4 & 1/6 & 7/12 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. De manera que $(A\mathbf{x})^T = [-4 \quad 2]$. También,

$$\mathbf{x}^TA^T = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 2 \end{bmatrix}$$

Las cantidades $(A\mathbf{x})^T$ y $\mathbf{x}^TA^T$ son iguales, por el teorema 3*d*). Después,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = [25 + 9] = 34$$

Una matriz de 1×1 como $\mathbf{x}^T\mathbf{x}$ generalmente se escribe sin corchetes. Por último, $A^T\mathbf{x}^T$ no está definida, ya que $\mathbf{x}^T$ no tiene dos filas que correspondan a las dos columnas de A^T.

2. La manera más rápida de calcular $A^2\mathbf{x}$ es calculando $A(A\mathbf{x})$. El producto $A\mathbf{x}$ requiere 16 multiplicaciones, 4 por cada entrada, y $A(A\mathbf{x})$ requiere 16 más. En contraste, el producto A^2 requiere 64 multiplicaciones, 4 por cada una de las 16 entradas en A^2. Después de eso, $A^2\mathbf{x}$ requiere 16 multiplicaciones más, para un total de 80.

2.2 LA INVERSA DE UNA MATRIZ

El álgebra de matrices brinda herramientas para manejar ecuaciones matriciales y crear diversas fórmulas útiles, de manera similar a lo que sucede en el álgebra con números reales. En esta sección se investiga el análogo matricial del recíproco, o inverso multiplicativo, de un número diferente de cero.

Recuerde que el inverso multiplicativo de un número como 5 es $1/5$ o 5^{-1}. Este inverso satisface la ecuación

$$5^{-1} \cdot 5 = 1 \quad \text{y} \quad 5 \cdot 5^{-1} = 1$$

La generalización matricial requiere *ambas* ecuaciones y evita la notación con diagonales (para división), ya que la multiplicación de matrices no es conmutativa. Además, una generalización completa solo es posible si las matrices implicadas son cuadradas.[1]

Se dice que una matriz A de $n \times n$ es **invertible** si existe otra matriz C de $n \times n$ tal que

$$CA = I \quad \text{y} \quad AC = 1$$

donde $I = I_n$, la matriz identidad de $n \times n$. En este caso, C es una **inversa** de A. En efecto, C está determinada únicamente por A, ya que si B fuera otra inversa de A, entonces $B = BI = B(AC) = (BA)C = IC = C$. Esta inversa única se denota mediante A^{-1}, tal que,

$$\boxed{A^{-1}A = I \quad \text{y} \quad AA^{-1} = I}$$

Una matriz que *no* es invertible en ocasiones se llama **matriz singular**, y una matriz invertible se llama **matriz no singular**.

EJEMPLO 1 Si $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ y $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$ entonces

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y}$$

$$CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Por lo tanto, $C = A^{-1}$. ■

A continuación se presenta una fórmula sencilla para la inversa de una matriz de 2×2, junto con una prueba para saber si existe la inversa.

TEOREMA 4

Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Si $ad - bc \neq 0$, entonces A es invertible y

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Si $ad - bc = 0$, entonces A *no* es invertible.

La sencilla demostración del teorema 4 se esboza en los ejercicios 25 y 26. La cantidad $ad - bc$ se llama **determinante** de A, y se escribe como

$$\boxed{\det A = ad - bc}$$

El teorema 4 establece que una matriz A de 2×2 es invertible si y solo si $\det A \neq 0$.

[1] Podría decirse que una matriz A de $m \times n$ es invertible si existen matrices C y D de $n \times m$, tales que $CA = I_n$ y $AD = I_m$. Sin embargo, estas ecuaciones implican que A es cuadrada y $C = D$. Por lo tanto, A es invertible como ya se definió. Véase los ejercicios 23 a 25 en la sección 2.1.

EJEMPLO 2 Encuentre la inversa de $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

SOLUCIÓN Ya que $\det A = 3(6) - 4(5) = -2 \neq 0$, A es invertible, y

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \qquad \blacksquare$$

Las matrices invertibles son indispensables en álgebra lineal, sobre todo para cálculos algebraicos y deducciones de fórmulas, como en el teorema siguiente. En ocasiones, una matriz inversa permite entender mejor un modelo matemático de alguna situación de la vida real, como en el ejemplo 3 que se presenta un poco más adelante.

TEOREMA 5

Si A es una matriz invertible de $n \times n$, entonces, para cada $\mathbf{b}$ en $\mathbb{R}^n$, la ecuación $A\mathbf{x} = \boldsymbol{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$.

DEMOSTRACIÓN Tome cualquier $\mathbf{b}$ en $\mathbb{R}^n$. Existe una solución porque cuando se sustituye $A^{-1}\mathbf{b}$ por $\mathbf{x}$, se tiene $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. Así que $A^{-1}\mathbf{b}$ es una solución. Para probar que la solución es única, demuestre que si $\mathbf{u}$ es cualquier solución, entonces $\mathbf{u}$ debe ser, de hecho, $A^{-1}\mathbf{b}$. En efecto, si $A\mathbf{u} = \mathbf{b}$, podemos multiplicar ambos miembros por A^{-1} y obtener

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \quad I\mathbf{u} = A^{-1}\mathbf{b} \quad \text{y} \quad \mathbf{u} = A^{-1}\mathbf{b} \qquad \blacksquare$$

EJEMPLO 3 Una viga elástica horizontal tiene soportes en cada extremo y está sujeta a fuerzas en los puntos 1, 2 y 3, como indica la figura 1. Sea $\mathbf{f}$ en $\mathbb{R}^3$ tal que liste las fuerzas en estos puntos, y sea $\mathbf{y}$ en $\mathbb{R}^3$ tal que liste las magnitudes de deflexión (es decir, de movimiento) de la viga en los tres puntos. Con base en la ley de Hooke de la física, se puede demostrar que

$$\mathbf{y} = D\mathbf{f}$$

donde D es una *matriz de flexibilidad*. Su inversa se denomina *matriz de rigidez*. Describa el significado físico de las columnas de D y D^{-1}.

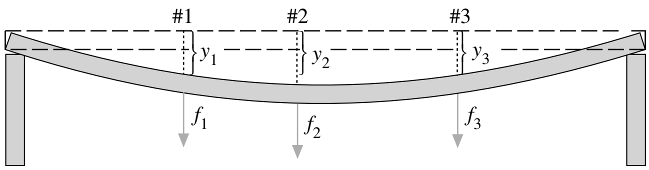


FIGURA 1 Deflexión de una viga elástica.

SOLUCIÓN Escriba $I_3 = [\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3]$ y observe que

$$D = DI_3 = [D\mathbf{e}_1 \quad D\mathbf{e}_2 \quad D\mathbf{e}_3]$$

Interprete el vector $\mathbf{e}_1 = (1, 0, 0)$ como una fuerza unitaria aplicada hacia abajo en el punto 1 (con fuerza cero en los otros dos puntos). De esta forma, $D\mathbf{e}_1$, la primera columna de D, lista las deflexiones de la viga debidas a una fuerza unitaria en el punto 1. Descripciones similares son válidas para la segunda y tercera columnas de D.

Para estudiar la matriz de rigidez D^{-1}, observe que la ecuación $\mathbf{f} = D^{-1}\mathbf{y}$ calcula un vector de fuerza $\mathbf{f}$ cuando se da un vector de deflexión $\mathbf{y}$. Escriba

$$D^{-1} = D^{-1}I_3 = [D^{-1}\mathbf{e}_1 \quad D^{-1}\mathbf{e}_2 \quad D^{-1}\mathbf{e}_3]$$

Ahora interprete $\mathbf{e}_1$ como un vector de deflexión. De esta forma, $D^{-1}\mathbf{e}_1$ lista las fuerzas que crean la deflexión. Es decir, la primera columna de D^{-1} lista las fuerzas que deben aplicarse en los tres puntos para producir una deflexión unitaria en el punto 1 y deflexión cero en los otros puntos. De manera similar, las columnas 2 y 3 de D^{-1} listan las fuerzas requeridas para producir deflexiones unitarias en los puntos 2 y 3, respectivamente. En cada columna, una o dos de las fuerzas deben ser negativas (apuntan hacia arriba) para producir una deflexión unitaria en el punto deseado y deflexiones cero en los otros dos puntos. Si se mide la flexibilidad, por ejemplo, en pulgadas de deflexión por libra de carga, entonces las entradas de la matriz de rigidez están dadas en libras de carga por pulgada de deflexión. ■

La fórmula del teorema 5 se utiliza muy pocas veces para resolver en forma numérica una ecuación $A\mathbf{x} = \mathbf{b}$ porque la reducción por filas de $[A \quad \mathbf{b}]$ casi siempre es más rápida. (La reducción por filas también es más precisa, en general, cuando los cálculos requieren el redondeo de los números). Una posible excepción es el caso 2×2, ya que los cálculos mentales para resolver $A\mathbf{x} = \mathbf{b}$ en ocasiones resultan más fáciles usando la fórmula para A^{-1}, como en el siguiente ejemplo.

EJEMPLO 4 Use la inversa de la matriz A del ejemplo 2 para resolver el sistema

$$\begin{aligned} 3x_1 + 4x_2 &= 3 \\ 5x_1 + 6x_2 &= 7 \end{aligned}$$

SOLUCIÓN Este sistema es equivalente a $A\mathbf{x} = \mathbf{b}$, por lo que

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$ ■

El siguiente teorema aporta tres datos útiles acerca de las matrices invertibles.

TEOREMA 6

a) Si A es una matriz invertible, entonces A^{-1} es invertible y

$$(A^{-1})^{-1} = A$$

b) Si A y B son matrices invertibles de $n \times n$, entonces también lo es AB, y la inversa de AB es el producto de las inversas de A y B en el orden opuesto. Es decir,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c) Si A es una matriz invertible, también lo es A^T, y la inversa de A^T es la transpuesta de A^{-1}. Es decir,

$$(A^T)^{-1} = (A^{-1})^T$$

DEMOSTRACIÓN Para comprobar el enunciado *a*), se debe encontrar una matriz C tal que

$$A^{-1}C = I \quad \text{y} \quad CA^{-1} = I$$

De hecho, estas ecuaciones se satisfacen colocando a A en lugar de C. Por lo tanto, A^{-1} es invertible y A es su inversa. A continuación, para demostrar el enunciado *b*), se calcula:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

Un cálculo similar indica que $(B^{-1}A^{-1})(AB) = I$. En el enunciado *c*) utilice el teorema 3*d*), lea de derecha a izquierda, $(A^{-1})^TA^T = (AA^{-1})^T = I^T = I$. De manera similar, $A^T(A^{-1})^T = I^T = I$. Por lo tanto, A^T es invertible, y su inversa es $(A^{-1})^T$. ■

La siguiente generalización del teorema 6*b*) se necesitará más adelante.

El producto de matrices invertibles de $n \times n$ es invertible, y la inversa es el producto de sus inversas en orden opuesto.

Existe una conexión importante entre las matrices invertibles y las operaciones de fila que conduce a un método para calcular inversas. Como se verá, una matriz invertible A es equivalente por filas a una matriz identidad, y es posible encontrar A^{-1} *observando la reducción por filas de A a I.*

Matrices elementales

Una **matriz elemental** es aquella que se obtiene al realizar una única operación elemental de fila sobre una matriz identidad. El siguiente ejemplo ilustra los tres tipos de matrices elementales.

EJEMPLO 5 Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$
$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Calcule E_1A, E_2A y E_3A, y describa cómo se pueden obtener estos productos por medio de operaciones elementales de fila sobre A.

SOLUCIÓN Compruebe que

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$
$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

Al sumar a la fila 3 la fila 1 de A multiplicada por -4, se obtiene E_1A. (Esta es una operación de remplazo de filas). Con un intercambio de las filas 1 y 2 de A se obtiene E_2A, y multiplicando la fila 3 de A por 5 se obtiene E_3A. ■

La multiplicación izquierda (es decir, multiplicación por la izquierda) por E_1 en el ejemplo 5 tiene el mismo efecto en cualquier matriz de $3 \times n$. Esta operación suma a la fila 3 la fila 1 multiplicada por -4. En particular, ya que $E_1 \cdot I = E_1$, vemos que E_1 *se autoproduce* mediante esta misma operación de fila sobre la identidad. Así, el ejemplo 5 ilustra la siguiente propiedad general de las matrices elementales. Véase los ejercicios 27 y 28.

> Si se realiza una operación elemental de fila con una matriz A de $m \times n$, la matriz resultante se puede escribir como EA, donde la matriz E de $m \times m$ se crea al realizar la misma operación de fila sobre I_m.

Puesto que las operaciones de fila son reversibles, como se demostró en la sección 1.1, las matrices elementales son invertibles, porque si E se produce aplicando una operación de fila sobre I, entonces existe otra operación de fila del mismo tipo que convierte a E de nuevo en I. Por lo tanto, existe una matriz elemental F tal que $FE = I$. Puesto que E y F corresponden a operaciones inversas, también $EF = I$.

Toda matriz elemental E es invertible. La inversa de E es la matriz elemental del mismo tipo que transforma a E de nuevo en I.

EJEMPLO 6 Encuentre la inversa de $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

SOLUCIÓN Para transformar E_1 en I, sume la fila 1 multiplicada por $+4$ a la fila 3. La matriz elemental que hace esto es

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$ ■

El siguiente teorema ofrece la mejor manera de "visualizar" una matriz invertible, y conduce de inmediato a un método para encontrar la inversa de una matriz.

TEOREMA 7

Una matriz A de $n \times n$ es invertible si y solo si A es equivalente por filas a I_n, y, en este caso, cualquier secuencia de operaciones elementales de fila que reduzca A a I_n también transforma a I_n en A^{-1}.

DEMOSTRACIÓN Suponga que A es invertible. Entonces, como la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ (teorema 5), A tiene una posición pivote en cada fila (teorema 4 de la sección 1.4). Puesto que A es cuadrada, las n posiciones pivote deben estar sobre la diagonal, lo que implica que la forma escalonada reducida de A es $\mathrm{I_n}$. Es decir, $A \sim I_n$.

A la inversa, ahora suponga que $A \sim I_n$. Entonces, puesto que cada paso de la reducción por filas de A corresponde a una multiplicación izquierda por una matriz elemental, existen matrices elementales $E_1,\dots, E_P$ tales que

$$A \sim E_1 A \sim E_2(E_1 A) \sim \cdots \sim E_p(E_{p-1} \cdots E_1 A) = I_n$$

Es decir,

$$E_p \cdots E_1 A = I_n \tag{1}$$

Puesto que el producto $E_p,\dots, E_1$ de matrices invertibles es invertible, (1) conduce a

$$\begin{aligned} (E_p \cdots E_1)^{-1}(E_p \cdots E_1)A &= (E_p \cdots E_1)^{-1} I_n \\ A &= (E_p \cdots E_1)^{-1} \end{aligned}$$

Por lo tanto, A es invertible, porque es la inversa de una matriz invertible (teorema 6). También,

$$A^{-1} = [\,(E_p \cdots E_1)^{-1}\,]^{-1} = E_p \cdots E_1$$

Así, $A^{-1} = E_p \cdots E_1 \cdot I_n$, lo que indica que A^{-1} resulta de aplicar $E_1,\dots, E_P$ sucesivamente a I_n. Esta es la misma secuencia en (1) que redujo A a I_n. ■

Un algoritmo para determinar A^{-1}

Si colocamos A e I lado a lado para formar una matriz aumentada $[A \quad I]$, entonces las operaciones de fila en esta matriz producen operaciones idénticas sobre A e I. De acuerdo con el teorema 7, hay operaciones de fila que transforman a A en I_n y a I_n en A^{-1}, o A no es invertible.

ALGORITMO PARA DETERMINAR A^{-1}

Reduzca por filas la matriz aumentada $[A \quad I]$. Si A es equivalente por filas a I, entonces $[A \quad I]$ es equivalente por filas a $[I \quad A^{-1}]$. De otra manera, A no tiene inversa.

EJEMPLO 7 Encuentre la inversa de la matriz $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, si acaso existe.

SOLUCIÓN

$$[A \quad I] = \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix}$$

El teorema 7 señala que, como $A \sim I$, A es invertible, y

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

Es buena idea comprobar la respuesta final:

$$AA^{-1} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix} \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

No es necesario comprobar que $A^{-1}A = I$, ya que A es invertible. ■

Otro punto de vista de la inversión de matrices

Denote las columnas de I_n por $\mathbf{e}_1, \ldots, \mathbf{e}_n$. De esta forma, la reducción por filas de $[A \quad I]$ a $[I \quad A^{-1}]$ se puede ver como la solución simultánea de los n sistemas

$$A\mathbf{x} = \mathbf{e}_1, \quad A\mathbf{x} = \mathbf{e}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{e}_n \tag{2}$$

donde todas las "columnas aumentadas" de estos sistemas se han colocado al lado de A para formar $[A \quad \mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n] = [A \quad I]$. La ecuación $AA^{-1} = I$ y la definición de multiplicación de matrices indican que las columnas de A^{-1} son precisamente las soluciones de los sistemas de (2). Esta observación es útil porque algunos problemas aplicados requieren encontrar solamente una o dos columnas de A^{-1}. En este caso, solo se necesita resolver los sistemas correspondientes en (2).

NOTAS NUMÉRICAS

WEB

En la práctica, rara vez se calcula A^{-1}, a menos que se necesiten las entradas de A^{-1}. Calcular tanto A^{-1} como $A^{-1}\mathbf{b}$ requiere aproximadamente tres veces más operaciones aritméticas que resolver con reducción por filas $A\mathbf{x} = \mathbf{b}$, y la reducción por filas quizá resulte más precisa.

PROBLEMAS DE PRÁCTICA

1. Utilice determinantes para establecer cuáles de las siguientes matrices son invertibles:

a) $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ *b*) $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ *c*) $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. Encuentre la inversa de la matriz $A = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$, si existe.

2.2 EJERCICIOS

Encuentre las inversas de las matrices en los ejercicios 1 a 4.

1. $\begin{bmatrix} 8 & 6 \\ 5 & 4 \end{bmatrix}$

2. $\begin{bmatrix} 3 & 2 \\ 8 & 5 \end{bmatrix}$

3. $\begin{bmatrix} 7 & 3 \\ -6 & -3 \end{bmatrix}$

4. $\begin{bmatrix} 2 & -4 \\ 4 & -6 \end{bmatrix}$

5. Utilice la inversa que encontró en el ejercicio 1 para resolver el sistema

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. Utilice la inversa que encontró en el ejercicio 3 para resolver el sistema

$$\begin{aligned} 7x_1 + 3x_2 &= -9 \\ -6x_1 - 3x_2 &= 4 \end{aligned}$$

7. Sean $A = \begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ y $\mathbf{b}_4 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

   a) Determine A^{-1} y utilícela para resolver las ecuaciones

   $$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \quad A\mathbf{x} = \mathbf{b}_4$$

   b) Las cuatro ecuaciones del inciso *a*) se pueden resolver con el *mismo* conjunto de operaciones de fila, ya que la matriz de coeficientes es la misma en cada caso. Resuelva las cuatro ecuaciones del inciso *a*) mediante la reducción por filas de la matriz aumentada $[A \quad \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \quad \mathbf{b}_4]$.

8. Suponga que P es invertible y $A = PBP^{-1}$. Determine B en términos de A.

En los ejercicios 9 y 10, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

9. *a*) Para que una matriz B sea la inversa de A, las ecuaciones $AB = I$ y $BA = I$ deben ser verdaderas.

   b) Si A y B son de $n \times n$ e invertibles, entonces $A^{-1}B^{-1}$ es la inversa de AB.

   c) Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y $ab - cd \neq 0$, entonces A es invertible.

   d) Si A es una matriz invertible de $n \times n$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para *toda* $\mathbf{b}$ en $\mathbb{R}^n$.

   e) Toda matriz elemental es invertible.

10. *a*) Si A es invertible, entonces las operaciones elementales de fila que reducen A a la identidad I_n también reducen A^{-1} a I_n.

    b) Si A es invertible, entonces la inversa A^{-1} es A misma.

    c) Un producto de matrices invertibles de $n \times n$ es invertible, y la inversa del producto es el producto de sus inversas en el mismo orden.

    d) Si A es una matriz de $n \times n$ y $A\mathbf{x} = \mathbf{e}_j$ es consistente para toda $j \in \{1, 2, \ldots, n\}$, entonces A es invertible. Nota: $\mathbf{e}_1, \ldots, \mathbf{e}_n$ representa las columnas de la matriz identidad.

    e) Si A puede reducirse por filas a la matriz identidad, entonces A debe ser invertible.

11. Sea A una matriz invertible de $n \times n$, y sea B una matriz de $n \times p$. Demuestre que la ecuación $AX = B$ tiene una solución única $A^{-1}B$.

12. Utilice álgebra de matrices para demostrar que si A es invertible y D satisface $AD = I$, entonces $D = A^{-1}$.

13. Suponga que $AB = AC$, donde B y C son matrices de $n \times p$ y A es invertible. Demuestre que $B = C$. ¿Esto es cierto, en general, cuando A no es invertible?

14. Suponga que $(B - C)D = 0$, donde B y C son matrices de $m \times n$ y D es invertible. Demuestre que $B = C$.

15. Sea A una matriz invertible de $n \times n$, y sea B una matriz de $n \times p$. Explique por qué $A^{-1}B$ se puede calcular por reducción de filas:

Si $[A \quad B] \sim \cdots \sim [I \quad X]$, entonces $X = A^{-1}B$.

Si A es mayor de 2×2, entonces la reducción por filas de $[A \quad B]$ es mucho más rápida que calcular a A^{-1} y a $A^{-1}B$.

16. Suponga que A y B son matrices de $n \times n$, B es invertible y AB es invertible. Demuestre que A es invertible. [*Sugerencia:* Considere que $C = AB$, y despeje A en esta ecuación].

17. Suponga que A, B y C son matrices invertibles de $n \times n$. Demuestre que ABC también es invertible al obtener una matriz D tal que $(ABC)D = I$ y $D(ABC) = I$.

18. Resuelva la ecuación $AB = BC$ para A, suponiendo que A, B y C son matrices cuadradas y B es invertible.

19. Si A, B y C son matrices invertibles $n \times n$, ¿la ecuación $C^{-1}(A + X)B^{-1} = I_n$ tiene una solución, X? Si es así, encuéntrela.

20. Suponga que A, B y X son matrices de $n \times n$ con A, X y $A - AX$ invertibles, y suponga que

$$(A - AX)^{-1} = X^{-1}B \tag{3}$$

a) Explique por qué B es invertible.

b) Despeje X en la ecuación (3). Si se necesita invertir una matriz, explique por qué esta matriz es invertible.

21. Explique por qué las columnas de una matriz A de $n \times n$ son linealmente independientes cuando A es invertible.

22. Explique por qué las columnas de una matriz A de $n \times n$ generan a $\mathbb{R}^n$ cuando A es invertible. [*Sugerencia:* Repase el teorema 4 de la sección 1.4].

23. Suponga que A es de $n \times n$ y que la ecuación $A\mathbf{x} = 0$ tiene solamente la solución trivial. Explique por qué A tiene n columnas pivote y es equivalente por filas a I_n. De acuerdo con el teorema 7, esto indica que A debe ser invertible. (Este ejercicio y el 24 se mencionarán en la sección 2.3).

24. Suponga que para una matriz A de $n \times n$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$. Explique por qué A debe ser invertible. [*Sugerencia*: ¿A es equivalente por filas a I_n?]

Los ejercicios 25 y 26 demuestran el teorema 4 para $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. Demuestre que si $ad - bc = 0$, entonces la ecuación $A\mathbf{x} = \mathbf{0}$ tiene más de una solución. ¿Por qué esto implica que A no es invertible? [*Sugerencia:* Primero, considere $a = b = 0$. Después, si a y b no son ambas cero, considere el vector $\mathbf{x} = \begin{bmatrix} -b \\ a \end{bmatrix}$].

26. Demuestre que si $ad - bc \neq 0$, la fórmula para A^{-1} funciona.

Los ejercicios 27 y 28 demuestran casos especiales de los hechos acerca de matrices elementales establecidos en el recuadro que sigue al ejemplo 5. Aquí A es una matriz de 3×3 e $I = I_3$. (Una demostración general requeriría un poco más de notación).

27. Sea A una matriz de 3×3.

a) Use la ecuación (2) de la sección 2.1 para demostrar que $\text{fila}_i(A) = \text{fila}_i(I) \cdot A$, para $i = 1, 2, 3$.

b) Demuestre que si las filas 1 y 2 de A se intercambian, entonces el resultado se puede escribir como EA, donde E es una matriz elemental formada al intercambiar las filas 1 y 2 de I.

c) Demuestre que si la fila 3 de A se multiplica por 5, entonces el resultado se puede escribir como EA, donde E se forma al multiplicar la fila 3 de I por 5.

28. Suponga que se remplaza la fila 2 de A por $\text{fila}_2(A) - 3 \cdot \text{fila}_i(A)$. Demuestre que el resultado es EA, donde E se forma a partir de I al remplazar $\text{fila}_2(I)$ por $\text{fila}_2(I) - 3 \cdot \text{fila}_1(A)$.

Encuentre las inversas de las matrices en los ejercicios 29 a 32, si existen. Use el algoritmo presentado en esta sección.

29. $\begin{bmatrix} 1 & -3 \\ 4 & -9 \end{bmatrix}$

30. $\begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix}$

31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$

32. $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 3 \\ -2 & -6 & 4 \end{bmatrix}$

33. Use el algoritmo de esta sección para encontrar las inversas de

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \text{ y } \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Sea A la matriz de $n \times n$ correspondiente, y sea B su inversa. Infiera la forma de B, y después demuestre que $AB = I$.

34. Repita la estrategia del ejercicio 33 para inferir la inversa B de

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 2 & 2 & 0 & & 0 \\ 3 & 3 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ n & n & n & \cdots & n \end{bmatrix}.$$

Demuestre que $AB = I$.

35. Sea $A = \begin{bmatrix} -1 & -7 & -3 \\ 2 & 15 & 6 \\ 1 & 3 & 2 \end{bmatrix}$. Encuentre la tercera columna de A^{-1} sin calcular las otras columnas.

36. **[M]** Sea $A = \begin{bmatrix} -25 & -9 & -27 \\ 536 & 185 & 537 \\ 154 & 52 & 143 \end{bmatrix}$. Encuentre la segunda y tercera columnas de A^{-1} sin calcular la primera columna.

37. Sea $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. Construya una matriz C de 2×3 (por prueba y error) usando sólo 1, -1 y 0 como entradas, de tal forma que $CA = I_2$. Calcule AC y observe que $AC \neq I_3$.

38. Sea $A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}$. Construya una matriz D, de 4×2, usando solo 1 y 0 como entradas, de tal forma que $AD = I_2$. ¿Es posible que $CA = I_4$ para alguna matriz C de 4×2? ¿Por qué?

39. **[M]** Sea

$$D = \begin{bmatrix} .011 & .003 & .001 \\ .003 & .009 & .003 \\ .001 & .003 & .011 \end{bmatrix}$$

una matriz de flexibilidad, con la flexibilidad medida en pulgadas por libra. Suponga que se aplican fuerzas de 40, 50 y 30 lb sobre los puntos 1, 2 y 3, respectivamente, en la figura 1 del ejemplo 3. Encuentre las deflexiones correspondientes.

40. **[M]** Encuentre la matriz de rigidez D^{-1} para la D del ejercicio 39. Liste las fuerzas que se necesitan para producir una deflexión de .04 pulgadas en el punto 3, con deflexión cero en los otros puntos.

41. **[M]** Sea

$$D = \begin{bmatrix} .0130 & .0050 & .0020 & .0010 \\ .0050 & .0100 & .0040 & .0020 \\ .0020 & .0040 & .0100 & .0050 \\ .0010 & .0020 & .0050 & .0130 \end{bmatrix}$$

la matriz de flexibilidad para una viga elástica, como la del ejemplo 3, con cuatro puntos en los que se aplican fuerzas. Las unidades son centímetros por newton de fuerza. Las mediciones en los cuatro puntos identifican deflexiones de .07, .12, .16 y .12 cm. Determine las fuerzas presentes en los cuatro puntos.

42. **[M]** Con D como en el ejercicio 41, determine las fuerzas que producen una deflexión de .22 cm en el segundo punto de la viga, con deflexión cero en los otros tres puntos. ¿Cómo están relacionadas la respuesta al problema y las entradas de D^{-1}? [*Sugerencia:* Primero conteste la pregunta para una deflexión de 1 cm en el segundo punto].

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *a*) $\det \begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix} = 3 \cdot 6 - (-9) \cdot 2 = 18 + 18 = 36$. El determinante es diferente de cero, de manera que la matriz es invertible.

b) $\det \begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix} = 4 \cdot 5 - (-9) \cdot 0 = 20 \neq 0$. La matriz es invertible.

c) $\det \begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} = 6 \cdot 6 - (-9)(-4) = 36 - 36 = 0$. La matriz no es invertible.

2.
$$[\,A \quad I\,] \sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ -1 & 5 & 6 & 0 & 1 & 0 \\ 5 & -4 & 5 & 0 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 6 & 10 & -5 & 0 & 1 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -7 & -2 & 1 \end{bmatrix}$$

Así, $[A \quad I]$ es ahora equivalente por filas a la matriz de la forma $[B\ D]$, donde B es una matriz cuadrada y tiene una fila de ceros. Las operaciones de fila adicionales no van a transformar a B en I, así que el proceso se detiene. A no tiene una inversa.

2.3 CARACTERIZACIONES DE MATRICES INVERTIBLES

Esta sección constituye un repaso de la mayoría de los conceptos estudiados en el capítulo 1, en relación con sistemas de n ecuaciones lineales de n incógnitas y con matrices *cuadradas*. El resultado principal es el teorema 8.

TEOREMA 8 **El teorema de la matriz invertible**

Sea A una matriz cuadrada de $n \times n$. Entonces, los siguientes enunciados son equivalentes. Es decir, para una A dada, los enunciados son todos ciertos o todos falsos.

a) A es una matriz invertible.

b) A es equivalente por filas a la matriz identidad de $n \times n$.

c) A tiene n posiciones pivote.

d) La ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.

e) Las columnas de A forman un conjunto linealmente independiente.

f) La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.

g) La ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$.

h) Las columnas de A generan $\mathbb{R}^n$.

i) La transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$.

j) Existe una matriz C de $n \times n$ tal que $CA = I$,

k) Existe una matriz D de $n \times n$ tal que $AD = I$.

l) A^T es una matriz invertible.

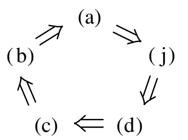
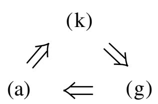


FIGURA 1

Primero, se necesita alguna notación. Si la veracidad del enunciado *a*) siempre implica que el enunciado *j*) sea cierto, se dice que *a*) *implica* a *j*), y esto se representa como *a*) ⇒ *j*). La demostración establecerá el "círculo" de implicaciones que se ilustra en la figura 1. Si cualquiera de estos cinco enunciados es cierto, entonces también lo son los demás. Por último, la demostración relacionará los enunciados restantes del teorema con los enunciados incluidos en este círculo.

DEMOSTRACIÓN Si el enunciado *a*) es cierto, entonces A^{-1} funciona para C en *j*), de manera que *a*) ⇒ *j*). Luego, *j*) ⇒ *d*) por el ejercicio 23 de la sección 2.1. (Regrese y lea el ejercicio). También, *d*) ⇒ *c*) por el ejercicio 23 de la sección 2.2. Si A es cuadrada y tiene n posiciones pivote, entonces los pivotes deben estar sobre la diagonal principal; en tal caso, la forma escalonada reducida de A es I_n. Por lo tanto, *c*) ⇒ *b*). También, *b*) ⇒ *a*) por el teorema 7 de la sección 2.2. Esto completa el círculo de la figura 1.



$$\text{(g)} \Longleftrightarrow \text{(h)} \Longleftrightarrow \text{(i)}$$

$$\text{(d)} \Longleftrightarrow \text{(e)} \Longleftrightarrow \text{(f)}$$

$$\text{(a)} \Longleftrightarrow \text{(l)}$$

Ahora, *a*) ⇒ *k*) porque A^{-1} funciona para D. También, *k*) ⇒ *g*) por el ejercicio 26 de la sección 2.1, y *g*) ⇒ *a*) por el ejercicio 24 de la sección 2.2. Así que *g*) y *k*) están vinculados al círculo. Además, *g*), *h*) e *i*) son equivalentes para cualquier matriz, de acuerdo con el teorema 4 de la sección 1.4 y el teorema 12*a*) de la sección 1.9. Por consiguiente, *h*) e *i*) están vinculados al círculo a través de *g*).

Como *d*) está vinculado al círculo, también lo están *e*) y *f*), porque *d*), *e*) y *f*) son todos equivalentes para *cualquier* matriz A. [Véase la sección 1.7 y el teorema 12*b*) de la sección 1.9]. Por último, *a*) ⇒ *l*) de acuerdo con el teorema 6*c*) de la sección 2.2, y *l*) ⇒ *a*) por el mismo teorema intercambiando A y A^T. Esto completa la demostración. ■

Según el teorema 5 de la sección 2.2, el enunciado *g*) del teorema 8 también se podría escribir como: "La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución *única* para toda $\mathbf{b}$ en $\mathbb{R}^n$". Este enunciado realmente implica a *b*) y, por lo tanto, implica que A es invertible.

El siguiente hecho es consecuencia del teorema 8 y del ejercicio 12 de la sección 2.2.

Sean A y B matrices cuadradas. Si $AB = I$, entonces A y B son invertibles, con $B = A^{-1}$ y $A = B^{-1}$.

El teorema de la matriz invertible divide al conjunto de todas las matrices de $n \times n$ en dos clases disjuntas: las matrices invertibles (no singulares) y las matrices no invertibles (singulares). Cada enunciado del teorema describe una propiedad de toda matriz de $n \times n$ invertible. La *negación* de un enunciado del teorema describe una propiedad de toda matriz singular de $n \times n$. Por ejemplo, una matriz singular de $n \times n$ *no* es equivalente por filas a I_n, *no* tiene n posiciones pivote, y tiene columnas linealmente *dependientes*. Las negaciones de los otros enunciados se consideran en los ejercicios.

EJEMPLO 1 Use el teorema de la matriz invertible para determinar si A es invertible:

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$$

SOLUCIÓN

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{bmatrix}$$

Por lo que A tiene tres posiciones pivote y, por lo tanto, es invertible, de acuerdo con el enunciado *c*) del teorema de la matriz invertible. ■

El poder del teorema de la matriz invertible radica en las relaciones que establece entre tantos conceptos importantes, tales como la independencia lineal de las columnas de una matriz A y la existencia de soluciones para ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$. Sin embargo, se debe enfatizar que el teorema de la matriz invertible *se aplica solo a matrices cuadradas*. Por ejemplo, si las columnas de una matriz de 4×3 son linealmente independientes, no puede usarse el teorema de la matriz invertible para obtener cualquier conclusión acerca de la existencia o inexistencia de soluciones a ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$.

Transformaciones lineales invertibles

Recuerde de la sección 2.1 que la multiplicación de matrices corresponde a la composición de transformaciones lineales. Cuando una matriz A es invertible, la ecuación $A^{-1}A\mathbf{x} = \mathbf{x}$ se puede ver como un enunciado acerca de transformaciones lineales. Véase la figura 2.

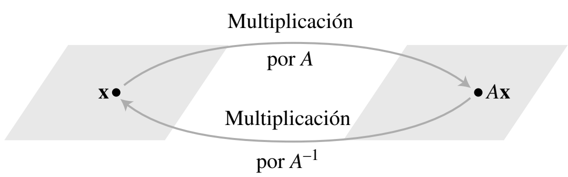


FIGURA 2 A^{-1} transforma $A\mathbf{x}$ de regreso a $\mathbf{x}$.

Se dice que una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ es **invertible** si existe una función $S : \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$S(T(\mathbf{x})) = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^n \tag{1}$$

$$T(S(\mathbf{x})) = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^n \tag{2}$$

El siguiente teorema establece que si dicha S existe, es única y debe ser una transformación lineal. Se dice que S es la **inversa** de T y se escribe como T^{-1}.

TEOREMA 9 Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal y sea A la matriz estándar para T. Así, T es invertible si y solo si A es una matriz invertible. En tal caso, la transformación lineal S dada por $S(\mathbf{x}) = A^{-1}\mathbf{x}$ es la única función que satisface las ecuaciones (1) y (2).

DEMOSTRACIÓN Suponga que T es invertible. Entonces (2) indica que T es sobre $\mathbb{R}^n$, porque si $\mathbf{b}$ está en $\mathbb{R}^n$ y $\mathbf{x} = S(\mathbf{b})$, entonces $T(\mathbf{x}) = T(S(\mathbf{b})) = \mathbf{b}$, así que toda $\mathbf{b}$ está en el rango de T. Por lo tanto, A es invertible, de acuerdo con el teorema de la matriz invertible, enunciado *i*).

Por el contrario, suponga que A es invertible y sea $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Entonces, S es una transformación lineal y evidentemente satisface (1) y (2). Por ejemplo,

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = \mathbf{x}$$

Por consiguiente, T es invertible. La demostración de que S es única se describe de manera general en el ejercicio 38. ■

EJEMPLO 2 ¿Qué se puede decir acerca de una transformación lineal T uno a uno de $\mathbb{R}^n$ en $\mathbb{R}^n$?

SOLUCIÓN Las columnas de la matriz estándar A de T son linealmente independientes (según el teorema 12 de la sección 1.9). Por lo que A es invertible, de acuerdo con el teorema de la matriz invertible, y T mapea $\mathbb{R}^n$ *sobre* $\mathbb{R}^n$. También, T es invertible, según el teorema 9. ■

NOTAS NUMÉRICAS

En la práctica, se puede encontrar ocasionalmente una matriz "casi singular" o **mal condicionada**: una matriz invertible que puede convertirse en singular si algunas de sus entradas se modifican ligeramente. En este caso, es posible que la reducción por filas produzca menos de n posiciones pivote, debido al error de redondeo. Además, los errores de redondeo, algunas veces, hacen que una matriz singular parezca invertible.

WEB

Algunos programas de matrices calculan un **número de condición** para una matriz cuadrada. Cuanto mayor sea el número de condición, más cerca estará la matriz de ser singular. El número de condición de la matriz identidad es 1. Una matriz singular tiene un número de condición infinito. En casos extremos, un programa de matrices podría no distinguir entre una matriz singular y una matriz mal condicionada.

Los ejercicios 41 a 45 ponen de manifiesto que los cálculos de matrices llegan a producir errores sustanciales cuando un número de condición es grande.

PROBLEMAS DE PRÁCTICA

1. Determine si $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ es invertible.
2. Suponga que para cierta matriz A de $n \times n$, el enunciado *g*) del teorema de la matriz invertible *no* es verdadero. ¿Qué puede decirse acerca de las ecuaciones de la forma $A\mathbf{x} = \mathbf{b}$?
3. Suponga que A y B son matrices de $n \times n$ y que la ecuación $AB\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. ¿Qué puede decirse acerca de la matriz AB?

2.3 EJERCICIOS

A menos que se especifique lo contrario, suponga que en estos ejercicios todas las matrices son de $n \times n$. En los ejercicios 1 a 10, determine cuáles de las matrices son invertibles. Use tan pocos cálculos como sea posible. Justifique sus respuestas.

1. $\begin{bmatrix} 5 & 7 \\ -3 & -6 \end{bmatrix}$

2. $\begin{bmatrix} -4 & 2 \\ 6 & -3 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 0 & 0 \\ -3 & -4 & 0 \\ 8 & 5 & -3 \end{bmatrix}$

4. $\begin{bmatrix} -5 & 1 & 4 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 0 & -3 \\ 2 & 0 & 4 \\ -4 & 0 & 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 & -3 & -6 \\ 0 & 4 & 3 \\ -3 & 6 & 0 \end{bmatrix}$

7. $\begin{bmatrix} -1 & -3 & 0 & 1 \\ 3 & 5 & 8 & -3 \\ -2 & -6 & 3 & 2 \\ 0 & -1 & 2 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 4 & 7 & 4 \\ 0 & 1 & 4 & 6 \\ 0 & 0 & 2 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

9. [**M**] $\begin{bmatrix} 4 & 0 & -3 & -7 \\ -6 & 9 & 9 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 4 & -1 \end{bmatrix}$

10. [**M**] $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

En los ejercicios 11 y 12, todas las matrices son $n \times n$. Cada inciso de estos ejercicios es una *implicación* de la forma "si (enunciado 1), entonces (enunciado 2)". Marque cada implicación como verdadera o falsa, considerando lo siguiente. Una implicación es verdadera si el enunciado 2 es verdadero *siempre* que el enunciado 1 sea cierto. Una implicación es falsa si existe un caso en el que el enunciado 2 es falso, pero el enunciado 1 es verdadero. Justifique sus respuestas.

11. *a*) Si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial, entonces A es equivalente por filas a la matriz identidad de $n \times n$.

b) Si las columnas de A generan a $\mathbb{R}^n$, entonces las columnas son linealmente independientes.

c) Si A es una matriz de $n \times n$, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$.

d) Si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, entonces A tiene menos de n posiciones pivote.

e) Si A^T no es invertible, entonces A no es invertible.

12. *a*) Si existe una matriz D de $n \times n$ tal que $AD = I$, entonces $DA = I$.

b) Si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ en $\mathbb{R}^n$, entonces la forma escalonada reducida de A es I.

c) Si las columnas de A son linealmente independientes, entonces las columnas de A generan a $\mathbb{R}^n$.

d) Si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es uno a uno.

e) Si existe una $\mathbf{b}$ en $\mathbb{R}^n$ tal que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces la solución es única.

13. Una **matriz triangular superior** de $m \times n$ es aquella cuyas entradas *debajo* de la diagonal principal son ceros (como en el ejercicio 8). ¿Cuándo es invertible una matriz triangular superior cuadrada? Justifique su respuesta.

14. Una **matriz triangular inferior** de $m \times n$ es aquella cuyas entradas *arriba* de la diagonal principal son ceros (como en el ejercicio 3). ¿Cuándo es invertible una matriz triangular inferior cuadrada? Justifique su respuesta.

15. ¿Puede ser invertible una matriz de 4×4, cuando sus columnas no generan a $\mathbb{R}^4$? ¿Por qué?

16. Si una matriz A de $n \times n$ es invertible, entonces las columnas de A^T son linealmente independientes. Explique por qué.

17. ¿Puede una matriz cuadrada con dos columnas idénticas ser invertible? ¿Por qué?

18. ¿Puede una matriz cuadrada con dos filas idénticas ser invertible? ¿Por qué?

19. Si las columnas de una matriz D, de 7×7, son linealmente independientes, ¿qué se puede decir acerca de las soluciones de $D\mathbf{x} = \mathbf{b}$? ¿Por qué?

20. Si A es una matriz de 5×5 y la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para toda $\mathbf{b}$ en $\mathbb{R}^5$, ¿es posible que, para alguna $\mathbf{b}$, la ecuación $A\mathbf{x} = \mathbf{b}$ tenga más de una solución? ¿Por qué?

21. Si la ecuación $C\mathbf{u} = \mathbf{v}$ tiene más de una solución para alguna $\mathbf{v}$ en $\mathbb{R}^n$, ¿pueden las columnas de la matriz C de $n \times n$ generar a $\mathbb{R}^n$? ¿Por qué?

22. Si las matrices E y F de $n \times n$ tienen la propiedad de que $EF = I$, entonces E y F conmutan. Explique por qué.

23. Suponga que F es una matriz de $n \times n$. Si la ecuación $F\mathbf{x} = \mathbf{y}$ es inconsistente para alguna $\mathbf{y}$ en $\mathbb{R}^n$, ¿qué se puede decir acerca de la ecuación $F\mathbf{x} = \mathbf{0}$? ¿Por qué?

24. Si una matriz G de $n \times n$ no se puede reducir por filas a I_n, ¿qué se puede decir acerca de las columnas de G? ¿Por qué?

25. Compruebe el enunciado del recuadro antes del ejemplo 1.

26. Explique por qué las columnas de A^2 generan a $\mathbb{R}^n$ siempre que las columnas de una matriz A de $n \times n$ son linealmente independientes.

27. Sean A y B matrices de $n \times n$. Demuestre que si AB es invertible, también lo es A. No podrá utilizar el teorema 6*b*), porque no es posible *suponer* que A y B son invertibles. [*Sugerencia:* Existe una matriz W tal que $ABW = I$. ¿Por qué?].

28. Sean A y B matrices de $n \times n$. Demuestre que si AB es invertible, también B lo es.

29. Si A es una matriz de $n \times n$ y la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno, ¿qué más puede decirse acerca de esta transformación? Justifique su respuesta.

30. Si A es una matriz de $n \times n$ y la ecuación $A\mathbf{x} = \mathbf{b}$ tiene más de una solución para alguna $\mathbf{b}$, entonces la transformación $\mathbf{x} \mapsto A\mathbf{x}$ no es uno a uno. ¿Qué más se puede decir acerca de esta transformación? Justifique su respuesta.

31. Suponga que A es una matriz de $n \times n$ con la propiedad de que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$. Sin utilizar los teoremas 5 u 8, explique por qué cada ecuación $A\mathbf{x} = \mathbf{b}$ tiene, en efecto, exactamente una solución.

32. Suponga que A es una matriz de $n \times n$ con la propiedad de que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial. Sin utilizar el teorema de la matriz invertible, explique directamente por qué la ecuación $A\mathbf{x} = \mathbf{b}$ debe tener una solución para toda $\mathbf{b}$ en $\mathbb{R}^n$.

En los ejercicios 33 y 34, T es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$. Demuestre que T es invertible y encuentre una fórmula para T^{-1}.

33. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

34. $T(x_1, x_2) = (2x_1 - 8x_2, -2x_1 + 7x_2)$

35. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal invertible. Explique por qué T es tanto uno a uno como sobre $\mathbb{R}^n$. Use las ecuaciones (1) y (2). Después, dé una segunda explicación usando uno o más teoremas.

36. Suponga una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ con la propiedad de que $T(\mathbf{u}) = T(\mathbf{v})$ para algún par de vectores distintos $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. ¿Puede T mapear $\mathbb{R}^n$ sobre $\mathbb{R}^n$? ¿Por qué?

37. Suponga que T y U son transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^n$ tales que $T(U(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$. ¿Es cierto que $U(T(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$? ¿Por qué?

38. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal invertible, y sean S y U funciones de $\mathbb{R}^n$ en $\mathbb{R}^n$ tales que $S(T(\mathbf{x})) = \mathbf{x}$ y $U(T(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$. Demuestre que $U(\mathbf{v}) = S(\mathbf{v})$ para toda $\mathbf{v}$ en $\mathbb{R}^n$. Esto demostrará que T tiene una inversa única, como se establece en el teorema 9. [*Sugerencia:* Dada cualquier $\mathbf{v}$ en $\mathbb{R}^n$, se puede escribir $\boldsymbol{v} = T(\mathbf{x})$ para alguna $\mathbf{x}$. ¿Por qué? Calcule $S(\mathbf{v})$ y $U(\mathbf{v})$].

39. Sea T una transformación lineal que mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$. Demuestre que T^{-1} existe y mapea $\mathbb{R}^n$ en $\mathbb{R}^n$. ¿T^{-1} es también uno a uno?

40. Suponga que T y S satisfacen las ecuaciones de invertibilidad (1) y (2), donde T es una transformación lineal. Demuestre directamente que S es una transformación lineal. [*Sugerencia:* Dadas $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, sea $\mathbf{x} = S(\mathbf{u})$, $\mathbf{y} = S(\mathbf{v})$. Entonces, $T(\mathbf{x}) = \boldsymbol{u}$, $T(\mathbf{y}) = \mathbf{v}$. ¿Por qué? Aplique S a ambos miembros de la ecuación $T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y})$. También, considere $T(c\mathbf{x}) = cT(\mathbf{x})$].

41. **[M]** Suponga que un experimento conduce al siguiente sistema de ecuaciones:

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.249 \\ 1.6x_1 + 1.1x_2 &= 6.843 \end{aligned} \tag{3}$$

a) Resuelva el sistema (3), y después resuelva el sistema (4) que se presenta a continuación, en el cual los datos a la derecha se redondearon a dos decimales. En cada caso, encuentre la solución *exacta*.

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.25 \\ 1.6x_1 + 1.1x_2 &= 6.84 \end{aligned} \tag{4}$$

b) Las entradas del sistema (4) difieren de las del sistema (3) en menos del .05%. Encuentre el porcentaje de error cuando se utiliza la solución de (4) como una aproximación a la solución de (3).

c) Use un programa de matrices para producir el número de condición de la matriz de coeficientes de (3).

Los ejercicios 42, 43 y 44 ilustran cómo utilizar el número de condición de una matriz A para estimar la exactitud de una solución calculada de $A\mathbf{x} = $ b. Si las entradas de A y $\mathbf{b}$ son exactas con aproximadamente r dígitos significativos, y si el número de condición de A es aproximadamente 10^k (siendo k un entero positivo), entonces la solución calculada de $A\mathbf{x} = \mathbf{b}$ debería ser exacta hasta al menos $r - k$ dígitos significativos.

42. **[M]** Sea A la matriz del ejercicio 9. Encuentre el número de condición de A. Construya un vector aleatorio $\mathbf{x}$ en $\mathbb{R}^4$ y calcule $\mathbf{b} = A\mathbf{x}$. Después use un programa de matrices para calcular la solución $\mathbf{x}_1$ de $A\mathbf{x} = \mathbf{b}$. ¿En cuántos dígitos concuerdan $\mathbf{x}$ y $\mathbf{x}_1$? Encuentre el número de dígitos que el programa de matrices almacena con precisión, e informe cuántos dígitos de exactitud se pierden cuando se usa $\mathbf{x}_1$ en lugar de la solución exacta $\mathbf{x}$.

43. **[M]** Repita el ejercicio 42 para la matriz del ejercicio 10.

44. **[M]** Resuelva la ecuación $A\mathbf{x} = \mathbf{b}$ para obtener una $\mathbf{b}$ que sirva para encontrar la última columna de la inversa de la *matriz de Hilbert de quinto orden*

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

¿Cuántos dígitos de cada entrada de $\mathbf{x}$ espera que sean correctos? Explique su respuesta. [*Nota:* La solución exacta es (630, -12600, 56700, -88200, 44100)].

45. **[M]** Algunos programas de matrices, como MATLAB, tienen una orden para crear matrices de Hilbert de varios tamaños. Si es posible, use un comando inverso para calcular la inversa de una matriz de Hilbert A de duodécimo orden o mayor. Calcule AA^{-1}. Realice un informe de sus hallazgos.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Es evidente que las columnas de A son linealmente dependientes, ya que las columnas 2 y 3 son múltiplos de la columna 1. Por lo tanto, A *no* puede ser invertible, de acuerdo con el teorema de la matriz invertible.
2. Si el enunciado *g*) *no* es verdadero, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ es inconsistente para al menos una $\mathbf{b}$ en $\mathbb{R}^n$.
3. Aplique el teorema de la matriz invertible a la matriz AB en lugar de A. Entonces, el enunciado *d*) se convierte en: "$AB\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial". Esto no es cierto. Por lo tanto, AB no es invertible.

2.4 MATRICES PARTICIONADAS

Una característica clave de nuestro trabajo con matrices ha sido la capacidad para considerar a una matriz A como una lista de vectores columna y no tan solo un arreglo rectangular de números. Este punto de vista ha resultado tan útil que sería deseable considerar otras **particiones** de A, indicadas por las líneas divisorias horizontales y verticales, como en el ejemplo 1 que se presenta a continuación. Las matrices particionadas se presentan en la mayoría de las aplicaciones modernas del álgebra lineal porque la notación resalta la estructura esencial de los cálculos matriciales, como se mostró en el ejemplo introductorio de este capítulo acerca del diseño de aeronaves. Esta sección ofrece una oportunidad para revisar el álgebra matricial y usar el teorema de la matriz invertible.

EJEMPLO 1 La matriz

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

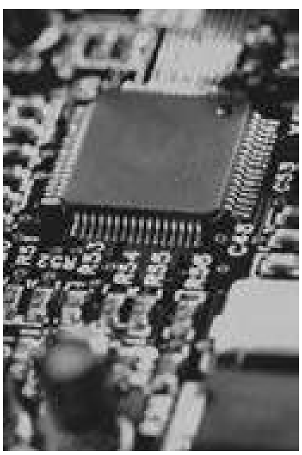
también se puede escribir como la **matriz particionada** de 2×3 (o **por bloques**) de

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

cuyas entradas son los *bloques* (o las *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$ ■



EJEMPLO 2 Cuando una matriz A se presenta en un modelo matemático de un sistema físico, como en una red eléctrica, un sistema de transporte o una gran compañía, tal vez resulte natural considerar A como una matriz particionada. Por ejemplo, si un tablero de circuitos de microcomputadora consta principalmente de tres microcircuitos VLSI (*very large-scale integrated*, es decir, integrados a escala muy grande), entonces la matriz para el tablero de circuitos podría tener la forma general

$$A = \left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right]$$

Las submatrices en la "diagonal" de A —a saber, A_{11}, A_{22} y A_{33}— se refieren a los tres circuitos VLSI, mientras que las otras submatrices dependen de las interconexiones que haya entre esos microcircuitos. ■

Suma y multiplicación escalar

Si las matrices A y B son del mismo tamaño y están particionadas exactamente en la misma forma, resulta natural efectuar una partición similar de la suma ordinaria matricial $A + B$. En este caso, cada bloque de $A + B$ es la suma (matricial) de los bloques correspondientes de A y B. La multiplicación por un escalar de una matriz particionada también se calcula bloque por bloque.

Multiplicación de matrices particionadas

Las matrices particionadas se pueden multiplicar utilizando la regla fila-columna como si las entradas del bloque fueran escalares, siempre que para un producto AB, la partición por columnas de A equivalga a la partición por filas de B.

EJEMPLO 3 Sean

$$A = \left[\begin{array}{rrr|rr} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{rr} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

Las cinco columnas de A están particionadas en un conjunto de tres columnas y, luego, en uno de dos columnas. Las cinco filas de B están particionadas de igual manera (en un conjunto de tres filas y después en uno de dos filas). Se dice que las particiones de A y B están **conformadas** para la **multiplicación por bloques**. Es posible demostrar que el producto común AB se escribe como

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{rr} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

Es importante escribir cada producto menor de la expresión para AB con la submatriz de A a la izquierda, ya que la multiplicación de matrices no es conmutativa. Por ejemplo,

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

Por lo tanto, el bloque superior es

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix}$$ ■

La regla fila-columna para la multiplicación de matrices por bloques ofrece la manera más general de considerar un producto de dos matrices. Cada una de las siguientes formas de ver un producto ya se describió mediante particiones sencillas de matrices: **1.** la definición de $A\mathbf{x}$ usando las columnas de A, **2.** la definición de columna de AB, **3.** la regla fila-columna para calcular AB, y **4.** las filas de AB como productos de las filas de A y la matriz B. Una quinta manera de ver AB, usando de nuevo particiones, se presentará más adelante en el teorema 10.

Los cálculos del siguiente ejemplo preparan el camino para el teorema 10. Aquí, $\text{col}_k(A)$ es la k-ésima columna de A, y la $\text{fila}_k\,(B)$ es la k-ésima fila de B.

EJEMPLO 4 Sea $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Compruebe que

$$AB = \text{col}_1(A)\ \text{fila}_1(B) + \text{col}_2(A)\ \text{fila}_2(B) + \text{col}_3(A)\ \text{fila}_3(B)$$

SOLUCIÓN Cada uno de los términos anteriores es un *producto externo*. (Véase los ejercicios 27 y 28 de la sección 2.1). Por la regla fila-columna para calcular un producto matricial,

$$\text{col}_1(A)\ \text{fila}_1(B) = \begin{bmatrix} -3 \\ 1 \end{bmatrix}\begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix}$$

$$\text{col}_2(A)\ \text{fila}_2(B) = \begin{bmatrix} 1 \\ -4 \end{bmatrix}\begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix}$$

$$\text{col}_3(A)\ \text{fila}_3(B) = \begin{bmatrix} 2 \\ 5 \end{bmatrix}\begin{bmatrix} e & f \end{bmatrix} = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix}$$

De modo que

$$\sum_{k=1}^{3} \text{col}_k(A)\ \text{fila}_k(B) = \begin{bmatrix} -3a + c + 2e & -3b + d + 2f \\ a - 4c + 5e & b - 4d + 5f \end{bmatrix}$$

Es evidente que esta matriz es AB. Observe que la entrada (1, 1) de AB es la suma de las entradas (1, 1) de los tres productos externos, la entrada (1, 2) en AB es la suma de las entradas (1, 2) de los tres productos externos, y así sucesivamente. ■

TEOREMA 10

Expansión columna-fila de AB

Si A es de $m \times n$ y B es de $n \times p$, entonces

$$\begin{aligned} AB &= \begin{bmatrix} \text{col}_1(A) & \text{col}_2(A) & \cdots & \text{col}_n(A) \end{bmatrix} \begin{bmatrix} \text{fila}_1(B) \\ \text{fila}_2(B) \\ \vdots \\ \text{fila}_n(B) \end{bmatrix} \qquad (1) \\ &= \text{col}_1(A)\ \text{fila}_1(B) + \cdots + \text{col}_n(A)\ \text{fila}_n(B) \end{aligned}$$

DEMOSTRACIÓN Para cada índice de fila i e índice columna j, la entrada (i, j) en $\text{col}_k(A)$ $\text{fila}_k(B)$ es el producto de a_{ik} de $\text{col}_k(A)$ y b_{kj} de $\text{fila}_k(B)$. Por lo tanto, la entrada (i, j) de la suma que se muestra en la ecuación (1) es

$$\underset{(k=1)}{a_{i1}b_{1j}} + \underset{(k=2)}{a_{i2}b_{2j}} + \cdots + \underset{(k=n)}{a_{in}b_{nj}}$$

Esta suma también es la entrada (i, j) de AB, por la regla fila-columna. ■

Inversas de matrices particionadas

El siguiente ejemplo ilustra los cálculos relacionados con inversas y matrices particionadas.

EJEMPLO 5 Se dice que una matriz de la forma

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

es *triangular superior por bloques*. Suponga que A_{11} es de $p \times p$, A_{22} es de $q \times q$, y A invertible. Encuentre una fórmula para A^{-1}.

SOLUCIÓN Denote A^{-1} con B, y efectúe una partición de B para que

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \tag{2}$$

Esta ecuación matricial proporciona cuatro ecuaciones que conducen a los bloques desconocidos $B_{11},\ldots, B_{22}$. Calcule el producto a la izquierda de (2), e iguale cada entrada con el bloque correspondiente en la matriz identidad a la derecha. Es decir, establezca

$$A_{11}B_{11} + A_{12}B_{21} = I_p \tag{3}$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \tag{4}$$

$$A_{22}B_{21} = 0 \tag{5}$$

$$A_{22}B_{22} = I_q \tag{6}$$

Por sí misma, (6) no establece que A_{22} sea invertible. Sin embargo, ya que A_{22} es cuadrada, el teorema de la matriz invertible y (6), juntos, indican que A_{22} es invertible y $B_{22} = A_{22}^{-1}$. Ahora, al multiplicar por la izquierda ambos lados de (5) por A_{22}^{-1}, se obtiene

$$B_{21} = A_{22}^{-1}0 = 0$$

por lo que (3) se simplifica a

$$A_{11}B_{11} + 0 = I_p$$

Ya que A_{11} es cuadrada, esto demuestra que A_{11} es invertible y $B_{11} = A_{11}^{-1}$. Por último, usando estos resultados con la ecuación (4) se encuentra que

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{y} \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Así,

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}$$ ■

Una matriz **diagonal por bloques** es una matriz particionada con bloques cero fuera de la diagonal (de bloques) principal. Una matriz de este tipo es invertible si y solo si cada bloque sobre la diagonal es invertible. Véase los ejercicios 13 y 14.

NOTAS NUMÉRICAS

1. Cuando las matrices son demasiado grandes para caber en la memoria de alta velocidad de una computadora, particionarlas permite trabajar solamente con dos o tres submatrices a la vez. Por ejemplo, un equipo de investigación de programación lineal simplificó un problema al particionar la matriz en 837 filas y 51 columnas. La solución del problema tardó aproximadamente cuatro minutos en una supercomputadora Cray.[1]
2. Algunas computadoras de alta velocidad, en particular aquellas con arquitectura de conducción vectorial, realizan cálculos matriciales con mayor eficiencia cuando los algoritmos usan matrices particionadas.[2]
3. Los programas de computadora profesionales para álgebra lineal numérica de alto desempeño, como LAPACK, utilizan de manera intensiva cálculos de matrices particionadas.

[1] El tiempo de solución no parece muy impresionante hasta saber que cada bloque de las 51 columnas contenía, aproximadamente, 250,000 columnas individuales. ¡El problema original tenía 837 ecuaciones y más de 12,750,000 variables! Casi 100 millones de las más de 10 mil millones de entradas eran diferentes de cero. Véase Robert E. Bixby *et al*., "Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods", *Operations Research*, 40, núm. 5 (1992): 885-897.

[2] La importancia de los algoritmos de matrices por bloques para cálculos de computadora se describe en *Matrix Computations*, 3a. ed., de Gene H. Golub y Charles F. van Loan (Baltimore: Johns Hopkins University Press, 1966).

Los siguientes ejercicios permiten obtener práctica con el álgebra matricial e ilustran los cálculos comunes que se encuentran en aplicaciones.

PROBLEMAS DE PRÁCTICA

1. Demuestre que $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ es invertible y encuentre su inversa.

2. Calcule X^TX, donde X está particionada como $[X_1 \quad X_2]$.

2.4 EJERCICIOS

En los ejercicios 1 a 9, suponga que las matrices están particionadas de manera conformada para la multiplicación por bloques. Calcule los productos que se indican en los ejercicios 1 a 4.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$

3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

4. $\begin{bmatrix} I & 0 \\ -E & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$

En los ejercicios 5 a 8, encuentre fórmulas para X, Y y Z, en términos de A, B y C, y justifique sus cálculos. En algunos casos, tendrá que hacer suposiciones acerca del tamaño de una matriz para obtener una fórmula. [*Sugerencia:* Calcule el producto a la izquierda e iguálelo al lado derecho].

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. Suponga que B_{11} es una matriz invertible. Encuentre las matrices A_{21} y A_{31} (en términos de los bloques de B) de tal manera que el producto que aparece a continuación tenga la forma indicada. Además, calcule C_{22} (en términos de los bloques de B). [*Sugerencia:* Calcule el producto a la izquierda e iguálelo al lado derecho].

$$\begin{bmatrix} I & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \\ 0 & C_{32} \end{bmatrix}$$

10. La inversa de

$$\begin{bmatrix} I & 0 & 0 \\ A & I & 0 \\ B & D & I \end{bmatrix} \text{ es } \begin{bmatrix} I & 0 & 0 \\ P & I & 0 \\ Q & R & I \end{bmatrix}$$

Encuentre P, Q y R.

En los ejercicios 11 y 12, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) Si $A = [A_1 \ \ A_2]$ y $B = [B_1 \ \ B_2]$, con A_1 y A_2 de las mismas dimensiones que B_1 y B_2, respectivamente, entonces $A + B = [A_1 + B_1 \ \ A_2 + B_2]$.

    b) Si $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ y $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, entonces las particiones de A y B están conformadas para la multiplicación por bloques.

12. *a*) Si A_1, A_2, B_1 y B_2 son matrices de $n \times n$, $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, y $B = [B_1 \ \ B_2]$, entonces el producto BA está definido, pero AB no lo está.

    b) Si $A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$, entonces la transpuesta de A es

    $$A^T = \begin{bmatrix} P^T & Q^T \\ R^T & S^T \end{bmatrix}.$$

13. Sea $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$, donde B y C son cuadradas. Demuestre que A es invertible si y solo si B y C son invertibles.

14. Demuestre que el bloque de matriz triangular superior A en el ejemplo 5 es invertible si y solo si tanto A_{11} como A_{22} son invertibles. [*Sugerencia:* Si A_{11} y A_{22} son invertibles, la fórmula para A^{-1}, que se dio en el ejemplo 5, en realidad funciona como la inversa de A]. Este hecho de A es una parte importante de varios algoritmos de computadora que estiman valores propios de matrices. Los valores propios se analizan en el capítulo 5.

15. Cuando se lanza una sonda espacial, es necesario hacer algunas correcciones para colocar la nave en una trayectoria exacta. La radio telemetría proporciona un flujo de vectores, $\mathbf{x}_1, \ldots, \mathbf{x}_k$, que arrojan información en diferentes momentos sobre cómo se compara la posición de la sonda con la trayectoria prevista. Sea X_k la matriz $[\mathbf{x}_1 \cdots \mathbf{x}_k]$. La matriz $G_k = X_k X_k^T$ se calcula conforme se analizan los datos del radar. Cuando llega $\mathbf{x}_{k+1}$, se debe calcular un nuevo G_{k+1}. Como los vectores de datos llegan a alta velocidad, la carga computacional podría ser grande. Pero la multiplicación de la matriz particionada ayuda muchísimo. Calcule las expansiones columna-fila de G_k y G_{k+1}, y describa lo que se debe calcular para *actualizar* G_k a la forma G_{k+1}.

La sonda Galileo fue lanzada el 18 de octubre de 1989 y llegó cerca de Júpiter a principios de diciembre de 1995.

16. Sea $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Si A_{11} es invertible, entonces la matriz $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$ se llama **complemento de Schur** de A_{11}. Por otra parte, si A_{22} es invertible, la matriz $A_{11} - A_{12}A_{11}^{-1}A_{21}$ se llama complemento de Schur de A_{22}. Suponga que A_{11} es invertible. Encuentre X y Y tal que

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \tag{7}$$

17. Suponga que la matriz por bloques A del lado izquierdo de (7) y A_{11} son invertibles. Demuestre que el complemento de Schur S de A_{11} es invertible. [*Sugerencia:* Los factores externos localizados en el lado derecho de (7) siempre son invertibles. Compruebe esto]. Cuando A y A_{11} son ambas invertibles, (7) conduce a una fórmula para A^{-1}, utilizando S^{-1}, A_{11}^{-1} y las otras entradas de A.

18. Sea X una matriz de datos de $m \times n$ tal que X^TX es invertible, y sea $M = I_m - X(X^TX)^{-1}X^T$. Añada una columna $\mathbf{x}_0$ a los datos y forme

$W = [X \quad \mathbf{x}_0]$.

Calcule W^TW. La entrada (1, 1) es X^TX. Demuestre que el complemento de Schur (ejercicio 16) de X^TX se puede escribir en la forma $\mathbf{x}_0^T M\mathbf{x}_0$. Es posible demostrar que la cantidad $(\mathbf{x}_0^T M\mathbf{x}_0)^{-1}$ es la entrada (2, 2) de $(W^TW)^{-1}$. Esta entrada tiene una interpretación estadística útil, a la luz de las hipótesis adecuadas.

En el estudio de ingeniería de control de sistemas físicos, un conjunto estándar de ecuaciones diferenciales se transforma en el siguiente sistema de ecuaciones lineales por medio de transformadas de Laplace:

$$\begin{bmatrix} A - sI_n & B \\ C & I_m \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{8}$$

donde A es de $n \times n$, B es de $n \times m$, C es de $m \times n$, y s es una variable. El vector $\mathbf{u}$ en $\mathbb{R}^m$ es la "entrada" del sistema, $\mathbf{y}$ en $\mathbb{R}^m$ es la "salida" del sistema, y $\mathbf{x}$ en $\mathbb{R}^n$ es el vector de "estado". (En realidad, los vectores $\mathbf{x}$, $\mathbf{u}$ y $\mathbf{y}$ son funciones de s, pero esto no afecta los cálculos algebraicos de los ejercicios 19 y 20).

19. Suponga que $A - sI_n$ es invertible y considere la ecuación (8) como un sistema de dos ecuaciones matriciales. Resuelva la ecuación superior para $\mathbf{x}$ y sustitúyala en la ecuación inferior. El resultado es una ecuación de la forma $W(s)\mathbf{u} = \mathbf{y}$, donde $W(s)$ es una matriz que depende de s. $W(s)$ se denomina *función de transferencia* del sistema porque transforma la entrada $\mathbf{u}$ en la salida $\mathbf{y}$. Encuentre $W(s)$ y describa cómo está relacionada con el *sistema de matriz* particionada del miembro izquierdo de (8). Véase el ejercicio 16.

20. Suponga que la función de transferencia $W(s)$ del ejercicio 19 es invertible para alguna s. Es posible demostrar que la función de transferencia inversa $W(s)^{-1}$, que transforma salidas en entradas, es el complemento de Schur de $A - BC - sI_n$ para la matriz que se presenta a continuación. Encuentre este complemento de Schur. Véase el ejercicio 16.

$$\begin{bmatrix} A - BC - sI_n & B \\ -C & I_m \end{bmatrix}$$

21. *a*) Compruebe que $A^2 = I$ cuando $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$.

b) Use matrices particionadas para demostrar que $M^2 = I$ cuando

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

22. Generalice la idea del ejercicio 21 al construir una matriz M de 6×6, $M = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ C & 0 & D \end{bmatrix}$ tal que $M^2 = I$. Haga a C una matriz de 2×2 no nula. Muestre que su construcción funciona.

23. Use matrices particionadas para demostrar, por inducción, que el producto de dos matrices triangulares inferiores también es triangular inferior. [*Sugerencia:* Una matriz A_1 de $(k + 1) \times (k + 1)$ se puede escribir en la forma presentada a continuación, donde a es un escalar, $\mathbf{v}$ está en $\mathbb{R}^k$, y A es una matriz triangular inferior de $k \times k$. [Véase la *Guía de estudio* para obtener ayuda con la inducción].

$$A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}$$

24. Use matrices particionadas para demostrar por inducción que para $n = 2, 3, \ldots$, la matriz A de $n \times n$ que se presenta a continuación es invertible y que B es su inversa.

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & & 0 \\ 1 & 1 & 1 & & 0 \\ \vdots & & & \ddots & \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \\ 0 & & \ldots & -1 & 1 \end{bmatrix}$$

Para el paso de inducción, suponga que A y B son matrices de $(k+1) \times (k+1)$, y particione A y B de una manera similar a la que se presenta en el ejercicio 23.

25. Sin utilizar reducción por filas, encuentre la inversa de

$$A = \begin{bmatrix} 1 & 2 & 0 & 0 & 0 \\ 3 & 5 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 \\ 0 & 0 & 0 & 5 & 6 \end{bmatrix}$$

26. **[M]** Para las operaciones de bloque, podría ser necesario introducir o recurrir a submatrices de una matriz grande. Describa las funciones o los comandos de un programa de matrices que realice las siguientes tareas. Suponga que A es una matriz de 20×30.

a) Presente la submatriz de A de las filas 5 a 10 y de las columnas 15 a 20.

b) Inserte una matriz B de 5×10 en una matriz A, comenzando en la fila 5 y la columna 10.

c) Construya una matriz de 50×50 de la forma $C = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$. [*Nota:* Tal vez no sea necesario especificar los bloques de ceros en C].

27. **[M]** Suponga que debido a restricciones de memoria o al tamaño su programa de matrices no puede trabajar con matrices de más de 32 filas y 32 columnas, y suponga que algún proyecto requiere las matrices A y B de 50×50. Describa los comandos o las operaciones de su programa para matrices que realizan las siguientes tareas.

a) Cálculo de $A + B$.

b) Cálculo de AB.

c) Resolución de $A\mathbf{x} = \mathbf{b}$ para algún vector $\mathbf{b}$ en $\mathbb{R}^{50}$, suponiendo que A se pueda particionar en una matriz por bloques de 2×2 $[A_{ij}]$, con A_{11} una matriz invertible de 20×20, A_{22} una matriz invertible de 30×30, y A_{12} una matriz cero. [*Sugerencia:* Describa sistemas adecuados más pequeños que puedan resolverse sin usar matrices inversas].

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Si $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ es invertible, su inversa tiene la forma $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. Compruebe que

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW + Y & AX + Z \end{bmatrix}$$

Así, W, X, Y, Z deben satisfacer $W = I$, $X = 0$, $AW + Y = 0$, y $AX + Z = I$. Como consecuencia, $Y = -A$ y $Z = I$. Por lo tanto,

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -A & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

El producto en el orden inverso también es la identidad, de modo que la matriz de bloque es invertible, y su inversa es $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (También podría recurrir al teorema de la matriz invertible).

2. $X^T X = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^T X_1 & X_1^T X_2 \\ X_2^T X_1 & X_2^T X_2 \end{bmatrix}$. Las particiones de X^T y X se conforman de manera automática para la multiplicación por bloques, ya que las columnas de X^T son las filas de X. Esta partición de X^TX se usa en varios algoritmos de computadora para cálculos de matrices.

2.5 FACTORIZACIONES DE MATRICES

Una *factorización* de una matriz A es una ecuación que expresa a A como un producto de dos o más matrices. Mientras que la multiplicación de matrices implica una *síntesis* de datos (combinando el efecto de dos o más transformaciones lineales en una sola matriz), la factorización de matrices es un *análisis* de datos. En el lenguaje de la ciencia computacional, la expresión de A como un producto equivale a un *procesamiento previo* de los datos de A, organizando esos datos en dos o más partes cuyas estructuras son más útiles de algún modo, quizá por ser más accesibles para realizar cálculos.

Las factorizaciones de matrices y, después, las factorizaciones de transformaciones lineales se presentarán en un gran número de secciones clave a lo largo de este libro. Esta sección se enfoca en una factorización que constituye el centro neurálgico de varios importantes programas de cómputo usados ampliamente en aplicaciones, como en el problema de la aeronave descrito en la introducción del capítulo. Algunas otras factorizaciones que se estudiarán después, se presentan en los ejercicios.

La factorización LU

La factorización LU, descrita a continuación, está motivada por el muy frecuente problema industrial y de negocios que consiste en resolver una sucesión de ecuaciones, todas con la misma matriz de coeficientes:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \ldots, \quad A\mathbf{x} = \mathbf{b}_p \tag{1}$$

Véase el ejercicio 32, por ejemplo. También vea la sección 5.8, donde se usa el método de la potencia inversa para estimar los valores propios de una matriz resolviendo una secuencia de ecuaciones, como en (1), una a la vez.

Cuando A es invertible, se podría calcular A^{-1} y después calcular $A^{-1}\mathbf{b}_1$, $A^{-1}\mathbf{b}_2$, y así sucesivamente. Sin embargo, resulta más eficiente resolver la primera ecuación en la secuencia (1) con reducción por filas y obtener una factorización LU de A al mismo tiempo. Después, las ecuaciones restantes de (1) se resuelven con la factorización LU.

Primero, suponga que A es una matriz de $m \times n$ que se puede reducir por filas a su forma escalonada *sin intercambios de fila*. (Más adelante, se tratará el caso general). Entonces, A se puede escribir en la forma $A = LU$, donde L es una matriz triangular inferior de $m \times m$ con números 1 en la diagonal, y U es una forma escalonada de $m \times n$ de A. Por ejemplo, véase la figura 1. Una factorización de este tipo se llama **factorización LU** de A. La matriz L es invertible y se llama matriz triangular inferior *unitaria*.

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U}$$

FIGURA 1 Una factorización LU.

Antes de estudiar la forma de construir L y U, es necesario examinar la razón de su utilidad. Cuando $A = LU$, la ecuación $A\mathbf{x} = \mathbf{b}$ se escribe como $L(U\mathbf{x}) = \mathbf{b}$. Escribiendo $\mathbf{y}$ en lugar de $U\mathbf{x}$, se puede encontrar $\mathbf{x}$ resolviendo el *par* de ecuaciones

$$\boxed{\begin{aligned} L\mathbf{y} &= \mathbf{b} \\ U\mathbf{x} &= \mathbf{y} \end{aligned}} \tag{2}$$

Primero se despeja $\mathbf{y}$ de $L\mathbf{y} = \mathbf{b}$, y luego se resuelve $U\mathbf{x} = \mathbf{y}$ para obtener $\mathbf{x}$. Véase la figura 2. Las dos ecuaciones resultan fáciles de resolver porque L y U son triangulares.

EJEMPLO 1 Es posible comprobar que

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} = LU$$

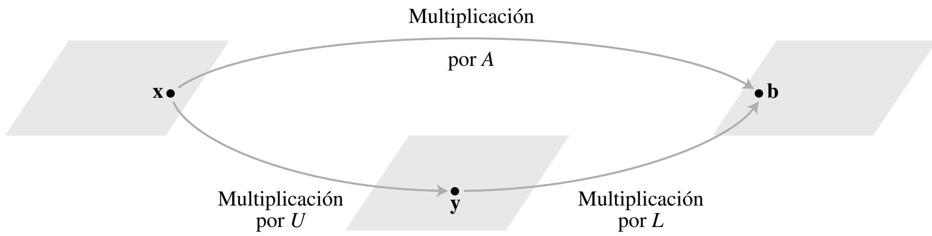


FIGURA 2 Factorización del mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

Use esta factorización LU de A para resolver $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

SOLUCIÓN La solución de $L\mathbf{y} = \mathbf{b}$ requiere únicamente de 6 multiplicaciones y 6 sumas, porque la aritmética ocurre solo en la columna 5. (En L, los ceros debajo de cada pivote se crean automáticamente con la elección de las operaciones de fila).

$$\begin{bmatrix} L & \mathbf{b} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{y} \end{bmatrix}$$

Entonces, para $U\mathbf{x} = \mathbf{y}$, la fase "regresiva" de la reducción por filas requiere de 4 divisiones, 6 multiplicaciones y 6 sumas. (Por ejemplo, para producir ceros en la columna 4 de $[U \quad \mathbf{y}]$ se requieren una división en la fila 4 y tres pares de multiplicación-suma para sumar múltiplos de la fila 4 a las filas de arriba).

$$\begin{bmatrix} U & \mathbf{y} \end{bmatrix} = \begin{bmatrix} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$$

Para encontrar $\mathbf{x}$ se requieren 28 operaciones aritméticas o "flops" (operaciones de punto flotante), excluyendo el costo de encontrar L y U. En contraste, la reducción por filas de $[A \quad \mathbf{b}]$ a $[I \quad \mathbf{x}]$ requiere de 62 operaciones. ■

La eficiencia computacional de la factorización LU depende de que se conozcan L y U. El siguiente algoritmo muestra que la reducción por filas de A a su forma escalonada U equivale a una factorización LU, porque produce L prácticamente sin trabajo extra. Después de la primera reducción por filas, L y U se obtienen al resolver ecuaciones adicionales cuya matriz de coeficientes es A.

Un algoritmo de factorización LU

Suponga que A se puede reducir a una forma escalonada U utilizando solo remplazos de filas que suman un múltiplo de una fila a otra situada *debajo de esta*. En este caso, existen matrices elementales triangulares inferiores unitarias $E_1, \ldots, E_p$ tales que

$$E_p \cdots E_1 A = U \tag{3}$$

Luego,

$$A = (E_p \cdots E_1)^{-1} U = LU$$

donde

$$L = (E_p \cdots E_1)^{-1} \tag{4}$$

Es posible demostrar que los productos y las inversas de las matrices triangulares inferiores unitarias también son triangulares inferiores unitarias. (Por ejemplo, véase el ejercicio 19). Así, L es triangular inferior unitaria.

Observe que las operaciones de fila en la ecuación (3), que reducen A *a* U, también reducen la L en la ecuación (4) a I, debido a que $E_p \cdots E_1 L = (E_p \cdots E_1)(E_p \cdots E_1)^{-1} = I$. Esta observación es la clave para *construir* L.

ALGORITMO PARA UNA FACTORIZACIÓN LU

1. Si es posible, reduzca A a una forma escalonada U con una sucesión de operaciones de remplazo de filas.
2. Coloque las entradas de L de tal manera que la *misma secuencia de operaciones de fila* reduzca L a I.

El paso 1 no siempre es posible, pero cuando lo es, el argumento anterior indica que existe una factorización LU. En el ejemplo 2 se mostrará cómo implementar el paso 2. Por construcción, L satisfará

$$(E_p \cdots E_1)L = I$$

donde se usan las mismas $E_1, \ldots, E_p$ que en la ecuación (3). Así, L será invertible, de acuerdo con el teorema de la matriz invertible, con $(E_p \cdots E_1) = L^{-1}$. A partir de (3), $L^{-1}A = U$, y $A = LU$. Por lo tanto, el paso 2 producirá una L aceptable.

EJEMPLO 2 Encuentre una factorización LU de

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$$

SOLUCIÓN Como A tiene cuatro filas, L debe ser de 4×4. La primera columna de L es la primera columna de A dividida entre la entrada pivote superior:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix}$$

Compare las primeras columnas de A y L. *Las operaciones de fila que crearon ceros en la primera columna de A también crearán ceros en la primera columna de L.* Para lograr que esta misma correspondencia de operaciones de fila sea válida para el resto de L, se examina una reducción por filas de A a una forma escalonada U. Es decir, *se resaltan las entradas* en cada una de las matrices que se utilizan para determinar la secuencia de las operaciones de fila que transforman A en U. [Véase las entradas resaltadas en la ecuación (5)].

$$A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix} = A_1 \tag{5}$$

$$\sim A_2 = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = U$$

Las entradas resaltadas de la ecuación (5) determinan la reducción por filas de A a U. En cada columna pivote, divida las entradas resaltadas entre el pivote y coloque el resultado en L:

$$\begin{bmatrix} 2 \\ -4 \\ 2 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -9 \\ 12 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \end{bmatrix}$$

$$\div 2 \quad \div 3 \quad \div 2 \quad \div 5$$

$$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow$$

$$\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix} \quad \text{y} \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix}$$

Un cálculo fácil comprueba que estas L y U satisfacen que $LU = A$. ■

En el trabajo práctico, casi siempre son necesarios los intercambios de fila, porque se usa el pivoteo parcial para lograr una precisión alta. (Recuerde que este procedimiento selecciona, entre las posibles opciones de pivote, una entrada en la columna que tenga el mayor valor absoluto). Para manejar los intercambios de fila, la factorización LU anterior se puede modificar con facilidad para producir una L que sea *triangular inferior permutada*, en el sentido de que un reordenamiento (llamado permutación) de las filas de L puede hacer que L sea triangular inferior (unitaria). La factorización *LU permutada* resultante resuelve $A\mathbf{x} = \mathbf{b}$ en la misma forma que antes, excepto que la reducción de $[L \quad \mathbf{b}]$ a $[I \quad \mathbf{y}]$ es consecuencia del orden de los pivotes de L de izquierda a derecha, empezando con el pivote de la primera columna. Una referencia a una "factorización LU" normalmente incluye la posibilidad de que L pueda ser triangular inferior permutada. Para mayores detalles, véase la *Guía de estudio*.

NOTAS NUMÉRICAS

Los siguientes conteos de operaciones corresponden a una matriz densa A de $n \times n$ (con la mayoría de sus entradas distintas de cero), donde n es moderadamente grande, por ejemplo, $n \geq 30$.[1]

1. El cálculo de una factorización LU de A requiere $2n^3/3$ flops (aproximadamente lo mismo que reducir por filas $[A \quad \mathbf{b}]$), mientras que encontrar A^{-1} requiere alrededor de $2n^3$ flops.
2. Resolver $L\mathbf{y} = \mathbf{b}$ y $U\mathbf{x} = \mathbf{y}$ requiere alrededor de $2n^2$ flops, ya que cualquier sistema triangular $n \times n$ se puede resolver en aproximadamente n^2 flops.
3. La multiplicación de $\mathbf{b}$ por A^{-1} también requiere cerca de $2n^2$ flops, pero el resultado quizá no sea tan preciso como el obtenido a partir de L y U (debido al error de redondeo cuando se calculan tanto a A^{-1} como a $A^{-1}\mathbf{b}$).
4. Si A es dispersa (la mayoría de sus entradas son cero), entonces L y U podrían ser dispersas también, pero es probable que A^{-1} sea densa. En este caso, una solución de $A\mathbf{x} = \mathbf{b}$ con una factorización LU es *mucho* más rápida que usar A^{-1}. Véase el ejercicio 31.

WEB

Factorización de matrices en ingeniería eléctrica

La factorización de matrices está íntimamente relacionada con el problema de construir una red eléctrica de propiedades específicas. El análisis que se presenta a continuación permite vislumbrar la relación entre factorización y diseño de circuitos.

[1] Véase la sección 3.8 de *Applied Linear Algebra*, 3a. ed., de Ben Noble y James W. Daniel (Englewood Cliffs, NJ: Prentice-Hall, 1988). Recuerde que para nuestros propósitos, un *flop* es $+$, $-$, $\times$ o $\div$.

Suponga que el cuadro de la figura 3 representa algún tipo de circuito eléctrico, con una entrada y una salida. El voltaje y la corriente de entrada se registran mediante $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$ (con el voltaje v en volts y la corriente i en amperes), y el voltaje y la corriente de salida se registran como $\begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$. Con frecuencia, la transformación $\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} \mapsto \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}$ es lineal. Es decir, existe una matriz A, que se llama *matriz de transferencia*, tal que

$$\begin{bmatrix} v_2 \\ i_2 \end{bmatrix} = A \begin{bmatrix} v_1 \\ i_1 \end{bmatrix}$$

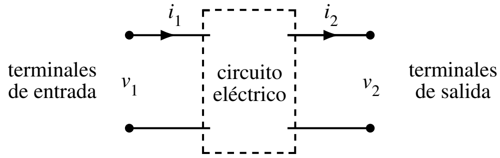


FIGURA 3 Un circuito con terminales de entrada y salida.

En la figura 4 se muestra una *red en escalera*, donde dos circuitos (podría haber más) están conectados en serie, de modo que la salida de un circuito sea la entrada del siguiente circuito. El circuito de la izquierda en la figura 4 es un *circuito en serie*, con resistencia R_1 (en ohms).

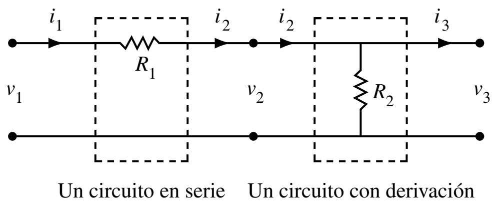


FIGURA 4 Una red en escalera.

El circuito de la derecha en la figura 4 es un *circuito con derivación*, con resistencia R_2. Con base en la ley de Ohm y las leyes de Kirchhoff, es posible demostrar que las matrices de transferencia de los circuitos en serie y con derivación, respectivamente, son,

$$\underset{\text{Matriz de transferencia del circuito en serie}}{\begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix}} \quad \text{y} \quad \underset{\text{Matriz de transferencia del circuito con derivación}}{\begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix}}$$

EJEMPLO 3

a) Calcule la matriz de transferencia para la red en escalera de la figura 4.

b) Diseñe una red en escalera cuya matriz es $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$.

SOLUCIÓN

a) Sean A_1 y A_2 las matrices de transferencia de los circuitos en serie y con derivación, respectivamente. Entonces, un vector de entrada $\mathbf{x}$ se transforma primero en $A_1\mathbf{x}$ y luego en $A_2(A_1\mathbf{x})$. La conexión en serie de los circuitos corresponde a la composición de transformaciones lineales, y la matriz de transferencia de la red en escalera es (observe el orden)

$$A_2A_1 = \begin{bmatrix} 1 & 0 \\ -1/R_2 & 1 \end{bmatrix} \begin{bmatrix} 1 & -R_1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} \tag{6}$$

b) Para factorizar la matriz $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$ en el producto de matrices de transferencia, como en la ecuación (6), se buscan las R_1 y R_2 de la figura 4 que satisfagan

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$$

De las entradas (1, 2), se tiene que $R_1 = 8$ ohms, y de las entradas (2, 1), $1/R_2 = .5$ ohm y $R_2 = 1/.5 = 2$ ohms. Con estos valores, la red de la figura 4 tiene la matriz de transferencia deseada. ∎

Una matriz de transferencia de red resume el comportamiento de entrada y salida (las especificaciones de diseño) de la red, sin referencia a los circuitos internos. Para construir físicamente una red con propiedades específicas, un ingeniero determina al principio si es posible construir (o *realizar*) dicha red. Después, trata de factorizar la matriz de transferencia para obtener matrices correspondientes a circuitos más pequeños que quizá ya fueron fabricados y estén listos para ensamblarse. En el caso común de la corriente alterna, las entradas de la matriz de transferencia normalmente son funciones con valores complejos. (Véase los ejercicios 19 y 20 de la sección 2.4 y el ejemplo 2 de la sección 3.3). Un problema estándar consiste en encontrar una *realización mínima* que use el menor número de componentes eléctricos.

PROBLEMA DE PRÁCTICA

Encuentre una factorización LU de $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$. [*Nota:* Resultará que A tiene solamente tres columnas pivote, de manera que el método del ejemplo 2 solo produce las tres primeras columnas de L. Las dos columnas restantes de L provienen de I_5].

2.5 EJERCICIOS

En los ejercicios 1 a 6, resuelva la ecuación $A\mathbf{x} = \mathbf{b}$ usando la factorización LU dada para A. En los ejercicios 1 y 2, resuelva también $A\mathbf{x} = \mathbf{b}$ por reducción ordinaria de columnas.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -5 & 1 \end{bmatrix}\begin{bmatrix} 3 & -7 & -2 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$

2. $A = \begin{bmatrix} 2 & -6 & 4 \\ -4 & 8 & 0 \\ 0 & -4 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -6 & 4 \\ 0 & -4 & 8 \\ 0 & 0 & -2 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & -4 & 2 \\ -4 & 5 & 2 \\ 6 & -9 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -1 & 1 \end{bmatrix}\begin{bmatrix} 2 & -4 & 2 \\ 0 & -3 & 6 \\ 0 & 0 & 1 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & -3 & 1 \\ 3 & 7 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ -5 \\ 7 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & -5 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ 0 & -2 & -1 \\ 0 & 0 & -6 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -2 & -2 & -3 \\ 3 & -9 & 0 & -9 \\ -1 & 2 & 4 & 7 \\ -3 & -6 & 26 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 0 \\ 3 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -3 & 4 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 & -2 & -3 \\ 0 & -3 & 6 & 0 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -2 & -3 & -4 & 12 \\ 3 & 0 & 4 & -36 \\ -5 & -3 & -8 & 49 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 3 & -3 & 1 & 0 \\ -5 & 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 2 & 0 \\ 0 & 3 & 0 & 12 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Encuentre una factorización LU de las matrices de los ejercicios 7 a 16 (con L triangular inferior unitaria). Observe que MATLAB generalmente producirá una factorización LU permutada porque utiliza pivoteo parcial para lograr exactitud numérica.

7. $\begin{bmatrix} 2 & 5 \\ -3 & -4 \end{bmatrix}$

8. $\begin{bmatrix} 6 & 4 \\ 12 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & 1 & 2 \\ -9 & 0 & -4 \\ 9 & 9 & 14 \end{bmatrix}$

10. $\begin{bmatrix} -5 & 0 & 4 \\ 10 & 2 & -5 \\ 10 & 10 & 16 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 7 & 2 \\ 6 & 19 & 4 \\ -3 & -2 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 2 & 3 & 2 \\ 4 & 13 & 9 \\ -6 & 5 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 3 & -5 & -3 \\ -1 & -5 & 8 & 4 \\ 4 & 2 & -5 & -7 \\ -2 & -4 & 7 & 5 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 3 & 1 & 5 \\ 5 & 20 & 6 & 31 \\ -2 & -1 & -1 & -4 \\ -1 & 7 & 1 & 7 \end{bmatrix}$

15. $\begin{bmatrix} 2 & 0 & 5 & 2 \\ -6 & 3 & -13 & -3 \\ 4 & 9 & 16 & 17 \end{bmatrix}$

16. $\begin{bmatrix} 2 & -3 & 4 \\ -4 & 8 & -7 \\ 6 & -5 & 14 \\ -6 & 9 & -12 \\ 8 & -6 & 19 \end{bmatrix}$

17. Cuando A es invertible, MATLAB encuentra A^{-1} al factorizar $A = LU$ (donde L puede ser triangular inferior permutada), invirtiendo L y U, y luego calculando $U^{-1}L^{-1}$. Use este método para calcular la inversa de A en el ejercicio 2. (Aplique el algoritmo de la sección 2.2 a L y a U).

18. Encuentre A^{-1} como en el ejercicio 17, usando A del ejercicio 3.

19. Sea A una matriz de $n \times n$ triangular inferior con entradas diferentes de cero en la diagonal. Demuestre que A es invertible y A^{-1} es triangular inferior. [*Sugerencia:* Explique por qué A puede convertirse en I usando solo remplazos de filas y escalamientos. (¿Dónde están los pivotes?). También, explique por qué las operaciones de fila que reducen A a I transforman a I en una matriz triangular inferior].

20. Sea $A = LU$ una factorización LU. Explique por qué A se puede reducir por filas a U utilizando solamente operaciones de remplazo. (Este hecho es el recíproco de lo que se demostró en el libro).

21. Suponga que $A = BC$, donde B es invertible. Demuestre que cualquier sucesión de operaciones de fila que reduzca B a I también reduce A a C. Lo contrario no es cierto, puesto que la matriz cero puede factorizarse como $0 = B \cdot 0$.

Los ejercicios 22 a 26 ofrecen una visualización de ciertas factorizaciones de matriz ampliamente utilizadas, algunas de las cuales se analizan posteriormente en el libro.

22. (*Factorización LU reducida*). Con A como en el problema de práctica, encuentre una matriz B de 5×3 y una matriz C de 3×4 tales que $A = BC$. Generalice esta idea para el caso donde A es $m \times n$, $A = LU$, y U tiene solamente tres filas diferentes de cero.

23. (*Factorización de rango*). Suponga que una matriz A de $m \times n$ admite una factorización $A = CD$, donde C es de $m \times 4$ y D es de $4 \times n$.

a) Demuestre que A es la suma de cuatro productos exteriores. (Véase la sección 2.4).

b) Sea $m = 400$ y $n = 100$. Explique por qué un programador de computadoras preferiría almacenar los datos de A en forma de dos matrices C y D.

24. (*Factorización QR*). Suponga que $A = QR$, donde Q y R son $n \times n$, R es invertible y triangular superior, y Q tiene la propiedad de que $Q^TQ = I$. Demuestre que para cada $\mathbf{b}$ en $\mathbb{R}^n$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución única. ¿Qué cálculos con Q y R producirán la solución?

WEB

25. (*Descomposición en valores singulares*). Suponga que $A = UDV^T$, donde U y V son matrices de $n \times n$ con la propiedad de que $U^TU = I$ y $V^TV = I$, y donde D es una matriz diagonal con números positivos $\sigma_1, \ldots, \sigma_n$ en la diagonal. Demuestre que A es invertible y encuentre una fórmula para A^{-1}.

26. (*Factorización espectral*). Suponga que una matriz A de 3×3 admite una factorización como $A = PDP^{-1}$, donde P es alguna matriz invertible de 3×3, y D es la matriz diagonal

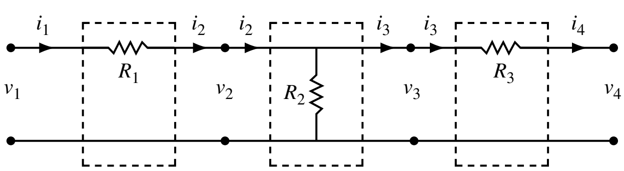
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Muestre que esta factorización es útil cuando se calculan potencias grandes de A. Encuentre fórmulas relativamente sencillas para A^2, A^3 y A^k (k es un entero positivo), usando P y las entradas en D.

27. Diseñe dos redes en escalera diferentes con salida de 9 volts y 4 amperes cuando la entrada sea de 12 volts y 6 amperes.

28. Demuestre que si tres circuitos con derivación (cuyas resistencias son R_1, R_2, R_3) se conectan en serie, la red resultante tiene la misma matriz de transferencia que un único circuito con derivación. Encuentre una fórmula para la resistencia que haya en ese circuito.

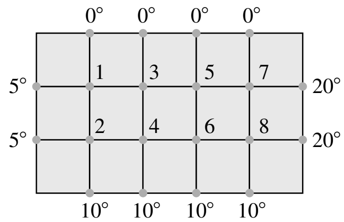
29. *a*) Encuentre la matriz de transferencia de la red que se ilustra en la figura.



b) Sea $A = \begin{bmatrix} 3 & -12 \\ -1/3 & 5/3 \end{bmatrix}$. Diseñe una red en escalera, cuya matriz de transferencia sea A, encontrando una factorización matricial adecuada de A.

30. Encuentre una factorización diferente de la matriz de transferencia A del ejercicio 29 y, a partir de ello, diseñe una red en escalera diferente cuya matriz de transferencia sea A.

31. **[M]** Considere la placa térmica en la siguiente figura (consulte el ejercicio 33 en la sección 1.1).



La solución al problema de flujo de calor en estado estable para la placa de la figura se aproxima al resolver la ecuación $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = (5, 15, 0, 10, 0, 10, 20, 30)$ y

$$A = \begin{bmatrix} 4 & -1 & -1 & & & & & \\ -1 & 4 & 0 & -1 & & & & \\ -1 & 0 & 4 & -1 & -1 & & & \\ & -1 & -1 & 4 & 0 & -1 & & \\ & & -1 & 0 & 4 & -1 & -1 & \\ & & & -1 & -1 & 4 & 0 & -1 \\ & & & & -1 & 0 & 4 & -1 \\ & & & & & -1 & -1 & 4 \end{bmatrix}$$

WEB

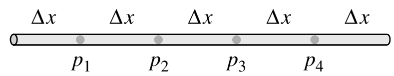
Las entradas faltantes en A son ceros. Las entradas de A diferentes de cero se encuentran dentro de una banda a lo largo de la diagonal principal. Tales *matrices de banda* se presentan en diversas aplicaciones, y con frecuencia son extremadamente grandes (con miles de filas y columnas, pero con bandas relativamente angostas).

a) Utilice el método del ejemplo 2 para construir una factorización LU de A, y observe que ambos factores son matrices de banda (con dos diagonales diferentes de cero abajo o arriba de la diagonal principal). Calcule $LU - A$ para comprobar su trabajo.

b) Use la factorización LU para resolver $A\mathbf{x} = \mathbf{b}$.

c) Obtenga A^{-1} y observe que A^{-1} es una matriz densa sin estructura de banda. Cuando A es grande, L y U se pueden almacenar en mucho menos espacio que A^{-1}. Este hecho es otra razón para preferir la factorización LU de A en lugar de A^{-1}.

32. **[M]** La matriz de banda A que se ilustra a continuación puede servir para calcular la conducción inestable de calor en una varilla para la cual las temperaturas en los puntos $p_1, \ldots, p_4$ cambian con el tiempo.[2]



La constante C de la matriz depende de la naturaleza física de la varilla, de la distancia Δx entre los puntos de la varilla, y del tiempo Δt que transcurre entre mediciones sucesivas de temperatura. Suponga que para $k = 0, 1, 2, \ldots$, un vector $\mathbf{t}_k$ en $\mathbb{R}^4$ lista las temperaturas en el tiempo $k\Delta t$. Si ambos extremos de la varilla se mantienen a 0°, entonces los vectores de temperatura satisfacen la ecuación $A\mathbf{t}_{k+1} = \mathrm{t}_k$ ($k = 0, 1, \ldots$), donde

$$A = \begin{bmatrix} (1+2C) & -C & & \\ -C & (1+2C) & -C & \\ & -C & (1+2C) & -C \\ & & -C & (1+2C) \end{bmatrix}$$

a) Encuentre la factorización LU de A cuando $C = 1$. Una matriz como A con tres diagonales diferentes de cero se denomina *matriz tridiagonal*. Los factores L y U son *matrices bidiagonales*.

b) Suponga que $C = 1$ y $\mathbf{t}_0 = (10, 15, 15, 10)^T$. Use la factorización LU de A para encontrar las distribuciones de temperatura $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$ y $\mathbf{t}_4$.

[2] Véase Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200-201.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 10 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

Divida las entradas de cada columna resaltada por el pivote en la parte superior. Las columnas resultantes forman las tres primeras columnas de la mitad inferior de L. Esto basta para hacer que la reducción por filas de L a I corresponda a la reducción de A a U. Use las dos últimas columnas de I_5 para hacer que L sea triangular inferior unitaria.

$$\begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix} \begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix}$$

$$\begin{array}{ccc} \div 2 & \div 3 & \div 5 \\ \downarrow & \downarrow & \downarrow \end{array}$$

$$\begin{bmatrix} 1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

2.6 EL MODELO DE LEONTIEF DE ENTRADA Y SALIDA

WEB

El álgebra lineal desempeñó un papel fundamental en el trabajo de Wassily Leontief ganador del Premio Nobel, como se mencionó al principio del capítulo 1. El modelo económico descrito en esta sección es la base de modelos más complejos usados en muchas partes del mundo.

Suponga que la economía de una nación se divide en n sectores que producen bienes o servicios, y sea $\mathbf{x}$ un **vector de producción** en $\mathbb{R}^n$ que lista la producción de cada sector en un año. También, suponga que otra parte de la economía (que se llama *sector abierto*) no produce bienes ni servicios, sino que solamente los consume, y sea $\mathbf{d}$ un **vector de demanda final** (o **cuenta de demandas finales**) que lista los valores de los bienes y servicios demandados a los diversos sectores por la parte no productiva de la economía. El vector $\mathbf{d}$ puede representar la demanda del consumidor, el consumo del gobierno, el superávit de producción, las exportaciones u otras demandas externas.

Conforme los diversos sectores elaboran bienes para satisfacer la demanda del consumidor, los productores crean por sí mismos una **demanda intermedia** adicional de bienes que necesitan como insumos para su propia producción. Las interrelaciones de los sectores son muy complejas, y la conexión entre la demanda final y la producción es poco clara. Leontief se preguntó si hay un nivel de producción $\mathbf{x}$ tal que las cantidades producidas (o "suministradas") equilibran exactamente la demanda total de esa producción, de modo que

$$\left\{\begin{matrix} \text{cantidad} \\ \text{producida} \\ \mathbf{x} \end{matrix}\right\} = \left\{\begin{matrix} \text{demanda} \\ \text{intermedia} \end{matrix}\right\} + \left\{\begin{matrix} \text{demanda} \\ \text{final} \\ \mathbf{d} \end{matrix}\right\} \tag{1}$$

La suposición básica del modelo de Leontief de entrada y salida es que, para cada sector, hay un **vector de consumo unitario** en $\mathbb{R}^n$ que lista los insumos necesarios *por unidad de producción* del sector. Todas las unidades de entrada y salida se miden en millones de dólares, y no en cantidades como toneladas o fanegas. (Los precios de los bienes y servicios se mantienen constantes).

Como un ejemplo sencillo, suponga que la economía consiste en tres sectores —manufactura, agricultura y servicios—, con los vectores unitarios de consumo $\mathbf{c}_1$, $\mathbf{c}_2$ y $\mathbf{c}_3$ que se muestran en la siguiente tabla.

	Insumos consumidos por unidad de producción		
Comprados por:	**Manufactura**	**Agricultura**	**Servicios**
Manufactura	.50	.40	.20
Agricultura	.20	.30	.10
Servicios	.10	.10	.30
	$\uparrow$ $\mathbf{c}_1$	$\uparrow$ $\mathbf{c}_2$	$\uparrow$ $\mathbf{c}_3$

EJEMPLO 1 ¿Qué cantidades consumirá el sector de manufactura si decide producir 100 unidades?

SOLUCIÓN Calcule

$$100\mathbf{c}_1 = 100\begin{bmatrix} .50 \\ .20 \\ .10 \end{bmatrix} = \begin{bmatrix} 50 \\ 20 \\ 10 \end{bmatrix}$$

Para producir 100 unidades, manufactura ordenará (es decir, "demandará") y consumirá 50 unidades de otras partes del sector de manufactura, 20 unidades de agricultura y 10 unidades de servicios. ■

Si manufactura decide producir x_1 unidades, entonces $x_1\mathbf{c}_1$ representa las *demandas intermedias* de manufactura, porque las cantidades de $x_1\mathbf{c}_1$ se consumirán en el proceso de creación de las x_1 unidades de producción. De la misma forma, si x_2 y x_3 denotan las producciones planeadas de los sectores de agricultura y servicios, $x_2\mathbf{c}_2$ y $x_3\mathbf{c}_3$, listan las demandas intermedias correspondientes. La demanda intermedia total de los tres sectores está dada por

$$\begin{aligned} \{\text{demanda intermedia}\} &= x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + x_3\mathbf{c}_3 \\ &= C\mathbf{x} \end{aligned} \tag{2}$$

donde C es la **matriz de consumo** $[\mathbf{c}_1 \quad \mathbf{c}_2 \quad \mathbf{c}_3]$, a saber,

$$C = \begin{bmatrix} .50 & .40 & .20 \\ .20 & .30 & .10 \\ .10 & .10 & .30 \end{bmatrix} \tag{3}$$

Las ecuaciones (1) y (2) producen el modelo de Leontief.

EL MODELO DE LEONTIEF DE ENTRADA Y SALIDA, O ECUACIÓN DE PRODUCCIÓN

$$\underset{\substack{\text{Cantidad}\\ \text{producida}}}{\mathbf{x}} = \underset{\substack{\text{Demanda}\\ \text{intermedia}}}{C\mathbf{x}} + \underset{\substack{\text{Demanda}\\ \text{final}}}{\mathbf{d}} \tag{4}$$

La ecuación (4) también se puede escribir como $I\mathbf{x} - C\mathbf{x} = \mathbf{d}$, o

$$(I - C)\mathbf{x} = \mathbf{d} \tag{5}$$

EJEMPLO 2 Considere la economía cuya matriz de consumo está dada por (3). Suponga que la demanda final es de 50 unidades para manufactura, 30 unidades para agricultura, y 20 unidades para servicios. Encuentre el nivel de producción $\mathbf{x}$ que satisfará esta demanda.

SOLUCIÓN La matriz de coeficientes en (5) es

$$I - C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} .5 & .4 & .2 \\ .2 & .3 & .1 \\ .1 & .1 & .3 \end{bmatrix} = \begin{bmatrix} .5 & -.4 & -.2 \\ -.2 & .7 & -.1 \\ -.1 & -.1 & .7 \end{bmatrix}$$

Para resolver (5), reduzca por filas la matriz aumentada

$$\begin{bmatrix} .5 & -.4 & -.2 & 50 \\ -.2 & .7 & -.1 & 30 \\ -.1 & -.1 & .7 & 20 \end{bmatrix} \sim \begin{bmatrix} 5 & -4 & -2 & 500 \\ -2 & 7 & -1 & 300 \\ -1 & -1 & 7 & 200 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 226 \\ 0 & 1 & 0 & 119 \\ 0 & 0 & 1 & 78 \end{bmatrix}$$

La última columna se redondea a la unidad más cercana. El área de manufactura debe producir aproximadamente 226 unidades, agricultura 119 unidades, y servicios únicamente 78 unidades. ■

Si la matriz $I - C$ es invertible, entonces se puede aplicar el teorema 5 de la sección 2.2 con A remplazada *por* $(I - C)$, y a partir de la ecuación $(I - C)\mathbf{x} = \mathbf{d}$ se obtiene $\mathbf{x} = (I - C)^{-1}\mathbf{d}$. El siguiente teorema indica que, en la mayoría de los casos prácticos, $I - C$ *es* invertible y el vector de producción $\mathbf{x}$ es económicamente factible, en el sentido de que las entradas de $\mathbf{x}$ son no negativas.

En el teorema, el término **suma de columna** denota la suma de las entradas en una columna de una matriz. En circunstancias ordinarias, las sumas de columna de una matriz de consumo son menores que 1 porque un sector debería requerir menos de una unidad de insumos para generar una unidad de producción.

TEOREMA 11

Sea C la matriz de consumo de una economía, y sea $\mathbf{d}$ la demanda final. Si C y $\mathbf{d}$ tienen entradas no negativas, y si cada suma de columna de C es menor que uno, entonces $(I - C)^{-1}$ existe y el vector de producción

$$\mathbf{x} = (I - C)^{-1}\mathbf{d}$$

tiene entradas no negativas y es la solución única de

$$\mathbf{x} = C\mathbf{x} + \mathbf{d}$$

El siguiente análisis sugerirá por qué el teorema es cierto, y conducirá a una nueva manera de calcular $(I - C)^{-1}$.

Una fórmula para $(I - C)^{-1}$

Imagine que la demanda representada por $\mathbf{d}$ se propone a las distintas industrias al inicio del año, y que estas responden estableciendo sus niveles de producción en $\mathbf{x} = \mathbf{d}$, lo cual satisfará exactamente la demanda final. Conforme las industrias se preparan para producir $\mathbf{d}$, emiten órdenes solicitando materia prima y otros insumos. Esto crea una demanda intermedia de insumos de $C\mathbf{d}$.

Para satisfacer la demanda adicional de $C\mathbf{d}$, las industrias necesitarán como insumos adicionales las cantidades de $C(C\mathbf{d}) = C^2\mathbf{d}$. Desde luego, esto crea una segunda ronda de demanda intermedia, y cuando las industrias deciden producir aún más para satisfacer esta nueva demanda, se genera una tercera ronda de demanda, a saber, $C(C^2\mathbf{d}) = C^3\mathbf{d}$, y así sucesivamente.

En teoría, este proceso podría continuar de manera indefinida, aunque en la vida real no ocurriría en una sucesión tan rígida de acontecimientos. Podemos elaborar un diagrama de esta situación hipotética de la siguiente forma:

	Demanda que debe satisfacerse	Insumos necesarios para satisfacer esta demanda
Demanda final	$\mathbf{d}$	$C\mathbf{d}$
Demanda intermedia		
1a. ronda	$C\mathbf{d}$	$C(C\mathbf{d}) = C^2\mathbf{d}$
2a. ronda	$C^2\mathbf{d}$	$C(C^2\mathbf{d}) = C^3\mathbf{d}$
3a. ronda	$C^3\mathbf{d}$	$C(C^3\mathbf{d}) = C^4\mathbf{d}$
	$\vdots$	$\vdots$

El nivel de producción $\mathbf{x}$ que satisfará toda esta demanda es

$$\begin{aligned} \mathbf{x} &= \mathbf{d} + C\mathbf{d} + C^2\mathbf{d} + C^3\mathbf{d} + \cdots \\ &= (I + C + C^2 + C^3 + \cdots)\mathbf{d} \end{aligned} \tag{6}$$

Para que la ecuación (6) tenga sentido, considere la siguiente identidad algebraica:

$$(I - C)(I + C + C^2 + \cdots + C^m) = I - C^{m+1} \tag{7}$$

Es posible demostrar que si las sumas de columna en C son todas menores que 1, entonces $1 - C$ es invertible, C^m se aproxima a la matriz cero cuando m crece de manera arbitraria, e $I - C^{m+1} \to I$. (Esto es análogo al hecho de que si un número positivo t es menor que 1, entonces $t^m \to 0$ conforme m aumenta). Con base en la ecuación (7), se escribe

$$\begin{gathered} (I - C)^{-1} \approx I + C + C^2 + C^3 + \cdots + C^m \\ \text{cuando las sumas de columna de } C \text{ son menores que 1.} \end{gathered} \tag{8}$$

La aproximación en (8) significa que el miembro derecho puede acercarse a $(I - C)^{-1}$ tanto como se desee al hacer a m suficientemente grande.

En los modelos de entrada y salida reales, las potencias de la matriz de consumo se aproximan a la matriz cero con cierta rapidez. Así, (8) realmente ofrece una manera práctica de calcular $(I - C)^{-1}$. De la misma forma, para cualquier $\mathbf{d}$, los vectores $C^m\mathbf{d}$ se aproximan al vector cero rápidamente, y (6) es una manera práctica de resolver $(I - C)\mathbf{x} = \mathbf{d}$. Si las entradas de C y $\mathbf{d}$ son no negativas, entonces (6) indica que las entradas de $\mathbf{x}$ también son no negativas.

Importancia económica de las entradas de $(I - C)^{-1}$

Las entradas de $(I - C)^{-1}$ son significativas porque pueden servir para predecir cómo tendrá que cambiar la producción $\mathbf{x}$ conforme cambie la demanda final $\mathbf{d}$. De hecho, las entradas de la columna j de $(I - C)^{-1}$ son las cantidades *aumentadas* que los diversos sectores tendrán que producir para satisfacer un *aumento de 1 unidad* en la demanda final de producción del sector j. Véase el ejercicio 8.

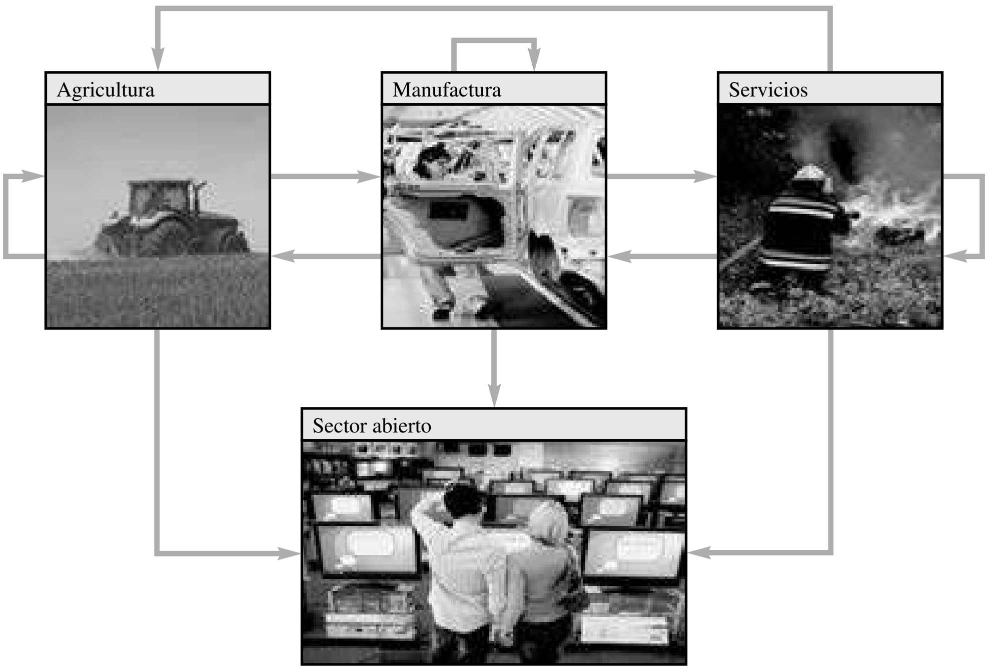
NOTA NUMÉRICA

En cualquier problema de aplicación (no solo en economía), una ecuación $A\mathbf{x} = \mathbf{b}$ se puede escribir siempre como $(I - C)\mathbf{x} = \mathbf{b}$, con $C = I - A$. Si el sistema es grande y *disperso* (con cero en la mayoría de sus entradas), es posible que las sumas de columna de los valores absolutos en C sean menores que 1. En este caso, $C^m \to 0$. Si C^m tiende a cero con la suficiente rapidez, (6) y (8) representarán fórmulas prácticas para resolver $A\mathbf{x} = \mathbf{b}$ y encontrar A^{-1}.

PROBLEMA DE PRÁCTICA

Suponga que una economía tiene dos sectores: bienes y servicios. Una unidad de producción de bienes requiere insumos de .2 unidades de bienes y .5 unidades de servicios. Una unidad de producción de servicios requiere insumos de .4 unidades de bienes y .3 unidades de servicios. Existe una demanda final de 20 unidades de bienes y 30 unidades de servicios. Implemente el modelo de Leontief de entrada y salida para esta situación.

2.6 EJERCICIOS



Los ejercicios 1 a 4 se refieren a una economía dividida en tres sectores: manufactura, agricultura y servicios. Por cada unidad de producción, manufactura requiere de .10 unidades de otras compañías pertenecientes a ese mismo sector, de .30 unidades del sector agricultura y de .30 unidades de servicios. Por cada unidad de producción, agricultura usa .20 unidades de su propia producción, .60 unidades de manufactura y .10 unidades de servicios. Por cada unidad de producción, el sector de servicios consume .10 unidades de servicios, .60 unidades de manufactura, pero ningún producto de agricultura.

1. Construya la matriz de consumo adecuada para esta economía, y determine cuáles demandas intermedias se crean si agricultura planea producir 100 unidades.

2. Determine los niveles de producción que se necesitan para satisfacer una demanda final de 20 unidades para agricultura, sin demanda final para los otros sectores. (No calcule una matriz inversa).

3. Determine los niveles de producción necesarios para satisfacer una demanda final de 20 unidades para manufactura, sin demanda final para los otros sectores. (No calcule una matriz inversa).

4. Determine los niveles de producción necesarios para satisfacer una demanda final de 20 unidades para manufactura, 20 para agricultura y 0 unidades para servicios.

5. Considere el modelo de producción $\mathbf{x} = C\mathbf{x} + \mathbf{d}$ para una economía con dos sectores, donde
$$C = \begin{bmatrix} .0 & .5 \\ .6 & .2 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 50 \\ 30 \end{bmatrix}$$
Use una matriz inversa y determine el nivel de producción necesario para satisfacer la demanda final.

6. Repita el ejercicio 5 con $C = \begin{bmatrix} .2 & .5 \\ .6 & .1 \end{bmatrix}$ y $\mathbf{d} = \begin{bmatrix} 16 \\ 12 \end{bmatrix}$.

7. Sean C y $\mathbf{d}$ como en el ejercicio 5.

   a) Determine el nivel de producción necesario para satisfacer una demanda final para una unidad de producción del sector 1.

   b) Con base en una matriz inversa, determine el nivel de producción necesario para satisfacer una demanda final de $\begin{bmatrix} 51 \\ 30 \end{bmatrix}$.

c) Con base en el hecho de que $\begin{bmatrix} 51 \\ 30 \end{bmatrix} = \begin{bmatrix} 50 \\ 30 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, explique cómo y por qué están relacionadas las respuestas a los incisos *a*) y *b*) y al ejercicio 5.

8. Sea C una matriz de consumo de $n \times n$ cuyas sumas de columna son menores a 1. Sea $\mathbf{x}$ el vector de producción que satisface la demanda final $\mathbf{d}$, y sea $\Delta\mathbf{x}$ un vector de producción para satisfacer una demanda final diferente $\Delta\mathbf{d}$.

a) Demuestre que si la demanda final cambia de $\mathbf{d}$ a $\mathbf{d} + \Delta\mathbf{d}$, entonces el nuevo nivel de producción debe ser $\mathbf{x} + \Delta\mathbf{x}$. Así, $\Delta\mathbf{x}$ indica las cantidades en que debe *cambiar* la producción para compensar el *cambio* $\Delta\mathbf{d}$ en la demanda.

b) Sea $\Delta\mathbf{d}$ el vector en $\mathbb{R}^n$ con 1 en la primera entrada y ceros en las demás. Explique por qué la producción correspondiente $\Delta\mathbf{x}$ es la primera columna de $(I - C)^{-1}$. Esto muestra que la primera columna de $(I - C)^{-1}$ indica las cantidades que deben producir los diversos sectores para satisfacer un aumento de una unidad en la demanda final para la producción del sector 1.

9. Resuelva la ecuación de producción de Leontief para una economía con tres sectores, considerando que

$$C = \begin{bmatrix} .2 & .2 & .0 \\ .3 & .1 & .3 \\ .1 & .0 & .2 \end{bmatrix} \quad \text{y} \quad \mathbf{d} = \begin{bmatrix} 40 \\ 60 \\ 80 \end{bmatrix}$$

10. La matriz de consumo C para la economía de Estados Unidos en 1972 tiene la propiedad de que *cada entrada* de la matriz $(I - C)^{-1}$ es diferente de cero (y positiva).[1] ¿Qué dice esto acerca del efecto de aumentar la demanda de la producción solamente en un sector de la economía?

11. La ecuación de producción de Leontief, $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, generalmente está acompañada por una **ecuación de precio** dual,

$$\mathbf{p} = C^T\mathbf{p} + \mathbf{v}$$

donde $\mathbf{p}$ es un **vector de precio** cuyas entradas listan el precio por unidad de producción de cada sector, y $\mathbf{v}$ es un **vector de valor agregado** cuyas entradas listan el valor agregado por unidad de producción. (El valor agregado incluye salarios, utilidades, depreciación, etcétera). Un hecho importante en economía es que el producto interno bruto (PIB) se puede expresar de dos maneras:

$$\{\text{producto interno bruto}\} = \mathbf{p}^T\mathbf{d} = \mathbf{v}^T\mathbf{x}$$

Compruebe la segunda igualdad. [*Sugerencia:* Calcule $\mathbf{p}^T\mathbf{x}$ de dos maneras].

[1] Wassily W. Leontief, "The World Economy of the Year 2000", *Scientific American*, septiembre de 1980, pp. 206-231.

12. Sea C una matriz de consumo tal que $C^m \to 0$ cuando $m \to \infty$, y para $m = 1, 2,\ldots,$ sea $D_m = I + C + \cdots + C^m$. Encuentre una ecuación en diferencias que relacione D_m y D_{m+1} y, a partir de ella, obtenga un procedimiento iterativo con la finalidad de calcular la fórmula (8) para $(I - C)^{-1}$.

13. **[M]** La siguiente matriz de consumo C está basada en datos de entrada y salida para la economía de Estados Unidos en 1958, con datos para 81 sectores agrupados en 7 sectores de mayores dimensiones: **1.** productos no metálicos personales y domésticos, **2.** productos metálicos finales (como vehículos motorizados), **3.** productos básicos de metal y minería, **4.** productos básicos no metálicos y de agricultura, **5.** energía, **6.** servicios, y **7.** entretenimiento y productos diversos.[2] Encuentre los niveles de producción necesarios para satisfacer la demanda final $\mathbf{d}$. (Las unidades están en millones de dólares).

$$\begin{bmatrix} .1588 & .0064 & .0025 & .0304 & .0014 & .0083 & .1594 \\ .0057 & .2645 & .0436 & .0099 & .0083 & .0201 & .3413 \\ .0264 & .1506 & .3557 & .0139 & .0142 & .0070 & .0236 \\ .3299 & .0565 & .0495 & .3636 & .0204 & .0483 & .0649 \\ .0089 & .0081 & .0333 & .0295 & .3412 & .0237 & .0020 \\ .1190 & .0901 & .0996 & .1260 & .1722 & .2368 & .3369 \\ .0063 & .0126 & .0196 & .0098 & .0064 & .0132 & .0012 \end{bmatrix},$$

$$\mathbf{d} = \begin{bmatrix} 74{,}000 \\ 56{,}000 \\ 10{,}500 \\ 25{,}000 \\ 17{,}500 \\ 196{,}000 \\ 5{,}000 \end{bmatrix}$$

14. **[M]** El vector de demanda del ejercicio 13 es razonable para los datos de 1958, pero el análisis de Leontief de la economía mencionado en el mismo ejercicio utilizó un vector de demanda más cercano a los datos de 1964:

$$\mathbf{d} = (99640, 75548, 14444, 33501, 23527, 263985, 6526)$$

Encuentre los niveles de producción necesarios para satisfacer esta demanda.

15. **[M]** Use la ecuación (6) para resolver el problema del ejercicio 13. Considere que $\mathbf{x}^{(0)} = \mathbf{d}$, y para $k = 1, 2,\ldots,$ calcule $\mathbf{x}^{(k)} = \mathbf{d} + Cx^{(k-1)}$. ¿Cuántos pasos se necesitan para obtener una respuesta al ejercicio 13 con cuatro cifras significativas?

[2] Wassily W. Leontief, "The Structure of the U.S. Economy", *Scientific American*, abril de 1965, pp. 30-32.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Se cuenta con los siguientes datos:

	Insumos necesarios por unidad de producción		
Comprados por:	**Bienes**	**Servicios**	**Demanda externa**
Bienes	.2	.4	20
Servicios	.5	.3	30

El modelo de entrada y salida de Leontief es $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, donde

$$C = \begin{bmatrix} .2 & .4 \\ .5 & .3 \end{bmatrix}, \qquad \mathbf{d} = \begin{bmatrix} 20 \\ 30 \end{bmatrix}$$

2.7 APLICACIONES A LOS GRÁFICOS POR COMPUTADORA

Los gráficos por computadora son imágenes desplegadas o animadas en una pantalla de computadora. Las aplicaciones de los gráficos por computadora están ampliamente difundidas y proliferan con rapidez. Por ejemplo, el diseño asistido por computadora (CAD) es una parte integral de muchos procesos de ingeniería, como el proceso de diseño de aeronaves descrito en la introducción de este capítulo. La industria del entretenimiento ha realizado el uso más espectacular de los gráficos por computadora: desde los efectos especiales en las películas *Matrix* a la Xbox de PlayStation 2.

La mayoría de los programas interactivos de cómputo diseñados para los negocios y la industria utilizan gráficos por computadora en los despliegues de pantalla y en otras funciones, como el despliegue gráfico de datos, la autoedición, y la producción de diapositivas para presentaciones comerciales y educativas. Por consiguiente, cualquier persona que estudie un lenguaje de computadora siempre pasa algún tiempo aprendiendo a usar gráficos de, por lo menos, dos dimensiones (2D).

En esta sección se examinará algo de las matemáticas básicas que se usan para manipular y desplegar imágenes gráficas, como el modelo de alambre de un avión. Una imagen (o dibujo) de ese tipo consta de varios puntos, líneas rectas o curvas conectados, e información sobre cómo llenar regiones cerradas delimitadas por esas rectas y curvas. Con frecuencia, las líneas curvas se aproximan utilizando segmentos de línea recta cortos, y una figura se define matemáticamente por medio de una lista de puntos.

Entre los símbolos gráficos más sencillos utilizados en 2D están las letras usadas como etiquetas en la pantalla. Algunas letras se guardan como objetos de alambre; otras, con porciones curvas, se almacenan con fórmulas matemáticas adicionales para las curvas.

EJEMPLO 1 La letra N mayúscula de la figura 1 está determinada por ocho puntos o *vértices*. Las coordenadas de los puntos se pueden almacenar en una matriz de datos, D.

$$\begin{array}{r c} & \text{Vértice:} \\ & \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{array}{r} \text{coordenada } x \\ \text{coordenada } y \end{array} & \begin{bmatrix} 0 & .5 & .5 & 6 & 6 & 5.5 & 5.5 & 0 \\ 0 & 0 & 6.42 & 0 & 8 & 8 & 1.58 & 8 \end{bmatrix} = D \end{array}$$

Además de D, es necesario especificar cuáles vértices están conectados con líneas, pero aquí se omite este detalle. ■

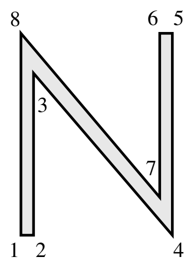


FIGURA 1
N regular.

La principal razón para describir los objetos gráficos por medio de segmentos de líneas rectas es que las transformaciones estándar en los gráficos por computadora mapean segmentos de línea sobre otros segmentos de línea. (Por ejemplo, véase el ejercicio 26 de la sección 1.8). Una vez transformados los vértices que describen un objeto, es posible conectar sus imágenes con las líneas rectas adecuadas para producir la imagen completa del objeto original.

EJEMPLO 2 A partir de $A = \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix}$, describa el efecto de la transformación de trasquilado $\mathbf{x} \mapsto A\mathbf{x}$ sobre la letra N del ejemplo 1.

SOLUCIÓN Por la definición de multiplicación de matrices, las columnas del producto AD contienen las imágenes de los vértices de la letra N.

$$AD = \begin{array}{c} \begin{matrix} 1 & \;2 & \;\;3 & \;\;4 & 5 & \;6 & \;\;7 & \;\;8 \end{matrix} \\ \begin{bmatrix} 0 & .5 & 2.105 & 6 & 8 & 7.5 & 5.895 & 2 \\ 0 & 0 & 6.420 & 0 & 8 & 8 & 1.580 & 8 \end{bmatrix} \end{array}$$

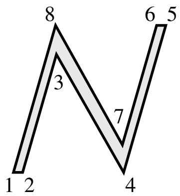


FIGURA 2
N inclinada.

Los vértices transformados se dibujan en la figura 2, junto con los segmentos de línea conectores que corresponden a los de la figura original. ■

La *N* cursiva de la figura 2 se ve demasiado ancha. Para compensar ese hecho, se puede reducir la anchura mediante una transformación de escala que afecta las coordenadas x de los puntos.

EJEMPLO 3 Calcule la matriz de la transformación que realiza una transformación de trasquilado, como en el ejemplo 2, y que después modifica a escala todas las coordenadas x por un factor de .75.

SOLUCIÓN La matriz que multiplica la coordenada x de un punto por .75 es

$$S = \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix}$$

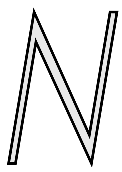


FIGURA 3
Transformación compuesta de *N*.

Así que la matriz de la transformación compuesta es

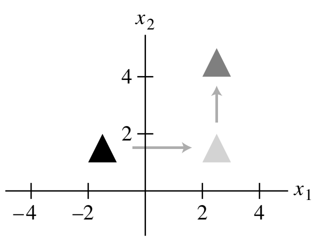
$$\begin{aligned} SA &= \begin{bmatrix} .75 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & .25 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} .75 & .1875 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

El resultado de esta transformación compuesta se muestra en la figura 3. ■

Las matemáticas de los gráficos por computadora están íntimamente relacionadas con la multiplicación de matrices. Por desgracia, trasladar un objeto a una pantalla no corresponde directamente a la multiplicación de matrices porque la traslación no es una transformación lineal. La manera estándar de evitar esta dificultad es introducir lo que se conoce como *coordenadas homogéneas*.

Coordenadas homogéneas

Cada punto (x, y) en $\mathbb{R}^2$ se puede identificar con el punto $(x, y, 1)$ en el plano en $\mathbb{R}^3$ que se encuentra una unidad por encima del plano xy. Se dice que (x, y) tiene *coordenadas homogéneas* $(x, y, 1)$. Por ejemplo, el punto $(0, 0)$ tiene coordenadas homogéneas $(0, 0, 1)$. Las coordenadas homogéneas de puntos no se suman ni se multiplican por escalares, pero se pueden transformar mediante multiplicación por matrices de 3×3.



Traslación por $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$.

EJEMPLO 4 Una traslación de la forma $(x, y) \mapsto (x + h, y + k)$ se escribe en coordenadas homogéneas como $(x, y, 1) \mapsto (x + h, y + k, 1)$. Esta transformación se puede calcular mediante la multiplicación de matrices:

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + h \\ y + k \\ 1 \end{bmatrix}$$ ■

EJEMPLO 5 Cualquier transformación lineal sobre $\mathbb{R}^2$ se representa con respecto a las coordenadas homogéneas por medio de una matriz particionada de la forma $\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$, donde A es una matriz de 2×2. Ejemplos típicos son:

$$\underset{\substack{\text{Rotación en sentido}\\ \text{antihorario en torno}\\ \text{al origen, ángulo } \varphi}}{\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}}, \quad \underset{\substack{\text{Reflexión}\\ \text{a través de } y = x}}{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}, \quad \underset{\substack{\text{Escala de } x \text{ por } s\\ \text{y de } y \text{ por } t}}{\begin{bmatrix} s & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

■

Transformaciones compuestas

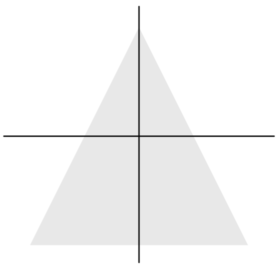
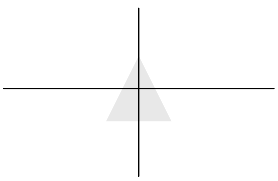


Figura original

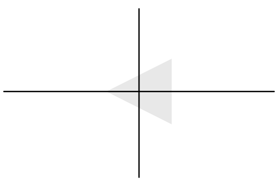
El movimiento de una figura en la pantalla de una computadora a menudo requiere de dos o más transformaciones básicas. La composición de tales transformaciones corresponde a la multiplicación de matrices cuando se usan coordenadas homogéneas.

EJEMPLO 6 Encuentre la matriz de 3×3 que corresponde a la transformación compuesta de un escalamiento por .3, una rotación de 90° en torno al origen y, por último, una traslación que suma $(-.5, 2)$ a cada punto de una figura.

SOLUCIÓN Si $\varphi = \pi/2$, entonces sen $\varphi = 1$ y cos $\varphi = 0$. A partir de los ejemplos 4 y 5, se tiene que



Después del escalamiento



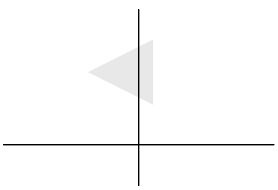
Después de la rotación

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \xrightarrow{\text{Escalamiento}} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\xrightarrow{\text{Rotación}} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\xrightarrow{\text{Traslación}} \begin{bmatrix} 1 & 0 & -.5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

La matriz para la transformación compuesta es



Después de la traslación

$$\begin{bmatrix} 1 & 0 & -.5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & -1 & -.5 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .3 & 0 & 0 \\ 0 & .3 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -.3 & -.5 \\ .3 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

■

Gráficos 3D por computadora

Algunos de los más recientes y estimulantes trabajos en gráficos por computadora se relacionan con el modelado molecular. Un biólogo tiene la posibilidad de examinar una molécula simulada de proteína utilizando gráficos tridimensionales (3D), y así buscar los sitios activos que pueden aceptar la molécula de un medicamento. El biólogo podrá hacer girar y trasladar un medicamento experimental para tratar de unirlo a la proteína. Esta capacidad de *visualizar* reacciones químicas potenciales es vital para la investigación de medicamentos modernos y del cáncer. De hecho, los avances en el diseño de medicamentos dependen, en cierta medida, del progreso que se logre en la capacidad de los gráficos por computadora para construir simulaciones realistas de las moléculas y sus interacciones.[1]

La investigación actual en el modelado de moléculas se enfoca en la *realidad virtual*, un entorno en el que un investigador puede ver y *sentir* la molécula de medicamento deslizarse dentro de la proteína. En la figura 4 se ilustra el proceso de retroalimentación táctil con un manipulador remoto que despliega la fuerza.

FIGURA 4 Modelado molecular en realidad virtual. (Departamento de Ciencias de la Computación, University of North Carolina en Chapel Hill. Fotografía de Bo Strain).

Otro diseño de realidad virtual consiste en un casco y un guante que detectan los movimientos de la cabeza, la mano y los dedos. El casco incluye dos pequeñas pantallas de computadora, una para cada ojo. Hacer que este medio virtual sea más realista es un desafío para ingenieros, científicos y matemáticos. Las matemáticas que se manejan aquí apenas abren la puerta a este campo de la investigación.

Coordenadas 3D homogéneas

Por analogía con el caso bidimensional, se dice que $(x, y, z, 1)$ son las coordenadas homogéneas para el punto (x, y, z) en $\mathbb{R}^3$. En general, (X, Y, Z, H) son las **coordenadas homogéneas** para (x, y, z) si $H \neq 0$ y

$$x = \frac{X}{H}, \qquad y = \frac{Y}{H} \quad \text{y} \quad z = \frac{Z}{H} \tag{1}$$

Cada múltiplo escalar diferente de cero de $(x, y, z, 1)$ proporciona un conjunto de coordenadas homogéneas para (x, y, z). Por ejemplo, $(10, -6, 14, 2)$ y $(-15, 9, -21, -3)$ son coordenadas homogéneas para $(5, -3, 7)$.

El siguiente ejemplo ilustra las transformaciones utilizadas en el modelado molecular para introducir un medicamento en una molécula de proteína.

EJEMPLO 7 Encuentre las matrices de 4×4 para las siguientes transformaciones:

a) Rotación en torno al eje y a través de un ángulo de 30°. (Por convención, un ángulo positivo es un giro en sentido antihorario cuando se ve hacia el origen desde la mitad positiva del eje de rotación, en este caso, el eje y).

[1] Robert Pool, "Computing in Science", *Science* **256**, 3 de abril de 1992, p. 45.

b) Traslación mediante el vector $\mathbf{p} = (-6, 4, 5)$.

SOLUCIÓN

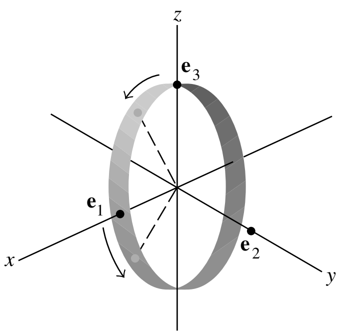


FIGURA 5

a) Primero, construya la matriz de 3×3 para la rotación. El vector $\mathbf{e}_1$ gira hacia abajo en la dirección del eje z negativo, y se detiene en $(\cos 30°, 0, -\text{sen } 30°) = (\sqrt{3}/2, 0, -.5)$. El vector $\mathbf{e}_2$ sobre el eje y no se mueve, pero $\mathbf{e}_3$ sobre el eje z gira hacia abajo en dirección del eje x positivo, hasta detenerse en $(\text{sen } 30°, 0, \cos 30°) = (.5, 0, \sqrt{3}/2)$. Véase la figura 5. De acuerdo con la sección 1.9, la matriz estándar para esta rotación es

$$\begin{bmatrix} \sqrt{3}/2 & 0 & .5 \\ 0 & 1 & 0 \\ -.5 & 0 & \sqrt{3}/2 \end{bmatrix}$$

Por lo tanto, la matriz de rotación para las coordenadas homogéneas es

$$A = \begin{bmatrix} \sqrt{3}/2 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ -.5 & 0 & \sqrt{3}/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

b) Se desea que $(x, y, z, 1)$ mapee a $(x - 6, y + 4, z + 5, 1)$. La matriz que logra esto es

$$\begin{bmatrix} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

■

Proyecciones en perspectiva

Un objeto tridimensional se representa en la pantalla de dos dimensiones de una computadora proyectándolo sobre un *plano visual*. (Se omiten otros pasos importantes, como la selección de la parte del plano visual que se desplegará en la pantalla). Para simplificar, considere que el plano xy representa la pantalla de la computadora e imagine que el ojo de un observador está sobre el eje positivo z, en un punto $(0, 0, d)$. Una *proyección en perspectiva* mapea cada punto (x, y, z) sobre un punto de imagen $(x^*, y^*, 0)$ de manera que los dos puntos y la posición del ojo, llamada **centro de proyección**, estén sobre una recta. Véase la figura 6*a*).

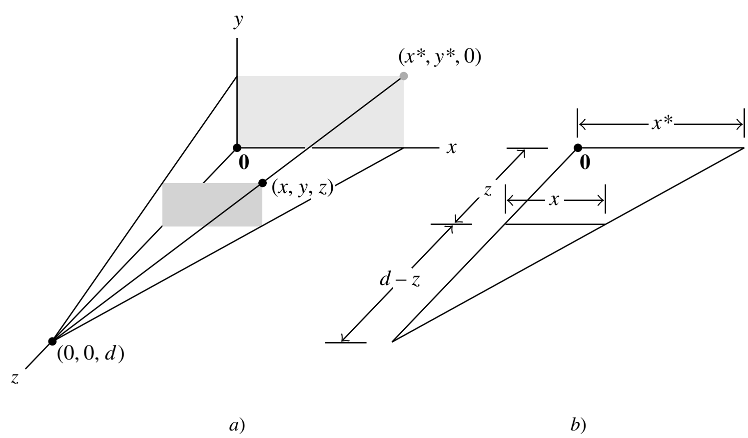


FIGURA 6 Proyección en perspectiva de (x, y, z) sobre $(x^*, y^*, 0)$.

El triángulo en el plano xz de la figura 6*a*) se vuelve a trazar en el inciso *b*) mostrando la longitud de los segmentos de recta. Con triángulos similares se muestra que

$$\frac{x^*}{d} = \frac{x}{d-z} \quad \text{y} \quad x^* = \frac{dx}{d-z} = \frac{x}{1-z/d}$$

De manera análoga,

$$y^* = \frac{y}{1-z/d}$$

Usando coordenadas homogéneas, es posible representar la proyección en perspectiva mediante una matriz, por ejemplo, P. Se desea que $(x, y, z, 1)$ se mapee en $\left(\dfrac{x}{1-z/d}, \dfrac{y}{1-z/d}, 0, 1\right)$. Al modificar a escala estas coordenadas por $1 - z/d$, también se puede utilizar $(x, y, 0, 1 - z/d)$ como coordenadas homogéneas para la imagen. Ahora resulta fácil desplegar P. De hecho,

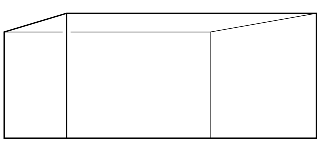
$$P\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/d & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \\ 1-z/d \end{bmatrix}$$

EJEMPLO 8 Sea S la caja con vértices (3, 1, 5), (5, 1, 5), (5, 0, 5), (3, 0, 5), (3, 1, 4), (5, 1, 4), (5, 0, 4) y (3, 0, 4). Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en (0, 0, 10).

SOLUCIÓN Sean P la matriz de proyección y D la matriz de datos para S usando coordenadas homogéneas. La matriz de datos para la imagen de S es

$$PD = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1/10 & 1 \end{bmatrix} \begin{array}{c} \text{Vértice:} \\ \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 5 & 4 & 4 & 4 & 4 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 3 & 5 & 5 & 3 & 3 & 5 & 5 & 3 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ .5 & .5 & .5 & .5 & .6 & .6 & .6 & .6 \end{bmatrix}$$



S bajo la transformación de perspectiva.

Para obtener las coordenadas en $\mathbb{R}^3$, use la ecuación (1) que está antes del ejemplo 7 y divida las tres entradas superiores de cada columna entre la entrada correspondiente de la cuarta fila:

$$\begin{array}{c} \text{Vértice:} \\ \begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\ \begin{bmatrix} 6 & 10 & 10 & 6 & 5 & 8.3 & 8.3 & 5 \\ 2 & 2 & 0 & 0 & 1.7 & 1.7 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

■

WEB

En el sitio Web de este libro se presentan algunas aplicaciones interesantes de los gráficos por computadora, incluyendo un análisis más profundo de las proyecciones en perspectiva. Uno de los proyectos para computadora presentados en el sitio Web implica una animación sencilla.

NOTA NUMÉRICA

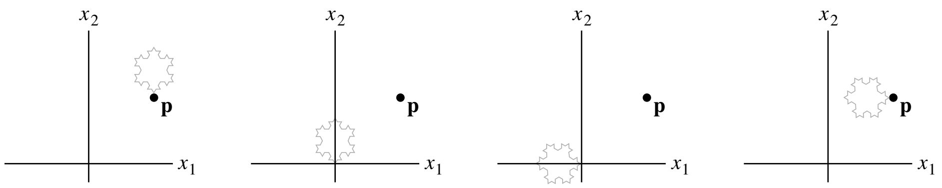
El movimiento continuo de objetos gráficos en 3D requiere de cálculo intenso con matrices de 4×4, en especial cuando las superficies deben parecer realistas, con la textura e iluminación adecuadas. Las *estaciones de trabajo para gráficos por computadora de acabado fino* realizan operaciones con matrices de 4×4 y algoritmos de gráficos integrados en sus microchips y circuitos. Estas estaciones son capaces de realizar miles de millones de multiplicaciones de matrices por segundo necesarias para presentar la animación realista en color en los programas de juegos 3D.[2]

Lectura adicional

James D. Foley, Andries van Dam, Steven K. Feiner y John F. Hughes, *Computer Graphics: Principles and Practice*, 3a. ed. (Boston, MA: Addison-Wesley, 2002), capítulos 5 y 6.

PROBLEMA DE PRÁCTICA

La rotación de una figura alrededor de un punto $\mathbf{p}$ en $\mathbb{R}^2$ se logra primero con una traslación de la figura mediante $-\mathbf{p}$, luego con una rotación alrededor del origen, y finalmente trasladándola de regreso mediante $\mathbf{p}$. Véase la figura 7. Construya la matriz de 3×3 que hace girar puntos en un ángulo de $-30°$ en torno al punto $(-2, 6)$, usando coordenadas homogéneas.



a) Figura original. *b*) Traslación al origen mediante $-\mathbf{p}$. *c*) Rotación en torno al origen. *d*) Traslación de regreso mediante $\mathbf{p}$.

FIGURA 7 Rotación de la figura con respecto al punto **p**.

2.7 EJERCICIOS

1. ¿Qué matriz de 3×3 tendrá el mismo efecto en las coordenadas homogéneas para $\mathbb{R}^2$ que la matriz de corte A del ejemplo 2?

2. Utilice multiplicación de matrices para encontrar la imagen del triángulo con datos $D = \begin{bmatrix} 4 & 2 & 5 \\ 0 & 2 & 3 \end{bmatrix}$ bajo la transformación que refleja los puntos a través del eje y. Bosqueje tanto el triángulo original como su imagen.

En los ejercicios 3 a 8, determine las matrices de 3×3 que producen las transformaciones 2D compuestas, usando coordenadas homogéneas.

3. Traslade con (2, 1), y después haga girar 90° con respecto al origen.

4. Traslade con $(-1, 4)$, y después modifique a escala la coordenada x por 1/2 y la coordenada y por 3/2.

5. Refleje los puntos a través del eje x, y después hágalos girar 45° con respecto al origen.

6. Haga girar los puntos 45° con respecto al origen, después refléjelos a través del eje x.

7. Haga girar los puntos un ángulo de 60° con respecto al punto (6, 8).

8. Haga girar los puntos un ángulo de 45° con respecto al punto (3, 7).

9. Una matriz de datos de 2×100 contiene las coordenadas de 100 puntos. Calcule el número de multiplicaciones necesarias para transformar estos puntos usando dos matrices arbitrarias A y B de 2×2. Considere las dos posibilidades $A(BD)$ y $(AB)D$. Analice las implicaciones de sus resultados con los cálculos de los gráficos por computadora.

[2] Véase Jan Ozer, "High-Performance Graphics Boards", *PC Magazine* **19**, 1 de septiembre de 2000, pp. 187-200. También "The Ultimate Upgrade Guide: Moving On Up", *PC Magazine* **21**, 29 de enero de 2002, pp. 82-91.

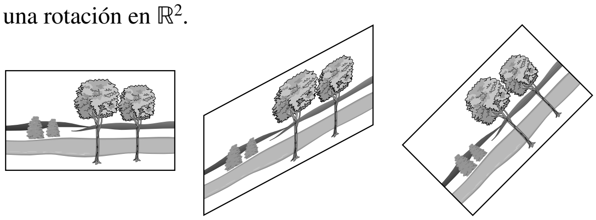
10. Considere las siguientes transformaciones geométricas 2D: D, una dilatación (en la que se modifica la escala de las coordenadas x y y por el mismo factor); R, una rotación; y T, una traslación. ¿*Conmuta* D con R? Es decir, ¿es $D(R(\mathbf{x})) = R(D(\mathbf{x}))$ para toda $\mathbf{x}$ en $\mathbb{R}^2$? ¿Conmuta D con T? ¿Conmuta R con T?

11. Una rotación en la pantalla de una computadora a veces se implementa como el producto de dos transformaciones de trasquilado y escalamiento, que pueden acelerar los cálculos para determinar cómo se presenta en realidad una imagen gráfica en términos de los pixeles de la pantalla. (La pantalla consiste en filas y columnas de puntos pequeños, llamados *pixeles*). La primera transformación A_1 trasquila verticalmente y después comprime cada columna de pixeles; la segunda transformación A_2 trasquila horizontalmente y después estira cada fila de pixeles. Sean

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \sec\varphi & -\tan\varphi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

y muestre que la composición de las dos transformaciones es una rotación en $\mathbb{R}^2$.



12. Una rotación en $\mathbb{R}^2$, por lo general, requiere cuatro multiplicaciones. Calcule el siguiente producto y demuestre que la matriz para una rotación se puede factorizar en tres transformaciones de trasquilado (cada una de las cuales requiere de tan solo una multiplicación).

$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \operatorname{sen}\varphi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

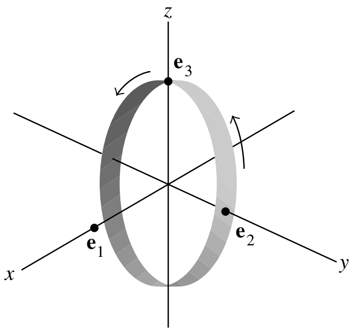
13. Las transformaciones usuales en coordenadas homogéneas para gráficos 2D por computadora implican matrices de 3×3 en la forma $\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}$ donde A es una matriz de 2×2 y $\mathbf{p}$ está en $\mathbb{R}^2$. Demuestre que una transformación de este tipo equivale a una transformación lineal en $\mathbb{R}^2$ seguida por una traslación. [*Sugerencia*: Encuentre una factorización de matrices adecuada en la que intervengan matrices particionadas].

14. Demuestre que la transformación del ejercicio 7 es equivalente a una rotación con respecto al origen seguida por una traslación mediante $\mathbf{p}$. Encuentre $\mathbf{p}$.

15. ¿Qué vector en $\mathbb{R}^3$ tiene las coordenadas homogéneas $(\frac{1}{2}, -\frac{1}{4}, -\frac{1}{8}, \frac{1}{24})$?

16. ¿Son $(1, -2, -3, 4)$ y $(10, -20, -30, 40)$ coordenadas homogéneas para el mismo punto en $\mathbb{R}^3$? ¿Por qué?

17. Proporcione la matriz de 4×4 que hace girar puntos en $\mathbb{R}^3$ con respecto al eje x a través de un ángulo de 60°. (Véase la figura).



18. Encuentre la matriz de 4×4 que hace girar puntos en $\mathbb{R}^3$ con respecto al eje z a través de un ángulo de $-30°$, y después los traslada mediante $\mathbf{p} = (5, -2, 1)$.

19. Sea S el triángulo con vértices $(4.2, 1.2, 4)$, $(6, 4, 2)$ y $(2, 2, 6)$. Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en $(0, 0, 10)$.

20. Sea S el triángulo con vértices $(7, 3, -5)$, $(12, 8, 2)$, $(1, 2, 1)$. Encuentre la imagen de S bajo la proyección en perspectiva con centro de proyección en $(0, 0, 10)$.

Los ejercicios 21 y 22 se refieren a la manera en que se especifica el color para mostrarse en gráficos por computadora. En una pantalla de computadora, el color se codifica utilizando tres números (R, G, B) para indicar la cantidad de energía que un cañón debe transmitir a los puntos fosforescentes rojos, verdes y azules sobre la pantalla de la computadora. (Un cuarto número especifica la luminosidad o intensidad del color).

21. **[M]** El color real que ve un observador en una pantalla está influido por el tipo específico y la cantidad de material fosforescente en la pantalla. De esta manera, cada fabricante de pantallas de computadora debe hacer la conversión entre los datos (R, G, B) y un estándar internacional para color, CIE, que usa tres colores primarios llamados X, Y y Z. Una conversión típica para el material fosforescente de persistencia breve es

$$\begin{bmatrix} .61 & .29 & .150 \\ .35 & .59 & .063 \\ .04 & .12 & .787 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Un programa de computadora envía un flujo de información acerca del color a la pantalla usando datos del estándar CIE (X, Y, Z). Encuentre la ecuación que convierte esta información en los datos (R, G, B) que necesita el cañón de electrones de la pantalla.

22. **[M]** La señal transmitida por la televisión comercial describe cada color por medio de un vector (Y, I, Q). Si la pantalla es en blanco y negro, solo se utiliza la coordenada Y. (Esto da una mejor imagen monocromática que si se usa el estándar CIE para los colores). La correspondencia entre YIQ y un color "estándar" RGB está dada por

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.528 & .311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(Un fabricante de pantallas cambiaría las entradas de la matriz para que funcionaran con sus pantallas *RGB*). Encuentre la ecuación que convierte los datos *YIQ* transmitidos por la estación de televisión a los datos *RGB* necesarios para la pantalla del televisor.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Acomode las matrices de derecha a izquierda para las tres operaciones. Usando $\mathbf{p} = (-2, 6)$, $\cos(-30°) = \sqrt{3}/2$ y $\operatorname{sen}(-30°) = -.5$, se tiene:

$$\overset{\substack{\text{Traslación de}\\ \text{regreso mediante } p}}{\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}} \overset{\substack{\text{Rotación con}\\ \text{respecto al origen}}}{\begin{bmatrix} \sqrt{3}/2 & 1/2 & 0 \\ -1/2 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \overset{\substack{\text{Traslación}\\ \text{mediante } -p}}{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}}$$

$$= \begin{bmatrix} \sqrt{3}/2 & 1/2 & \sqrt{3}-5 \\ -1/2 & \sqrt{3}/2 & -3\sqrt{3}+5 \\ 0 & 0 & 1 \end{bmatrix}$$

2.8 SUBESPACIOS DE $\mathbb{R}^n$

Esta sección se concentra en los importantes conjuntos de vectores en $\mathbb{R}^n$ llamados *subespacios*. Con frecuencia surgen subespacios conectados con alguna matriz A, los cuales brindan información útil acerca de la ecuación $A\mathbf{x} = \mathbf{b}$. Los conceptos y la terminología de esta sección se usarán repetidamente a lo largo del libro.[1]

DEFINICIÓN

Un **subespacio** de $\mathbb{R}^n$ es cualquier conjunto H en $\mathbb{R}^n$ que tenga tres propiedades:

a) El vector cero está en H.

b) Para cada $\mathbf{u}$ y $\mathbf{v}$ en H, la suma $\mathbf{u} + \mathbf{v}$ está en H.

c) Para cada $\mathbf{u}$ en H y cada escalar c, el vector $c\mathbf{u}$ está en H.

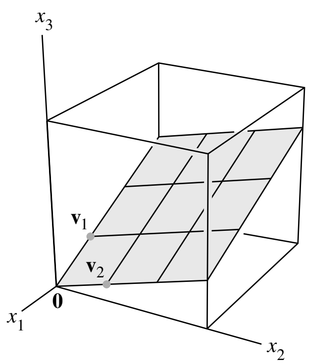


FIGURA 1 Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ tiene un plano a través del origen.

Dicho con palabras, un subespacio es *cerrado* bajo la suma y la multiplicación escalar. Como se verá en los siguientes ejemplos, casi todos los conjuntos de vectores analizados en el capítulo 1 son subespacios. Por ejemplo, un plano que pasa por el origen es la manera estándar de visualizar el subespacio del ejemplo 1. Véase la figura 1.

EJEMPLO 1 Si $\mathbf{v}_1$ y $\mathbf{v}_2$ están en $\mathbb{R}^n$ y $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, entonces H es un subespacio de $\mathbb{R}^n$. Para comprobar este enunciado, observe que el vector cero está en H (porque $0\mathbf{v}_1 + 0\mathbf{v}_2$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$). Ahora tome dos vectores arbitrarios en H, por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

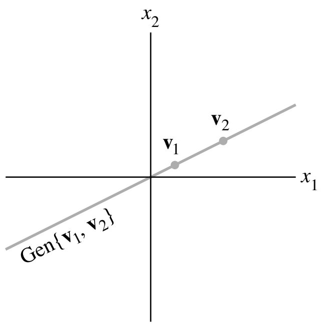
Luego,

$$\mathbf{u} + \mathbf{v} = (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2$$

lo que demuestra que $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y, por lo tanto, está en H. Además, para cualquier escalar c, el vector $c\mathbf{u}$ está en H, ya que $c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$. ■

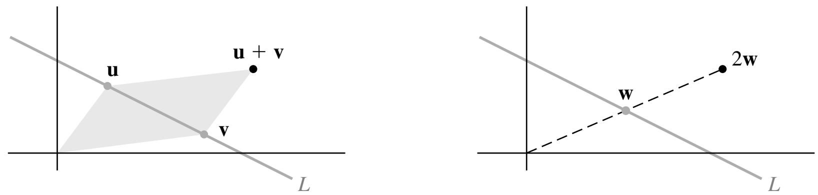
[1] Las secciones 2.8 y 2.9 se incluyen aquí para permitir que los lectores pospongan el estudio de la mayor parte o el total de los siguientes dos capítulos y vayan directamente al capítulo 5, si así lo desean. *Omita* estas dos secciones si planea estudiar el capítulo 4 antes que el capítulo 5.

Si $\mathbf{v}_1$ no es cero, y si $\mathbf{v}_2$ es un múltiplo de $\mathbf{v}_1$, entonces $\mathbf{v}_1$ y $\mathbf{v}_2$ simplemente generan una *recta* que pasa a través del origen. Por lo tanto, una recta que pasa por el origen es otro ejemplo de un subespacio.



$\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2 = k\mathbf{v}_1$.

EJEMPLO 2 Una recta L que *no* pasa por el origen *no* es un subespacio, porque no contiene al origen, como se requiere. Además, la figura 2 muestra que L no es cerrada bajo la suma o la multiplicación escalar. ■



$\mathbf{u} + \mathbf{v}$ no está sobre L $\qquad$ $2\mathbf{w}$ no está sobre L

FIGURA 2

EJEMPLO 3 Para $\mathbf{v}_1, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$, el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ es un subespacio de $\mathbb{R}^n$. La comprobación de este enunciado es similar al argumento dado en el ejemplo 1. Ahora nos referimos a Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ como el **subespacio generado** por $\mathbf{v}_1, \ldots, \mathbf{v}_p$. ■

Observe que $\mathbb{R}^n$ es un subespacio de sí mismo porque tiene las tres propiedades requeridas para un subespacio. Otro subespacio especial es el conjunto que consta exclusivamente del vector cero en $\mathbb{R}^n$. Este conjunto, llamado **subespacio cero**, también satisface las condiciones de un subespacio.

Espacio de columnas y espacio nulo de una matriz

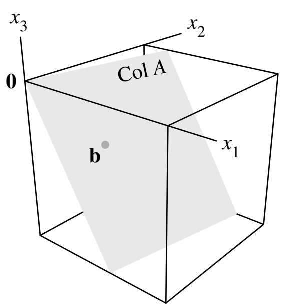
Los subespacios de $\mathbb{R}^n$ generalmente se presentan en aplicaciones y en la teoría en una de dos formas. En ambos casos, es posible relacionar el subespacio con una matriz.

DEFINICIÓN

El **espacio de columnas** de una matriz A es el conjunto Col A de todas las combinaciones de las columnas de A.

Si $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$, con las columnas en $\mathbb{R}^m$, entonces Col A es lo mismo que Gen $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$. En el ejemplo 4 se muestra que el **espacio columna de una matriz de** $\boldsymbol{m \times n}$ **es un subespacio de** $\mathbb{R}^{\boldsymbol{m}}$. Observe que Col A es igual a $\mathbb{R}^m$ solo cuando las columnas de A generan a $\mathbb{R}^m$. Si no la generan, Col A es solo una parte de $\mathbb{R}^m$.

EJEMPLO 4 Sean $A = \begin{bmatrix} 1 & -3 & -4 \\ -4 & 6 & -2 \\ -3 & 7 & 6 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix}$. Determine si $\mathbf{b}$ es el espacio columna de A.



SOLUCIÓN El vector $\mathbf{b}$ es una combinación lineal de las columnas de A si y solo si $\mathbf{b}$ se puede escribir como $A\mathbf{x}$ para alguna $\mathbf{x}$, es decir, si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución. Reduciendo por filas a la matriz aumentada $[A \quad \mathbf{b}]$, se tiene

$$\begin{bmatrix} 1 & -3 & -4 & 3 \\ -4 & 6 & -2 & 3 \\ -3 & 7 & 6 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & -2 & -6 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & -4 & 3 \\ 0 & -6 & -18 & 15 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

se concluye que $A\mathbf{x} = \mathbf{b}$ es consistente, y $\mathbf{b}$ está en Col A. ■

La solución del ejemplo 4 indica que cuando un sistema de ecuaciones lineales está escrito en la forma $A\mathbf{x} = \mathbf{b}$, el espacio columna de A es el conjunto de todas las $\mathbf{b}$ para las que el sistema tiene una solución.

DEFINICIÓN El **espacio nulo** de una matriz A es el conjunto Nul A de todas las soluciones posibles para la ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

Cuando A tiene n columnas, las soluciones de $A\mathbf{x} = \mathbf{0}$ pertenecen a $\mathbb{R}^n$, y el espacio nulo de A es un subconjunto de $\mathbb{R}^n$. De hecho, Nul A tiene las propiedades de un *subespacio* de matrices de $\mathbb{R}^n$.

TEOREMA 12 El espacio nulo de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^n$. De manera equivalente, el conjunto de todas las soluciones posibles para un sistema $A\mathbf{x} = \mathbf{0}$ de m ecuaciones lineales homogéneas con n incógnitas es un subespacio de $\mathbb{R}^n$.

DEMOSTRACIÓN El vector cero está en Nul A (porque $A\mathbf{0} = \mathbf{0}$). Para demostrar que Nul A satisface las otras dos propiedades requeridas para conformar un subespacio, tome cualesquiera $\mathbf{u}$ y $\mathbf{v}$ en Nul A. Es decir, se supone que $A\mathbf{u} = \mathbf{0}$ y $A\mathbf{v} = \mathbf{0}$. Así, por una propiedad de la multiplicación de matrices,

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Por lo tanto, $\mathbf{u} + \mathbf{v}$ satisface $A\mathbf{x} = \mathbf{0}$, y así $\mathbf{u} + \mathbf{v}$ está en Nul A. Además, para cualquier escalar c, $A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0}$, lo que demuestra que $c\mathbf{u}$ está en Nul A. ■

Para saber si un vector dado $\mathbf{v}$ está en Nul A, solo se calcula $A\mathbf{v}$ para ver si $A\mathbf{v}$ es el vector cero. Puesto que Nul A se describe por medio de una condición que debe comprobarse para cada vector, se dice que el espacio nulo está definido *implícitamente*. En contraste, el espacio columna se define *explícitamente*, ya que los vectores en Col A se pueden construir (con combinaciones lineales) a partir de las columnas de A. Para crear una descripción explícita de Nul A, se resuelve la ecuación $A\mathbf{x} = \mathbf{0}$ y se escribe la solución en forma vectorial paramétrica. (Véase el ejemplo 6 que se presenta más adelante).[2]

Base para un subespacio

Como, por lo general, un subespacio contiene un número infinito de vectores, algunos problemas relacionados con subespacios se manejan mejor trabajando con un conjunto finito y pequeño de vectores que genere el subespacio. Cuanto menor sea el conjunto, será mejor. Es factible demostrar que el conjunto generador más pequeño posible debe ser linealmente independiente.

DEFINICIÓN Una **base** de un subespacio H de $\mathbb{R}^n$ es un conjunto linealmente independiente en H, que genera a H.

EJEMPLO 5 Las columnas de una matriz invertible de $n \times n$ forman una base para toda $\mathbb{R}^n$ ya que son linealmente independientes y generan $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible. Una matriz de este tipo es la matriz identidad de $n \times n$. Sus columnas se denotan mediante $\mathbf{e}_1, \ldots, \mathbf{e}_n$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es la base estándar para $\mathbb{R}^n$. Véase la figura 3. ■

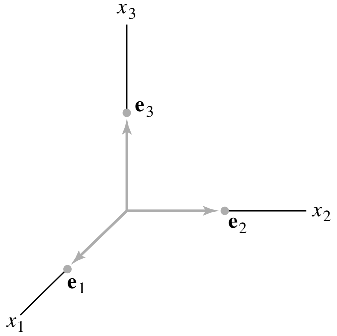


FIGURA 3
La base estándar para $\mathbb{R}^3$.

[2] El contraste entre Nul A y Col A se analiza con mayor detalle en la sección 4.2.

El siguiente ejemplo muestra que el procedimiento estándar para escribir el conjunto solución de $A\mathbf{x} = \mathbf{0}$ en la forma vectorial paramétrica en realidad identifica una base para Nul A. Este hecho se utilizará en todo el capítulo 5.

EJEMPLO 6 Encuentre una base para el espacio nulo de la matriz

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

SOLUCIÓN Primero, escriba la solución de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica:

$$\begin{bmatrix} A & \mathbf{0} \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 \qquad - x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

La solución general es $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, con x_2, x_4 y x_5 libres.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\uparrow \atop \mathbf{u}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\uparrow \atop \mathbf{v}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\uparrow \atop \mathbf{w}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \tag{1}$$

La ecuación (1) indica que Nul A coincide con el conjunto de todas las combinaciones lineales de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$. Es decir, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ genera Nul A. De hecho, esta construcción de $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ automáticamente las hace linealmente independientes, ya que (1) muestra que $\mathbf{0} = x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ solamente si todos los pesos x_2, x_4 y x_5 son cero. (Examine las entradas 2, 4 y 5 del vector $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$). Así, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es una *base* para Nul A. ∎

Encontrar una base para el espacio columna de una matriz, en realidad, representa menos trabajo que encontrar una base para el espacio nulo. Sin embargo, el método requiere cierta explicación. Comencemos con un caso sencillo.

EJEMPLO 7 Encuentre una base para el espacio columna de la matriz

$$B = \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN Denote las columnas de B mediante $\mathbf{b}_1, \ldots, \mathbf{b}_5$ y observe que $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$, y que $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. El hecho de que $\mathbf{b}_3$ y $\mathbf{b}_4$ sean combinaciones de las columnas pivote significa que cualquier combinación de $\mathbf{b}_1, \ldots, \mathbf{b}_5$ es en realidad solo una combinación de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. Efectivamente, si $\mathbf{v}$ es cualquier vector en Col B, por ejemplo,

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 + c_5\mathbf{b}_5$$

entonces, al sustituir $\mathbf{b}_3$ y $\mathbf{b}_4$, se puede escribir $\mathbf{v}$ en la forma

$$\mathbf{v} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3(-3\mathbf{b}_1 + 2\mathbf{b}_2) + c_4(5\mathbf{b}_1 - \mathbf{b}_2) + c_5\mathbf{b}_5$$

que es una combinación lineal de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. Así, $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5]$ genera Col B. También, $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$ son linealmente independientes, porque son columnas de una matriz identidad. Por lo tanto, las columnas pivote de B forman una base para Col B. ∎

La matriz B del ejemplo 7 está en forma escalonada reducida. Para manejar una matriz general A, recuerde que las relaciones de dependencia lineal entre las columnas de A se pueden expresar en la forma $A\mathbf{x} = \mathbf{0}$ para alguna $\mathbf{x}$. (Si algunas columnas no intervienen en una relación de dependencia particular, entonces las entradas correspondientes de $\mathbf{x}$ son cero). Cuando A se reduce por filas a la forma escalonada B, las columnas cambian en forma drástica, pero las ecuaciones $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen el mismo conjunto de soluciones. Es decir, las columnas de A tienen *exactamente las mismas relaciones de dependencia lineal* que las columnas de B.

EJEMPLO 8 Es posible comprobar que la matriz

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_5\,] = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$$

es equivalente por filas a la matriz B del ejemplo 7. Encuentre una base para Col A.

SOLUCIÓN A partir del ejemplo 7, las columnas pivote de A son las columnas 1, 2 y 5. También, $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$ y $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. Puesto que las operaciones de fila no afectan las relaciones de dependencia lineal entre las columnas de la matriz, deberíamos tener que

$$\mathbf{a}_3 = -3\mathbf{a}_1 + 2\mathbf{a}_2 \quad \text{y} \quad \mathbf{a}_4 = 5\mathbf{a}_1 - \mathbf{a}_2$$

Compruebe que ¡esto es cierto! Por el argumento del ejemplo 7, $\mathbf{a}_3$ y $\mathbf{a}_4$ no son necesarias para generar el espacio columna de A. Además, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ debe ser linealmente independiente, ya que cualquier relación de dependencia entre $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_5$ implicaría la misma relación de dependencia entre $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_5$. Puesto que $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ es linealmente independiente, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ también lo es y, por lo tanto, es una base para Col A. ■

El argumento del ejemplo 8 se puede adaptar para demostrar el siguiente teorema.

TEOREMA 13

Las columnas pivote de una matriz A forman una base para el espacio columna de A.

Advertencia: Tenga cuidado de utilizar las *columnas pivote de la misma* A para la base de Col A. Con frecuencia, las columnas de una forma escalonada B no están en el espacio columna de A. (Como se observa en los ejemplos 7 y 8, todas las columnas de B tienen ceros en sus últimas entradas y no pueden generar las columnas de A).

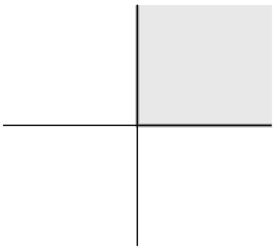
PROBLEMAS DE PRÁCTICA

1. Sea $A = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix}$. ¿Está $\mathbf{u}$ en Nul A? ¿Está $\mathbf{u}$ en Col A? Justifique sus respuestas.
2. Dada $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, encuentre un vector en Nul A y un vector en Col A.
3. Suponga que una matriz A de $n \times n$ es invertible. ¿Qué se puede decir acerca de Col A? ¿Y de Nul A?

2.8 EJERCICIOS

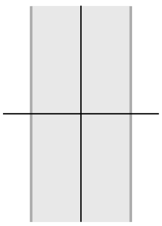
En los ejercicios 1 a 4 se muestran conjuntos en $\mathbb{R}^2$. Suponga que los conjuntos incluyen las líneas de frontera. En cada caso, dé una razón específica por la cual el conjunto H *no* es un subespacio de $\mathbb{R}^2$. (Por ejemplo, encuentre dos vectores en H cuya suma *no* esté en H, o encuentre un vector en H con un múltiplo escalar que no esté en H. Trace un dibujo).

1.

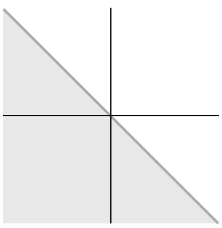


2.

3.



4.



5. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} -3 \\ -3 \\ 10 \end{bmatrix}$. Determine si $\mathbf{w}$ está en el subespacio de $\mathbb{R}^3$ generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

6. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ -4 \\ 5 \\ 7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ -3 \\ 6 \\ 5 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} -1 \\ -7 \\ -1 \\ 2 \end{bmatrix}$. Determine si $\mathbf{u}$ está en el subespacio de $\mathbb{R}^4$ generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

7. Sean

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 8 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -4 \\ 6 \\ -7 \end{bmatrix},$$

$$\mathbf{p} = \begin{bmatrix} 6 \\ -10 \\ 11 \end{bmatrix} \text{ y } A = [\,\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\,].$$

a) ¿Cuántos vectores hay en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

b) ¿Cuántos vectores hay en Col A?

c) ¿Está $\mathbf{p}$ en Col A? ¿Por qué?

8. Sean

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 0 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix},$$

y $\mathbf{p} = \begin{bmatrix} -6 \\ 1 \\ 17 \end{bmatrix}$. Determine si $\mathbf{p}$ está en Col A, donde $A = [\mathbf{v}_1\ \ \mathbf{v}_2\ \ \mathbf{v}_3]$.

9. Con A y $\mathbf{p}$ como en el ejercicio 7, determine si $\mathbf{p}$ está en Nul A.

10. Con $\mathbf{u} = \begin{bmatrix} -5 \\ 5 \\ 3 \end{bmatrix}$ y A como en el ejercicio 8, determine si $\mathbf{u}$ está en Nul A.

En los ejercicios 11 y 12, dé enteros p y q tales que Nul A sea un subespacio de $\mathbb{R}^p$ y Col A sea un subespacio de $\mathbb{R}^q$.

11. $A = \begin{bmatrix} 3 & 2 & 1 & -5 \\ -9 & -4 & 1 & 7 \\ 9 & 2 & -5 & 1 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -5 & -1 & 0 \\ 2 & 7 & 11 \\ 3 & 3 & 4 \end{bmatrix}$

13. Para A como en el ejercicio 11, encuentre un vector diferente de cero en Nul A y un vector diferente de cero en Col A.

14. Para A como en el ejercicio 12, encuentre un vector diferente de cero en Nul A y un vector diferente de cero en Col A.

Determine cuáles conjuntos de los ejercicios 15 a 20 son bases para $\mathbb{R}^2$ o $\mathbb{R}^3$. Justifique sus respuestas.

15. $\begin{bmatrix} 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 16 \\ -3 \end{bmatrix}$

16. $\begin{bmatrix} -2 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ -10 \end{bmatrix}$

17. $\begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 2 \end{bmatrix}$

18. $\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix}$

19. $\begin{bmatrix} 3 \\ -8 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -5 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ -6 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ 7 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 9 \end{bmatrix}$

En los ejercicios 21 y 22 marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Un subespacio de $\mathbb{R}^n$ es cualquier conjunto H tal que: **i.** el vector cero está en H, **ii.** $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u}+\mathbf{v}$ están en H, y **iii.** c es un escalar y $c\mathbf{u}$ está en H.

b) Si $\mathbf{v}_1,\ldots,\mathbf{v}_p$ están en $\mathbb{R}^n$, entonces Gen $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$ es el mismo que el espacio columna de la matriz $[\mathbf{v}_1 \ \cdots \ \mathbf{v}_p]$.

c) El conjunto de todas las soluciones de un sistema de m ecuaciones homogéneas con n incógnitas es un subespacio de $\mathbb{R}^m$.

d) Las columnas de una matriz invertible de $n \times n$ forman una base para $\mathbb{R}^n$.

e) Las operaciones de fila no afectan las relaciones de dependencia lineal entre las columnas de una matriz.

22. *a*) Un subconjunto H de $\mathbb{R}^n$ es un subespacio si el vector cero está en H.

b) Si B es la forma escalonada de una matriz A, entonces las columnas pivote de B forman una base para Col A.

c) Dados los vectores $\mathbf{v}_1,\ldots,\mathbf{v}_p$ en $\mathbb{R}^n$, el conjunto de todas las combinaciones lineales de estos vectores es un subespacio de $\mathbb{R}^n$.

d) Sea H un subespacio de $\mathbb{R}^n$. Si $\mathbf{x}$ está en H, y $\mathbf{y}$ está en $\mathbb{R}^n$, entonces $\mathbf{x}+\mathbf{y}$ está en H.

e) El espacio columna de una matriz A es el conjunto de soluciones de $A\mathbf{x} = \mathbf{b}$.

En los ejercicios 23 a 26 se presenta una matriz A y una forma escalonada de A. Encuentre una base para Col A y una base para Nul A.

23. $A = \begin{bmatrix} 4 & 5 & 9 & -2 \\ 6 & 5 & 1 & 12 \\ 3 & 4 & 8 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 6 & -5 \\ 0 & 1 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

24. $A = \begin{bmatrix} 3 & -6 & 9 & 0 \\ 2 & -4 & 7 & 2 \\ 3 & -6 & 6 & -6 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 5 & 4 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 1 & 4 & 8 & -3 & -7 \\ -1 & 2 & 7 & 3 & 4 \\ -2 & 2 & 9 & 5 & 5 \\ 3 & 6 & 9 & -5 & -2 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 4 & 8 & 0 & 5 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

26. $A = \begin{bmatrix} 3 & -1 & -3 & -1 & 8 \\ 3 & 1 & 3 & 0 & 2 \\ 0 & 3 & 9 & -1 & -4 \\ 6 & 3 & 9 & -2 & 6 \end{bmatrix}$

$\sim \begin{bmatrix} 3 & -1 & -3 & 0 & 6 \\ 0 & 2 & 6 & 0 & -4 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

27. Construya una matriz A de 3×3 y un vector $\mathbf{b}$ diferente de cero en forma tal que $\mathbf{b}$ esté en Col A, pero $\mathbf{b}$ no sea igual que alguna de las columnas de A.

28. Construya una matriz A de 3×3 y un vector $\mathbf{b}$ en forma tal que $\mathbf{b}$ *no* esté en Col A.

29. Construya una matriz A de 3×3 no nula y un vector $\mathbf{b}$ diferente de cero en forma tal que $\mathbf{b}$ esté en Nul A.

30. Suponga que las columnas de una matriz $A = [\mathbf{a}_1 \cdots \mathbf{a}_p]$ son linealmente independientes. Explique por qué $\{\mathbf{a}_1,\ldots,\mathbf{a}_p\}$ es una base para Col A.

En los ejercicios 31 a 36, responda de manera tan amplia como sea posible y justifique sus respuestas.

31. Suponga que F es una matriz de 5×5 cuyo espacio columna no es igual a $\mathbb{R}^5$. ¿Qué se puede decir acerca de Nul F?

32. Si B es una matriz de 7×7 y Col $B = \mathbb{R}^7$, ¿qué se puede decir acerca de las soluciones de las ecuaciones de la forma $B\mathbf{x} = \mathbf{b}$ para $\mathbf{b}$ en $\mathbb{R}^7$?

33. Si C es una matriz de 6×6 y Nul C es el subespacio cero, ¿qué se puede decir acerca de las soluciones a ecuaciones de la forma $C\mathbf{x} = \mathbf{b}$ para $\mathbf{b}$ en $\mathbb{R}^6$?

34. ¿Qué se puede decir acerca de la forma de una matriz A de $m \times n$ cuando las columnas de A forman una base para $\mathbb{R}^m$?

35. Si B es una matriz de 5×5 y Nul B *no* es el subespacio cero, ¿qué se puede decir acerca de Col B?

36. ¿Qué se puede decir acerca de Nul C cuando C es una matriz de 6×4 con columnas linealmente independientes?

[M] En los ejercicios 37 y 38, construya bases para el espacio columna y para el espacio nulo de la matriz A dada. Justifique su trabajo.

37. $A = \begin{bmatrix} 3 & -5 & 0 & -1 & 3 \\ -7 & 9 & -4 & 9 & -11 \\ -5 & 7 & -2 & 5 & -7 \\ 3 & -7 & -3 & 4 & 0 \end{bmatrix}$

38. $A = \begin{bmatrix} 5 & 3 & 2 & -6 & -8 \\ 4 & 1 & 3 & -8 & -7 \\ 5 & 1 & 4 & 5 & 19 \\ -7 & -5 & -2 & 8 & 5 \end{bmatrix}$

WEB Espacio columna y espacio nulo

WEB Una base para Col A

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Para determinar si **u** está en Nul A, simplemente calcule

$$A\mathbf{u} = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix} \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

El resultado indica que **u** está en Nul A. Para decidir si **u** está en Col A se requiere más trabajo. Reduzca la matriz aumentada $[A \quad \mathbf{u}]$ a la forma escalonada para determinar si la ecuación $A\mathbf{x} = \mathbf{u}$ es consistente:

$$\begin{bmatrix} 1 & -1 & 5 & -7 \\ 2 & 0 & 7 & 3 \\ -3 & -5 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & -8 & 12 & -19 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

La ecuación $A\mathbf{x} = \mathbf{u}$ no tiene solución, por lo tanto, **u** no está en Col A.

2. En contraste con el problema de práctica 1, encontrar un vector en Nul A requiere más trabajo que probar si un vector específico está en Nul A. Sin embargo, como A ya está en forma escalonada reducida, la ecuación $A\mathbf{x} = \mathbf{0}$ indica que si $\mathbf{x} = (x_1, x_2, x_3)$, entonces $x_2 = 0$, $x_3 = 0$, y x_1 es una variable libre. Por lo tanto, una base para Nul A es $\mathbf{v} = (1, 0, 0)$. Encontrar solo un vector en Col A es trivial, puesto que cada columna de A está en Col A. En este caso particular, el mismo vector **v** se encuentra tanto en Nul A como en Col A. Para la mayoría de las matrices de $n \times n$, el vector cero de $\mathbb{R}^n$ es el único vector que se encuentra tanto en Nul A como en Col A.

3. Si A es invertible, entonces las columnas de A generan $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible. Por definición, las columnas de cualquier matriz siempre generan el espacio columna, así que, en este caso, Col A es toda de $\mathbb{R}^n$. En forma simbólica, Col $A = \mathbb{R}^n$. Además, como A es invertible, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial. Esto significa que Nul A es el subespacio cero. En forma simbólica, Nul $A = \{\mathbf{0}\}$.

2.9 DIMENSIÓN Y RANGO

En esta sección se continúa el análisis de los subespacios y las bases para subespacios comenzando con el concepto de un sistema coordenado. La definición y el ejemplo presentados a continuación pretenden que un término nuevo y útil, *dimensión*, parezca bastante natural, al menos para los subespacios de $\mathbb{R}^3$.

Sistemas coordenados

La razón principal para seleccionar la base de un subespacio H, en lugar de simplemente un conjunto generador, es que cada vector en H se puede escribir solo de una manera como combinación lineal de los vectores de la base. Para ver por qué, suponga que $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base de H, y que un vector **x** en H se puede generar de dos maneras, por ejemplo,

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p \quad \text{y} \quad \mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_p\mathbf{b}_p \tag{1}$$

Después, al restar se obtiene

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_p - d_p)\mathbf{b}_p \tag{2}$$

Como $\mathcal{B}$ es linealmente independiente, los pesos en (2) deben ser todos cero. Es decir, $c_j = d_j$ para $1 \leq j \leq p$, lo que indica que las dos representaciones en (1), en realidad, son iguales.

DEFINICIÓN

Suponga, que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es la base de un subespacio H. Para cada $\mathbf{x}$ en H, **las coordenadas de x respecto de la base** $\mathcal{B}$ son los pesos $c_1, \ldots, c_p$ tales que $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p$, y el vector en $\mathbb{R}^p$

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

se llama **vector de coordenadas de x (respecto de** $\mathcal{B}$**)** o **vector de** $\mathcal{B}$**-coordenadas de x**.[1]

EJEMPLO 1 Sea $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathcal{B}$ es una base de $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ porque $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes. Determine si $\mathbf{x}$ está en H y, si lo está, encuentre el vector de coordenadas de $\mathbf{x}$ respecto de $\mathcal{B}$.

SOLUCIÓN Si $\mathbf{x}$ está en H, entonces la siguiente ecuación vectorial es consistente:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

Los escalares c_1 y c_2, si existen, son las $\mathcal{B}$-coordenadas de $\mathbf{x}$. Usando operaciones de fila, se tiene que

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Por lo tanto, $c_1 = 2$, $c_2 = 3$ y $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. La base $\mathcal{B}$ determina un "sistema de coordenadas" en H, lo que puede visualizarse por medio de la malla que se ilustra en la figura 1. ■

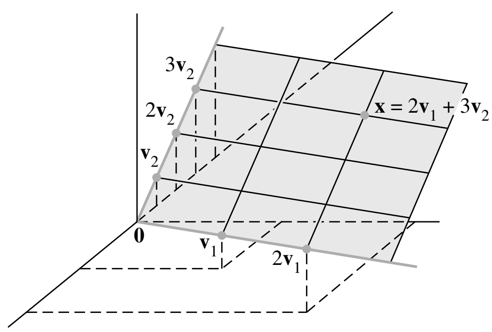


FIGURA 1 Un sistema de coordenadas sobre un plano H en $\mathbb{R}^3$.

[1] Es importante que los elementos de $\mathcal{B}$ estén numerados, ya que las entradas de $[\mathbf{x}]_{\mathcal{B}}$ dependen del orden de los vectores en $\mathcal{B}$.

Observe que a pesar de que los puntos de H también se encuentran en $\mathbb{R}^3$, están completamente determinados por sus vectores de coordenadas, los cuales pertenecen a $\mathbb{R}^2$. La malla en el plano de la figura 1 hace que H "se vea" como $\mathbb{R}^2$. La correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una correspondencia uno a uno entre H y $\mathbb{R}^2$ que conserva las combinaciones lineales. A una correspondencia de este tipo se le llama *isomorfismo*, y se dice que H es *isomorfo* a $\mathbb{R}^2$.

En general, si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base para H, entonces el mapeo $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una correspondencia uno a uno que permite a H verse y actuar igual que R^p (aunque los propios vectores de H puedan tener más de p entradas). (En la sección 4.4 se presentan más detalles).

Dimensión de un subespacio

Es posible demostrar que si un subespacio H tiene una base de p vectores, entonces cualquier base de H debe consistir en exactamente p vectores. (Véase los ejercicios 27 y 28). Por lo tanto, la siguiente definición tiene sentido.

DEFINICIÓN

La **dimensión** de un subespacio H diferente de cero, que se denota mediante dim H, es el número de vectores en cualquier base para H. La dimensión del subespacio cero $\{\mathbf{0}\}$ es, por definición, cero.[2]

El espacio $\mathbb{R}^n$ tiene dimensión n. Cada base para $\mathbb{R}^n$ consiste en n vectores. Un plano a través de $\mathbf{0}$ en $\mathbb{R}^3$ es bidimensional, y una recta que pasa a través de $\mathbf{0}$ es unidimensional.

EJEMPLO 2 Recuerde que el espacio nulo de la matriz A vista en el ejemplo 6, sección 2.8, tenía una base de tres vectores. Así que la dimensión de Nul A en este caso es 3. Observe cómo cada vector básico corresponde a una variable libre en la ecuación $A\mathbf{x} = \mathbf{0}$. La construcción realizada aquí siempre produce una base de este modo. Así, para encontrar la dimensión de Nul A, basta con identificar y contar el número de variables libres en $A\mathbf{x} = \mathbf{0}$. ■

DEFINICIÓN

El **rango** de una matriz A, que se denota como rango A, es la dimensión del espacio columna de A.

Puesto que las columnas pivote de A forman una base para Col A, el rango de A es simplemente el número de columnas pivote en A.

EJEMPLO 3 Determine el rango de la matriz

$$A = \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 4 & 7 & -4 & -3 & 9 \\ 6 & 9 & -5 & 2 & 4 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix}$$

SOLUCIÓN Reduzca A a la forma escalonada:

$$A \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & -6 & 4 & 14 & -20 \\ 0 & -9 & 6 & 5 & -6 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 2 & 5 & -3 & -4 & 8 \\ 0 & -3 & 2 & 5 & -7 \\ 0 & 0 & 0 & 4 & -6 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Columnas pivote

La matriz A tiene 3 columnas pivote, así que rango $A = 3$. ■

[2] El subespacio cero *no* tiene base (porque el vector cero forma, por sí mismo, un conjunto linealmente dependiente).

La reducción por filas del ejemplo 3 revela que hay dos variables libres en $A\mathbf{x} = \mathbf{0}$, porque dos de las cinco columnas de A *no* son columnas pivote. (Las columnas que no son pivote corresponden a las variables libres de $A\mathbf{x} = \mathbf{0}$). Como el número de columnas pivote más el número de columnas que no son pivote es exactamente el número de columnas, las dimensiones de Col A y Nul A guardan la siguiente relación útil. (Para más detalles, véase el teorema del rango en la sección 4.6).

TEOREMA 14

El teorema del rango

Si una matriz A tiene n columnas, entonces rango A + dim Nul $A = n$.

El siguiente teorema es importante para las aplicaciones y se necesitará en los capítulos 5 y 6. El teorema (demostrado en la sección 4.5) es verdaderamente superior, si se piensa en un subespacio p-dimensional como isomorfo a $\mathbb{R}^p$. El teorema de la matriz invertible establece que p vectores de $\mathbb{R}^p$ son linealmente independientes si y solo si también generan a $\mathbb{R}^p$.

TEOREMA 15

El teorema de la base

Sea H un subespacio p-dimensional de $\mathbb{R}^n$. Cualquier conjunto linealmente independiente de exactamente p elementos en H de forma automática es una base de H. Además, cualquier conjunto de p elementos de H que genere a H es automáticamente una base para H.

Rango y el teorema de la matriz invertible

Los diversos conceptos de espacio vectorial asociados con una matriz proporcionan varios enunciados más para el teorema de la matriz invertible. Estos enunciados se presentan como una continuación del teorema original presentado en la sección 2.3.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. En tal caso, cada uno de los siguientes enunciados es equivalente al enunciado de que A es una matriz invertible.

m) Las columnas de A forman una base de $\mathbb{R}^n$.

n) Col $A = \mathbb{R}^n$

o) dim Col $A = n$

p) rango $A = n$

q) Nul $A = \{\mathbf{0}\}$

r) dim Nul $A = 0$

DEMOSTRACIÓN El enunciado *m*) es, por lógica, equivalente a los enunciados *e*) y *h*) que consideran independencia lineal y generación. Los otros cinco enunciados se vinculan a los primeros del teorema por medio de la siguiente cadena de implicaciones casi triviales.

$$g) \Rightarrow n) \Rightarrow o) \Rightarrow p) \Rightarrow r) \Rightarrow q) \Rightarrow d)$$

El enunciado *g*), que establece que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$, implica al enunciado *n*), porque Col A es precisamente el conjunto de todas las $\mathbf{b}$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ sea consistente. Las implicaciones $n) \Rightarrow o) \Rightarrow p)$ se desprenden de las definiciones de *dimensión* y *rango*. Si el rango de A es n, el número de columnas de A, entonces dim Nul $A = 0$, de acuerdo con el teorema del rango, y así Nul $A = \{\mathbf{0}\}$. Por lo tanto,

$p) \Rightarrow r) \Rightarrow q)$. Además, el enunciado q) implica que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene únicamente la solución trivial, que es el enunciado d). Puesto que ya se sabe que los enunciados d) y g) son equivalentes al enunciado de que A es invertible, la prueba está completa. ■

NOTAS NUMÉRICAS

Muchos de los algoritmos analizados en este libro resultan útiles para entender conceptos y realizar a mano cálculos sencillos. Sin embargo, con frecuencia los algoritmos no son aplicables a problemas de gran escala en la vida real.

Un buen ejemplo de lo anterior es la determinación del rango. Tal vez parezca sencillo reducir una matriz a su forma escalonada y contar los pivotes. Pero, a menos que se realicen cálculos matemáticos precisos sobre una matriz cuyas entradas estén especificadas de manera exacta, las operaciones de fila pueden cambiar el rango aparente de una matriz. Por ejemplo, si el valor de x en la matriz $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ no se almacena exactamente como 7 en una computadora, entonces el rango puede ser 1 o 2, dependiendo de si la computadora considera a $x - 7$ como cero.

WEB

En las aplicaciones prácticas, es frecuente que el rango efectivo de una matriz A se determine a partir de la descomposición del valor singular de A, que se estudiará en la sección 7.4.

PROBLEMAS DE PRÁCTICA

1. Determine la dimensión del subespacio H de $\mathbb{R}^3$ generado por los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. (Primero, encuentre una base para H).

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 6 \\ -7 \end{bmatrix}$$

2. Considere la base

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ .2 \end{bmatrix}, \begin{bmatrix} .2 \\ 1 \end{bmatrix} \right\}$$

para $\mathbb{R}^2$. Si $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, ¿qué es $\mathbf{x}$?

3. ¿Podría $\mathbb{R}^3$ contener un subespacio cuatridimensional (o tetradimensional)? Explique su respuesta.

2.9 EJERCICIOS

En los ejercicios 1 y 2, encuentre el vector $\mathbf{x}$ determinado por el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ y la base $\mathcal{B}$ dados. Ilustre cada respuesta con una figura, como en la solución al problema de práctica 2.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

En los ejercicios 3 a 6, el vector $\mathbf{x}$ está en un subespacio H con una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$.

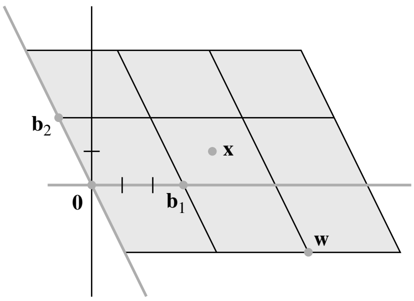
3. $\mathbf{b}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 0 \\ 7 \end{bmatrix}$

4. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 9 \end{bmatrix}$

5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 4 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ -7 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 2 \\ 9 \\ -7 \end{bmatrix}$

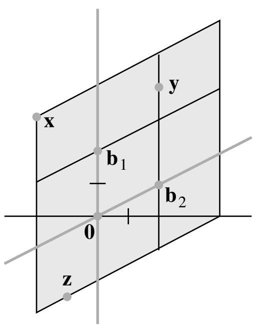
6. $\mathbf{b}_1 = \begin{bmatrix} -3 \\ 2 \\ -4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 7 \\ -3 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}$

7. Sean $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use la figura para estimar $[\mathbf{w}]_{\mathcal{B}}$ y $[\mathbf{x}]_{\mathcal{B}}$. Confirme su estimación de $[\mathbf{x}]_{\mathcal{B}}$ usándola junto con $\{\mathbf{b}_1, \mathbf{b}_2\}$ para calcular $\mathbf{x}$.



8. Sean $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} -2 \\ 3 \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} 2 \\ 4 \end{bmatrix},$

$\mathbf{z} = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use la figura para estimar $[\mathbf{x}]_{\mathcal{B}}$, $[\mathbf{y}]_{\mathcal{B}}$ y $[\mathbf{z}]_{\mathcal{B}}$. Confirme sus estimaciones de $[\mathbf{y}]_{\mathcal{B}}$ y $[\mathbf{z}]_{\mathcal{B}}$ usándolas junto con $\{\mathbf{b}_1, \mathbf{b}_2)$ para calcular $\mathbf{y}$ y $\mathbf{z}$.



En los ejercicios 9 a 12 se presenta una matriz A y una forma escalonada de A. Encuentre bases para Col A y Nul A, y después establezca las dimensiones de estos subespacios.

9. $A = \begin{bmatrix} 1 & 3 & 2 & -6 \\ 3 & 9 & 1 & 5 \\ 2 & 6 & -1 & 9 \\ 5 & 15 & 0 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 5 & -7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & -2 & -1 & 5 & 4 \\ 2 & -1 & 1 & 5 & 6 \\ -2 & 0 & -2 & 1 & -6 \\ 3 & 1 & 4 & 1 & 5 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & -2 & -1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

11. $A = \begin{bmatrix} 2 & 4 & -5 & 2 & -3 \\ 3 & 6 & -8 & 3 & -5 \\ 0 & 0 & 9 & 0 & 9 \\ -3 & -6 & -7 & -3 & -10 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 2 & -5 & 1 & -4 \\ 0 & 0 & 5 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 2 & -4 & 4 & 6 \\ 5 & 1 & -9 & 2 & 10 \\ 4 & 6 & -9 & 12 & 15 \\ 3 & 4 & -5 & 8 & 9 \end{bmatrix}$

$\sim \begin{bmatrix} 1 & 2 & 8 & 4 & -6 \\ 0 & 2 & 3 & 4 & -1 \\ 0 & 0 & 5 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 13 y 14, encuentre una base para el subespacio que generan los vectores dados. ¿Cuál es la dimensión del subespacio?

13. $\begin{bmatrix} 1 \\ -3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \\ -6 \\ 12 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -3 \\ 7 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ -7 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ -7 \\ 6 \\ -9 \end{bmatrix}$

15. Suponga una matriz A de 4×6 tiene cuatro columnas pivote. ¿Es Col $A = \mathbb{R}^4$? ¿Es Nul $A = \mathbb{R}^2$? Explique sus respuestas.

16. Suponga una matriz A de 4×7 tiene tres columnas pivote. ¿Es Col $A = \mathbb{R}^3$? ¿Cuál es la dimensión de Nul A? Explique sus respuestas.

En los ejercicios 17 y 18, marque cada enunciado como verdadero o falso. Justifique sus respuestas. Aquí A es una matriz de $m \times n$.

17. *a*) Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base para un subespacio H, y si $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, entonces $c_1, \ldots, c_p$ son las coordenadas de $\mathbf{x}$ respecto de la base $\mathcal{B}$.

b) Cada recta en $\mathbb{R}^n$ es un subespacio unidimensional de $\mathbb{R}^n$.

c) La dimensión de Col A es el número de columnas pivote de A.

d) Las dimensiones de Col A y Nul A suman el número de columnas de A.

e) Si un conjunto de p vectores generan un subespacio p-dimensional H de $\mathbb{R}^n$, entonces estos vectores forman una base para H.

18. *a*) Si $\mathcal{B}$ es una base para un subespacio H, entonces cada vector en H se puede escribir solamente de una forma como combinación lineal de los vectores en $\mathcal{B}$.

b) La dimensión de Nul A es el número de variables en la ecuación $A\mathbf{x} = \mathbf{0}$.

c) La dimensión del espacio columna de A es rango A.

d) Si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base del subespacio H de $\mathbb{R}^n$, entonces la correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ hace que H sea vea y actúe como $\mathbb{R}^p$.

e) Si H es un subespacio p-dimensional de R^n, entonces un conjunto linealmente independiente de p vectores en H es una base para H.

En los ejercicios 19 a 24 justifique cada respuesta o construcción.

19. Si el subespacio de todas las soluciones de $A\mathbf{x} = \mathbf{0}$ tiene una base que consiste en tres vectores, y si A es una matriz de 5×7, ¿cuál es el rango de A?

20. ¿Cuál es el rango de una matriz de 6×8 cuyo espacio nulo es tridimensional?

21. Si el rango de una matriz A de 9×8 es 7, ¿cuál es la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$?

22. Demuestre que un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_5\}$ en $\mathbb{R}^n$ es linealmente dependiente si dim Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_5\} = 4$.

23. Si es posible, construya una matriz A de 3×5 tal que dim Nul $A = 3$ y dim Col $A = 2$.

24. Construya una matriz de 3×4 con rango 1.

25. Sea A una matriz de $n \times p$ cuyo espacio columna es p-dimensional. Explique por qué las columnas de A deben ser linealmente independientes.

26. Suponga que las columnas 1, 3, 4, 5 y 7 de una matriz A son linealmente independientes (pero no son necesariamente columnas pivote), y que el rango de A es 5. Explique por qué las cinco columnas mencionadas deben ser una base para el espacio columna de A.

27. Suponga que los vectores $\mathbf{b}_1, \ldots, \mathbf{b}_p$ generan un subespacio W, y sea $\{\mathbf{a}_1, \ldots, \mathbf{a}_q\}$ cualquier conjunto en W que contenga más de p vectores. Complete los detalles del siguiente argumento para demostrar que $\{\mathbf{a}_1, \ldots, \mathbf{a}_q\}$ debe ser linealmente dependiente. Primero, sea $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$ y $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_q]$.

a) Explique por qué para cada vector $\mathbf{a}_j$, existe un vector $\mathbf{c}_j$ en $\mathbb{R}^p$ tal que $\mathbf{a}_j = B\mathbf{c}_j$.

b) Sea $C = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_q]$. Explique por qué existe un vector diferente de cero tal que $C\mathbf{u} = \mathbf{0}$.

c) Utilice B y C para demostrar que $A\mathbf{u} = \mathbf{0}$. Esto muestra que las columnas de A son linealmente dependientes.

28. Use el ejercicio 27 para demostrar que si $\mathcal{A}$ y $\mathcal{B}$ son bases para un subespacio W de $\mathbb{R}^n$, entonces $\mathcal{A}$ no puede contener más vectores que $\mathcal{B}$ y, a la inversa, que $\mathcal{B}$ no puede contener más vectores que $\mathcal{A}$.

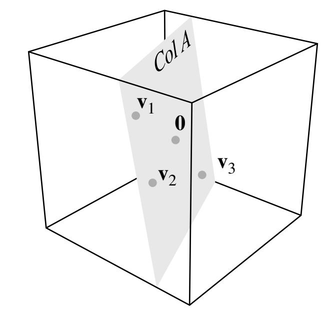
29. **[M]** Sean $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que $\mathbf{x}$ está en H, y encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$, cuando

$$\mathbf{v}_1 = \begin{bmatrix} 15 \\ -5 \\ 12 \\ 7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 14 \\ -10 \\ 13 \\ 17 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 16 \\ 0 \\ 11 \\ -3 \end{bmatrix}$$

30. **[M]** Sean $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Demuestre que $\mathcal{B}$ es una base para H y que $\mathbf{x}$ está en H, y encuentre el vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$, cuando

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 3 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ 0 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 4 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 11 \\ -2 \\ 17 \\ -8 \end{bmatrix}$$

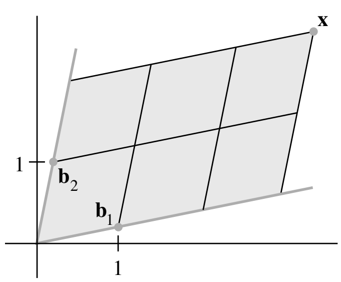
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA



1. Construya $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ de manera que el subespacio generado por $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sea el espacio columna de A. Las columnas pivote de A proporcionan una base para este espacio.

$$A = \begin{bmatrix} 2 & 3 & -1 \\ -8 & -7 & 6 \\ 6 & -1 & -7 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & -10 & -4 \end{bmatrix} \sim \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Las primeras dos columnas de A son columnas pivote y forman una base para H. Por lo tanto, dim $H = 2$.



2. Si $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, entonces $\mathbf{x}$ se forma a partir de una combinación lineal de los vectores básicos usando los pesos 3 y 2:

$$\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ .2 \end{bmatrix} + 2\begin{bmatrix} .2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 2.6 \end{bmatrix}$$

La base $\{\mathbf{b}_1, \mathbf{b}_2\}$ determina un *sistema de coordenadas* para $\mathbb{R}^2$, que se ilustra con la malla de la figura. Observe cómo $\mathbf{x}$ tiene 3 unidades en la dirección $\mathbf{b}_1$ y 2 unidades en la dirección $\mathbf{b}_2$.

3. Un subespacio cuatridimensional contendría una base de cuatro vectores linealmente independientes. Esto es imposible en $\mathbb{R}^3$. Como cualquier conjunto linealmente independiente en $\mathbb{R}^3$ no contiene más de tres vectores, cualquier subespacio de $\mathbb{R}^3$ tiene una dimensión no mayor a 3. El propio espacio $\mathbb{R}^3$ es el único subespacio tridimensional de $\mathbb{R}^3$. Los otros subespacios de $\mathbb{R}^3$ tienen dimensión 2, 1 o 0.

CAPÍTULO 2 EJERCICIOS COMPLEMENTARIOS

1. Suponga que las matrices mencionadas en los siguientes enunciados tienen los tamaños adecuados. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

a) Si A y B son de $m \times n$, entonces tanto AB^T como A^TB están definidas.

b) Si $AB = C$ y C tiene dos columnas, entonces A tiene dos columnas.

c) Al multiplicar por la izquierda una matriz B por una matriz diagonal A, con entradas distintas de cero en la diagonal, se modifica la escala de las filas de B.

d) Si $BC = BD$, entonces $C = D$.

e) Si $AC = 0$, entonces $A = 0$ o $C = 0$.

f) Si A y B son de $n \times n$, entonces $(A + B)(A - B) = A^2 - B^2$.

g) Una matriz elemental de $n \times n$ tiene n o $n + 1$ entradas diferentes de cero.

h) La transpuesta de una matriz elemental es una matriz elemental.

i) Una matriz elemental debe ser cuadrada.

j) Toda matriz cuadrada es un producto de matrices elementales.

k) Si A es una matriz de 3×3 con tres posiciones pivote, existen matrices elementales $E_1, \ldots, E_p$ tales que $E_p \cdots E_1 A = I$.

l) Si $AB = I$, entonces A es invertible.

m) Si A y B son cuadradas e invertibles, entonces AB es invertible, y $(AB)^{-1} = A^{-1}B^{-1}$

n) Si $AB = BA$ y A es invertible, entonces $A^{-1}B = BA^{-1}$.

o) Si A es invertible y $r \neq 0$, entonces $(rA)^{-1} = rA^{-1}$.

p) Si A es una matriz de 3×3 y la ecuación $A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ tiene una sola solución, entonces A es invertible.

2. Encuentre la matriz C cuya inversa es $C^{-1} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$.

3. Sea $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Demuestre que $A^3 = 0$. Utilice álgebra de matrices para calcular el producto $(I - A)(I + A + A^2)$.

4. Suponga que $A^n = 0$ para alguna $n > 1$. Encuentre una inversa para $I - A$.

5. Suponga que una matriz A de $n \times n$ satisface la ecuación $A^2 - 2A + I = 0$. Demuestre que $A^3 = 3A - 2I$, y que $A^4 = 4A - 3I$.

6. Sea $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Estas son *matrices de espín de Pauli* y se usan en el estudio del espín de electrones en mecánica cuántica. Demuestre que $A^2 = I$, $B^2 = I$ y $AB = -BA$. Las matrices del tipo $AB = -BA$ se llaman *anticonmutativas*.

7. Sea $A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 4 & 11 \\ 1 & 2 & 5 \end{bmatrix}$ y $B = \begin{bmatrix} -3 & 5 \\ 1 & 5 \\ 3 & 4 \end{bmatrix}$. Calcule $A^{-1}B$ sin calcular A^{-1}. [*Sugerencia:* $A^{-1}B$ es la solución de la ecuación $AX = B$].

8. Encuentre una matriz A tal que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ mapee $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ en $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectivamente. [*Sugerencia:* Escriba una ecuación de matrices que implique a A, y despeje A].

9. Suponga que $AB = \begin{bmatrix} 5 & 4 \\ -2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$. Encuentre A.

10. Suponga que A es invertible. Explique por qué A^TA también es invertible. Después demuestre que $A^{-1} = (A^TA)^{-1}A^T$.

11. Sean $x_1, \ldots, x_n$ números fijos. La siguiente matriz, llamada una *matriz de Vandermonde*, se presenta en aplicaciones como procesamiento de señales, códigos correctores de errores e interpolación de polinomios.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

A partir de $\mathbf{y} = (y_1, \ldots, y_n)$ en $\mathbb{R}^n$, suponga que $\mathbf{c} = (c_0, \ldots, c_{n-1})$ en $\mathbb{R}^n$ satisface $V\mathbf{c} = \mathbf{y}$, y defina el polinomio

$p(t) = c_0 + c_1t + c_2t^2 + \cdots + c_{n-1}t^{n-1}$

a) Demuestre que $p(x_1) = y_1, \ldots, p(x_n) = y_n$. Denominamos a $p(t)$ un *polinomio de interpolación para los puntos* $(x_1, y_1), \ldots, (x_n, y_n)$ porque la gráfica de $p(t)$ pasa por estos puntos.

b) Suponga que $x_1, \ldots, x_n$ son números distintos. Demuestre que las columnas de V son linealmente independientes. [*Sugerencia:* Piense en cuántos ceros puede tener un polinomio de grado $n - 1$].

c) Demuestre que: "Si $x_1, \ldots, x_n$, son números distintos $y_1, \ldots, y_n$ son números arbitrarios, entonces hay un polinomio de interpolación de grado $\leq n - 1$ para $(x_1, y_1), \ldots, (x_n, y_n)$".

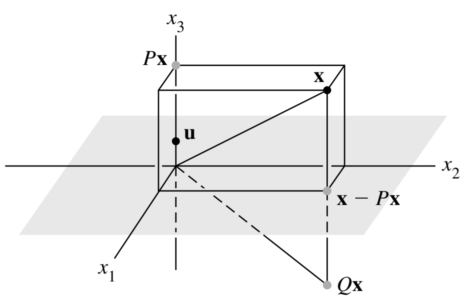
12. Sea $A = LU$, donde L es una matriz triangular inferior invertible y U es triangular superior. Explique por qué la primera columna de A es un múltiplo de la primera columna de L. ¿Cómo se relaciona la segunda columna de A con las columnas de L?

13. Dada $\mathbf{u}$ en $\mathbb{R}^n$ con $\mathbf{u}^T\mathbf{u} = 1$, sea $P = \mathbf{u}\mathbf{u}^T$ (un producto exterior) y $Q = 1 - 2P$. Justifique los enunciados *a*), *b*) y *c*).

a) $P^2 = P$ *b*) $P^T = P$ *c*) $Q^2 = I$

La transformación $\mathbf{x} \mapsto P\mathbf{x}$ es una *proyección*, y $\mathbf{x} \mapsto Q\mathbf{x}$ se llama *reflexión de Householder*. Dichas reflexiones se usan en programas de computadora para crear múltiples ceros en un vector (por lo general, una columna de una matriz).

14. Sean $\mathbf{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$. Determine P y Q como en el ejercicio 13, y calcule $P\mathbf{x}$ y $Q\mathbf{x}$. La figura muestra que $Q\mathbf{x}$ es la reflexión de $\mathbf{x}$ a través del plano x_1x_2.



Una reflexión de Householder a través del plano $x_3 = 0$.

15. Suponga que $C = E_3E_2E_1B$, donde E_1, E_2 y E_3 son matrices elementales. Explique por qué C es equivalente por filas a B.

16. Sea A una matriz singular de $n \times n$. Describa cómo se puede construir una matriz B de $n \times n$ no nula tal que $AB = 0$.

17. Sean A una matriz de 6×4 y B una matriz de 4×6. Demuestre que la matriz AB de 6×6 no puede ser invertible.

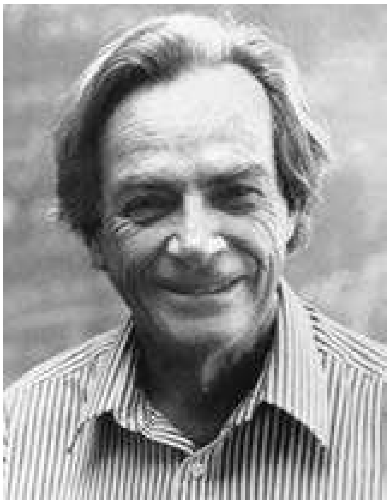
18. Suponga que A es una matriz de 5×3 y que existe una matriz C de 3×5 tal que $CA = I_3$. Suponga además que para alguna $\mathbf{b}$ dada en $\mathbb{R}^5$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene por lo menos una solución. Demuestre que esta solución es única.

19. **[M]** Ciertos sistemas dinámicos se pueden estudiar examinando las potencias de una matriz, como las que se presentan a continuación. Determine qué ocurre a A^k y B^k conforme k se incrementa (por ejemplo, pruebe con $k = 2, \ldots, 16$). Trate de identificar qué tienen de especial A y B. Investigue potencias grandes de otras matrices de este tipo y haga una conjetura acerca de dichas matrices.

$$A = \begin{bmatrix} .4 & .2 & .3 \\ .3 & .6 & .3 \\ .3 & .2 & .4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & .2 & .3 \\ .1 & .6 & .3 \\ .9 & .2 & .4 \end{bmatrix}$$

20. **[M]** Sea A_n una matriz de $n \times n$ con ceros en la diagonal principal y números 1 en el resto. Calcule A_n^{-1} para $n = 4$, 5 y 6, y haga una conjetura acerca de la forma general de A_n^{-1} para valores más grandes de n.

3 Determinantes



WEB

EJEMPLO INTRODUCTORIO

Trayectorias aleatorias y distorsión

En su libro autobiográfico *¿Está usted de broma, Sr. Feynman?*, el Premio Nobel de Física 1965, Richard P. Feynman, comenta que en su estancia universitaria de posgrado en Princeton solía observar a las hormigas. Estudió su comportamiento al proporcionarles un "transbordador" de pedazos de papel que podía trasladarlas hacia el terrón de azúcar que colgaba de una cuerda, donde las hormigas jamás podrían encontrarlo accidentalmente. Cuando una hormiga se paraba sobre el papel, entonces Feynman la transportaba hacia la comida y la traía de regreso al punto inicial. Después de que las hormigas aprendieron a utilizar el transbordador de papel, Feynman reubicó el punto de aterrizaje de retorno. Esto pronto generó confusión en la colonia de hormigas, lo que indicaba que el "aprendizaje" de las hormigas había consistido en crear y seguir rutas. Feynman confirmó esta conjetura colocando placas de vidrio sobre el piso. Una vez que las hormigas establecieron rutas sobre los trozos de vidrio, Feynman los reacomodó y, por consiguiente, las rutas que estos indicaban. Las hormigas siguieron las rutas reposicionadas y así Feynman podía dirigir a las hormigas adonde él quisiera.

Supongamos que Feynman hubiera deseado realizar investigaciones adicionales utilizando un globo, fabricado con una malla de alambre; las hormigas deberían seguir cada alambre y elegir entre dirigirse a la izquierda o a la derecha en cada intersección. Si varias hormigas y un número igual de fuentes de alimento se colocan sobre el globo, ¿qué tan probable sería que cada hormiga encontrara su propia fuente de alimento en lugar de utilizar la ruta de otra hormiga y seguirla hasta una fuente compartida?[1]

Para registrar las rutas reales de las hormigas y comunicar los resultados a otros, es conveniente utilizar un mapa rectangular del globo. Existen muchas maneras de trazar dichos mapas. Una manera sencilla es utilizar la latitud y longitud sobre el globo como coordenadas x y y en el mapa. Como ocurre con todos los mapas, el resultado no sería una representación totalmente fiel del globo. Los detalles cerca del "ecuador" se ven prácticamente iguales tanto sobre el globo como en el mapa, pero las regiones cercanas a los "polos" del globo están distorsionadas. Las imágenes de regiones polares son mucho más grandes que las imágenes de regiones ecuatoriales de tamaño similar. Para ajustarse a sus alrededores en el mapa, la imagen de una hormiga cerca de uno de los polos debería ser más grande que la de una cercana al ecuador. ¿Cuánto más de grande?

De manera sorprendente, los problemas de la trayectoria de las hormigas y la distorsión del área se contestan mejor utilizando determinantes, el tema de este capítulo. De hecho, el determinante tiene tantos usos que un resumen de sus aplicaciones a principios del siglo XX ocupó cuatro volúmenes del tratado que escribió Thomas Muir. Ante la trascendencia y el tamaño crecientes de las matrices en las aplicaciones modernas, muchos usos de los determinantes que antes eran importantes ahora ya no lo son. Sin embargo, los determinantes aún desempeñan un papel importante.

[1] La solución al problema de la trayectoria de las hormigas (y otras dos aplicaciones) se puede encontrar en el artículo de Arthur Benjamin y Naomi Cameron, publicado en la edición de *Mathematical Monthly*, de junio 2005.

Además de introducir el tema de determinantes en la sección 3.1, este capítulo presenta dos ideas importantes. La sección 3.2 deduce, para una matriz cuadrada, un criterio de invertibilidad que desempeña un papel importante en el capítulo 5. La sección 3.3 muestra cómo el determinante mide cuánto cambia una transformación lineal al área de una figura. Cuando esta técnica se aplica localmente, entonces se responde a la pregunta de la tasa de expansión de un mapa cerca de los polos. Esta idea desempeña un papel fundamental en cálculo multivariado en la forma del jacobiano.

3.1 INTRODUCCIÓN A LOS DETERMINANTES

Recuerde de la sección 2.2 que una matriz de 2×2 es invertible si y solo si su determinante es diferente de cero. Para extender este útil resultado a matrices más grandes, se necesita una definición para el determinante de una matriz de $n \times n$. Se puede descubrir la definición para el caso 3×3 observando qué ocurre cuando una matriz invertible A de 3×3 se reduce por filas.

Considere $A = [a_{ij}]$ con $a_{11} \neq 0$. Si la segunda y tercera filas de A se multiplican por a_{11}, y luego se restan múltiplos adecuados de la primera fila de las otras dos filas, se encuentra que A es equivalente por filas a las dos matrices siguientes:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{bmatrix} \tag{1}$$

Puesto que A es invertible, entonces la entrada (2, 2) o la entrada (3, 2) en el lado derecho de (1) es diferente de cero. Supongamos que la entrada (2, 2) es diferente de cero. (De otra forma, se puede realizar un intercambio de filas antes de proceder). Se multiplica la fila 3 por $a_{11}a_{22} - a_{12}a_{21}$, y luego a la nueva fila 3 se le suma la fila 2 multiplicada por $-(a_{11}a_{32} - a_{12}a_{31})$. Esto mostrará que

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

donde

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \tag{2}$$

Puesto que A es invertible, Δ debe ser diferente de cero. En la sección 3.2 se verá que lo contrario también es verdad. A Δ de la ecuación (2) se le llama el **determinante** de la matriz A de 3×3.

Recuerde que el determinante de una matriz de 2×2, $A = [a_{ij}]$, es el número

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

Para una matriz de 1×1, por ejemplo, $A = [a_{11}]$, se define $\det A = a_{11}$. Para generalizar la definición del determinante a matrices más grandes, se utilizarán determinantes de 2×2 para rescribir el determinante Δ de 3×3 descrito anteriormente. Puesto que los términos en Δ se pueden agrupar como $(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31})$,

$$\Delta = a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

Por brevedad, se escribe

$$\Delta = a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13} \tag{3}$$

donde A_{11}, A_{12} y A_{13} se obtienen de A eliminando la primera fila y una de las tres columnas. Para cualquier matriz cuadrada A, sea A_{ij} la submatriz formada al eliminar la i-ésima fila y la

j-ésima columna de A. Por ejemplo, si

$$A = \begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

entonces A_{32} se obtiene eliminando la fila 3 y la columna 2,

$$\begin{bmatrix} 1 & -2 & 5 & 0 \\ 2 & 0 & 4 & -1 \\ 3 & 1 & 0 & 7 \\ 0 & 4 & -2 & 0 \end{bmatrix}$$

de manera que

$$A_{32} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

Ahora se puede dar una definición *recursiva* de un determinante. Cuando $n = 3$, $\det A$ se define utilizando los determinantes de las submatrices A_{1j} de 2×2, como en la ecuación (3). Cuando $n = 4$, $\det A$ utiliza los determinantes de las submatrices A_{1j} de 3×3. En general, un determinante $n \times n$ se define mediante determinantes de submatrices de $(n - 1) \times (n - 1)$.

DEFINICIÓN

Para $n \geq 2$, el **determinante** de una matriz $A = [a_{ij}]$ de $n \times n$ es la suma de n términos de la forma $\pm a_{1j} \det A_{1j}$, con signos más y menos alternados, donde las entradas a_{11}, $a_{12}, \ldots, a_{1n}$ son de la primera fila de A. En símbolos,

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned}$$

EJEMPLO 1 Calcule el determinante de

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

SOLUCIÓN Calcule $\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$:

$$\begin{aligned} \det A &= 1 \cdot \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1(0 - 2) - 5(0 - 0) + 0(-4 - 0) = -2 \end{aligned}$$

■

Otra notación común para el determinante de una matriz utiliza un par de rectas verticales en lugar de corchetes. Así, el cálculo del ejemplo 1 se puede representar como

$$\det A = 1 \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} = \cdots = -2$$

Para establecer el siguiente teorema, es conveniente escribir la definición de $\det A$ en una forma ligeramente distinta. Dada $A = [a_{ij}]$, el **cofactor (*i*, *j*)** de A es el número C_{ij} definido por

$$C_{ij} = (-1)^{i+j} \det A_{ij} \tag{4}$$

Entonces

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$

Esta fórmula se conoce como **desarrollo por cofactores a lo largo de la primera fila** de A. Se omite la demostración del siguiente teorema fundamental para así evitar una larga digresión.

TEOREMA 1 El determinante de una matriz A de $n \times n$ se puede calcular mediante un desarrollo por cofactores a lo largo de cualquier fila o columna. El desarrollo a lo largo de la i-ésima fila utilizando los cofactores de la ecuación (4) es

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

El desarrollo por cofactores a lo largo de la j-ésima columna es

$$\det A = a_{ij}C_{ij} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

El signo más o menos en el cofactor (i, j) depende de la posición de a_{ij} en la matriz, sin importar el signo de a_{ij}. El factor $(-1)^{i+j}$ genera el siguiente patrón de signos:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

EJEMPLO 2 Utilice un desarrollo por cofactores a lo largo de la tercera fila para calcular el det A, donde

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

SOLUCIÓN Calcule

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= (-1)^{3+1}a_{31} \det A_{31} + (-1)^{3+2}a_{32} \det A_{32} + (-1)^{3+3}a_{33} \det A_{33} \\ &= 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= 0 + 2(-1) + 0 = -2 \end{aligned}$$

■

El teorema 1 es útil para calcular el determinante de una matriz que contiene muchos ceros. Por ejemplo, si en una fila hay una mayoría de ceros, entonces el desarrollo por cofactores a lo largo de esa fila tiene muchos términos iguales a cero, y no se necesita calcular los cofactores en estos términos. El mismo enfoque funciona con una columna que contiene muchos ceros.

EJEMPLO 3 Calcule det A, donde

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$$

SOLUCIÓN El desarrollo por cofactores a lo largo de la primera columna tiene todos los términos iguales a cero, excepto el primero. Así,

$$\det A = 3 \cdot \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0 \cdot C_{21} + 0 \cdot C_{31} - 0 \cdot C_{41} + 0 \cdot C_{51}$$

De aquí en adelante se omitirán los términos iguales a cero en el desarrollo por cofactores. Después, desarrolle este determinante de 4×4 a lo largo de la primera columna para tomar ventaja de los ceros que están ahí. Se tiene

$$\det A = 3 \cdot 2 \cdot \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

Este determinante de 3×3 se calculó en el ejemplo 1 y se encontró que su valor es -2. Por lo tanto, $\det A = 3 \cdot 2 \cdot (-2) = -12$. ■

La matriz en el ejemplo 3 era casi triangular. El método en ese ejemplo se adapta fácilmente para demostrar el siguiente teorema.

TEOREMA 2

Si A es una matriz triangular, entonces $\det A$ es el producto de las entradas sobre la diagonal principal de A.

La estrategia en el ejemplo 3 para detectar ceros funciona muy bien cuando una fila o columna completa consiste en ceros. En tal caso, el desarrollo por cofactores a lo largo de tal fila o columna ¡es una suma de ceros! Así, el determinante es cero. Por desgracia, la mayoría de los desarrollos por cofactores no son tan rápidos de evaluar.

NOTA NUMÉRICA

En la actualidad, una matriz de 25×25 se considera pequeña. Sin embargo, sería imposible calcular un determinante de 25×25 con un desarrollo por cofactores. En general, un desarrollo por cofactores requiere alrededor de $n!$ multiplicaciones, y $25!$ es aproximadamente 1.5×10^{25}.

Si una computadora efectúa un billón de multiplicaciones por segundo, tendría que trabajar unos 500,000 años para calcular un determinante de 25×25 utilizando este método. Por fortuna, hay métodos más rápidos, como pronto se verá.

Los ejercicios 19 a 38 exploran importantes propiedades de los determinantes, sobre todo para el caso de 2×2. Los resultados de los ejercicios 33 a 36 se utilizarán en la próxima sección con la finalidad de deducir propiedades análogas para matrices de $n \times n$.

PROBLEMA DE PRÁCTICA

Calcule $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

3.1 EJERCICIOS

En los ejercicios 1 a 8 obtenga los determinantes utilizando un desarrollo por cofactores a lo largo de la primera fila. En los ejercicios 1 a 4, también calcule el determinante mediante un desarrollo por cofactores a lo largo de la segunda columna.

1. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$

2. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

3. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

5. $\begin{vmatrix} 2 & 3 & -4 \\ 4 & 0 & 5 \\ 5 & 1 & 6 \end{vmatrix}$

6. $\begin{vmatrix} 5 & -2 & 4 \\ 0 & 3 & -5 \\ 2 & -4 & 7 \end{vmatrix}$

7. $\begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$ 8. $\begin{vmatrix} 8 & 1 & 6 \\ 4 & 0 & 3 \\ 3 & -2 & 5 \end{vmatrix}$

En los ejercicios 9 a 14 calcule los determinantes mediante desarrollo por cofactores. En cada paso, seleccione una fila o columna que implique la menor cantidad de operaciones.

9. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$ 10. $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

11. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$ 12. $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

13. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

14. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

El desarrollo de un determinante de 3×3 se puede recordar mediante el siguiente recurso. A la derecha de la matriz escriba una segunda copia de las primeras dos columnas, y calcule el determinante multiplicando las entradas sobre las seis diagonales:

$$\begin{array}{ccc|cc} & & & - \quad - & \quad - \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ & & & + \quad + & \quad + \end{array}$$

Sume los productos de las diagonales hacia abajo y reste los productos de las diagonales hacia arriba. Utilice este método para obtener los determinantes en los ejercicios 15 a 18. ***Advertencia:*** *Este truco no se generaliza de manera razonable a matrices* de 4×4 *o mayores*.

15. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$ 16. $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

17. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$ 18. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

En los ejercicios 19 a 24, explore el efecto de una operación elemental de fila sobre el determinante de una matriz. En cada caso, establezca la operación de fila y describa cómo afecta al determinante.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ a & b \end{bmatrix}$ 20. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$

21. $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 5+3k & 6+4k \end{bmatrix}$

22. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$

23. $\begin{bmatrix} 1 & 1 & 1 \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}, \begin{bmatrix} k & k & k \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}$

24. $\begin{bmatrix} a & b & c \\ 3 & 2 & 2 \\ 6 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 2 & 2 \\ a & b & c \\ 6 & 5 & 6 \end{bmatrix}$

En los ejercicios 25 a 30, calcule los determinantes de las matrices elementales dadas. (Véase la sección 2.2).

25. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$ 26. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$

27. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 28. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$

29. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 30. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

Con base en los ejercicios 25 a 28, conteste las preguntas de los ejercicios 31 y 32. Exprese las razones de sus respuestas.

31. ¿Cuál es el determinante de una matriz con remplazo elemental de fila?

32. ¿Cuál es el determinante de una matriz de escalamiento elemental con k en la diagonal?

En los ejercicios 33 a 36, compruebe que $\det EA = (\det E)(\det A)$, donde E es la matriz elemental mostrada y $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 34. $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

35. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ 36. $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

37. Sea $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$. Escriba $5A$. ¿$\det 5A = 5 \det A$?

38. Sean $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ y k un escalar. Encuentre una fórmula que relacione al $\det kA$ con k y $\det A$.

En los ejercicios 39 y 40, A es una matriz de $n \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

39. *a*) Un determinante de $n \times n$ está definido por determinantes de submatrices de $(n-1) \times (n-1)$.

b) El cofactor (i, j) de una matriz A es la matriz A_{ij} obtenida al eliminar de A la i-ésima fila y la j-ésima columna.

40. *a*) El desarrollo por cofactores de $\det A$ a lo largo de una columna es el negativo del desarrollo por cofactores a lo largo de una fila.

b) El determinante de una matriz triangular es la suma de las entradas sobre la diagonal principal.

41. Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Calcule el área del paralelogramo determinado por **u**, **v**, **u** + **v** y **0**, y obtenga el determinante de [**u** **v**]. ¿Hay comparación entre ambos resultados? Remplace la primera entrada de **v** por un número arbitrario x, y repita el problema. Realice un esquema y explique lo que haya encontrado.

42. Sean $u = \begin{bmatrix} a \\ b \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} c \\ 0 \end{bmatrix}$, donde a, b, c son positivos (para simplificar). Calcule el área del paralelogramo definido por **u**, **v**, **u** + **v** y **0**, y obtenga los determinantes de las matrices [**u** **v**] y [**v** **u**]. Realice un esquema y explique sus resultados.

43. [**M**] ¿Es verdad que $\det(A + B) = \det A + \det B$? Para averiguarlo, genere matrices aleatorias A y B de 5×5, y calcule $\det(A + B) - \det A - \det B$. (Consulte el ejercicio 37 de la sección 2.1). Repita los cálculos para otros tres pares de matrices de $n \times n$, con diversos valores de n. Informe sus resultados.

44. [**M**] ¿Es cierto que $\det AB = (\det A)(\det B)$? Experimente con cuatro pares de matrices aleatorias, como en el ejercicio 43, y haga una conjetura.

45. [**M**] Construya una matriz aleatoria A de 4×4, con entradas enteras entre -9 y 9, y compare $\det A$ con $\det A^T$, $\det(-A)$, $\det(2A)$ y $\det(10A)$. Repita el ejercicio con otras dos matrices aleatorias de 4×4, y haga conjeturas acerca de cómo se relacionan esos determinantes. (Véase el ejercicio 36 de la sección 2.1). Luego, compruebe sus conjeturas con varias matrices enteras aleatorias de 5×5 y de 6×6. Si es necesario, modifique sus conjeturas e informe sus resultados.

46. [**M**] ¿Cómo se relaciona $\det A^{-1}$ con $\det A$? Experimente con matrices enteras aleatorias de $n \times n$, para $n = 4$, 5 y 6, y haga una conjetura. *Nota:* En el improbable caso de que se encuentre con una matriz con determinante cero, redúzcala a una forma escalonada y analice su resultado.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Aproveche los ceros. Inicie con un desarrollo por cofactores a lo largo de la tercera columna para obtener una matriz de 3×3, que se puede evaluar con un desarrollo a lo largo de su primera columna.

$$\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix} = (-1)^{1+3}2 \begin{vmatrix} 0 & 3 & -4 \\ -5 & -8 & 3 \\ 0 & 5 & -6 \end{vmatrix}$$

$$= 2 \cdot (-1)^{2+1}(-5) \begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 20$$

El $(-1)^{2+1}$ en el penúltimo cálculo viene de la posición (2, 1) del -5 en el determinante de 3×3.

3.2 PROPIEDADES DE LOS DETERMINANTES

El secreto de los determinantes reside en cómo cambian cuando se efectúan operaciones de fila. El siguiente teorema generaliza los resultados de los ejercicios 19 a 24 de la sección 3.1. La demostración se encuentra al final de esta sección.

TEOREMA 3 **Operaciones de fila**

Sea A una matriz cuadrada.

a) Si un múltiplo de una fila de A se suma a otra fila para producir una matriz B, entonces $\det B = \det A$.

b) Si dos filas de A se intercambian para producir B, entonces $\det B = -\det A$.

c) Si una fila de A se multiplica por k para producir B, entonces $\det B = k \cdot \det A$.

Los siguientes ejemplos muestran cómo utilizar el teorema 3 para calcular determinantes con eficiencia.

EJEMPLO 1 Calcule det A, donde $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

SOLUCIÓN La estrategia es reducir A a una forma escalonada y luego utilizar el hecho de que el determinante de una matriz triangular es el producto de las entradas diagonales. Los primeros dos remplazos de fila en la columna 1 no alteran al determinante:

$$\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix}$$

Un intercambio de las filas 2 y 3 invierte el signo del determinante, de manera que

$$\det A = -\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -(1)(3)(-5) = 15 \qquad \blacksquare$$

Un uso común del teorema 3c) en cálculos a mano es *factorizar un múltiplo común de una fila* de una matriz. Por ejemplo,

$$\begin{vmatrix} * & * & * \\ 5k & -2k & 3k \\ * & * & * \end{vmatrix} = k\begin{vmatrix} * & * & * \\ 5 & -2 & 3 \\ * & * & * \end{vmatrix}$$

donde las entradas con asterisco quedan inalteradas. Este paso se emplea en el siguiente ejemplo.

EJEMPLO 2 Calcule det A, donde $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

SOLUCIÓN Para simplificar la aritmética, se desea un 1 en la esquina superior izquierda. Se podrían intercambiar las filas 1 y 4. Pero, en vez de ello, se saca el factor 2 de la fila superior, y luego se procede con remplazos de fila en la primera columna:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Después, se podría sacar otro factor 2 de la fila 3 o utilizar como pivote el 3 en la segunda columna. Seleccionamos la última operación, sumando a la fila 3 la fila 2 multiplicada por 4:

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Por último, sumando la fila 3 multiplicada por $-1/2$ a la fila 4, y calculando el determinante "triangular", se encuentra que

$$\det A = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2 \cdot (1)(3)(-6)(1) = -36 \qquad \blacksquare$$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix}$$
$\det U \neq 0$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$\det U = 0$

FIGURA 1
Formas escalonadas típicas de matrices cuadradas.

Suponga que una matriz cuadrada A se redujo a una forma escalonada U mediante remplazos de fila e intercambios de fila. (Esto siempre es posible. Véase el algoritmo de reducción por filas en la sección 1.2). Si hay r intercambios, entonces el teorema 3 indica que:

$$\det A = (-1)^r \det U$$

Como U está en forma escalonada, es triangular, y así $\det U$ es el producto de las entradas diagonales $u_{11},\ldots, u_{nn}$. Si A es invertible, las entradas u_{ii} son todas pivotes (porque $A \sim I_n$ y las u_{ii} no se han escalado a 1). De otra forma, al menos u_{nn} es cero, y el producto $u_{11} \cdots u_{nn}$ es cero. Véase la figura 1. Por lo tanto,

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{producto de} \\ \text{pivotes en } U \end{pmatrix} & \text{cuando } A \text{ es invertible} \\ 0 & \text{cuando } A \text{ no es invertible} \end{cases} \tag{1}$$

Es interesante hacer notar que aunque la forma escalonada U que se acaba de describir no es única (porque no está completamente reducida por filas), y los pivotes no son únicos, el *producto* de los pivotes *es* único, excepto por un posible signo menos.

La fórmula (1) no solo da una interpretación concreta de $\det A$, sino que también demuestra el principal teorema de esta sección:

TEOREMA 4 Una matriz cuadrada A es invertible si y solo si $\det A \neq 0$.

El teorema 4 agrega el enunciado “$\det A \neq 0$” al teorema de la matriz invertible. Un útil corolario es que $\det A = 0$ cuando las columnas de A son linealmente dependientes. Además, $\det A = 0$ cuando las *filas* de A son linealmente dependientes. (Filas de A son columnas de A^T, y columnas linealmente dependientes de A^T hacen que A^T sea singular. Cuando A^T es singular, también lo es A, de acuerdo con el teorema de la matriz invertible). En la práctica, la dependencia lineal es evidente cuando dos columnas o dos filas son iguales, o una columna o fila es cero.

EJEMPLO 3 Calcule $\det A$, donde $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

SOLUCIÓN Sume la fila 1 multiplicada por 2 a la fila 3 para obtener

$$\det A = \det \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$

ya que la segunda y tercera filas de la segunda matriz son iguales. ■

NOTA NUMÉRICA

1. La mayoría de los programas de cómputo que calculan $\det A$ para una matriz general A utilizan el método de la fórmula (1) anterior.
2. Es posible demostrar que la evaluación de un determinante de $n \times n$, utilizando operaciones de fila, requiere cerca de $2n^3/3$ operaciones aritméticas. Cualquier microcomputadora moderna es capaz de calcular un determinante de 25×25 en una fracción de segundo, porque solo necesita realizar unas 10,000 operaciones.

WEB

Las computadoras también pueden manejar grandes matrices "dispersas", con rutinas especiales que aprovechan la presencia de muchos ceros. Desde luego, las entradas cero también aceleran los cálculos a mano. Los cálculos en el siguiente ejemplo combinan el poder de las operaciones de fila con la estrategia de la sección 3.1, consistente en utilizar entradas nulas en los desarrollos por cofactores.

EJEMPLO 4 Calcule det A, donde $A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$.

SOLUCIÓN Una buena forma de comenzar es utilizar como pivote el 2 en la columna 1, eliminando el -2 que está debajo de este. Después se utiliza un desarrollo por cofactores para reducir el tamaño del determinante, y luego otra operación de remplazo de fila. De esta forma,

$$\det A = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix}$$

Un intercambio de las filas 2 y 3 producirían un "determinante triangular". Otro enfoque consiste en efectuar un desarrollo en cofactores a lo largo de la primera columna:

$$\det A = (-2)(1) \begin{vmatrix} 0 & 5 \\ -3 & 1 \end{vmatrix} = -2 \cdot (15) = -30 \qquad \blacksquare$$

Operaciones de columna

Es posible realizar operaciones sobre las columnas de una matriz en forma análoga a las operaciones de fila que hemos estudiado. El siguiente teorema indica que las operaciones de columna y las operaciones de fila tienen los mismos efectos sobre los determinantes.

TEOREMA 5 Si A es una matriz de $n \times n$, entonces $\det A^T = \det A$.

DEMOSTRACIÓN El teorema es evidente para $n = 1$. Suponga que el teorema es verdadero para determinantes de $k \times k$, y sea $n = k + 1$. Entonces el cofactor de a_{1j} en A es igual al cofactor de a_{j1} en A^T, porque los cofactores implican determinantes de $k \times k$. Así que el desarrollo por cofactores de det A a lo largo de la primera *fila* es igual al desarrollo por cofactores de det A^T a lo largo de la primera *columna*. Es decir, A y A^T tienen determinantes iguales. Por lo tanto, el teorema es válido para $n = 1$, y la veracidad del teorema para un valor de n implica su veracidad para el siguiente valor de n. Por el principio de inducción, el teorema es verdadero para toda $n \geq 1$. $\blacksquare$

De acuerdo con el teorema 5, cada enunciado en el teorema 3 es válido cuando en todas partes la palabra *fila* se remplaza por *columna.* Para comprobar esta propiedad, basta aplicar a A^T el teorema 3 original. Una operación de fila sobre A^T significa una operación de columna sobre A.

Las operaciones de columna son útiles tanto para fines teóricos como para realizar cálculos a mano. Sin embargo, para simplificar solo se efectuarán operaciones de fila en cálculos numéricos.

Determinantes y productos matriciales

La demostración del siguiente útil teorema se encuentra al final de la sección. Las aplicaciones se exponen en los ejercicios.

TEOREMA 6

Propiedad multiplicativa

Si A y B son matrices de $n \times n$, entonces $\det AB = (\det A)(\det B)$.

EJEMPLO 5 Compruebe el teorema 6 para $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$.

SOLUCIÓN

$$AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}$$

y

$$\det AB = 25 \cdot 13 - 20 \cdot 14 = 325 - 280 = 45$$

Como $\det A = 9$ y $\det B = 5$,

$$(\det A)(\det B) = 9 \cdot 5 = 45 = \det AB$$

■

Advertencia: Un error común es pensar que el teorema 6 tiene un análogo para *sumas* de matrices. Sin embargo, en general, $\det(A + B)$ *no es* igual a $\det A + \det B$.

Propiedad de linealidad de la función determinante

Para una matriz A de $n \times n$, podemos considerar a $\det A$ como una función de los n vectores columna en A. Se demostrará que si todas las columnas se mantienen fijas, excepto una, entonces $\det A$ es una *función lineal* de una variable (vectorial).

Suponga que a la j-ésima columna de A se le permite variar, lo que se escribe como

$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{j-1} & \mathbf{x} & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{bmatrix}$$

Defina una transformación T de $\mathbb{R}^n$ a $\mathbb{R}$ mediante

$$T(\mathbf{x}) = \det \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{j-1} & \mathbf{x} & \mathbf{a}_{j+1} & \cdots & \mathbf{a}_n \end{bmatrix}$$

Así,

$$T(c\mathbf{x}) = cT(\mathbf{x}) \quad \text{para todos los escalares } c, \text{ y todas las } \mathbf{x} \text{ en } \mathbb{R}^n. \tag{2}$$

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{para toda } \mathbf{u}, \mathbf{v} \text{ en } \mathbb{R}^n. \tag{3}$$

La propiedad (2) es el teorema 3*c*) aplicado a las columnas de A. Una demostración de la propiedad (3) es consecuencia de un desarrollo por cofactores a lo largo de la j-ésima columna de $\det A$. (Véase el ejercicio 43). Esta propiedad de (multi)linealidad del determinante tiene muchas consecuencias útiles que se estudian en cursos más avanzados.

Demostraciones de los teoremas 3 y 6

Es conveniente someter a prueba el teorema 3 cuando se enuncia en términos de las matrices elementales analizadas en la sección 2.2. Una matriz elemental E se denomina *matriz de remplazo de fila* si E se obtiene a partir de la identidad I al sumar un múltiplo de una fila a otra; E es un *intercambio* si E se obtiene mediante el intercambio de dos filas de I; y E es una *escala por r* si E se obtiene al multiplicar una fila de I por un escalar r diferente de cero. Con esta terminología, el teorema 3 se puede reformular de la siguiente manera:

Si A es una matriz de $n \times n$ y E es una matriz elemental de $n \times n$, entonces

$$\det EA = (\det E)(\det A)$$

donde

$$\det E = \begin{cases} 1 & \text{si } E \text{ es un remplazo de fila} \\ -1 & \text{si } E \text{ es un intercambio} \\ r & \text{si } E \text{ es una escala por } r \end{cases}$$

DEMOSTRACIÓN DEL TEOREMA 3 La demostración es por inducción sobre el tamaño de A. El caso de una matriz de 2×2 se comprobó en los ejercicios 33 a 36 de la sección 3.1. Suponga que el teorema se ha comprobado para determinantes de matrices de $k \times k$, con $k \geq 2$; sea $n = k + 1$ y sea A de $n \times n$. La acción de E sobre A implica a dos filas o solamente a una. Así, se puede desarrollar det EA a lo largo de una fila que es inalterada por la acción de E, por ejemplo, la fila i. Sea A_{ij} (respectivamente, B_{ij}) la matriz obtenida al eliminar la fila i y la columna j de A (respectivamente, EA). Luego, las filas de B_{ij} se obtienen de las filas de A_{ij} por el mismo tipo de operación elemental de fila que E efectúa sobre A. Como esas submatrices son únicamente de $k \times k$, la suposición de inducción implica que

$$\det B_{ij} = \alpha \cdot \det A_{ij}$$

donde $\alpha = 1, -1$, o r, dependiendo de la naturaleza de E. El desarrollo por cofactores a lo largo de la fila i es

$$\begin{aligned} \det EA &= a_{i1}(-1)^{i+1} \det B_{i1} + \cdots + a_{in}(-1)^{i+n} \det B_{in} \\ &= \alpha a_{i1}(-1)^{i+1} \det A_{i1} + \cdots + \alpha a_{in}(-1)^{i+n} \det A_{in} \\ &= \alpha \cdot \det A \end{aligned}$$

En particular, tomando $A = I_n$, se observa que det $E = 1, -1$, o r, dependiendo de la naturaleza de E. Así, el teorema es válido para $n = 2$, y la veracidad del teorema para un valor de n implica su veracidad para el siguiente valor de n. Por el principio de inducción, el teorema debe ser válido para $n \geq 2$. El teorema es trivialmente verdadero para $n = 1$. ■

DEMOSTRACIÓN DEL TEOREMA 6 Si A no es invertible, entonces AB tampoco lo es, de acuerdo con el ejercicio 27 de la sección 2.3. En este caso, det AB = (det A)(det B), porque ambos lados valen cero, por el teorema 4. Si A es invertible, entonces A y la matriz identidad I_n son equivalentes por filas de acuerdo con el teorema de matriz invertible. Así, existen matrices elementales $E_1, \ldots, E_p$ tales que

$$A = E_p E_{p-1} \cdots E_1 \cdot I_n = E_p E_{p-1} \cdots E_1$$

Por brevedad, escribimos $|A|$ en vez de det A. De esta forma, la repetida aplicación del teorema 3, como se acaba de reformular, muestra que

$$\begin{aligned} |AB| &= |E_p \cdots E_1 B| = |E_p||E_{p-1} \cdots E_1 B| = \cdots \\ &= |E_p| \cdots |E_1||B| = \cdots = |E_p \cdots E_1||B| \\ &= |A||B| \end{aligned}$$

■

PROBLEMAS DE PRÁCTICA

1. Calcule $\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix}$ en el menor número de pasos que sea posible.

2. Utilice un determinante para decidir si $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son linealmente independientes, cuando

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ -7 \\ 9 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 3 \\ -5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -7 \\ 5 \end{bmatrix}$$

3.2 EJERCICIOS

En los ejercicios 1 a 4 cada ecuación ilustra una propiedad de los determinantes. Enuncie la propiedad.

1. $\begin{vmatrix} 0 & 5 & -2 \\ 1 & -3 & 6 \\ 4 & -1 & 8 \end{vmatrix} = -\begin{vmatrix} 1 & -3 & 6 \\ 0 & 5 & -2 \\ 4 & -1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 2 & -6 & 4 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix} = 2\begin{vmatrix} 1 & -3 & 2 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix}$

3. $\begin{vmatrix} 1 & 3 & -4 \\ 2 & 0 & -3 \\ 5 & -4 & 7 \end{vmatrix} = \begin{vmatrix} 1 & 3 & -4 \\ 0 & -6 & 5 \\ 5 & -4 & 7 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 3 & 7 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & -4 \\ 0 & 1 & -5 \end{vmatrix}$

En los ejercicios 5 a 10 encuentre los determinantes por reducción de filas a una forma escalonada.

5. $\begin{vmatrix} 1 & 5 & -6 \\ -1 & -4 & 4 \\ -2 & -7 & 9 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 5 & -3 \\ 3 & -3 & 3 \\ 2 & 13 & -7 \end{vmatrix}$

7. $\begin{vmatrix} 1 & 3 & 0 & 2 \\ -2 & -5 & 7 & 4 \\ 3 & 5 & 2 & 1 \\ 1 & -1 & 2 & -3 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 3 & 3 & -4 \\ 0 & 1 & 2 & -5 \\ 2 & 5 & 4 & -3 \\ -3 & -7 & -5 & 2 \end{vmatrix}$

9. $\begin{vmatrix} 1 & -1 & -3 & 0 \\ 0 & 1 & 5 & 4 \\ -1 & 2 & 8 & 5 \\ 3 & -1 & -2 & 3 \end{vmatrix}$

10. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

En los ejercicios 11 a 14 calcule los determinantes combinando los métodos de reducción por filas y desarrollo por cofactores.

11. $\begin{vmatrix} 2 & 5 & -3 & -1 \\ 3 & 0 & 1 & -3 \\ -6 & 0 & -4 & 9 \\ 4 & 10 & -4 & -1 \end{vmatrix}$

12. $\begin{vmatrix} -1 & 2 & 3 & 0 \\ 3 & 4 & 3 & 0 \\ 5 & 4 & 6 & 6 \\ 4 & 2 & 4 & 3 \end{vmatrix}$

13. $\begin{vmatrix} 2 & 5 & 4 & 1 \\ 4 & 7 & 6 & 2 \\ 6 & -2 & -4 & 0 \\ -6 & 7 & 7 & 0 \end{vmatrix}$

14. $\begin{vmatrix} -3 & -2 & 1 & -4 \\ 1 & 3 & 0 & -3 \\ -3 & 4 & -2 & 8 \\ 3 & -4 & 0 & 4 \end{vmatrix}$

En los ejercicios del 15 al 20, calcule los determinantes, donde

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 7.$$

15. $\begin{vmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{vmatrix}$

16. $\begin{vmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{vmatrix}$

17. $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$

18. $\begin{vmatrix} g & h & i \\ a & b & c \\ d & e & f \end{vmatrix}$

19. $\begin{vmatrix} a & b & c \\ 2d+a & 2e+b & 2f+c \\ g & h & i \end{vmatrix}$

20. $\begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}$

En los ejercicios 21 a 23, use determinantes para saber si la matriz es invertible.

21. $\begin{bmatrix} 2 & 3 & 0 \\ 1 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix}$

22. $\begin{bmatrix} 5 & 0 & -1 \\ 1 & -3 & -2 \\ 0 & 5 & 3 \end{bmatrix}$

23. $\begin{bmatrix} 2 & 0 & 0 & 8 \\ 1 & -7 & -5 & 0 \\ 3 & 8 & 6 & 0 \\ 0 & 7 & 5 & 4 \end{bmatrix}$

En los ejercicios 24 a 26, utilice determinantes para saber si el conjunto de vectores es linealmente independiente.

24. $\begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ -5 \\ 6 \end{bmatrix}$

25. $\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

26. $\begin{bmatrix} 3 \\ 5 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 0 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ -3 \end{bmatrix}$

En los ejercicios 27 y 28, A y B son matrices de $n \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

27. *a*) Una operación de remplazo por filas no afecta al determinante de una matriz.

b) El determinante de A es el producto de los pivotes en cualquier forma escalonada U de A, multiplicada por $(-1)^r$, donde r es el número de intercambios de fila realizados durante la reducción por filas de A a U.

c) Si las columnas de A son linealmente dependientes, entonces $\det A = 0$.

d) $\det(A + B) = \det A + \det B$.

28. *a*) Si dos intercambios de fila se realizan en secuencia, entonces el nuevo determinante es igual al determinante original.

b) El determinante de A es el producto de las entradas diagonales en A.

c) Si $\det A$ es cero, entonces dos filas o dos columnas son iguales, o una fila o una columna es cero.

d) $\det A^T = (-1)\det A$.

29. Calcule $\det B^5$, donde $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.

30. Utilice el teorema 3 (pero no el teorema 4) para demostrar que si dos filas de una matriz cuadrada A son iguales, entonces $\det A = 0$. Lo mismo es válido para dos columnas. ¿Por qué?

En los ejercicios 31 a 36, su explicación debe mencionar un teorema pertinente.

31. Demuestre que si A es invertible, entonces $\det A^{-1} = \dfrac{1}{\det A}$.

32. Obtenga una fórmula para $\det(rA)$ cuando A es una matriz de $n \times n$.

33. Sean A y B matrices cuadradas. Demuestre que aun cuando AB y BA pueden no ser iguales, siempre es verdad que $\det AB = \det BA$.

34. Sean A y P matrices cuadradas, con P invertible. Demuestre que $\det(PAP^{-1}) = \det A$.

35. Sea U una matriz cuadrada tal que $U^TU = I$. Demuestre que $\det U = \pm 1$.

36. Suponga que A es una matriz cuadrada tal que $\det A^4 = 0$. Explique por qué A no puede ser invertible.

En los ejercicios 37 y 38, compruebe que $\det AB = (\det A)(\det B)$ para las matrices dadas. (No utilice el teorema 6).

37. $A = \begin{bmatrix} 3 & 0 \\ 6 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 5 & 4 \end{bmatrix}$

38. $A = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}, B = \begin{bmatrix} 4 & 2 \\ -1 & -1 \end{bmatrix}$

39. Sean A y B matrices de 3×3, con $\det A = 4$ y $\det B = -3$. Con base en propiedades de determinantes (en el libro y en los ejercicios anteriores), calcule:

a) $\det AB$ *b*) $\det 5A$ *c*) $\det B^T$

d) $\det A^{-1}$ *e*) $\det A^3$

40. Sean A y B matrices de 4×4, con $\det A = -1$ y $\det B = 2$. Calcule:

a) $\det AB$ *b*) $\det B^5$ *c*) $\det 2A$

d) $\det A^TA$ *e*) $\det B^{-1}AB$

41. Compruebe que $\det A = \det B + \det C$, donde

$$A = \begin{bmatrix} a+e & b+f \\ c & d \end{bmatrix}, \; B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \; C = \begin{bmatrix} e & f \\ c & d \end{bmatrix}$$

42. Sean $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Demuestre que $\det(A + B) = \det A + \det B$ si y solo si $a + d = 0$.

43. Compruebe que $\det A = \det B + \det C$, donde

$$A = \begin{bmatrix} a_{11} & a_{12} & u_1 + v_1 \\ a_{21} & a_{22} & u_2 + v_2 \\ a_{31} & a_{32} & u_3 + v_3 \end{bmatrix},$$

$$B = \begin{bmatrix} a_{11} & a_{12} & u_1 \\ a_{21} & a_{22} & u_2 \\ a_{31} & a_{32} & u_3 \end{bmatrix}, \; C = \begin{bmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ a_{31} & a_{32} & v_3 \end{bmatrix}$$

Sin embargo, observe, que A *no* es igual a $B + C$.

44. La multiplicación por la derecha por una matriz elemental E afecta a las *columnas* de A en la misma forma que la multiplicación por la izquierda afecta a las *filas*. Utilice los teoremas 3 y 5, así como el evidente resultado de que E^T es otra matriz elemental, para demostrar que

$\det AE = (\det E)(\det A)$

No utilice el teorema 6.

45. **[M]** Calcule $\det A^TA$ y $\det AA^T$ para varias matrices de 4×5 aleatorias y diversas matrices de 5×6, también aleatorias. ¿Qué puede decirse acerca de A^TA y AA^T cuando A tiene más columnas que filas?

46. **[M]** Si $\det A$ es cercano a cero, ¿la matriz A es casi singular? Experimente con la matriz A de 4×4 casi singular del ejercicio 9 de la sección 2.3. Calcule los determinantes de A, $10A$ y $0.1A$. Por otra parte, calcule los números de condición de esas matrices. Repita los cálculos cuando A es la matriz identidad de 4×4. Analice sus resultados.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Realice remplazos de fila para crear ceros en la primera columna y así obtener una fila de ceros.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

2. $$\det[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3] = \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix} \qquad \text{Fila 1 sumada a la fila 2}$$

$$= -(-3)\begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5)\begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix} \qquad \text{Cofactores de la columna 2}$$

$$= 3 \cdot (35) + 5 \cdot (-21) = 0$$

Según el teorema 4, la matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$ no es invertible. Las columnas son linealmente dependientes, de acuerdo con el teorema de la matriz invertible.

3.3 REGLA DE CRAMER, VOLUMEN Y TRANSFORMACIONES LINEALES

Esta sección aplica la teoría de las secciones anteriores para obtener importantes fórmulas teóricas y una interpretación geométrica del determinante.

Regla de Cramer

La regla de Cramer es necesaria en una variedad de cálculos teóricos. Por ejemplo, se puede utilizar para estudiar cómo resulta afectada la solución de $A\mathbf{x} = \mathbf{b}$ por cambios en las entradas de $\mathbf{b}$. Sin embargo, la fórmula es ineficiente para cálculos a mano, excepto para matrices de 2×2, o quizá de 3×3.

Para cualquier matriz A de $n \times n$ y cualquier $\mathbf{b}$ en $\mathbb{R}^n$, sea $A_i(\mathbf{b})$ la matriz obtenida a partir de A al remplazar la columna i por el vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{col } i}}{\mathbf{b}} \quad \cdots \quad \mathbf{a}_n]$$

TEOREMA 7 **Regla de Cramer**

Sea A una matriz invertible *de* $n \times n$. Para cualquier $\mathbf{b}$ en $\mathbb{R}^n$, la única solución $\mathbf{x}$ de $A\mathbf{x} = \mathbf{b}$ tiene entradas dadas por

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \quad i = 1, 2, \ldots, n \tag{1}$$

DEMOSTRACIÓN Denote las columnas de A por $\mathbf{a}_1, \ldots, \mathbf{a}_n$ y las columnas de la matriz identidad I de $n \times n$ por $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Si $A\mathbf{x} = \mathbf{b}$, la definición de multiplicación matricial indica que

$$\begin{aligned} A \cdot I_i(\mathbf{x}) &= A[\mathbf{e}_1 \quad \cdots \quad \mathbf{x} \quad \cdots \quad \mathbf{e}_n] = [A\mathbf{e}_1 \quad \cdots \quad A\mathbf{x} \quad \cdots \quad A\mathbf{e}_n] \\ &= [\mathbf{a}_1 \quad \cdots \quad \mathbf{b} \quad \cdots \quad \mathbf{a}_n] = A_i(\mathbf{b}) \end{aligned}$$

Por la propiedad multiplicativa de determinantes,

$$(\det A)(\det I_i(\mathbf{x})) = \det A_i(\mathbf{b})$$

El segundo determinante a la izquierda es simplemente x_i. (Realice un desarrollo por cofactores a lo largo de la i-ésima fila). Así, $(\det A) \cdot x_i = \det A_i(\mathbf{b})$. Esto demuestra (1), ya que A es invertible y $\det A \neq 0$. ■

EJEMPLO 1 Use la regla de Cramer para resolver el sistema

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8 \end{aligned}$$

SOLUCIÓN Vea el sistema como $A\mathbf{x} = \mathbf{b}$. Utilizando la notación ya presentada,

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \qquad A_1(\mathbf{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \qquad A_2(\mathbf{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Como $\det A = 2$, el sistema tiene una solución única. Por la regla de Cramer,

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{24 + 16}{2} = 20$$
$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{24 + 30}{2} = 27$$

■

Aplicación a la ingeniería

Un gran número de importantes problemas de ingeniería, particularmente en teoría del control e ingeniería eléctrica, se pueden analizar con las *transformadas de Laplace*. Este enfoque convierte un adecuado sistema de ecuaciones diferenciales lineales a un sistema de ecuaciones algebraicas lineales cuyos coeficientes implican un parámetro. El siguiente ejemplo ilustra el tipo de sistema algebraico que puede presentarse.

EJEMPLO 2 Considere el siguiente sistema, en el cual s es un parámetro no especificado. Determine los valores de s para los cuales el sistema tiene una solución única, y utilice la regla de Cramer para describir la solución.

$$\begin{aligned} 3sx_1 - 2x_2 &= 4 \\ -6x_1 + sx_2 &= 1 \end{aligned}$$

SOLUCIÓN Vea el sistema como $A\mathbf{x} = \mathbf{b}$. De esta forma,

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \quad A_1(\mathbf{b}) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \quad A_2(\mathbf{b}) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

Puesto que

$$\det A = 3s^2 - 12 = 3(s + 2)(s - 2)$$

el sistema tiene una solución única precisamente cuando $s \neq \pm 2$. Para tal s, la solución es (x_1, x_2), donde

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{4s + 2}{3(s + 2)(s - 2)}$$
$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{3s + 24}{3(s + 2)(s - 2)} = \frac{s + 8}{(s + 2)(s - 2)}$$

■

Una fórmula para A^{-1}

La regla de Cramer conduce fácilmente a una fórmula general para la inversa de una matriz A de $n \times n$. La j-ésima columna de A^{-1} es un vector $\mathbf{x}$ que satisface

$$A\mathbf{x} = \mathbf{e}_j$$

donde $\mathbf{e}_j$ es la j-ésima columna de la matriz identidad, y la i-ésima entrada de $\mathbf{x}$ es la entrada (i, j) de A^{-1}. De acuerdo con la regla de Cramer,

$$\{\text{entrada } (i, j) \text{ de } A^{-1}\} = x_i = \frac{\det A_i(\mathbf{e}_j)}{\det A} \tag{2}$$

Recuerde que A_{ji} denota la submatriz de A formada al eliminar la fila j y la columna i. Un desarrollo por cofactores a lo largo de la columna i de $A_i(\mathbf{e}_j)$ muestra que

$$\det A_i\,(\mathbf{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \tag{3}$$

donde C_{ji} es un cofactor de A. De acuerdo con la expresión (2), la entrada (i, j) de A^{-1} es el cofactor C_{ji} dividido entre $\det A$. [Observe que los subíndices en C_{ji} son los inversos de (i, j)]. Por consiguiente,

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \tag{4}$$

La matriz de cofactores en el miembro derecho de (4) se llama la **adjunta** de A, que se denota con adj A. (El término *adjunta* también tiene otro significado en libros avanzados de transformaciones lineales). El siguiente teorema reformula la expresión (4).

TEOREMA 8 **Una fórmula para la inversa**

Sea A una matriz invertible de $n \times n$. Así,

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

EJEMPLO 3 Encuentre la inversa de la matriz $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

SOLUCIÓN Los nueve cofactores son

$$C_{11} = +\begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = -\begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = +\begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = -\begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = +\begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = -\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = +\begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = -\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = +\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

La matriz adjunta es la *transpuesta* de la matriz de cofactores. [Por ejemplo, C_{12} va a la posición (2, 1).] Así,

$$\operatorname{adj} A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

Se podría calcular $\det A$ directamente, pero el siguiente cálculo ofrece una comprobación de las operaciones anteriores *y* produce $\det A$:

$$(\operatorname{adj} A) \cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Puesto que $(\operatorname{adj} A)A = 14I$, el teorema 8 indica que $\det A = 14$ y

$$A^{-1} = \frac{1}{14} \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix}$$

■

NOTA NUMÉRICA

El teorema 8 es útil principalmente para cálculos teóricos. La fórmula para A^{-1} permite deducir propiedades de la inversa sin calcularla en realidad. Excepto para casos especiales, el algoritmo de la sección 2.2 ofrece una forma mucho mejor de calcular A^{-1}, si la inversa es realmente necesaria.

La regla de Cramer también es una herramienta teórica. Se puede emplear para estudiar qué tan sensible es la solución de $A\mathbf{x} = \mathbf{b}$ ante cambios en una entrada de $\mathbf{b}$ o de $\mathbf{A}$ (quizá debido al error experimental cuando se obtienen las entradas para $\mathbf{b}$ o A). Cuando A es una matriz de 3×3 con entradas *complejas*, entonces algunas veces la regla de Cramer se utiliza en cálculos a mano porque la reducción por filas de $[A \quad \mathbf{b}]$ con aritmética compleja puede resultar confusa, y los determinantes son bastante fáciles de calcular. Para grandes matrices de $n \times n$ (reales o complejas), la regla de Cramer es irremediablemente ineficiente. Para calcular solo *un* determinante se requiere tanto trabajo como resolver $A\mathbf{x} = \mathbf{b}$ por reducción de filas.

Determinantes como área o volumen

En la siguiente aplicación, se comprueba la interpretación geométrica de determinantes descrita en la introducción de este capítulo. Aunque en el capítulo 6 se hará un análisis general de longitud y distancia en $\mathbb{R}^n$, aquí se supone que los conceptos euclidianos usuales de longitud, área y volumen ya se entienden para $\mathbb{R}^2$ y $\mathbb{R}^3$.

TEOREMA 9

Si A es una matriz de 2×2, el área del paralelogramo definido por las columnas de A es $|\det A|$. Si A es una matriz de 3×3, el volumen del paralelepípedo definido por las columnas de A es $|\det A|$.

DEMOSTRACIÓN Como es evidente, el teorema es cierto para cualquier matriz diagonal de 2×2:

$$\left|\det\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right| = |ad| = \begin{Bmatrix} \text{área del} \\ \text{rectángulo} \end{Bmatrix}$$

Véase la figura 1. Será suficiente demostrar que cualquier matriz de 2×2, $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, se puede transformar a una matriz diagonal de tal manera que no cambie el área del paralelogramo asociado ni tampoco $|\det A|$. De la sección 3.2, se conoce que el valor absoluto del determinante es inalterado cuando dos columnas se intercambian, o un múltiplo de una columna se suma a otra. Es fácil ver que tales operaciones son suficientes para transformar a A en una matriz diagonal. Los intercambios de columnas no modifican el paralelogramo. Así, es suficiente probar la siguiente sencilla observación geométrica que se aplica a vectores en $\mathbb{R}^2$ o $\mathbb{R}^3$:

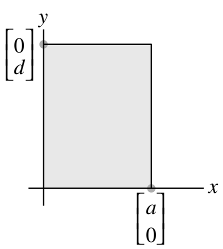


FIGURA 1
Área $= |ad|$.

Sean $\mathbf{a}_1$ y $\mathbf{a}_2$ vectores diferentes de cero (no nulos). Luego, para cualquier escalar c, el área del paralelogramo definido por $\mathbf{a}_1$ y $\mathbf{a}_2$ es igual al área del paralelogramo determinado por $\mathbf{a}_1$ y $\mathbf{a}_2 + c\mathbf{a}_1$.

Para demostrar este enunciado, se supone que $\mathbf{a}_2$ no es un múltiplo de $\mathbf{a}_1$, ya que, de otra forma, los dos paralelogramos serían degenerados y tendrían área igual a cero. Si L es la recta que pasa por $\mathbf{0}$ y $\mathbf{a}_1$, entonces $\mathbf{a}_2 + L$ es la recta que pasa por $\mathbf{a}_2$ paralela a L, y $\mathbf{a}_2 + c\mathbf{a}_1$ está sobre esta recta. Véase la figura 2. Los puntos $\mathbf{a}_2$ y $\mathbf{a}_2 + c\mathbf{a}_1$ tienen la misma distancia perpendicular a L. Por eso, los dos paralelogramos en la figura 2 tienen la misma área porque comparten la base de $\mathbf{0}$ a $\mathbf{a}_1$. Esto completa la demostración para $\mathbb{R}^2$.

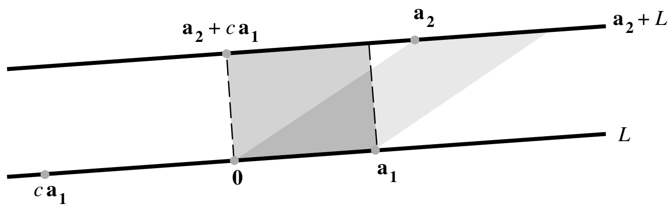


FIGURA 2 Dos paralelogramos de igual área.

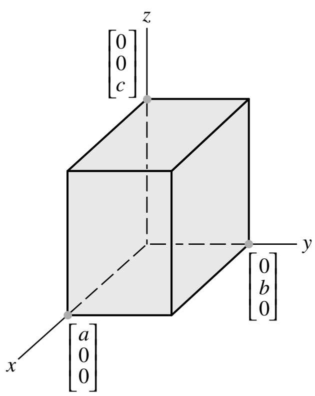


FIGURA 3
Volumen $= |abc|$.

La demostración para $\mathbb{R}^3$ es similar. Como es evidente, el teorema es cierto para una matriz diagonal de 3×3. Véase la figura 3. Y cualquier matriz A de 3×3 se puede transformar en una matriz diagonal utilizando operaciones de columna que no cambian a $|\det A|$. (Piense en efectuar operaciones de fila sobre A^T). Así, es suficiente demostrar que esas operaciones no afectan el volumen del paralelepípedo definido por las columnas de A.

La figura 4 muestra un paralelepípedo como una caja sombreada con dos lados inclinados. Su volumen es el área de la base en el plano Gen $\{\mathbf{a}_1, \mathbf{a}_3\}$ multiplicada por la altura de $\mathbf{a}_2$ sobre Gen $\{\mathbf{a}_1, \mathbf{a}_3\}$. Cualquier vector $\mathbf{a}_2 + c\mathbf{a}_1$ tiene la misma altura porque $\mathbf{a}_2 + c\mathbf{a}_1$ está en el plano $\mathbf{a}_2$ + Gen $\{\mathbf{a}_1, \mathbf{a}_3\}$, el cual es paralelo a Gen $\{\mathbf{a}_1, \mathbf{a}_3\}$. Así que el volumen del paralelepípedo queda inalterado cuando $[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ se cambia a $[\mathbf{a}_1 \quad \mathbf{a}_2 + c\mathbf{a}_1 \quad \mathbf{a}_3]$. Por consiguiente, una operación de remplazo de columna no afecta el volumen del paralelepípedo. Como el intercambio de columnas no tiene efecto sobre el volumen, la demostración se completa. ■

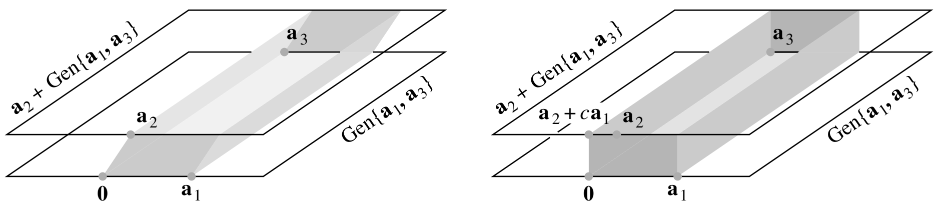


FIGURA 4 Dos paralelepípedos de igual volumen.

EJEMPLO 4 Calcule el área del paralelogramo definido por los puntos $(-2, -2)$, $(0, 3)$, $(4, -1)$ y $(6, 4)$. Véase la figura 5*a*).

SOLUCIÓN Primero el paralelogramo se traslada de manera que un vértice esté en el origen. Por ejemplo, reste el vértice $(-2, -2)$ de cada uno de los cuatro vértices. El nuevo paralelogramo tiene la misma área, y sus vértices son $(0, 0)$, $(2, 5)$, $(6, 1)$ y $(8, 6)$. Véase la figura 5*b*).

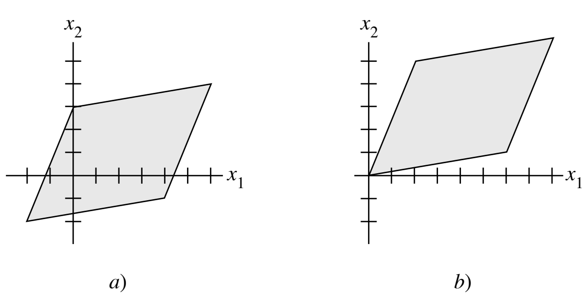


FIGURA 5 Trasladar un paralelogramo no cambia su área.

Este paralelogramo está definido por las columnas de

$$A = \begin{bmatrix} 2 & 6 \\ 5 & 1 \end{bmatrix}$$

Como $|\det A| = |-28|$, entonces el área del paralelogramo es 28. ■

Transformaciones lineales

Los determinantes se pueden usar para describir una importante propiedad geométrica de transformaciones lineales en el plano y en $\mathbb{R}^3$. Si T es una transformación lineal y S es un conjunto en el dominio de T, entonces $T(S)$ denota el conjunto de imágenes de puntos en S. Nos interesa conocer cómo se compara el área (o volumen) de $T(S)$ con el área (o volumen) del conjunto original S. Por conveniencia, cuando S es una región acotada por un paralelogramo, también nos referimos a S como un paralelogramo.

TEOREMA 10

Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ una transformación lineal determinada por una matriz A de 2×2. Si S es un paralelogramo en $\mathbb{R}^2$, entonces

$$\{\text{área de } T(S)\} = |\det A| \cdot \{\text{área de } S\} \tag{5}$$

Si T está determinada por una matriz A de 3×3, y si S es un paralelepípedo en $\mathbb{R}^3$, entonces

$$\{\text{volumen de } T(S)\} = |\det A| \cdot \{\text{volumen de } S\} \tag{6}$$

DEMOSTRACIÓN Considere el caso de 2×2, con $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$. Un paralelogramo en el origen en $\mathbb{R}^2$ definido por los vectores $\mathbf{b}_1$ y $\mathbf{b}_2$ tiene la forma

$$S = \{s_1\mathbf{b}_1 + s_2\mathbf{b}_2 : 0 \le s_1 \le 1,\ 0 \le s_2 \le 1\}$$

La imagen de S bajo T consiste en puntos de la forma

$$\begin{aligned} T(s_1\mathbf{b}_1 + s_2\mathbf{b}_2) &= s_1T(\mathbf{b}_1) + s_2T(\mathbf{b}_2) \\ &= s_1A\mathbf{b}_1 + s_2A\mathbf{b}_2 \end{aligned}$$

donde $0 \le s_1 \le 1$, $0 \le s_2 \le 1$. De ello se sigue que $T(S)$ es el paralelogramo determinado por las columnas de la matriz $[A\mathbf{b}_1 \ \ A\mathbf{b}_2]$. Esta matriz se puede escribir como AB, donde $B = [\mathbf{b}_1 \ \ \mathbf{b}_2]$. De acuerdo con el teorema 9 y el teorema del producto para determinantes,

$$\begin{aligned} \{\text{área de } T(S)\} &= |\det AB| = |\det A| \cdot |\det B| \\ &= |\det A| \cdot \{\text{área de } S\} \end{aligned} \tag{7}$$

Un paralelogramo arbitrario tiene la forma $\mathbf{p} + S$, donde $\mathbf{p}$ es un vector y S es un paralelogramo en el origen, como antes. Es fácil ver que T transforma a $\mathbf{p} + S$ en $T(\mathbf{p}) + T(S)$. (Véase el ejercicio 26). Puesto que la traslación no afecta el área de un conjunto,

$$\begin{aligned} \{\text{área de } T(\mathbf{p} + S\} &= \{\text{área de } T(\mathbf{p}) + T(S)\} \\ &= \{\text{área de } T(S)\} && \text{Traslación} \\ &= |\det A| \cdot \{\text{área de } S\} && \text{Por la ecuación (7)} \\ &= |\det A| \cdot \{\text{área de } \mathbf{p} + S\} && \text{Traslación} \end{aligned}$$

Esto demuestra que (5) es válida para todos los paralelogramos en $\mathbb{R}^2$. Es análoga la demostración de (6) para el caso 3×3. ■

Cuando se intenta generalizar el teorema 10 a una región en $\mathbb{R}^2$ o $\mathbb{R}^3$ que no está acotada por líneas rectas o planos, se debe enfrentar el problema de cómo definir y calcular su área o volumen. Este es un tema estudiado en cálculo, y solamente se indicará la idea básica para $\mathbb{R}^2$. Si R es una región plana que tiene área finita, entonces R se puede aproximar por una rejilla de pequeños cuadrados que están dentro de R. Haciendo los cuadrados suficientemente pequeños, el área de R se puede aproximar tanto como se quiera mediante la suma de las áreas de los pequeños cuadrados. Véase la figura 6.

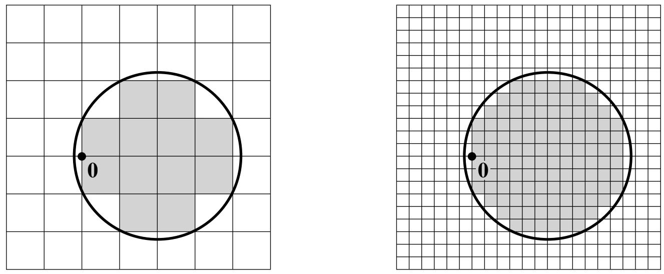


FIGURA 6 Aproximación de una región plana mediante la unión de cuadrados. La aproximación mejora conforme la rejilla se hace más fina.

Si T es una transformación lineal asociada con una matriz A de 2×2, entonces la imagen de la región plana R bajo T se aproxima mediante las imágenes de los pequeños cuadrados dentro de R. La demostración del teorema 10 señala que cada imagen es un paralelogramo cuya área es $|\det A|$ por el área del cuadrado. Si R' es la unión de los cuadrados dentro de R, entonces el área de $T(R')$ es $|\det A|$ por el área de R'. Véase la figura 7. También, el área de $T(R')$ es cercana al área de $T(R)$. Se puede dar un argumento que implique un proceso de cálculo de un límite, para justificar la siguiente generalización del teorema 10.

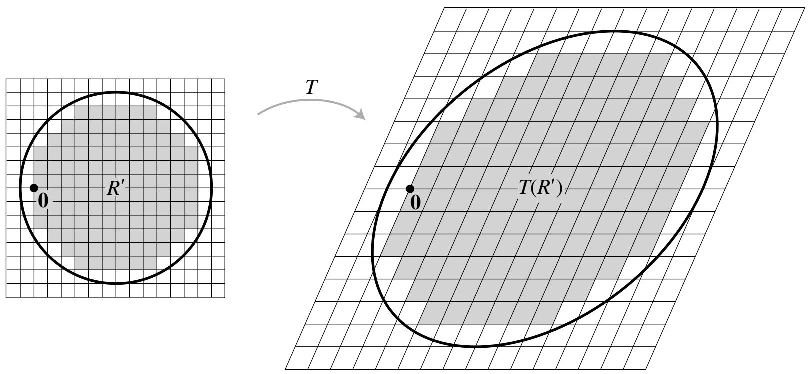
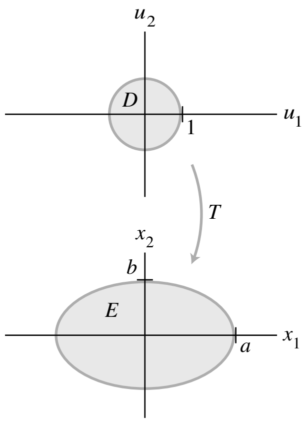


FIGURA 7 Aproximación de $T(R)$ mediante la unión de paralelogramos.

> Las conclusiones del teorema 10 son válidas si S es una región en $\mathbb{R}^2$ con área finita o una región en $\mathbb{R}^3$ con volumen finito.

EJEMPLO 5 Sean a y b números positivos. Encuentre el área de la región E acotada por la elipse cuya ecuación es

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$



SOLUCIÓN Se afirma que E es la imagen del disco unitario D bajo la transformación lineal T determinada por la matriz $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, porque si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $\mathbf{x} = A\mathbf{u}$, entonces

$$u_1 = \frac{x_1}{a} \quad \text{y} \quad u_2 = \frac{x_2}{b}$$

De ello se sigue que $\mathbf{u}$ está en el disco unitario, con $u_1^2 + u_2^2 \leq 1$, si y solo si $\mathbf{x}$ está en E, con $(x_1/a)^2 + (x_2/b)^2 \leq 1$. Al generalizar el teorema 10,

$$\begin{aligned} \{\text{área de la elipse}\} &= \{\text{área de } T(D)\} \\ &= |\det A| \cdot \{\text{área de } D\} \\ &= ab \cdot \pi(1)^2 = \pi ab \end{aligned}$$

■

PROBLEMA DE PRÁCTICA

Sea S el paralelogramo definido por los vectores $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, y sea $A = \begin{bmatrix} 1 & -.1 \\ 0 & 2 \end{bmatrix}$. Calcule el área de la imagen de S bajo el mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

3.3 EJERCICIOS

En los ejercicios 1 a 6, utilice la regla de Cramer para calcular las soluciones de los sistemas.

1. $\begin{aligned} 5x_1 + 7x_2 &= 3 \\ 2x_1 + 4x_2 &= 1 \end{aligned}$

2. $\begin{aligned} 4x_1 + x_2 &= 6 \\ 5x_1 + 2x_2 &= 7 \end{aligned}$

3. $\begin{aligned} 3x_1 - 2x_2 &= 7 \\ -5x_1 + 6x_2 &= -5 \end{aligned}$

4. $\begin{aligned} -5x_1 + 3x_2 &= 9 \\ 3x_1 - x_2 &= -5 \end{aligned}$

5. $\begin{aligned} 2x_1 + x_2 &= 7 \\ -3x_1 + x_3 &= -8 \\ x_2 + 2x_3 &= -3 \end{aligned}$

6. $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ -x_1 + 2x_3 &= 2 \\ 3x_1 + x_2 + 3x_3 &= -2 \end{aligned}$

En los ejercicios 7 a 10, determine los valores del parámetro s para los cuales el sistema tiene una solución única, y describa la solución.

7. $\begin{aligned} 6sx_1 + 4x_2 &= 5 \\ 9x_1 + 2sx_2 &= -2 \end{aligned}$

8. $\begin{aligned} 3sx_1 - 5x_2 &= 3 \\ 9x_1 + 5sx_2 &= 2 \end{aligned}$

9. $\begin{aligned} sx_1 - 2sx_2 &= -1 \\ 3x_1 + 6sx_2 &= 4 \end{aligned}$

10. $\begin{aligned} 2sx_1 + x_2 &= 1 \\ 3sx_1 + 6sx_2 &= 2 \end{aligned}$

En los ejercicios 11 a 16, calcule la adjunta de la matriz dada, y utilice el teorema 8 para dar la inversa de la matriz.

11. $\begin{bmatrix} 0 & -2 & -1 \\ 3 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 1 & 3 \\ 2 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

13. $\begin{bmatrix} 3 & 5 & 4 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 3 & 6 & 7 \\ 0 & 2 & 1 \\ 2 & 3 & 4 \end{bmatrix}$

15. $\begin{bmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 3 & 2 \end{bmatrix}$

16. $\begin{bmatrix} 1 & 2 & 4 \\ 0 & -3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

17. Demuestre que si A es 2×2, entonces el teorema 8 da la misma fórmula para A^{-1} que el teorema 4 de la sección 2.2.

18. Suponga que todas las entradas en A son enteros y $\det A = 1$. Explique por qué todas las entradas en A^{-1} son enteros.

En los ejercicios 19 a 22, encuentre el área del paralelogramo cuyos vértices se indican.

19. (0, 0), (5, 2), (6, 4), (11, 6)

20. (0, 0), (−1, 3), (4, −5), (3, −2)

21. (−1, 0), (0, 5), (1, −4), (2, 1)

22. (0, −2), (6, −1), (−3, 1), (3, 2)

23. Encuentre el volumen del paralelepípedo con un vértice en el origen y vértices adyacentes en (1, 0, −2), (1, 2, 4) y (7, 1, 0).

24. Encuentre el volumen del paralelepípedo con un vértice en el origen y vértices adyacentes en (1, 4, 0), (−2, −5, 2) y (−1, 2, −1).

25. Utilice el concepto de volumen para explicar por qué el determinante de una matriz A de 3×3 es cero si y solo si A no es invertible. No recurra al teorema 4 de la sección 3.2. [*Sugerencia:* Piense en las columnas de A].

26. Sean $T : \mathbb{R}^m \to \mathbb{R}^n$ una transformación lineal, $\mathbf{p}$ un vector y S un conjunto en $\mathbb{R}^m$. Demuestre que la imagen de $\mathbf{p} + S$ bajo T es el conjunto trasladado $T(\mathbf{p}) + T(S)$ en $\mathbb{R}^n$.

27. Sea S el paralelogramo determinado por los vectores $\mathbf{b}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, y sea $A = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Calcule el área de la imagen de S bajo el mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

28. Repita el ejercicio 27 con $\mathbf{b}_1 = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ y $A = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$.

29. Encuentre una fórmula para el área del triángulo cuyos vértices son $\mathbf{0}$, $\mathbf{v}_1$, y $\mathbf{v}_2$ en $\mathbb{R}^2$.

30. Sea R el triángulo con vértices en (x_1, y_1), (x_2, y_2) y (x_3, y_3). Demuestre que

$$\{\text{área del triángulo}\} = \frac{1}{2}\det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$$

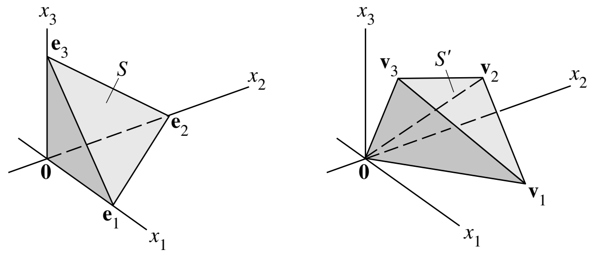
[*Sugerencia:* Traslade R al origen restando uno de los vértices, y considere el ejercicio 29].

31. Sea $T : \mathbb{R}^3 \to \mathbb{R}^3$ la transformación lineal determinada por la matriz $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, donde a, b y c son números positivos. Sea S la bola unitaria, cuya superficie frontera tiene la ecuación $x_1^2 + x_2^2 + x_3^2 = 1$.

a) Demuestre que $T(S)$ está acotada por el elipsoide con la ecuación $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} + \dfrac{x_3^2}{c^2} = 1$.

b) Con base en el hecho de que el volumen de la bola unitaria es $4\pi/3$, obtenga el volumen de la región acotada por el elipsoide en el inciso *a*).

32. Sea S el tetraedro en $\mathbb{R}^3$ con vértices en los vectores $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$, y sea S' el tetraedro con vértices en los vectores $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Véase la figura.



a) Describa una transformación lineal que mapee a S sobre S'.

b) Encuentre una fórmula para el volumen del tetraedro S' considerando el hecho de que

$\{\text{volumen de } S\} = (1/3)\{\text{área de la base}\} \cdot \{\text{altura}\}$

33. [**M**] Pruebe la fórmula del teorema 8 para la inversa con una matriz aleatoria A de 4×4. Utilice un programa de matrices para calcular los cofactores de las submatrices 3×3, construya la adjunta, y establezca que $B = (\text{adj}\, A)/(\det A)$. Luego, encuentre $B - \text{inv}(A)$, donde $\text{inv}(A)$ es la inversa de A calculada por el programa de matrices. Utilice aritmética de punto flotante con el número máximo posible de lugares decimales. Informe sus resultados.

34. [**M**] Pruebe la regla de Cramer con una matriz A de 4×4 y un vector aleatorio $\mathbf{b}$, 4×1. Calcule cada entrada en la solución de $A\mathbf{x} = \mathbf{b}$, y compare esas entradas con las correspondientes en $A^{-1}\mathbf{b}$. Escriba el comando (u oprima las teclas) para su programa de matrices que emplea la regla de Cramer para producir la segunda entrada de $\mathbf{x}$.

35. [**M**] Si su versión de MATLAB tiene el comando `flops`, úselo para contar el número de operaciones de punto flotante para calcular A^{-1} considerando una matriz aleatoria de 30×30. Compare este número con el número de flops necesarios para construir $(\text{adj}\, A)/(\det A)$.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

El área de S es $\left|\det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right| = 14$, y $\det A = 2$. De acuerdo con el teorema 10, el área de la imagen de S bajo el mapeo $\mathbf{x} \mapsto A\mathbf{x}$ es

$$|\det A| \cdot \{\text{área de } S\} = 2 \cdot 14 = 28$$

CAPÍTULO 3 EJERCICIOS COMPLEMENTARIOS

1. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. Suponga que todas las matrices son cuadradas.

a) Si A es una matriz de 2×2 con determinante cero, entonces una columna de A es un múltiplo de la otra.

b) Si dos filas de una matriz A de 3×3 son iguales, entonces $\det A = 0$.

c) Si A es una matriz de 3×3, entonces $\det 5A = 5 \det A$.

d) Si A y B son matrices de $n \times n$, con $\det A = 2$ y $\det B = 3$, entonces $\det(A + B) = 5$.

e) Si A es de $n \times n$ con $\det A = 2$, entonces $\det A^3 = 6$.

f) Si B se obtiene al intercambiar dos filas de A, entonces $\det B = \det A$.

g) Si B se obtiene multiplicando la fila 3 de A por 5, entonces $\det B = 5 \cdot \det A$.

h) Si B se forma sumando a una fila de A una combinación lineal de las otras filas, entonces $\det B = \det A$.

i) $\det A^T = -\det A$.

j) $\det(-A) = -\det A$.

k) $\det(A^TA) \geq 0$.

l) Cualquier sistema de n ecuaciones lineales en n variables se puede resolver con la regla de Cramer.

m) Si $\mathbf{u}$ y $\mathbf{v}$ están en $\mathbb{R}^2$ y $\det[\mathbf{u} \quad \mathbf{v}] = 10$, entonces el área del triángulo en el plano con vértices en $\mathbf{0}$, $\mathbf{u}$ y $\mathbf{v}$ es 10.

n) Si $A^3 = 0$, entonces $\det A = 0$.

o) Si A es invertible, entonces $\det A^{-1} = \det A$.

p) Si A es invertible, entonces $(\det A)(\det A^{-1}) = 1$.

En los ejercicios 2 a 4, utilice operaciones de fila para demostrar que todos los determinantes son cero.

2. $\begin{vmatrix} 12 & 13 & 14 \\ 15 & 16 & 17 \\ 18 & 19 & 20 \end{vmatrix}$ **3.** $\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix}$

4. $\begin{vmatrix} a & b & c \\ a+x & b+x & c+x \\ a+y & b+y & c+y \end{vmatrix}$

En los ejercicios 5 y 6, calcule los determinantes.

5. $\begin{vmatrix} 9 & 1 & 9 & 9 & 9 \\ 9 & 0 & 9 & 9 & 2 \\ 4 & 0 & 0 & 5 & 0 \\ 9 & 0 & 3 & 9 & 0 \\ 6 & 0 & 0 & 7 & 0 \end{vmatrix}$

6. $\begin{vmatrix} 4 & 8 & 8 & 8 & 5 \\ 0 & 1 & 0 & 0 & 0 \\ 6 & 8 & 8 & 8 & 7 \\ 0 & 8 & 8 & 3 & 0 \\ 0 & 8 & 2 & 0 & 0 \end{vmatrix}$

7. Demuestre que la ecuación de la recta en $\mathbb{R}^2$ que pasa por los puntos distintos (x_1, y_1) y (x_2, y_2) se puede escribir como

$$\det \begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix} = 0$$

8. Encuentre una ecuación de determinantes de 3×3, similar a la del ejercicio 7, que describa la ecuación de la recta con pendiente m y que pasa por (x_1, y_1).

Los ejercicios 9 y 10 se relacionan con los determinantes de las siguientes *matrices de Vandermonde*.

$$T = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}, \quad V(t) = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{bmatrix}$$

9. Utilice operaciones de fila para demostrar que

$\det T = (b-a)(c-a)(c-b)$

10. Sea $f(t) = \det V$, con x_1, x_2, x_3 todos distintos. Explique por qué $f(t)$ es un polinomio cúbico, demuestre que el coeficiente de t^3 es diferente de cero, y encuentre tres puntos sobre la gráfica de f.

11. Calcule el área del paralelogramo definido por los puntos $(1, 4)$, $(-1, 5)$, $(3, 9)$ y $(5, 8)$. ¿Cómo comprobar que el cuadrilátero determinado por estos puntos es realmente un paralelogramo?

12. Con base en el concepto de área de un paralelogramo, escriba un enunciado sobre una matriz A de 2×2, que sea válido si y solo si A es invertible.

13. Demuestre que si A es invertible, entonces adj A es invertible, y

$$(\operatorname{adj} A)^{-1} = \frac{1}{\det A} A$$

[*Sugerencia:* Dadas las matrices B y C, ¿con qué cálculo(s) se demostraría que C es la inversa de B?].

14. Sean A, B, C, D e I matrices de $n \times n$. Utilice la definición o las propiedades de un determinante para justificar las siguientes fórmulas. El inciso *c*) es útil en aplicaciones de valores propios (capítulo 5).

a) $\det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = \det A$ *b*) $\det \begin{bmatrix} I & 0 \\ C & D \end{bmatrix} = \det D$

c) $\det \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = (\det A)(\det D) = \det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$

15. Sean A, B, C y D matrices de $n \times n$ con A invertible.

a) Encuentre las matrices X y Y para producir el bloque de factorización LU

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & Y \end{bmatrix}$$

y entonces demuestre que

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (\det A) \cdot \det(D - CA^{-1}B)$$

b) Demuestre que si $AC = CA$, entonces

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB)$$

16. Sea J la matriz de $n \times n$ con solo números 1, y considere $A = (a-b)I + bJ$; es decir,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Confirme que $\det A = (a-b)^{n-1}[a + (n-1)b]$ como sigue:

a) Reste la fila 2 de la fila 1, la fila 3 de la fila 2, y así sucesivamente, y explique por qué esto no cambia el determinante de la matriz.

b) Con la matriz resultante del inciso a), sume la columna 1 a la columna 2, después sume esta nueva columna 2 a la columna 3, y así sucesivamente, y explique por qué esto no cambia el determinante.

c) Encuentre el determinante de la matriz resultante en b).

17. Sea A la matriz original del ejercicio 16, y sean

$$B = \begin{bmatrix} a-b & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \cdots & a \end{bmatrix}$$

$$C = \begin{bmatrix} b & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Observe que A, B y C son muy semejantes excepto que la primera columna de A es igual a la suma de las primeras columnas de B y C. Una *propiedad de linealidad* de la función determinante, analizada en la sección 3.2, dice que $\det A = \det B + \det C$. Utilice este hecho para probar la fórmula del ejercicio 16 mediante inducción sobre el tamaño de la matriz A.

18. [**M**] Aplique el resultado del ejercicio 16 para encontrar los determinantes de las siguientes matrices, y confirme sus respuestas utilizando un programa de matrices.

$$\begin{bmatrix} 3 & 8 & 8 & 8 \\ 8 & 3 & 8 & 8 \\ 8 & 8 & 3 & 8 \\ 8 & 8 & 8 & 3 \end{bmatrix} \qquad \begin{bmatrix} 8 & 3 & 3 & 3 & 3 \\ 3 & 8 & 3 & 3 & 3 \\ 3 & 3 & 8 & 3 & 3 \\ 3 & 3 & 3 & 8 & 3 \\ 3 & 3 & 3 & 3 & 8 \end{bmatrix}$$

19. [**M**] Utilice un programa de matrices para calcular los determinantes de las siguientes matrices.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$$

Utilice los resultados para hacer una conjetura sobre el determinante de la matriz que se presenta a continuación, y confirme la conjetura utilizando operaciones de fila para evaluar ese determinante.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$$

20. [**M**] Aplique el método del ejercicio 19 para hacer una conjetura sobre el determinante de

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 3 & \cdots & 3 \\ 1 & 3 & 6 & \cdots & 6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 6 & \cdots & 3(n-1) \end{bmatrix}$$

Justifique su conjetura. [*Sugerencia:* Considere el ejercicio 14c) y el resultado del ejercicio 19].

4 Espacios vectoriales

EJEMPLO INTRODUCTORIO

Vuelo espacial y sistemas de control

Con sus 12 pisos de altura y un peso de 75 toneladas, el *Columbia* se elevó majestuosamente de la plataforma de lanzamiento en una fresca mañana de Domingo de Ramos en abril de 1981. Producto de 10 años de una intensa labor de investigación y desarrollo, el primer transbordador espacial de Estados Unidos fue un triunfo de diseño de la ingeniería de sistemas de control, en el que participaron varias ramas de la ingeniería: aeronáutica, química, eléctrica, hidráulica y mecánica.

Los sistemas de control de la nave espacial son absolutamente esenciales para el vuelo. Como el transbordador tiene un fuselaje inestable, requiere de una monitorización computarizada constante durante el vuelo atmosférico. El sistema de control de vuelo envía una secuencia de comandos a las superficies de control aerodinámico y a 44 pequeños impulsores de propulsión a chorro. En la figura 1 se muestra un típico sistema de circuito cerrado de retroalimentación que controla el cabeceo del transbordador durante el vuelo. (El cabeceo es el ángulo de elevación del cono de proa). Los símbolos de unión (⊗) indican dónde se añaden las señales de varios sensores a las señales de la computadora que fluyen a través de la parte superior de la figura.

Matemáticamente, las señales de entrada y salida a un sistema de ingeniería son funciones. En las aplicaciones, es importante que estas funciones se puedan sumar, como se muestra en la figura 1, y multiplicarse por escalares. Estas dos operaciones con las funciones tienen propiedades algebraicas que son completamente análogas a las operaciones de suma de vectores en $\mathbb{R}^n$ y a la multiplicación de un vector por un escalar, como se verá en las secciones 4.1 y 4.8. Por esta razón, el conjunto de todas las entradas posibles (funciones) se denomina *espacio vectorial*. Los fundamentos matemáticos para la ingeniería de sistemas se basan en espacios

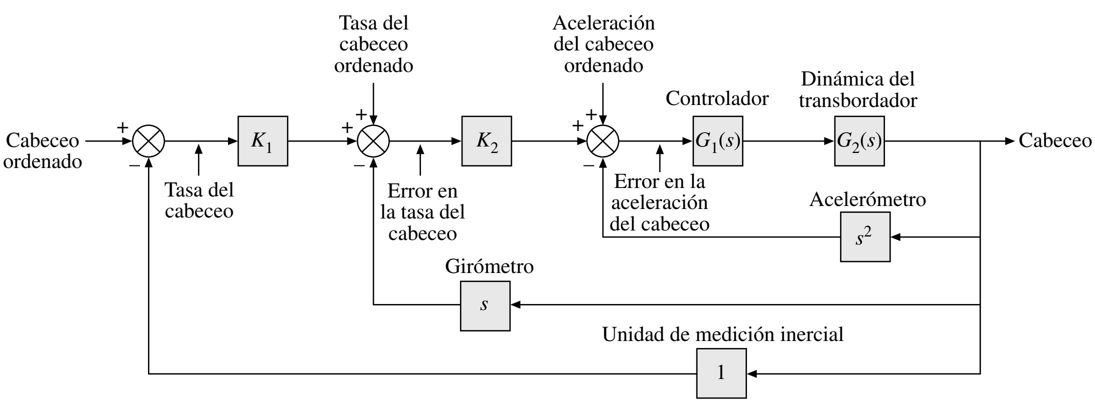


FIGURA 1 Sistema de control del cabeceo para el transbordador espacial. (*Fuente:* Adaptado de *Space Shuttle GN&C Operations Manual*, Rockwell International, © 1988).

vectoriales de funciones, y en este capítulo 4 se amplía la teoría de los vectores en $\mathbb{R}^n$ para incluir dichas funciones. Más adelante, veremos cómo surgen otros espacios vectoriales en ingeniería, física y estadística.

WEB

Las "semillas" matemáticas sembradas en los capítulos 1 y 2 germinan y comienzan a florecer en este capítulo. La belleza y el poder del álgebra lineal se apreciarán más claramente cuando el lector considere a $\mathbb{R}^n$ solo como uno de tantos espacios vectoriales que surgen de forma natural en problemas aplicados. En realidad, el estudio de espacios vectoriales no es muy diferente del propio estudio de $\mathbb{R}^n$, porque usted podrá utilizar su experiencia geométrica con $\mathbb{R}^2$ y $\mathbb{R}^3$ para visualizar muchos conceptos generales.

Comenzando con las definiciones básicas en la sección 4.1, el marco teórico de los espacios vectoriales se desarrolla gradualmente a lo largo del capítulo. Uno de los objetivos de las secciones 4.3 a 4.5 es demostrar cuán estrechamente se parecen otros espacios vectoriales a $\mathbb{R}^n$. En la sección 4.6 se estudiará el rango, uno de los temas más importantes de este capítulo, utilizando la terminología de espacio vectorial para vincular datos importantes de las matrices rectangulares. En la sección 4.8 se aplica la teoría del capítulo a las señales discretas y ecuaciones en diferencias utilizadas en los sistemas de control digital, como en el transbordador espacial. Las cadenas de Markov, en la sección 4.9, representan un cambio de ritmo respecto de las secciones más teóricas del capítulo y dan buenos ejemplos para los conceptos que se introducirán en el capítulo 5.

4.1 ESPACIOS Y SUBESPACIOS VECTORIALES

Gran parte de la teoría en los capítulos 1 y 2 se basa en ciertas propiedades algebraicas sencillas y evidentes de $\mathbb{R}^n$, que se listan en la sección 1.3. De hecho, muchos otros sistemas matemáticos tienen las mismas propiedades. Las propiedades específicas de interés se incluyen en la siguiente definición.

DEFINICIÓN

Un **espacio vectorial** es un conjunto no vacío V de objetos, llamados *vectores*, en el que están definidas dos operaciones, llamadas *suma* y *multiplicación por escalares* (números reales), sujetas a los 10 axiomas (o reglas) que se listan a continuación.[1] Los axiomas deben ser válidos para todos los vectores $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ en V, y para todos los escalares c y d.

1. La suma de $\mathbf{u}$ y $\mathbf{v}$, denotada con $\mathbf{u} + \mathbf{v}$, está en V.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. Hay un vector **cero** ($\mathbf{0}$) en V tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. Para cada $\mathbf{u}$ en V, existe un vector $-\mathbf{u}$ en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. El múltiplo escalar de $\mathbf{u}$ por c, que se denota con $c\mathbf{u}$, está en V.
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

[1] Técnicamente, V es un *espacio vectorial real*. Toda la teoría en este capítulo también es válida para un *espacio vectorial complejo* en el que los escalares son números complejos. Trataremos brevemente este asunto en el capítulo 5. Hasta entonces, se supone que todos los escalares son reales.

Utilizando tan solo estos axiomas, es posible demostrar que el vector cero en el axioma 4 es único, y el vector $-\mathbf{u}$, llamado el **negativo** de **u**, en el axioma 5 es único para toda **u** en V. Véase los ejercicios 25 y 26. Demostraciones de los hechos que se presentan a continuación se describen en los ejercicios:

Para cada **u** en V y escalar c,

$$0\mathbf{u} = \mathbf{0} \tag{1}$$

$$c\mathbf{0} = \mathbf{0} \tag{2}$$

$$-\mathbf{u} = (-1)\mathbf{u} \tag{3}$$

EJEMPLO 1 Los espacios $\mathbb{R}^n$, donde $n \geq 1$, son los ejemplos principales de espacios vectoriales. La intuición geométrica desarrollada para $\mathbb{R}^3$ le ayudará a entender y visualizar muchos conceptos en todo el capítulo. ■

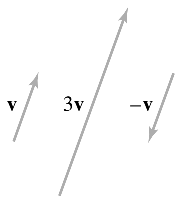


FIGURA 1

EJEMPLO 2 Sea V el conjunto de todas las flechas (segmentos de recta dirigidos) en el espacio tridimensional; se considera que dos flechas son iguales si tienen la misma longitud y la punta en la misma dirección. Defina la suma por la regla del paralelogramo (de la sección 1.3), y para cada **v** en V, defina $c\mathbf{v}$ como la flecha cuya longitud es $|c|$ por la longitud de **v**, apuntando en la misma dirección que **v** si $c \geq 0$ y, en caso contrario, apuntando en la dirección opuesta. (Véase la figura 1). Demuestre que V es un espacio vectorial. Este espacio es un modelo común de los problemas físicos de varias fuerzas.

SOLUCIÓN La definición de V es geométrica, utilizando los conceptos de longitud y dirección. Ningún sistema de coordenadas xyz está implicado. Una flecha de longitud cero es un punto único y representa el vector cero. El negativo de **v** es $(-1)\mathbf{v}$. Por lo tanto, los axiomas 1, 4, 5, 6 y 10 son evidentes. El resto se comprobará con geometría. Por ejemplo, véase las figuras 2 y 3. ■

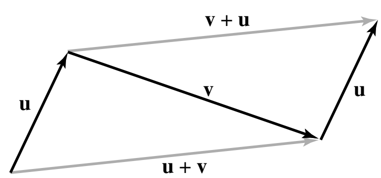


FIGURA 2 $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

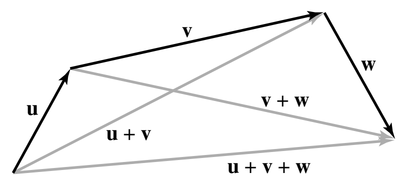


FIGURA 3 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

EJEMPLO 3 Sea $\mathbb{S}$ el espacio de todas las secuencias doblemente infinitas de números (que generalmente se anotan en una fila más que en una columna):

$$\{y_k\} = (\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots)$$

Si $\{z_k\}$ es otro elemento de $\mathbb{S}$, entonces la suma $\{y_k\} + \{z_k\}$ es la sucesión $\{y_k + z_k\}$ formada por la suma de los términos correspondientes de $\{y_k\}$ y $\{z_k\}$. El múltiplo escalar $c\{y_k\}$ es la sucesión $\{cy_k\}$. Los axiomas de espacio vectorial se comprueban en la misma forma que para $\mathbb{R}^n$.

Los elementos de $\mathbb{S}$ surgen en ingeniería, por ejemplo, cuando una señal se mide (o se muestrea) en tiempos discretos. Una señal puede ser eléctrica, mecánica, óptica, etcétera. Los sistemas de control principal del transbordador espacial, que se mencionaron en la introducción del capítulo, usan señales discretas (o digitales). Por conveniencia, llamaremos a $\mathbb{S}$ el espacio de **señales** (discretas de tiempo). Una señal se puede visualizar con una gráfica como la que se ilustra en la figura 4. ■

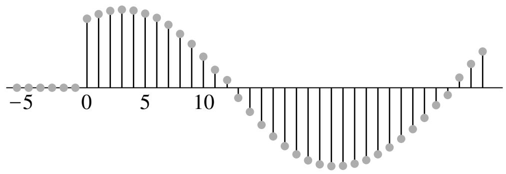


FIGURA 4 Una señal discreta de tiempo.

EJEMPLO 4 Para $n \geq 0$, el conjunto $\mathbb{P}_n$ de polinomios de grado n o menor consiste en todos los polinomios de la forma

$$\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n \tag{4}$$

donde los coeficientes $a_0, \ldots, a_n$ y la variable t son números reales. El *grado* de $\mathbf{p}$ es la mayor potencia de t en (4) cuyo coeficiente no es cero. Si $\mathbf{p}(t) = a_0 \neq 0$, el grado de $\mathbf{p}$ es cero. Si todos los coeficientes son iguales a cero, $\mathbf{p}$ se llama el *polinomio cero*. El polinomio cero está incluido en $\mathbb{P}_n$, a pesar de que su grado, por razones técnicas, no esté definido.

Si $\mathbf{p}$ está dada por la ecuación (4) y si $\mathbf{q}(t) = b_0 + b_1 t + \cdots + b_n t^n$, entonces la suma $\mathbf{p} + \mathbf{q}$ se define mediante

$$\begin{aligned}(\mathbf{p} + \mathbf{q})(t) &= \mathbf{p}(t) + \mathbf{q}(t) \\ &= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n\end{aligned}$$

El múltiplo escalar $c\mathbf{p}$ es el polinomio definido por

$$(c\mathbf{p})(t) = c\mathbf{p}(t) = ca_0 + (ca_1)t + \cdots + (ca_n)t^n$$

Estas definiciones satisfacen los axiomas 1 y 6, ya que $\mathbf{p} + \mathbf{q}$ y $c\mathbf{p}$ son polinomios de grado igual o menor que n. Los axiomas 2, 3, y 7 a 10 son consecuencias de las propiedades de los números reales. Evidentemente, el polinomio cero actúa como vector cero en el axioma 4. Por último, $(-1)\mathbf{p}$ actúa como el negativo de $\mathbf{p}$, por lo que el axioma 5 está satisfecho. Así, $\mathbb{P}_n$ es un espacio vectorial.

Se utilizan espacios vectoriales $\mathbb{P}_n$ para varias n, por ejemplo, en el análisis de tendencia estadística de datos, que se analiza en la sección 6.8. ■

EJEMPLO 5 Sea V el conjunto de todas las funciones de valores reales definidas en un conjunto $\mathbb{D}$. (Por lo general, $\mathbb{D}$ es el conjunto de números reales o algún intervalo en la recta real). Las funciones se suman en la forma habitual: $\mathbf{f} + \mathbf{g}$ es la función cuyo valor en t en el dominio $\mathbb{D}$ es $\mathbf{f}(t) + \mathbf{g}(t)$. Asimismo, para un escalar c y una $\mathbf{f}$ en V, el múltiplo escalar $c\mathbf{f}$ es la función cuyo valor en t es $c\mathbf{f}(t)$. Por ejemplo, si $\mathbb{D} = \mathbb{R}$, $\mathbf{f}(t) = 1 + \text{sen}\, 2t$, y $\mathbf{g}(t) = 2 + .5t$, entonces

$$(\mathbf{f} + \mathbf{g})(t) = 3 + \text{sen}\, 2t + .5t \quad \text{y} \quad (2\mathbf{g})(t) = 4 + t$$

Dos funciones en V son iguales si y solo si sus valores son iguales para toda t en $\mathbb{D}$. Por lo tanto, el vector cero en V es la función igual a cero, $\mathbf{f}(t) = 0$ para toda t, y el negativo de $\mathbf{f}$ es $(-1)\mathbf{f}$. Como es evidente, los axiomas 1 y 6 son ciertos, y los otros axiomas se deducen de las propiedades de los números reales, por lo que V es un espacio vectorial. ■

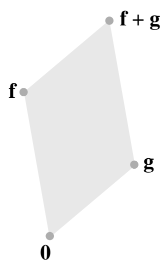


FIGURA 5
La suma de dos vectores (funciones).

Es importante pensar en cada función en el espacio vectorial V del ejemplo 5 como un objeto único, un solo "punto" o un vector en el espacio vectorial. La suma de dos vectores $\mathbf{f}$ y $\mathbf{g}$ (funciones en V, o elementos de *cualquier* espacio vectorial) se puede visualizar como en la figura 5, ya que esto le ayudará a aplicar a un espacio vectorial general la intuición geométrica que ha desarrollado al trabajar con el espacio vectorial $\mathbb{R}^n$. Consulte la *Guía de estudio* como ayuda conforme aprenda a adoptar este punto de vista más general.

Subespacios

En muchos problemas, un espacio vectorial se compone de un subconjunto adecuado de vectores de un espacio vectorial más grande. En este caso, solo se deben comprobar tres de los 10 axiomas de espacio vectorial, y el resto se satisfacen de manera automática.

DEFINICIÓN

Un **subespacio** de un espacio vectorial V es un subconjunto H de V que tiene tres propiedades:

a) El vector cero de V está en H.[2]

b) H es cerrado bajo la suma de vectores. Es decir, por cada $\mathbf{u}$ y $\mathbf{v}$ en H, la suma $\mathbf{u} + \mathbf{v}$ está en H.

c) H es cerrado bajo la multiplicación por escalares. Es decir, para cada $\mathbf{u}$ en H y cada escalar c, el vector $c\mathbf{u}$ está en H.

Las propiedades *a*), *b*) y *c*) garantizan que un subespacio H de V sea en sí mismo un *espacio vectorial* bajo las operaciones de espacio vectorial ya definidas en V. Para comprobar esto, observe que las propiedades *a*), *b*) y *c*) corresponden a los axiomas 1, 4 y 6. Los axiomas 2, 3 y 7 a 10 son automáticamente verdaderos en H, ya que se aplican a todos los elementos de V, incluidos los de H. El axioma 5 también es verdadero en H, ya que si $\mathbf{u}$ está en H, entonces $(-1)\mathbf{u}$ está en H de acuerdo con la propiedad *c*), y sabemos a partir de la ecuación (3) de la página 191 que $(-1)\mathbf{u}$ es el vector $-\mathbf{u}$ en el axioma 5.

Así, cada subespacio es un espacio vectorial. Por el contrario, todo espacio vectorial es un subespacio (de sí mismo y, posiblemente, de otros espacios más grandes). El término *subespacio* se utiliza cuando al menos dos espacios vectoriales están en mente, con uno dentro del otro, y la frase *subespacio de V* identifica a V como el espacio más grande. (Véase la figura 6).

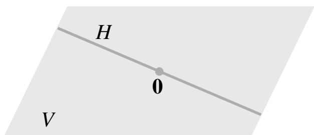


FIGURA 6
Un subespacio de V.

EJEMPLO 6 El conjunto que consta solo del vector cero en un espacio vectorial V es un subespacio de V, llamado **subespacio cero**, y se representa como $\{\mathbf{0}\}$. ■

EJEMPLO 7 Sea $\mathbb{P}$ el conjunto de todos los polinomios con coeficientes reales, con operaciones en $\mathbb{P}$ definidas como en las funciones. De esta forma, $\mathbb{P}$ es un subespacio del espacio de todas las funciones de valores reales definidas en $\mathbb{R}$. Además, para cada $n \geq 0$, $\mathbb{P}_n$ es un subespacio de $\mathbb{P}$, ya que $\mathbb{P}_n$ es un subconjunto de $\mathbb{P}$ que contiene el polinomio cero; la suma de dos polinomios en $\mathbb{P}_n$ también está en $\mathbb{P}_n$, y un múltiplo escalar de un polinomio en $\mathbb{P}_n$ también está en $\mathbb{P}_n$. ■

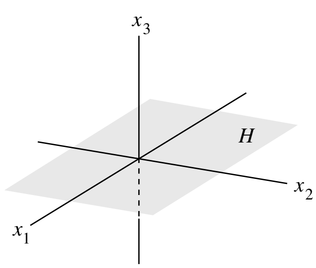


FIGURA 7
El plano x_1x_2 como un subespacio de $\mathbb{R}^3$.

EJEMPLO 8 El espacio vectorial $\mathbb{R}^2$ *no* es un subespacio de $\mathbb{R}^3$ porque $\mathbb{R}^2$ ni siquiera es un subconjunto de $\mathbb{R}^3$. (Todos los vectores en $\mathbb{R}^3$ tienen tres entradas, mientras que los vectores de $\mathbb{R}^2$ tienen solo dos). El conjunto

$$H = \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} : s \text{ y } t \text{ son reales} \right\}$$

es un subconjunto de $\mathbb{R}^3$ que se "ve" y "actúa" como $\mathbb{R}^2$, aunque lógicamente es distinto de $\mathbb{R}^2$. Véase la figura 7. Demuestre que H es un subespacio de $\mathbb{R}^3$.

SOLUCIÓN El vector cero está en H, y H es cerrado bajo la suma de vectores y la multiplicación escalar debido a que estas operaciones sobre los vectores de H siempre producen vectores cuyas terceras entradas son iguales a cero (y, por lo tanto, pertenecen a H). Por consiguiente, H es un subespacio de $\mathbb{R}^3$. ■

[2] Algunos libros remplazan la propiedad *a*) en esta definición por el supuesto de que H no es vacío. Así, *a*) se puede deducir a partir de *c*) y el hecho de que $0\mathbf{u} = \mathbf{0}$. Pero la mejor manera de someter a prueba un subespacio es buscar primero al vector cero. Si $\mathbf{0}$ está en H, entonces se deben revisar las propiedades *b*) y *c*). Si $\mathbf{0}$ *no* está en H, entonces H no puede ser un subespacio y no se necesita comprobar las otras propiedades.

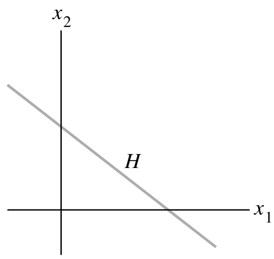


FIGURA 8
Una recta que no es un espacio vectorial.

EJEMPLO 9 Un plano en $\mathbb{R}^3$ que *no* pasa por el origen no es un subespacio de $\mathbb{R}^3$, porque el plano no contiene el vector cero de $\mathbb{R}^3$. Del mismo modo, una recta en $\mathbb{R}^2$ que *no* pasa por el origen, como en la figura 8, *no* es un subespacio de $\mathbb{R}^2$. ■

Un subespacio generado por un conjunto

El siguiente ejemplo ilustra una de las formas más comunes de describir un subespacio. Como en el capítulo 1, el término **combinación lineal** se refiere a cualquier suma de los múltiplos escalares de vectores, y Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ denota el conjunto de todos los vectores que se pueden escribir como combinaciones lineales de $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

EJEMPLO 10 Dados $\mathbf{v}_1$ y $\mathbf{v}_2$ en un espacio vectorial V, sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que H es un subespacio de V.

SOLUCIÓN El vector cero está en H, ya que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Para demostrar que H es cerrado bajo la suma de vectores, tome dos vectores arbitrarios en H, por ejemplo,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{w} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Por los axiomas 2, 3 y 8 para el espacio vectorial V,

$$\begin{aligned}\mathbf{u} + \mathbf{w} &= (s_1\mathbf{v}_1 + s_2\mathbf{v}_2) + (t_1\mathbf{v}_1 + t_2\mathbf{v}_2)\\ &= (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2\end{aligned}$$

Así que $\mathbf{u} + \mathbf{w}$ está en H. Además, si c es un escalar cualquiera, entonces por los axiomas 7 y 9,

$$c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$$

lo que demuestra que $c\mathbf{u}$ está en H, y H es cerrado bajo la multiplicación por escalares. Por lo tanto, H es un subespacio de V. ■

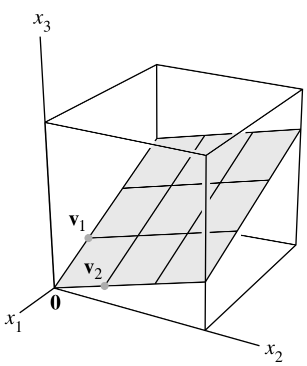


FIGURA 9
Un ejemplo de un subespacio.

En la sección 4.5, veremos que todo subespacio de $\mathbb{R}^3$ distinto de cero, que no sea $\mathbb{R}^3$ mismo, es Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ para algunos $\mathbf{v}_1$ y $\mathbf{v}_2$ linealmente independientes o Gen $\{\mathbf{v}\}$ para $\mathbf{v} \neq \mathbf{0}$. En el primer caso, el subespacio es un plano que pasa por el origen; en el segundo caso, se trata de una recta que pasa por el origen. (Véase la figura 9). Es útil tener en mente estas imágenes geométricas, incluso para un espacio vectorial abstracto.

El argumento en el ejemplo 10 se puede generalizar fácilmente para comprobar el siguiente teorema.

TEOREMA 1

> Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en un espacio vectorial V, entonces Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es un subespacio de V.

Se llama Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ al **subespacio generado** por $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Dado un subespacio H de V, un **conjunto generador** para H es un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en H tal que $H = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

El siguiente ejemplo muestra cómo utilizar el teorema 1.

EJEMPLO 11 Sea H el conjunto de todos los vectores de la forma $(a - 3b, b - a, a, b)$, donde a y b son escalares arbitrarios. Es decir, sea $H = \{(a - 3b, b - a, a, b)\text{: } a \text{ y } b \text{ en } \mathbb{R}\}$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

SOLUCIÓN Represente los vectores en H como vectores columna. Así, un vector arbitrario en H tiene la forma

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_1} + b\underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_2}$$

Este cálculo muestra que $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1$ y $\mathbf{v}_2$ son los vectores indicados anteriormente. Por lo tanto, H es un subespacio de $\mathbb{R}^4$ de acuerdo con el teorema 1. ■

El ejemplo 11 muestra una técnica útil para expresar un subespacio H como el conjunto de combinaciones lineales de una pequeña colección de vectores. Si $H = \text{Gen}\,\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$, podemos pensar en los vectores $\mathbf{v}_1,\ldots, \mathbf{v}_p$ del conjunto generador como "asas" que nos permiten sujetar el subespacio H. Cálculos con un número infinito de vectores en H con frecuencia se reducen a operaciones con un número finito de vectores en el conjunto generador.

EJEMPLO 12 Determine qué valor(es) de h hará(n) que $\mathbf{y}$ sea un subespacio de $\mathbb{R}^3$ generado por $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, si

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

SOLUCIÓN Esta pregunta es el problema de práctica 2 en la sección 1.3, que aquí se modificó utilizando el término *subespacio* en vez de Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. La solución obtenida ahí muestra que $\mathbf{y}$ está en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ si y solo si $h = 5$. Vale la pena revisar esta solución ahora, junto con los ejercicios 11 a 16 y 19 a 21 de la sección 1.3. ■

Aunque muchos espacios vectoriales en este capítulo son subespacios de $\mathbb{R}^n$, es importante considerar que la teoría abstracta se aplica a otros espacios vectoriales. Espacios vectoriales de funciones surgen en muchas aplicaciones, y más adelante se tratarán con más detalle.

PROBLEMAS DE PRÁCTICA

1. Demuestre que el conjunto H de todos los puntos en $\mathbb{R}^2$ de la forma $(3s, 2 + 5s)$ no es un espacio vectorial, al mostrar que no es cerrado bajo la multiplicación escalar. (Encuentre un vector específico $\mathbf{u}$ en H y un escalar c tal que $c\mathbf{u}$ no está en H).
2. Sea $W = \text{Gen}\,\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$, donde $\mathbf{v}_1,\ldots, \mathbf{v}_p$ se encuentran en un espacio vectorial V. Demuestre que $\mathbf{v}_k$ está en W para $1 \leq k \leq p$. [*Sugerencia:* Escriba primero una ecuación que demuestre que $\mathbf{v}_1$ está en W. Después, ajuste su notación para el caso general].

WEB

4.1 EJERCICIOS

1. Sea V el primer cuadrante en el plano xy; es decir, sea
$$V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0, y \geq 0 \right\}$$
   a) Si $\mathbf{u}$ y $\mathbf{v}$ están en V, ¿está $\mathbf{u} + \mathbf{v}$ en V? ¿Por qué?
   b) Encuentre un vector específico $\mathbf{u}$ en V y un escalar específico c tal que $c\mathbf{u}$ *no* esté en V. (Esto es suficiente para demostrar que V *no* es un espacio vectorial).

2. Sea W la unión del primer y tercer cuadrantes en el plano xy. Es decir, sea $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : xy \geq 0 \right\}$.
   a) Si $\mathbf{u}$ está en W y c es cualquier escalar, ¿está $c\mathbf{u}$ en W? ¿Por qué?

b) Encuentre dos vectores específicos $\mathbf{u}$ y $\mathbf{v}$ tales que $\mathbf{u} + \mathbf{v}$ no esté en W. Esto es suficiente para demostrar que W *no* es un espacio vectorial.

3. Sea H el conjunto de puntos en el interior y sobre el círculo unitario en el plano xy. Es decir, sea $H = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + y^2 \leq 1 \right\}$. Encuentre un ejemplo específico —dos vectores o un vector y un escalar—, que demuestre que H no es un subespacio de $\mathbb{R}^2$.

4. Construya una figura geométrica que ilustre por qué una recta en $\mathbb{R}^2$ que *no* pasa por el origen no es cerrada bajo la suma de vectores.

En los ejercicios 5 a 8, determine si el conjunto dado es un subespacio de $\mathbb{P}_n$ para un valor adecuado de n. Justifique sus respuestas.

5. Todos los polinomios de la forma $\mathbf{p}(t) = at^2$, donde a se encuentra en $\mathbb{R}$.

6. Todos los polinomios de la forma $\mathbf{p}(t) = a + t^2$, donde a se encuentra en $\mathbb{R}$.

7. Todos los polinomios de grado 3 como máximo, con números enteros como coeficientes.

8. Todos los polinomios en $\mathbb{P}_n$ tales que $\mathbf{p}(0) = 0$.

9. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} -2t \\ 5t \\ 3t \end{bmatrix}$. Encuentre un vector $\mathbf{v}$ en $\mathbb{R}^3$ tal que $H = \text{Gen}\,\{\mathbf{v}\}$. ¿Por qué esto demuestra que H es un subespacio de $\mathbb{R}^3$?

10. Sea H el conjunto de todos los vectores de la forma $\begin{bmatrix} 3t \\ 0 \\ -7t \end{bmatrix}$, donde t es cualquier número real. Demuestre que H es un subespacio de $\mathbb{R}^3$. (Utilice el método del ejercicio 9).

11. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} 2b + 3c \\ -b \\ 2c \end{bmatrix}$, donde b y c son arbitrarios. Encuentre los vectores $\mathbf{u}$ y $\mathbf{v}$ tales que $W = \text{Gen}\,\{\mathbf{u}, \mathbf{v}\}$. ¿Por qué esto demuestra que W es un subespacio de $\mathbb{R}^3$?

12. Sea W el conjunto de todos los vectores de la forma $\begin{bmatrix} 2s + 4t \\ 2s \\ 2s - 3t \\ 5t \end{bmatrix}$. Demuestre que W es un subespacio de $\mathbb{R}^4$. (Utilice el método del ejercicio 11).

13. Sea $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

a) ¿Está $\mathbf{w}$ en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Cuántos vectores se encuentran en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

b) ¿Cuántos vectores están en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

c) ¿$\mathbf{w}$ es el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

14. Sea $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ como en el ejercicio 13, y sea $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$. ¿Está $\mathbf{w}$ en el subespacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? ¿Por qué?

En los ejercicios 15 a 18, sea W el conjunto de todos los vectores de la forma indicada, donde a, b y c representan números reales arbitrarios. En cada caso, encuentre un conjunto S de vectores que genere W o dé un ejemplo para demostrar que W *no* es un espacio vectorial.

15. $\begin{bmatrix} 2a + 3b \\ -1 \\ 2a - 5b \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 3a - 5b \\ 3b + 2a \end{bmatrix}$

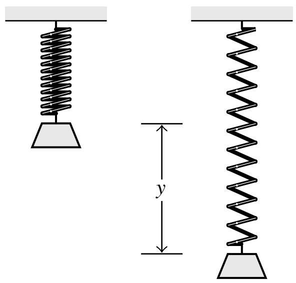
17. $\begin{bmatrix} 2a - b \\ 3b - c \\ 3c - a \\ 3b \end{bmatrix}$

18. $\begin{bmatrix} 4a + 3b \\ 0 \\ a + 3b + c \\ 3b - 2c \end{bmatrix}$

19. Si una masa m se coloca en el extremo de un resorte, y si se tira de la masa hacia abajo y luego se suelta, el sistema masa-resorte comenzará a oscilar. El desplazamiento y de la masa desde su posición de reposo está dado por una función de la forma

$$y(t) = c_1 \cos \omega t + c_2 \operatorname{sen} \omega t \tag{5}$$

donde ω es una constante que depende del resorte y la masa. (Véase la figura que se muestra a continuación). Demuestre que el conjunto de todas las funciones descritas en la ecuación (5) (con ω fija y c_1 y c_2 arbitrarias) es un espacio vectorial.



20. El conjunto de todas las funciones de valor real continuas y definidas en un intervalo cerrado $[a, b]$ en $\mathbb{R}$ se denota mediante $C[a, b]$. Este conjunto es un subespacio del espacio vectorial de todas las funciones de valor real definidas en $[a, b]$.

a) ¿Qué hechos acerca de las funciones continuas se deberían someter a prueba para demostrar que $C[a, b]$ es en realidad el subespacio que se afirma? (Estos hechos se suelen tratar en una clase de cálculo).

b) Demuestre que $\{\mathbf{f} \text{ en } C[a, b] : \mathbf{f}(a) = \mathbf{f}(b)\}$ es un subespacio de $C[a, b]$.

Para enteros positivos fijos m y n, el conjunto $M_{m\times n}$ de todas las matrices de $m \times n$ es un espacio vectorial, bajo las operaciones habituales de suma de matrices y multiplicación por escalares reales.

21. Determine si el conjunto H de todas las matrices de la forma $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ es un subespacio de $M_{2\times 2}$.

22. Sea F una matriz fija de 3×2, y sea H el conjunto de todas las matrices A en $M_{2\times 4}$ con la propiedad de que $FA = 0$ (la matriz cero en $M_{3\times 4}$). Determine si H es un subespacio de $M_{2\times 4}$.

En los ejercicios 23 y 24, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) Si $\mathbf{f}$ es una función en el espacio vectorial V de todas las funciones de valores reales en $\mathbb{R}$ y si $\mathbf{f}(t) = 0$ para alguna t, entonces $\mathbf{f}$ es el vector cero en V.

b) Un vector es una flecha en el espacio tridimensional.

c) Un subconjunto H de un espacio vectorial V es un subespacio de V si el vector cero está en H.

d) Un subespacio también es un espacio vectorial.

e) Las señales analógicas se utilizan en los sistemas de control principal del transbordador espacial, mencionado en la introducción al capítulo.

24. *a*) Un vector es cualquier elemento de un espacio vectorial.

b) Si $\mathbf{u}$ es un vector en un espacio vectorial V, entonces $(-1)\mathbf{u}$ es igual al negativo de $\mathbf{u}$.

c) Un espacio vectorial es un subespacio.

d) $\mathbb{R}^2$ es un subespacio de $\mathbb{R}^3$.

e) Un subconjunto H de un espacio vectorial V es un subespacio de V si se cumplen las siguientes condiciones: **i.** el vector cero de V está en H, **ii.** $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{u} + \mathbf{v}$ en H, y **iii.** c es un escalar y $c\mathbf{u}$ está en H.

Los ejercicios 25 a 29 muestran cómo se pueden utilizar los axiomas de un espacio vectorial V para demostrar las propiedades elementales descritas después de la definición de un espacio vectorial. Complete los espacios en blanco con los números del axioma adecuado. De acuerdo con el axioma 2, los axiomas 4 y 5 implican, respectivamente, que $\mathbf{0} + \mathbf{u} = \mathbf{u}$ y $-\mathbf{u} + \mathbf{u} = \mathbf{0}$ para toda $\mathbf{u}$.

25. Complete la siguiente demostración de que el vector cero es único. Suponga que $\mathbf{w}$ en V tiene la propiedad de que $\mathbf{u} + \mathbf{w} = \mathbf{w} + \mathbf{u} = \mathbf{u}$ para toda $\mathbf{u}$ en V. En particular, $\mathbf{0} + \mathbf{w} = \mathbf{0}$. Sin embargo, $\mathbf{0} + \mathbf{w} = \mathbf{w}$, de acuerdo con el axioma ______. Por lo tanto, $\mathbf{w} = \mathbf{0} + \mathbf{w} = \mathbf{0}$.

26. Complete la siguiente demostración de que $-\mathbf{u}$ es el *único vector* en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. Suponga que $\mathbf{w}$ satisface $\mathbf{u} + \mathbf{w} = \mathbf{0}$. Sumando $-\mathbf{u}$ en ambos lados, se tiene

$(-\mathbf{u}) + [\mathbf{u} + \mathbf{w}] = (-\mathbf{u}) + \mathbf{0}$

$[(-\mathbf{u}) + \mathbf{u}] + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ según el axioma _____ *a*)

$\mathbf{0} + \mathbf{w} = (-\mathbf{u}) + \mathbf{0}$ según el axioma _____ *b*)

$\mathbf{w} = -\mathbf{u}$ según el axioma _____ *c*)

27. Complete los números de axioma que faltan en la siguiente demostración de que $0\mathbf{u} = \mathbf{0}$ para cada $\mathbf{u}$ en V.

$0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$ según el axioma _____ *a*)

Sume el negativo de $0\mathbf{u}$ en ambos lados:

$0\mathbf{u} + (-0\mathbf{u}) = [0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u})$

$0\mathbf{u} + (-0\mathbf{u}) = 0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})]$ según el axioma _____ *b*)

$\mathbf{0} = 0\mathbf{u} + \mathbf{0}$ según el axioma _____ *c*)

$\mathbf{0} = 0\mathbf{u}$ según el axioma _____ *d*)

28. Complete los números de axioma que faltan en la siguiente demostración de que

$c\mathbf{0} = \mathbf{0}$ para cada escalar c.

$c\mathbf{0} = c(\mathbf{0} + \mathbf{0})$ según el axioma _____ *a*)

$= c\mathbf{0} + c\mathbf{0}$ según el axioma _____ *b*)

Sume el negativo de $c\mathbf{0}$ en ambos lados:

$c\mathbf{0} + (-c\mathbf{0}) = [c\mathbf{0} + c\mathbf{0}] + (-c\mathbf{0})$

$c\mathbf{0} + (-c\mathbf{0}) = c\mathbf{0} + [c\mathbf{0} + (-c\mathbf{0})]$ según el axioma _____ *c*)

$\mathbf{0} = c\mathbf{0} + \mathbf{0}$ según el axioma _____ *d*)

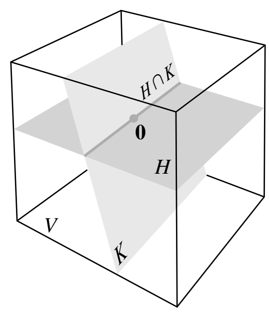
$\mathbf{0} = c\mathbf{0}$ según el axioma _____ *e*)

29. Demuestre que $(-1)\mathbf{u} = -\mathbf{u}$. [*Sugerencia:* Demuestre que $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. Utilice algunos axiomas y los resultados de los ejercicios 27 y 26].

30. Suponga que $c\mathbf{u} = \mathbf{0}$ para algún escalar c distinto de cero. Demuestre que $\mathbf{u} = \mathbf{0}$. Mencione los axiomas o las propiedades que utilice.

31. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio vectorial V, y sea H cualquier subespacio de V que contiene a $\mathbf{u}$ y $\mathbf{v}$. Explique por qué H también contiene a Gen $\{\mathbf{u}, \mathbf{v}\}$. Esto demuestra que Gen $\{\mathbf{u}, \mathbf{v}\}$ es el menor subespacio de V que contiene a $\mathbf{u}$ y $\mathbf{v}$.

32. Sean H y K subespacios de un espacio vectorial V. La **intersección** de H y K, que se representa como $H \cap K$, es el conjunto de $\mathbf{v}$ en V que pertenece tanto a H como K. Demuestre que $H \cap K$ es un subespacio de V. (Véase la figura). Dé un ejemplo en $\mathbb{R}^2$ para demostrar que la unión de dos subespacios no es, en general, un subespacio.



33. Dados los subespacios H y K de un espacio vectorial V, la **suma** de H y K, que se escribe como $H + K$, es el conjunto de todos los vectores en V que se pueden representar como la suma de dos vectores, uno en H y el otro en K; es decir,

$$H + K = \{\mathbf{w} : \mathbf{w} = \mathbf{u} + \mathbf{v} \text{ para alguna } \mathbf{u} \text{ en } H \text{ y alguna } \mathbf{v} \text{ en } K\}$$

a) Demuestre que $H + K$ es un subespacio de V.

b) Demuestre que H es un subespacio de $H + K$ y K es un subespacio de $H + K$.

34. Suponga que $\mathbf{u}_1, \ldots, \mathbf{u}_p$ y $\mathbf{v}_1, \ldots, \mathbf{v}_q$ son vectores en un espacio vectorial V, y sea

$H = \text{Gen}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ y $K = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$

Demuestre que $H + K = \text{Gen}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$.

35. [**M**] Demuestre que $\mathbf{w}$ está en el subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, donde

$$\mathbf{w} = \begin{bmatrix} 9 \\ -4 \\ -4 \\ 7 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 8 \\ -4 \\ -3 \\ 9 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \\ -2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 6 \\ -5 \\ -18 \end{bmatrix}$$

36. [**M**] Determine si $\mathbf{y}$ está en el subespacio de $\mathbb{R}^4$ generado por las columnas de A, donde

$$\mathbf{y} = \begin{bmatrix} -4 \\ -8 \\ 6 \\ -5 \end{bmatrix}, \quad A = \begin{bmatrix} 3 & -5 & -9 \\ 8 & 7 & -6 \\ -5 & -8 & 3 \\ 2 & -2 & -9 \end{bmatrix}$$

37. [**M**] El espacio vectorial $H = \text{Gen}\,\{1, \cos^2 t, \cos^4 t, \cos^6 t\}$ contiene al menos dos funciones interesantes que se utilizarán en un ejercicio posterior:

$$\mathbf{f}(t) = 1 - 8\cos^2 t + 8\cos^4 t$$
$$\mathbf{g}(t) = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$$

Estudie la gráfica de $\mathbf{f}$ para $0 \leq t \leq 2\pi$, e infiera una fórmula sencilla para $\mathbf{f}(t)$. Verifique su suposición mediante la representación gráfica de la diferencia entre $1 + \mathbf{f}(t)$ y su fórmula para $\mathbf{f}(t)$. (Se espera que vea la función constante 1). Repita el procedimiento para $\mathbf{g}$.

38. [**M**] Repita el ejercicio 37 para las funciones

$\mathbf{f}(t) = 3\,\text{sen}\,t - 4\,\text{sen}^3 t$

$\mathbf{g}(t) = 1 - 8\,\text{sen}^2 t + 8\,\text{sen}^4 t$

$\mathbf{h}(t) = 5\,\text{sen}\,t - 20\,\text{sen}^3 t + 16\,\text{sen}^5 t$

en el espacio vectorial Gen $\{1, \text{sen}\,t, \text{sen}^2 t, \ldots, \text{sen}^5 t\}$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Tome cualquier $\mathbf{u}$ en H, digamos, $\mathbf{u} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, y tome cualquier $c \neq 1$, por ejemplo, $c = 2$. Así, $c\mathbf{u} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$. Si esto se encuentra en H, entonces hay alguna s tal que

$$\begin{bmatrix} 3s \\ 2 + 5s \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

Es decir, $s = 2$ y $s = 12/5$, lo que es imposible. Así que $2\mathbf{u}$ no está en H, y H no es un espacio vectorial.

2. $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. Esta ecuación expresa a $\mathbf{v}_1$ como una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_p$, por lo que $\mathbf{v}_1$ se encuentra en W. En general, $\mathbf{v}_k$ está en W, ya que

$$\mathbf{v}_k = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{k-1} + 1\mathbf{v}_k + 0\mathbf{v}_{k+1} + \cdots + 0\mathbf{v}_p$$

4.2 ESPACIOS NULOS, ESPACIOS COLUMNA Y TRANSFORMACIONES LINEALES

En las aplicaciones del álgebra lineal, los subespacios de $\mathbb{R}^n$ generalmente se presentan en una de dos maneras: **1.** como el conjunto de todas las soluciones a un sistema de ecuaciones lineales homogéneas, o **2.** como el conjunto de todas las combinaciones lineales de ciertos vectores específicos. En esta sección, se comparan estas dos descripciones de subespacios, lo que nos permitirá practicar el uso del concepto de subespacio. En realidad, como pronto descubrirá, hemos trabajado con subespacios desde la sección 1.3. La principal novedad aquí es la terminología. La sección concluye con un análisis acerca del núcleo y el rango de una transformación lineal.

El espacio nulo de una matriz

Considere el siguiente sistema de ecuaciones homogéneas:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \qquad (1)$$

En forma de matriz, este sistema se representa como $A\mathbf{x} = \mathbf{0}$, donde

$$A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \tag{2}$$

Recordemos que el conjunto de todas las $\mathbf{x}$ que satisfacen (1) se llama el **conjunto solución** del sistema (1). Con frecuencia es conveniente relacionar este conjunto directamente con la matriz A y la ecuación $A\mathbf{x} = \mathbf{0}$. El conjunto de $\mathbf{x}$ que satisfacen $A\mathbf{x} = \mathbf{0}$ se denomina el **espacio nulo** de la matriz A.

DEFINICIÓN

El **espacio nulo** de una matriz A de $m \times n$, que se denota como Nul A, es el conjunto de todas las soluciones de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$. En notación de conjuntos,

$$\text{Nul } A = \{\mathbf{x} : \mathbf{x} \text{ está en } \mathbb{R}^n \text{ y } A\mathbf{x} = \mathbf{0}\}$$

Una descripción más dinámica de Nul A es el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^n$ que se mapean en el vector cero de $\mathbb{R}^m$ a través de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Véase la figura 1.

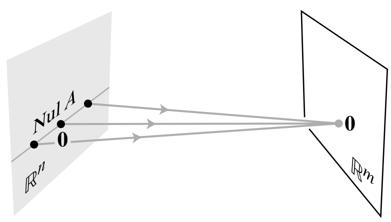


FIGURA 1

EJEMPLO 1 Sea A la matriz en la ecuación (2) anterior, y sea $\mathbf{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine si $\mathbf{u}$ pertenece al espacio nulo de A.

SOLUCIÓN Para probar si $\mathbf{u}$ satisface $A\mathbf{u} = \mathbf{0}$, simplemente se calcula

$$A\mathbf{u} = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 5 - 9 + 4 \\ -25 + 27 - 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Por lo tanto, $\mathbf{u}$ está en Nul A. ■

El término *espacio* en el *espacio nulo* es adecuado ya que el espacio nulo de una matriz es un espacio vectorial, como se muestra en el siguiente teorema.

TEOREMA 2

El espacio nulo de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^n$. De manera equivalente, el conjunto de todas las soluciones a un sistema $A\mathbf{x} = \mathbf{0}$ de m ecuaciones lineales homogéneas con n incógnitas es un subespacio de $\mathbb{R}^n$.

DEMOSTRACIÓN Sin duda, Nul A es un subconjunto de $\mathbb{R}^n$ porque A tiene n columnas. Debemos demostrar que Nul A satisface las tres propiedades de un subespacio. Desde luego, $\mathbf{0}$ está en Nul A. Ahora, considere que $\mathbf{u}$ y $\mathbf{v}$ representan los dos vectores de Nul A. De esta forma,

$$A\mathbf{u} = \mathbf{0} \quad \text{y} \quad A\mathbf{v} = \mathbf{0}$$

Para demostrar que $\mathbf{u} + \mathbf{v}$ está en Nul A, se tiene que demostrar que $A(\mathbf{u} + \mathbf{v}) = \mathbf{0}$. Utilizando una propiedad de multiplicación de matrices, calcule

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Por lo tanto, $\mathbf{u} + \mathbf{v}$ está en Nul A, y Nul A es cerrado bajo la suma de vectores. Por último, si c es un escalar cualquiera, entonces

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c(\mathbf{0}) = \mathbf{0},$$

lo que demuestra que $c\mathbf{u}$ está en Nul A. Así, Nul A es un subespacio de $\mathbb{R}^n$. ■

EJEMPLO 2 Sea H el conjunto de los vectores en $\mathbb{R}^4$ cuyas coordenadas a, b, c, d satisfacen las ecuaciones $a - 2b + 5c = d$, y $c - a = b$. Demuestre que H es un subespacio de $\mathbb{R}^4$.

SOLUCIÓN Al reordenar las ecuaciones que describen los elementos de H, y al observar que H es el conjunto de todas las soluciones del siguiente sistema de ecuaciones lineales homogéneas:

$$\begin{aligned} a - 2b + 5c - d &= 0 \\ -a - \ \ b + \ \ c \qquad &= 0 \end{aligned}$$

De acuerdo con el teorema 2, H es un subespacio de $\mathbb{R}^4$. ■

Es importante hacer notar que las ecuaciones lineales que definen el conjunto H son homogéneas. De lo contrario, el conjunto de soluciones *no* será definitivamente un subespacio (porque el vector cero no es una solución de un sistema no homogéneo). Además, en algunos casos, el conjunto de soluciones podría estar vacío.

Una descripción explícita de Nul A

No hay una relación evidente entre los vectores en Nul A y las entradas de A. Se dice que Nul A se define de forma *implícita*, ya que se define por una condición que se debe comprobar. No existe una lista o una descripción explícitas de los elementos de Nul A. Sin embargo, *resolver* la ecuación $A\mathbf{x} = \mathbf{0}$ equivale a hacer una descripción *explícita* de Nul A. El siguiente ejemplo revisa el procedimiento de la sección 1.5.

EJEMPLO 3 Determine un conjunto generador del espacio nulo de la matriz

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

SOLUCIÓN El primer paso es encontrar la solución general de $A\mathbf{x} = \mathbf{0}$ en términos de variables libres. Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a la forma escalonada *reducida* para escribir las variables básicas en términos de las variables libres:

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad \begin{aligned} x_1 - 2x_2 \qquad - \ \ x_4 + 3x_5 &= 0 \\ x_3 + 2x_4 - 2x_5 &= 0 \\ 0 &= 0 \end{aligned}$$

La solución general es $x_1 = 2x_2 + x_4 - 3x_5$, $x_3 = -2x_4 + 2x_5$, con x_2, x_4 y x_5 libres. Luego, hay que descomponer el vector que da la solución general en una combinación lineal de vectores donde *los pesos son las variables libres*. Es decir,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \underset{\substack{\uparrow \\ \mathbf{u}}}{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}} + x_4 \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}} + x_5 \underset{\substack{\uparrow \\ \mathbf{w}}}{\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}}$$

$$= x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w} \tag{3}$$

Cada combinación lineal de **u**, **v** y **w** es un elemento de Nul A. Por lo tanto, $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es un conjunto generador para Nul A. ■

Hay que hacer notar dos hechos acerca de la solución del ejemplo 3 que se aplican a todos los problemas de este tipo, donde Nul A contiene vectores distintos de cero. Vamos a considerar estos datos más adelante.

1. El conjunto generador producido por el método del ejemplo 3 es linealmente independiente de manera automática debido a que las variables libres son los pesos de los vectores generados. Por ejemplo, vea la segunda, cuarta y quinta entradas en el vector solución en la ecuación (3) y observe que $x_2\mathbf{u} + x_4\mathbf{v} + x_5\mathbf{w}$ puede ser $\mathbf{0}$ solo si los pesos x_2, x_4 y x_5 son todos cero.

2. Cuando Nul A contiene vectores distintos de cero, el número de vectores en el conjunto generador para Nul A es igual al número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$.

El espacio columna de una matriz

Otro subespacio importante asociado con una matriz es el espacio columna. A diferencia del espacio nulo, el espacio columna se define de forma explícita a través de combinaciones lineales.

DEFINICIÓN

El **espacio columna** de una matriz A de $m \times n$, que se denota como Col A, es el conjunto de todas las combinaciones lineales de las columnas de A. Si $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, entonces

$$\text{Col } A = \text{Gen }\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$$

Ya que Gen $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ es un subespacio, de acuerdo con el teorema 1, el siguiente teorema se deduce de la definición de Col A y del hecho de que las columnas de A están en $\mathbb{R}^m$.

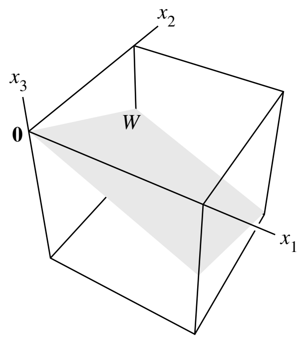
TEOREMA 3

El espacio columna de una matriz A de $m \times n$ es un subespacio de $\mathbb{R}^m$.

Observe que un vector típico de Col A se puede escribir como $A\mathbf{x}$ para alguna $\mathbf{x}$, ya que la notación $A\mathbf{x}$ representa una combinación lineal de las columnas de A. Es decir,

$$\boxed{\text{Col } A = \{\mathbf{b} : \mathbf{b} = A\mathbf{x} \text{ para alguna } \mathbf{x} \text{ en } \mathbb{R}^n\}}$$

La notación $A\mathbf{x}$ para los vectores en Col A también muestra que A es el *rango* de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$. Regresaremos a este punto de vista al final de la sección.



EJEMPLO 4 Encuentre una matriz A tal que $W = \text{Col } A$.

$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\}$$

SOLUCIÓN En primer lugar, escriba W como un conjunto de combinaciones lineales.

$$W = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\} = \text{Gen} \left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

En segundo lugar, utilice los vectores en el conjunto generador como las columnas de A. Sea $A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ -7 & 0 \end{bmatrix}$. De esta forma, $W = \text{Col } A$, como se desea. ■

Recuerde el teorema 4 de la sección 1.4, según el cual las columnas de A generan a $\mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$. Podemos enunciar este hecho con palabras de la siguiente manera:

> El espacio columna de una matriz A de $m \times n$ es $\mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^m$.

Contraste entre Nul A y Col A

Es natural preguntarse cómo se relacionan el espacio nulo y el espacio columna de una matriz. De hecho, los dos espacios son muy diferentes, como se constata en los ejemplos 5 a 7. Sin embargo, una sorprendente conexión entre el espacio nulo y el espacio columna se verá en la sección 4.6, cuando hayamos estudiado más teoría.

EJEMPLO 5 Sea

$$A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$$

a) Si el espacio columna de A es un subespacio de $\mathbb{R}^k$, ¿a qué es igual k?

b) Si el espacio nulo de A es un subespacio de $\mathbb{R}^k$, ¿a qué es igual k?

SOLUCIÓN

a) Cada una de las columnas de A tiene tres entradas, por lo que Col A es un subespacio de $\mathbb{R}^k$, donde $k = 3$.

b) Un vector $\mathbf{x}$ tal que $A\mathbf{x}$ esté definido debe tener cuatro entradas, por lo que Nul A es un subespacio de $\mathbb{R}^k$, donde $k = 4$. ■

Cuando una matriz no es cuadrada, como en el ejemplo 5, los vectores en Nul A y Col A viven en "universos" totalmente diferentes. Por ejemplo, ninguna combinación lineal de vectores en $\mathbb{R}^3$ puede producir un vector en $\mathbb{R}^4$. Cuando A es cuadrada, Nul A y Col A tienen el vector cero en común, y en casos especiales es posible que algunos vectores diferentes de cero pertenezcan tanto a Nul A como a Col A.

EJEMPLO 6 Con A como en el ejemplo 5, encuentre un vector diferente de cero en Col A y otro en Nul A.

SOLUCIÓN Es fácil encontrar un vector en Col A. Cualquier columna de A lo permite, por ejemplo, $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$. Para encontrar un vector distinto de cero en Nul A, al reducir por filas la matriz aumentada $[A \quad \mathbf{0}]$ se obtiene

$$[\,A \quad \mathbf{0}\,] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Por lo tanto, si $\mathbf{x}$ satisface $A\mathbf{x} = 0$, entonces $x_1 = -9x_3$, $x_2 = 5x_3$, $x_4 = 0$, y x_3 es libre. Al asignar un valor distinto de cero para x_3, por ejemplo, $x_3 = 1$, se obtiene un vector de Nul A, a saber, $\mathbf{x} = (-9, 5, 1, 0)$. ■

EJEMPLO 7 Con A como en el ejemplo 5, sean $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$.

a) Determine si $\mathbf{u}$ está en Nul A. ¿Podría $\mathbf{u}$ estar en Col A?

b) Determine si $\mathbf{v}$ está en Col A. ¿Podría $\mathbf{v}$ estar en Nul A?

SOLUCIÓN

a) Una descripción explícita de Nul A no es necesaria aquí. Basta con calcular el producto $A\mathbf{u}$.

$$A\mathbf{u} = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Evidentemente, $\mathbf{u}$ *no* es una solución de $A\mathbf{x} = \mathbf{0}$, por lo que $\mathbf{u}$ no está en Nul A. Además, con cuatro entradas, $\mathbf{u}$ no podría estar en Col A, ya que Col A es un subespacio de $\mathbb{R}^3$.

b) Se reduce $[A \quad \mathbf{v}]$ a una forma escalonada.

$$[\,A \quad \mathbf{v}\,] = \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 & -2 & 1 & 3 \\ 0 & 1 & -5 & -4 & -2 \\ 0 & 0 & 0 & 17 & 1 \end{bmatrix}$$

En este punto, es claro que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente, por lo que $\mathbf{v}$ está en Col A. Con solo tres entradas, $\mathbf{v}$ no podría estar en Nul A, ya que Nul A es un subespacio de $\mathbb{R}^4$. ■

La tabla de la página 204 resume lo que hemos aprendido acerca de Nul A y Col A. El punto 8 es una nueva formulación de los teoremas 11 y 12*a*) de la sección 1.9.

Núcleo y rango de una transformación lineal

Otros subespacios de espacios vectoriales diferentes de $\mathbb{R}^n$ con frecuencia se describen en términos de una transformación lineal y no de una matriz. Con la finalidad de precisar esto, generalizamos la definición que dimos en la sección 1.8.

Comparación entre Nul A y Col A para una matriz A de $m \times n$

Nul A	Col A
1. Nul A es un subespacio de $\mathbb{R}^n$.	**1.** Col A es un subespacio de $\mathbb{R}^m$.
2. Nul A se define implícitamente, es decir, solo estableciendo una condición ($A\mathbf{x} = \mathbf{0}$) que los vectores de Nul A deben satisfacer.	**2.** Col A se define explícitamente, es decir, se indica cómo construir vectores en Col A.
3. Se necesita tiempo para encontrar vectores de Nul A. Se requiere efectuar operaciones de fila en $[A \quad \mathbf{0}]$.	**3.** Es fácil encontrar vectores en Col A. Las columnas de A se muestran, y las otras se forman a partir de ellas.
4. No hay una relación evidente entre Nul A y las entradas de A.	**4.** Existe una relación evidente entre Col A y las entradas de A, ya que cada columna de A está en Col A.
5. Un vector típico $\mathbf{v}$ de Nul A tiene la propiedad de que $A\mathbf{v} = \mathbf{0}$.	**5.** Un vector típico $\mathbf{v}$ en Col A tiene la propiedad de que la ecuación $A\mathbf{x} = \mathbf{v}$ es consistente.
6. Dado un vector específico $\mathbf{v}$, es fácil saber si $\mathbf{v}$ está en Nul A. Solo hay que calcular $A\mathbf{v}$.	**6.** Dado un vector específico $\mathbf{v}$, puede tomar algún tiempo saber si $\mathbf{v}$ está en Col A. Se requieren operaciones de fila en $[A \quad \mathbf{v}]$.
7. Nul $A = \{\mathbf{0}\}$ si y solo si la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.	**7.** Col $A = \mathbb{R}^m$ si y solo si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ si y solo si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es de uno a uno.	**8.** Col $A = \mathbb{R}^m$ si y solo si la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$.

DEFINICIÓN

Una **transformación lineal** T de un espacio vectorial V en un espacio vectorial W es una regla que asigna a cada vector $\mathbf{x}$ en V un único vector $T(\mathbf{x})$ en W, tal que

i. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para toda $\mathbf{u}$, $\mathbf{v}$ en V, y

ii. $T(c\mathbf{u}) = cT(\mathbf{u})$ para toda $\mathbf{u}$ en V y todo escalar c.

El **núcleo** (o **espacio nulo**) de dicha T es el conjunto de todas las $\mathbf{u}$ en V tales que $T(\mathbf{u}) = \mathbf{0}$ (el vector cero en W). El **rango** de T es el conjunto de todos los vectores en W de la forma $T(\mathbf{x})$ para alguna $\mathbf{x}$ en V. Si ocurre que T surge como una transformación matricial, por ejemplo, $T(\mathbf{x}) = A\mathbf{x}$ para alguna matriz A, entonces el núcleo y la imagen de T son solo el espacio nulo y el espacio columna de A, como se definió anteriormente.

No es difícil demostrar que el núcleo de T es un subespacio de V. La demostración es, en esencia, la misma que para el teorema 2. Además, el rango de T es un subespacio de W. Véase la figura 2 y el ejercicio 30.

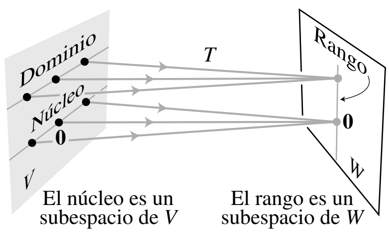


FIGURA 2 Subespacios asociados con una transformación lineal.

En aplicaciones, un subespacio suele surgir ya sea como el núcleo, o bien, como el rango de una transformación lineal adecuada. Por ejemplo, el conjunto de todas las soluciones de una ecuación diferencial lineal homogénea resulta ser el núcleo de una transformación lineal.

En general, una transformación lineal de este tipo se describe en términos de una o más derivadas de una función. Explicar esto en detalle nos llevaría demasiado lejos en este momento, por lo que consideramos solo dos ejemplos. El primero explica por qué la operación de diferenciación es una transformación lineal.

EJEMPLO 8 (Requiere conocimientos de cálculo) Sea V el espacio vectorial de todas las funciones de valores reales f definidas en un intervalo $[a, b]$ con la propiedad de que son diferenciables, y sus derivadas son funciones continuas en $[a, b]$. Sea W el espacio vectorial $C[a, b]$ de todas las funciones continuas en $[a, b]$, y sea $D : V \to W$ la transformación que cambia a f en V en su derivada f'. En cálculo, dos reglas sencillas de la diferenciación son

$$D(f + g) = D(f) + D(g) \quad \text{y} \quad D(cf) = cD(f)$$

Es decir, D es una transformación lineal. Es posible demostrar que el núcleo de D es el conjunto de las funciones constantes en $[a, b]$ y el rango de D es el conjunto W de todas las funciones continuas en $[a, b]$. ■

EJEMPLO 9 (Requiere conocimientos de cálculo) La ecuación diferencial

$$y'' + \omega^2 y = 0, \tag{4}$$

donde ω es una constante, se utiliza para describir una variedad de sistemas físicos, tales como la vibración de un resorte del que pende un peso, el movimiento de un péndulo, y el voltaje en un circuito eléctrico de inductancia-capacitancia. El conjunto de soluciones de (4) es precisamente el núcleo de la transformación lineal que mapea una función $y = f(t)$ en la función de $f''(t) + \omega^2 f(t)$. Encontrar una descripción explícita de este espacio vectorial es un problema de ecuaciones diferenciales. El conjunto solución resulta ser el espacio descrito en el ejercicio 19 de la sección 4.1. ■

PROBLEMAS DE PRÁCTICA

1. Sea $W = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. Demuestre de dos maneras diferentes que W es un subespacio de $\mathbb{R}^3$. (Utilice dos teoremas).

2. Sea $A = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$. Suponga que sabe que las ecuaciones $A\mathbf{x} = \mathbf{v}$ y $A\mathbf{x} = \mathbf{w}$ son consistentes. ¿Qué puede decir acerca de la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$?

4.2 EJERCICIOS

1. Determine si $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ está en Nul A, donde

$$A = \begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix}.$$

2. Determine si $\mathbf{w} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ está en Nul A, donde

$$A = \begin{bmatrix} 2 & 6 & 4 \\ -3 & 2 & 5 \\ -5 & -4 & 1 \end{bmatrix}.$$

En los ejercicios 3 a 6, encuentre una descripción explícita de Nul A, haciendo una lista de vectores que generan el espacio nulo.

3. $A = \begin{bmatrix} 1 & 2 & 4 & 0 \\ 0 & 1 & 3 & -2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -3 & 2 & 0 \\ 0 & 0 & 3 & 0 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & -4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 3 & -4 & -3 & 1 \\ 0 & 1 & -3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 7 a 14, utilice un teorema adecuado para demostrar que el conjunto dado, W, es un espacio vectorial, o bien, encuentre un ejemplo específico de lo contrario.

7. $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 2 \right\}$ **8.** $\left\{ \begin{bmatrix} r \\ s \\ t \end{bmatrix} : 3r - 2 = 3s + t \right\}$

9. $\left\{ \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} : \begin{matrix} p - 3q = 4s \\ 2p = s + 5r \end{matrix} \right\}$ **10.** $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} 3a + b = c \\ a + b + 2c = 2d \end{matrix} \right\}$

11. $\left\{ \begin{bmatrix} s - 2t \\ 3 + 3s \\ 3s + t \\ 2s \end{bmatrix} : s, t \text{ reales} \right\}$ **12.** $\left\{ \begin{bmatrix} 3p - 5q \\ 4q \\ p \\ q + 1 \end{bmatrix} : p, q \text{ reales} \right\}$

13. $\left\{ \begin{bmatrix} c - 6d \\ d \\ c \end{bmatrix} : c, d \text{ reales} \right\}$ **14.** $\left\{ \begin{bmatrix} -s + 3t \\ s - 2t \\ 5s - t \end{bmatrix} : s, t \text{ reales} \right\}$

En los ejercicios 15 y 16, encuentre A tal que el conjunto dado sea Col A.

15. $\left\{ \begin{bmatrix} 2s + t \\ r - s + 2t \\ 3r + s \\ 2r - s - t \end{bmatrix} : r, s, t \text{ reales} \right\}$

16. $\left\{ \begin{bmatrix} b - c \\ 2b + 3d \\ b + 3c - 3d \\ c + d \end{bmatrix} : b, c, d \text{ reales} \right\}$

Para las matrices de los ejercicios 17 a 20, *a*) encuentre k tal que Nul A sea un subespacio de $\mathbb{R}^k$, y *b*) encuentre k tal que Col A sea un subespacio de $\mathbb{R}^k$.

17. $A = \begin{bmatrix} 6 & -4 \\ -3 & 2 \\ -9 & 6 \\ 9 & -6 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 5 & -2 & 3 \\ -1 & 0 & -1 \\ 0 & -2 & -2 \\ -5 & 7 & 2 \end{bmatrix}$

19. $A = \begin{bmatrix} 4 & 5 & -2 & 6 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & -3 & 2 & 0 & -5 \end{bmatrix}$

21. Con A como en el ejercicio 17, encuentre un vector distinto de cero en Nul A y otro en Col A.

22. Con A como en el ejercicio 18, encuentre un vector distinto de cero en Nul A y otro en Col A.

23. Sea $A = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

24. Sea $A = \begin{bmatrix} 10 & -8 & -2 & -2 \\ 0 & 2 & 2 & -2 \\ 1 & -1 & 6 & 0 \\ 1 & 1 & 0 & -2 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 2 \end{bmatrix}$. Determine si $\mathbf{w}$ está en Col A. ¿Está $\mathbf{w}$ en Nul A?

En los ejercicios 25 y 26, A denota una matriz de $m \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

25. *a*) El espacio nulo de A es el conjunto solución de la ecuación $A\mathbf{x} = \mathbf{0}$.

b) El espacio nulo de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c) El espacio columna de A es el rango del mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

d) Si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces Col A es $\mathbb{R}^m$.

e) El núcleo de una transformación lineal es un espacio vectorial.

f) Col A es el conjunto de todos los vectores que se pueden escribir como $A\mathbf{x}$ para alguna $\mathbf{x}$.

26. *a*) Un espacio nulo es un espacio vectorial.

b) El espacio columna de una matriz de $m \times n$ está en $\mathbb{R}^m$.

c) Col A es el conjunto de todas las soluciones de $A\mathbf{x} = \mathbf{b}$.

d) Nul A es el núcleo del mapeo $\mathbf{x} \mapsto A\mathbf{x}$.

e) El rango de una transformación lineal es un espacio vectorial.

f) El conjunto de todas las soluciones de una ecuación diferencial lineal homogénea es el núcleo de una transformación lineal.

27. Es posible demostrar que una solución del sistema que se muestra a continuación es $x_1 = 3$, $x_2 = 2$, y $x_3 = -1$. Con base en este hecho y en la teoría de esta sección, explique por qué otra solución es $x_1 = 30$, $x_2 = 20$, y $x_3 = -10$. (Observe cómo están relacionadas las soluciones, pero no realice otros cálculos).

$$\begin{aligned} x_1 - 3x_2 - 3x_3 &= 0 \\ -2x_1 + 4x_2 + 2x_3 &= 0 \\ -x_1 + 5x_2 + 7x_3 &= 0 \end{aligned}$$

28. Considere los siguientes dos sistemas de ecuaciones:

$$\begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 1 \\ 4x_1 + x_2 - 6x_3 &= 9 \end{aligned} \qquad \begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 5 \\ 4x_1 + x_2 - 6x_3 &= 45 \end{aligned}$$

Es posible demostrar que el primer sistema tiene una solución. Con base en este hecho y en la teoría de esta sección, explique por qué el segundo sistema también debe tener una solución. (No realice operaciones de fila).

29. Demuestre el teorema 3 de la siguiente manera: dada una matriz A de $m \times n$, un elemento de Col A tiene la forma $A\mathbf{x}$ para alguna $\mathbf{x}$ en $\mathbb{R}^n$. Sean $A\mathbf{x}$ y $A\mathbf{w}$ dos vectores cualesquiera en Col A.

a) Explique por qué el vector cero está en Col A.

b) Demuestre que el vector $A\mathbf{x} + A\mathbf{w}$ está en Col A.

c) Dado un escalar c, demuestre que $c(A\mathbf{x})$ está en Col A.

30. Sea $T : V \to W$ una transformación lineal de un espacio vectorial V en un espacio vectorial W. Demuestre que el rango de T es un subespacio de W. [*Sugerencia:* Considere que los elementos típicos del rango tienen la forma $T(\mathbf{x})$ y $T(\mathbf{w})$ para alguna $\mathbf{x}$, $\mathbf{w}$ en V].

31. Defina $T : \mathbb{P}_2 \to \mathbb{R}^2$ mediante $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$. Por ejemplo, si $\mathbf{p}(t) = 3 + 5t + 7t^2$, entonces $T(\mathbf{p}) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

a) Demuestre que T es una transformación lineal. [*Sugerencia:* Para polinomios arbitrarios $\mathbf{p}$, $\mathbf{q}$ en $\mathbb{P}_2$, calcule $T(\mathbf{p} + \mathbf{q})$ y $T(c\mathbf{p})$].

b) Encuentre un polinomio $\mathbf{p}$ en $\mathbb{P}_2$ que genere el núcleo de T y describa el rango de T.

32. Defina una transformación lineal $T : \mathbb{P}_2 \to \mathbb{R}^2$ mediante $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(0) \end{bmatrix}$. Determine los polinomios $\mathbf{p}_1$ y $\mathbf{p}_2$ en $\mathbb{P}_2$ que generan el núcleo de T, y describa el rango de T.

33. Sea $M_{2\times2}$ el espacio vectorial de todas las matrices de 2×2, y defina $T : M_{2\times2} \to M_{2\times2}$ mediante $T(A) = A + A^T$, donde $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

a) Demuestre que T es una transformación lineal.

b) Sea B cualquier elemento de $M_{2\times2}$ tal que $B^{\mathrm{T}} = B$. Determine una A en $M_{2\times2}$ tal que $T(A) = B$.

c) Demuestre que el rango de T es el conjunto de B en $M_{2\times2}$ con la propiedad de que $B^{\mathrm{T}} = B$.

d) Describa el núcleo de T.

34. (*Se requiere cálculo*) Defina a $T : C[0, 1] \to C[0, 1]$ como sigue: Para $\mathbf{f}$ en $C[0, 1]$, sea $T(\mathbf{f})$ la antiderivada $\mathbf{F}$ de $\mathbf{f}$ tal que $\mathbf{F}(0) = 0$. Demuestre que T es una transformación lineal, y describa el núcleo de T. (Véase la notación en el ejercicio 20 de la sección 4.1).

35. Sean V y W espacios vectoriales, y sea $T : V \to W$ una transformación lineal. Dado un subespacio U de V, sea $T(U)$ el conjunto de todas las imágenes de la forma $T(\mathbf{x})$, donde $\mathbf{x}$ está en U. Demuestre que $T(U)$ es un subespacio de W.

36. Dada $T : V \to W$ como en el ejercicio 35, y dado un subespacio Z de W, sea U el conjunto de todas las $\mathbf{x}$ en V tales que $T(\mathbf{x})$ está en Z. Demuestre que U es un subespacio de V.

37. [**M**] Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde,

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, \quad A = \begin{bmatrix} 7 & 6 & -4 & 1 \\ -5 & -1 & 0 & -2 \\ 9 & -11 & 7 & -3 \\ 19 & -9 & 7 & 1 \end{bmatrix}$$

38. [**M**] Determine si $\mathbf{w}$ está en el espacio columna de A, en el espacio nulo de A, o en ambos, donde,

$$\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$$

39. [**M**] Considere que $\mathbf{a}_1, \ldots, \mathbf{a}_5$ denotan las columnas de la matriz A, donde

$$A = \begin{bmatrix} 5 & 1 & 2 & 2 & 0 \\ 3 & 3 & 2 & -1 & -12 \\ 8 & 4 & 4 & -5 & 12 \\ 2 & 1 & 1 & 0 & -2 \end{bmatrix}, \quad B = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_4\,]$$

a) Explique por qué $\mathbf{a}_3$ y $\mathbf{a}_5$ están en el espacio columna de B.

b) Encuentre un conjunto de vectores que genere a Nul A.

c) Sea $T : \mathbb{R}^5 \to \mathbb{R}^4$ definida por $T(\mathbf{x}) = A\mathbf{x}$. Explique por qué T no es uno a uno ni sobre.

40. [**M**] Sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ y $K = \text{Gen}\,\{\mathbf{v}_3, \mathbf{v}_4\}$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ -12 \\ -28 \end{bmatrix}.$$

Entonces, H y K son subespacios de $\mathbb{R}^3$. De hecho, H y K son planos en $\mathbb{R}^3$ que pasan por el origen, y que se cruzan en una recta que pasa por $\mathbf{0}$. Encuentre un vector $\mathbf{w}$ distinto de cero que genere esa recta. [*Sugerencia*: $\mathbf{w}$ se puede escribir como $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ y también como $c_3\mathbf{v}_3 + c_4\mathbf{v}_4$. Para construir $\mathbf{w}$, resuelva la ecuación $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_3\mathbf{v}_3 + c_4\mathbf{v}_4$ para las incógnitas c_j].

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *Primer método:* W es un subespacio de $\mathbb{R}^3$ de acuerdo con el teorema 2, ya que W es el conjunto de todas las soluciones de un sistema de ecuaciones lineales homogéneas (donde el sistema solamente tiene una ecuación). De manera equivalente, W es el espacio nulo de la matriz de 1×3, $A = [1 \quad -3 \quad -1]$.

Segundo método: Resuelva la ecuación $a - 3b - c = 0$ para la variable principal a en términos de las variables libres b y c. Cualquier solución tiene la forma $\begin{bmatrix} 3b+c \\ b \\ c \end{bmatrix}$, donde b y c son arbitrarias, y

$$\begin{bmatrix} 3b+c \\ b \\ c \end{bmatrix} = b\underbrace{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}}_{\uparrow\,\mathbf{v}_1} + c\underbrace{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}_{\uparrow\,\mathbf{v}_2}$$

Este cálculo muestra que $W = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Por lo tanto, W es un subespacio de $\mathbb{R}^3$ de acuerdo con el teorema 1. También se podría despejar b o c de la ecuación $a - 3b - c = 0$ y obtener descripciones alternativas de W como un conjunto de combinaciones lineales de dos vectores.

2. Tanto $\mathbf{v}$ como $\mathbf{w}$ están en Col A. Ya que Col A es un espacio vectorial, $\mathbf{v} + \mathbf{w}$ debe estar en Col A. Es decir, la ecuación $A\mathbf{x} = \mathbf{v} + \mathbf{w}$ es consistente.

4.3 CONJUNTOS LINEALMENTE INDEPENDIENTES; BASES

En esta sección identificaremos y estudiaremos los subgrupos que generan un espacio vectorial V o un subespacio H tan "eficientemente" como sea posible. La idea clave es la de independencia lineal, definida como en $\mathbb{R}^n$.

Se dice que un conjunto indexado de vectores $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V es **linealmente independiente** si la ecuación vectorial

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{1}$$

tiene *solamente* la solución trivial, $c_1 = 0,\ldots, c_p = 0$.[1]

Se dice que el conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es **linealmente dependiente** si (1) tiene una solución no trivial, es decir, si hay algunos pesos, $c_1,\ldots, c_p$, *no todos cero*, tales que la ecuación (1) sea válida. En tal caso, la ecuación (1) se llama una **relación de dependencia lineal** entre $\mathbf{v}_1,\ldots, \mathbf{v}_p$.

Al igual que en $\mathbb{R}^n$, un conjunto que contiene un único vector $\mathbf{v}$ es linealmente independiente si y solo si $\mathbf{v} \neq 0$. Además, un conjunto de dos vectores es linealmente dependiente si y solo si uno de los vectores es un múltiplo del otro. Y cualquier conjunto que contenga al vector cero es linealmente dependiente. El siguiente teorema tiene la misma demostración que el teorema 7 de la sección 1.7.

TEOREMA 4

Un conjunto indexado $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ de dos o más vectores, con $\mathbf{v}_1 \neq \mathbf{0}$, es linealmente dependiente si y solo si alguna $\mathbf{v}_j$ (con $j > 1$) es una combinación lineal de los vectores anteriores, $\mathbf{v}_1,\ldots, \mathbf{v}_{j-1}$.

La principal diferencia entre la dependencia lineal en $\mathbb{R}^n$ y en un espacio vectorial general es que cuando los vectores no son n-adas, la ecuación homogénea (1), por lo general, no se puede escribir como un sistema de n ecuaciones lineales. Es decir, los vectores no se pueden convertir en las columnas de una matriz A con la finalidad de estudiar la ecuación $A\mathbf{x} = \mathbf{0}$. En vez de ello, debemos utilizar la definición de dependencia lineal y el teorema 4.

EJEMPLO 1 Sea $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$ y $\mathbf{p}_3(t) = 4 - t$. Entonces, $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es linealmente dependiente en $\mathbb{P}$ debido a $\mathbf{p}_3 = 4\mathbf{p}_1 - \mathbf{p}_2$. ■

[1] Es conveniente utilizar $c_1,\ldots, c_p$ en la ecuación (1) para los escalares en vez de $x_1,\ldots, x_p$, como lo hicimos en el capítulo 1.

EJEMPLO 2 El conjunto {sen t, cos t} es linealmente independiente en $C[0, 1]$, el espacio de todas las funciones continuas en $0 \le t \le 1$, porque sen t y cos t no son múltiplos entre sí *como vectores en* $C[0, 1]$. Es decir, no hay escalar c tal que cos $t = c \cdot$ sen t para toda t en $[0, 1]$. (Véase las gráficas de sen t y cos t). Sin embargo, {sen t cos t, sen $2t$} es linealmente dependiente debido a la identidad: sen $2t = 2$ sen t cos t, para toda t. ■

DEFINICIÓN

> Sea H un subespacio de un espacio vectorial V. Un conjunto indexado de vectores $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ en V es una **base** de H si
>
> **i.** $\mathcal{B}$ es un conjunto linealmente independiente, y
>
> **ii.** el subespacio generado por $\mathcal{B}$ coincide con H, es decir,
>
> $$H = \text{Gen}\,\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$$

La definición de una base se aplica al caso en que $H = V$, ya que cualquier espacio vectorial es un subespacio de sí mismo. Así, una base de V es un conjunto linealmente independiente que genera a V. Observe que cuando $H \neq V$, la condición **ii** incluye el requisito de que cada uno de los vectores $\mathbf{b}_1, \ldots, \mathbf{b}_p$ debe pertenecer a H, porque Gen $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ contiene a $\mathbf{b}_1, \ldots, \mathbf{b}_p$, como se muestra en la sección 4.1.

EJEMPLO 3 Sea A una matriz invertible de $n \times n$, por ejemplo, $A = \{\mathbf{a}_1 \cdots\ \mathbf{a}_n]$. Entonces, las columnas de A forman una base para $\mathbb{R}^n$, ya que son linealmente independientes y generan a $\mathbb{R}^n$, por el teorema de la matriz invertible. ■

EJEMPLO 4 Sean $\mathbf{e}_1, \ldots, \mathbf{e}_n$, las columnas de la matriz identidad I_n de $n \times n$. Es decir,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ se llama la **base estándar** para $\mathbb{R}^n$ (figura 1). ■

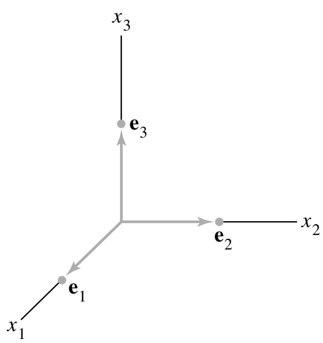


FIGURA 1
La base estándar para $\mathbb{R}^3$.

EJEMPLO 5 Sean $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$.

SOLUCIÓN Puesto que hay exactamente tres vectores aquí en $\mathbb{R}^3$, se puede utilizar cualquiera de los diversos métodos para determinar si la matriz $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ es invertible. Por ejemplo, dos remplazos de fila revelan que A tiene tres posiciones pivote. Por lo tanto, A es invertible. Como en el ejemplo 3, las columnas de A forman una base para $\mathbb{R}^3$. ■

EJEMPLO 6 Sea $S = \{1, t, t^2, \ldots, t^n\}$. Compruebe que S es una base para $\mathbb{P}_n$. Esta base se llama la **base estándar** para $\mathbb{P}_n$.

SOLUCIÓN Sin duda, S genera a $\mathbb{P}_n$. Para demostrar que S es linealmente independiente, suponga que $c_0, \ldots, c_n$ satisfacen

$$c_0 \cdot 1 + c_1 t + c_2 t^2 + \cdots + c_n t^n = \mathbf{0}(t) \tag{2}$$

Esta igualdad significa que el polinomio de la izquierda tiene los mismos valores que el polinomio cero de la derecha. Un teorema fundamental del álgebra dice que el único polinomio en $\mathbb{P}_n$, con más de n ceros es el polinomio cero. Es decir, la ecuación (2) se cumple para todas las t solo si $c_0 = \cdots = c_n = 0$. Esto demuestra que S es linealmente independiente y, por lo tanto, es una base para P_n. Véase la figura 2. ■

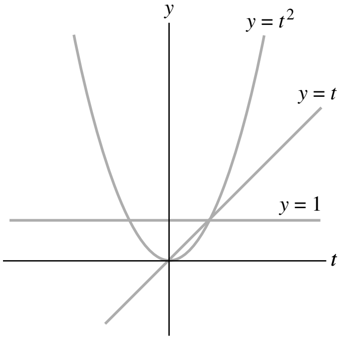


FIGURA 2
La base estándar para $\mathbb{P}_2$.

Problemas relacionados con la independencia lineal y la generación en $\mathbb{P}_n$ se manejan mejor con una técnica que se analizará en la sección 4.4.

El teorema del conjunto generador

Como se verá, una base es un "eficiente" conjunto generador que no contiene vectores innecesarios. De hecho, una base se construye a partir de un conjunto generador descartando los vectores innecesarios.

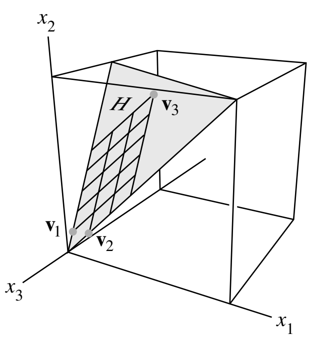
EJEMPLO 7 Sean

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 16 \\ -5 \end{bmatrix} \quad \text{y} \quad H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}.$$

Considere que $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, y demuestre que Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ = Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$. Luego, encuentre una base para el subespacio H.

SOLUCIÓN Cada vector en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ pertenece a H porque

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + 0\mathbf{v}_3$$



Ahora sea $\mathbf{x}$ cualquier vector en H, por ejemplo, $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Puesto que $\mathbf{v}_3 = 5\mathbf{v}_1 + 3\mathbf{v}_2$, podemos sustituir

$$\begin{aligned}\mathbf{x} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(5\mathbf{v}_1 + 3\mathbf{v}_2) \\ &= (c_1 + 5c_3)\mathbf{v}_1 + (c_2 + 3c_3)\mathbf{v}_2\end{aligned}$$

Así, $\mathbf{x}$ está en Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$, por lo que cada vector en H ya pertenece a Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$. Llegamos a la conclusión de que H y Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ en realidad son el mismo conjunto de vectores. De lo que se deduce que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base de H, ya que $\{\mathbf{v}_1, \mathbf{v}_2\}$ es, sin duda, linealmente independiente. ■

El siguiente teorema generaliza el ejemplo 7.

TEOREMA 5 **El teorema del conjunto generador**

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un conjunto en V, y sea $H = \text{Gen}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a) Si uno de los vectores de S, por ejemplo $\mathbf{v}_k$, es una combinación lineal de los vectores restantes en S, entonces el conjunto formado a partir de S al eliminar $\mathbf{v}_k$ aún genera a H.

b) Si $H \neq \{\mathbf{0}\}$, algún subconjunto de S es una base para H.

DEMOSTRACIÓN

a) Al reordenar la lista de vectores en S, si es necesario, podemos suponer que $\mathbf{v}_p$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$, por ejemplo,

$$\mathbf{v}_p = a_1\mathbf{v}_1 + \cdots + a_{p-1}\mathbf{v}_{p-1} \tag{3}$$

Dada cualquier $\mathbf{x}$ en H, podemos escribir

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_{p-1}\mathbf{v}_{p-1} + c_p\mathbf{v}_p \tag{4}$$

para escalares adecuados $c_1, \ldots, c_p$. Sustituyendo la expresión para $\mathbf{v}_p$ de la ecuación (3) en (4), es fácil ver que $\mathbf{x}$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}$. Por lo tanto, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p-1}\}$ genera a H, ya que $\mathbf{x}$ es un elemento arbitrario de H.

b) Si el conjunto generador original S es linealmente independiente, entonces ya es una base para H. De lo contrario, uno de los vectores de S depende de los demás y se puede eliminar, de acuerdo con el inciso *a*). Siempre y cuando haya dos o más vectores en el conjunto generado, podemos repetir el proceso hasta que el conjunto generador sea linealmente independiente y, por lo tanto, sea una base para H. Si el conjunto generador finalmente se reduce a un vector, ese vector será distinto de cero (y, por lo tanto, linealmente independiente), ya que $H \neq \{\mathbf{0}\}$. ■

Bases para Nul A y Col A

Ya sabemos cómo encontrar los vectores que generan el espacio nulo de una matriz A. El análisis de la sección 4.2 indicó que nuestro método siempre produce un conjunto linealmente independiente cuando Nul A contiene vectores distintos de cero. Por lo tanto, en este caso, ese método produce una *base* para Nul A.

Los dos ejemplos siguientes describen un sencillo algoritmo que permite encontrar una base para el espacio columna.

EJEMPLO 8 Encuentre una base para Col B, donde

$$B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

SOLUCIÓN Cada columna que no es pivote de B es una combinación lineal de las columnas pivote. De hecho, $\mathbf{b}_2 = 4\mathbf{b}_1$ y $\mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$. De acuerdo con el teorema del conjunto generador, podemos descartar a $\mathbf{b}_2$ y $\mathbf{b}_4$, y $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ aún generará a Col B. Sea

$$S = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Ya que $\mathbf{b}_1 \neq 0$ y ningún vector en S es una combinación lineal de los vectores que lo preceden, S es linealmente independiente (teorema 4). Por lo tanto, S es una base de Col B. ■

¿Qué pasa con una matriz A que *no* está en forma escalonada reducida? Recuerde que toda relación de dependencia lineal entre las columnas de A se puede expresar en la forma $A\mathbf{x} = \mathbf{0}$, donde $\mathbf{x}$ es una columna de pesos. (Si algunas columnas no están implicadas en una relación de dependencia en particular, entonces sus pesos son iguales a cero). Cuando A se reduce por filas a una matriz B, las columnas de B con frecuencia son totalmente diferentes de las columnas de A. Sin embargo, las ecuaciones $A\mathbf{x} = \mathbf{0}$ y $B\mathbf{x} = \mathbf{0}$ tienen exactamente el mismo conjunto de soluciones. Si $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$ y $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$, entonces las ecuaciones vectoriales

$$x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \quad \text{y} \quad x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n = \mathbf{0}$$

también tienen el mismo conjunto de soluciones. Es decir, las columnas de A tienen *exactamente la misma relación de dependencia lineal* que las columnas de B.

EJEMPLO 9 Es posible demostrar que la matriz

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$$

es equivalente por filas a la matriz B del ejemplo 8. Encuentre una base para Col A.

SOLUCIÓN En el ejemplo 8 vimos que

$$\mathbf{b}_2 = 4\mathbf{b}_1 \quad \text{y} \quad \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3$$

por lo que podemos esperar que

$$\mathbf{a}_2 = 4\mathbf{a}_1 \quad \text{y} \quad \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3$$

¡Compruebe que este es el caso! Por lo tanto, podemos descartar a $\mathbf{a}_2$ y $\mathbf{a}_4$ cuando seleccionamos un conjunto generador mínimo de Col A. En efecto, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ debe ser linealmente independiente porque cualquier relación de dependencia lineal entre $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ implicaría una relación de dependencia lineal entre $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$. Pero sabemos que $\{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5\}$ es un conjunto linealmente independiente. Por lo tanto, $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ es una base para Col A. Las columnas que se han utilizado para esta base son las columnas pivote de A. ■

Los ejemplos 8 y 9 ilustran el siguiente hecho que resulta de utilidad.

TEOREMA 6 Las columnas pivote de una matriz A forman una base para Col A.

DEMOSTRACIÓN La demostración general utiliza los argumentos analizados antes. Sea B la forma escalonada reducida de A. El conjunto de columnas pivote de B es linealmente independiente, pues ningún vector del conjunto es una combinación lineal de los vectores que le preceden. Puesto que A es equivalente por filas, también las columnas pivote de A son linealmente independientes, ya que cualquier relación de dependencia lineal entre las columnas de A corresponde a una relación de dependencia lineal entre las columnas de B. Por esta misma razón, todas las columnas que no sean pivote de A son una combinación lineal de las columnas pivote de A. Así, las columnas de A que no son pivote se pueden descartar del conjunto generador de Col A, de acuerdo con el teorema del conjunto generador. Esto deja a las columnas pivote de A como base para Col A. ■

Advertencia: Las columnas pivote de un matriz A son evidentes cuando A se ha reducido solamente a la forma *escalonada*. Sin embargo, tenga cuidado al usar las *columnas pivote de la misma* A como la base de Col A. Las operaciones de fila pueden cambiar el espacio columna de una matriz. Las columnas de una forma escalonada B de A con frecuencia no están en el espacio columna de A. Por ejemplo, todas las columnas de la matriz B en el ejemplo 8 tienen ceros en sus últimas entradas, por lo que no pueden generar el espacio columna de la matriz A en el ejemplo 9.

Dos perspectivas de una base

Cuando se usa el teorema del conjunto generador, la eliminación de los vectores de un conjunto generador se debe detener cuando el conjunto se convierte en linealmente independiente. Si se elimina un vector adicional, no será una combinación lineal de los vectores restantes y, por lo tanto, el conjunto más pequeño ya no generará a V. Así, una base es un conjunto generador que es lo más pequeño posible.

Una base también es un conjunto linealmente independiente lo más grande posible. Si S es una base para V, y si S se amplía con un vector (por ejemplo, $\mathbf{w}$) de V, entonces el nuevo conjunto no puede ser linealmente independiente, ya que S genera a V, y $\mathbf{w}$ es, por lo tanto, una combinación lineal de los elementos en S.

EJEMPLO 10 Los siguientes tres conjuntos de $\mathbb{R}^3$ muestran cómo se puede ampliar un conjunto linealmente independiente a una base y cómo esta ampliación adicional destruye la independencia lineal del conjunto. Además, un conjunto generador se puede reducir a una base, pero una reducción adicional destruye la propiedad de generación.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

Linealmente independiente, pero no genera a $\mathbb{R}^3$ — Una base para $\mathbb{R}^3$ — Genera a $\mathbb{R}^3$, pero es linealmente dependiente ■

PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\-2\\3\end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix}-2\\7\\-9\end{bmatrix}$. Determine si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para $\mathbb{R}^3$. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para $\mathbb{R}^2$?

2. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\-3\\4\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}6\\2\\-1\end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix}2\\-2\\3\end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix}-4\\-8\\9\end{bmatrix}$. Determine una base para el subespacio W generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

3. Sea $\mathbf{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ y $H = \left\{\begin{bmatrix}s\\s\\0\end{bmatrix} : s \text{ en } \mathbb{R}\right\}$. Entonces cada vector en H es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ porque
$$\begin{bmatrix}s\\s\\0\end{bmatrix} = s\begin{bmatrix}1\\0\\0\end{bmatrix} + s\begin{bmatrix}0\\1\\0\end{bmatrix}$$
¿Es $\{\mathbf{v}_1, \mathbf{v}_2\}$ una base para H?

4.3 EJERCICIOS

Determine si los conjuntos de los ejercicios 1 a 8 son bases para $\mathbb{R}^3$. De los conjuntos que *no* son bases, determine cuáles son linealmente independientes y cuáles generan a $\mathbb{R}^3$. Justifique sus respuestas.

1. $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}$

2. $\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}$

3. $\begin{bmatrix}1\\0\\-3\end{bmatrix},\begin{bmatrix}3\\1\\-4\end{bmatrix},\begin{bmatrix}-2\\-1\\1\end{bmatrix}$

4. $\begin{bmatrix}2\\-1\\1\end{bmatrix},\begin{bmatrix}2\\-3\\2\end{bmatrix},\begin{bmatrix}-8\\5\\4\end{bmatrix}$

5. $\begin{bmatrix}3\\-3\\0\end{bmatrix},\begin{bmatrix}-3\\7\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

6. $\begin{bmatrix}1\\2\\-4\end{bmatrix},\begin{bmatrix}-4\\3\\6\end{bmatrix}$

7. $\begin{bmatrix}-2\\3\\0\end{bmatrix},\begin{bmatrix}6\\-1\\5\end{bmatrix}$

8. $\begin{bmatrix}1\\-2\\3\end{bmatrix},\begin{bmatrix}0\\3\\-1\end{bmatrix},\begin{bmatrix}2\\-1\\5\end{bmatrix},\begin{bmatrix}0\\0\\-1\end{bmatrix}$

Encuentre las bases para los espacios nulos de las matrices en los ejercicios 9 y 10. Consulte los comentarios que siguen el ejemplo 3 en la sección 4.2.

9. $\begin{bmatrix}1 & 0 & -2 & -2\\0 & 1 & 1 & 4\\3 & -1 & -7 & 3\end{bmatrix}$

10. $\begin{bmatrix}1 & 1 & -2 & 1 & 5\\0 & 1 & 0 & -1 & -2\\0 & 0 & -8 & 0 & 16\end{bmatrix}$

11. Encuentre una base para el conjunto de vectores en $\mathbb{R}^3$ en el plano $x - 3y + 2z = 0$. [*Sugerencia:* Piense en la ecuación como en un “sistema” de ecuaciones homogéneas].

12. Encuentre una base para el conjunto de vectores de $\mathbb{R}^2$ en la recta $y = -3x$.

En los ejercicios 13 y 14, suponga que A es equivalente por filas a B. Encuentre las bases de Nul A y Col A.

13. $A = \begin{bmatrix}-2 & 4 & -2 & -4\\2 & -6 & -3 & 1\\-3 & 8 & 2 & -3\end{bmatrix}$, $B = \begin{bmatrix}1 & 0 & 6 & 5\\0 & 2 & 5 & 3\\0 & 0 & 0 & 0\end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 2 & 3 & -4 & 8 \\ 1 & 2 & 0 & 2 & 8 \\ 2 & 4 & -3 & 10 & 9 \\ 3 & 6 & 0 & 6 & 9 \end{bmatrix}$,

$B = \begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ 0 & 0 & 3 & -6 & 3 \\ 0 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 15 a 18, determine una base para el espacio generado por los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_5$ dados.

15. $\begin{bmatrix} 1 \\ 0 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -8 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 10 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -6 \\ 9 \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$

17. [**M**] $\begin{bmatrix} 2 \\ 0 \\ -4 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 2 \\ -4 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 0 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} 8 \\ 4 \\ 8 \\ -3 \\ 15 \end{bmatrix}, \begin{bmatrix} -8 \\ 4 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

18. [**M**] $\begin{bmatrix} -3 \\ 2 \\ 6 \\ 0 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -9 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -4 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \\ -14 \\ 0 \\ 13 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \\ 0 \\ -1 \\ 0 \end{bmatrix}$

19. Sea $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -3 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 9 \\ -2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 7 \\ 11 \\ 6 \end{bmatrix}$, y también sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Es posible comprobar que $4\mathbf{v}_1 + 5\mathbf{v}_2 - 3\mathbf{v}_3 = \mathbf{0}$. Utilice esta información para encontrar una base para H. Hay más de una respuesta.

20. Sea $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \\ -2 \\ -5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 4 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 5 \\ -6 \\ -14 \end{bmatrix}$. Es posible comprobar que $2\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. Utilice esta información para encontrar una base para $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Un solo vector por sí mismo es linealmente dependiente.

b) Si $H = \text{Gen}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$, entonces $\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ es una base para H.

c) Las columnas de una matriz invertible de $n \times n$ forman una base para $\mathbb{R}^n$.

d) Una base es un conjunto generador que es tan grande como sea posible.

e) En algunos casos, las relaciones de dependencia lineal entre las columnas de una matriz se pueden ver afectadas por algunas operaciones elementales de fila en la matriz.

22. *a*) Un conjunto linealmente independiente en un subespacio H es una base para H.

b) Si un conjunto finito S de vectores diferentes de cero genera un espacio vectorial V, entonces algún subconjunto de S es una base para V.

c) Una base es un conjunto linealmente independiente que es lo más grande posible.

d) El método estándar para la obtención de un conjunto generador para Nul A, que se describe en la sección 4.2, a veces falla al producir una base para Nul A.

e) Si B es una forma escalonada de una matriz A, entonces las columnas pivote de B forman una base para Col A.

23. Suponga que $\mathbb{R}^4 = \text{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$. Explique por qué $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es una base para $\mathbb{R}^4$.

24. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente en $\mathbb{R}^n$. Explique por qué $\mathcal{B}$ debe ser una base para $\mathbb{R}^n$.

25. Sea $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ y sea H el conjunto de vectores en $\mathbb{R}^3$ cuyas segunda y tercera entradas son iguales. Entonces, todo vector de H tiene una expansión única como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, porque

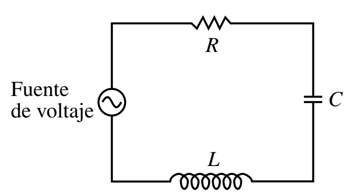
$$\begin{bmatrix} s \\ t \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (t - s)\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

para cualquier s y t. ¿Es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ una base para H? ¿Por qué?

26. En el espacio vectorial de todas las funciones de valores reales, encuentre una base para el subespacio generado por $\{\text{sen}\ t, \text{sen}\ 2t, \text{sen}\ t \cos t\}$.

27. Sea V el espacio vectorial de funciones que describen la vibración de un sistema masa-resorte. (Consulte el ejercicio 19 de la sección 4.1). Encuentre una base para V.

28. (*Circuito RLC*) El circuito de la figura consiste en una resistencia (R, ohms), un inductor (L, henrys), un condensador (C, farads), y una fuente de voltaje inicial. Sea $b = R/(2L)$, y suponga que R, L y C se han seleccionado de manera que b también es igual a $1/\sqrt{LC}$. (Esto se hace, por ejemplo, cuando el circuito se utiliza en un voltímetro). Sea $v(t)$ el voltaje (en volts) en el tiempo t, medido a través del condensador. Es posible demostrar que v está en el espacio nulo H de la transformación lineal que mapea $v(t)$ en $Lv''(t) + Rv'(t) + (1/C)v(t)$, y H se compone de todas las funciones de la forma $v(t) = e^{-bt}(c_1 + c_2 t)$. Encuentre una base para H.



Los ejercicios 29 y 30 indican que cada base de $\mathbb{R}^n$ debe contener exactamente n vectores.

29. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, con $k < n$. Utilice un teorema de la sección 1.4 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

30. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto de k vectores en $\mathbb{R}^n$, con $k > n$. Utilice un teorema del capítulo 1 para explicar por qué S no puede ser una base para $\mathbb{R}^n$.

Los ejercicios 31 y 32 revelan una importante conexión entre la independencia lineal y las transformaciones lineales, y le permiten practicar el uso de la definición de dependencia lineal. Sean V y W espacios vectoriales, sea $T : V \to W$ una transformación lineal, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ un subconjunto de V.

31. Demuestre que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente dependiente en V, entonces el conjunto de imágenes, $\{T(\mathbf{v}_1), \ldots, T(v_p)\}$ es linealmente dependiente en W. Este hecho demuestra que si una transformación lineal mapea un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en un conjunto linealmente *independiente* $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$, entonces el conjunto original es linealmente independiente también (porque no puede ser linealmente dependiente).

32. Suponga que T es una transformación uno a uno, de modo que la ecuación $T(\mathbf{u}) = T(\mathbf{v})$ siempre implica que $\mathbf{u} = \mathbf{v}$. Demuestre que si el conjunto de imágenes $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es linealmente dependiente. Este hecho demuestra que *una transformación lineal uno a uno mapea un conjunto linealmente independiente sobre un conjunto linealmente independiente* (porque en este caso el conjunto de las imágenes no puede ser linealmente dependiente).

33. Considere los polinomios $\mathbf{p}_1(t) = 1 + t^2$ y $\mathbf{p}_2(t) = 1 - t^2$. ¿Es $\{\mathbf{p}_1, \mathbf{p}_2\}$ un conjunto linealmente independiente en $\mathbb{P}_3$? ¿Por qué?

34. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, y $\mathbf{p}_3(t) = 2$ (para toda t). Por inspección, escriba una relación de dependencia lineal entre $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Después encuentre una base para Gen $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

35. Sea V un espacio vectorial que contiene un conjunto linealmente independiente $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Explique cómo construir un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ en V tal que $\{\mathbf{v}_1, \mathbf{v}_3\}$ sea una base para Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

36. [**M**] Sea $H = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y $K = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, donde

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ -4 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ -2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 2 \\ -6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 6 \\ -2 \end{bmatrix}$$

Encuentre las bases para H, K y $H + K$. (Véase los ejercicios 33 y 34 en la sección 4.1).

37. [**M**] Demuestre que $\{t, \text{sen}\, t, \cos 2t, \text{sen}\, t \cos t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Comience suponiendo que

$$c_1 \cdot t + c_2 \cdot \text{sen}\, t + c_3 \cdot \cos 2t + c_4 \cdot \text{sen}\, t \cos t = 0 \qquad (5)$$

La ecuación (5) debe ser válida para toda t real, así que elija varios valores específicos de t (es decir, $t = 0, .1, .2$) hasta obtener un sistema de ecuaciones suficientes para determinar que todas las c_j deben ser cero.

38. [**M**] Demuestre que $\{1, \cos t, \cos^2 t, \ldots, \cos^6 t\}$ es un conjunto linealmente independiente de funciones definidas en $\mathbb{R}$. Utilice el método del ejercicio 37. (Este resultado se necesitará en el ejercicio 34 de la sección 4.5).

WEB

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $A = [\mathbf{v}_1 \quad \mathbf{v}_2]$. Las operaciones de fila muestran que

$$A = \begin{bmatrix} 1 & -2 \\ -2 & 7 \\ 3 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$$

No todas las filas de A contienen una posición pivote. De manera que las columnas de A no generan a $\mathbb{R}^3$, de acuerdo con el teorema 4 de la sección 1.4. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ no es una base para $\mathbb{R}^3$. Ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ no están en $\mathbb{R}^2$, no pueden ser la base de $\mathbb{R}^2$. Sin embargo, puesto que es evidente que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, son una base para un subespacio de $\mathbb{R}^3$, a saber, Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$.

2. Construya una matriz cuyo espacio columna sea el espacio generado por $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, y después reduzca por filas a A para encontrar sus columnas pivote.

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las dos primeras columnas de A son las columnas pivote y, por consiguiente, forman una base de Col A = W. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W. Observe que no se necesita la forma escalonada reducida de A para localizar las columnas pivote.

3. Ni $\mathbf{v}_1$ ni $\mathbf{v}_2$ están en H, por lo que $\{\mathbf{v}_1, \mathbf{v}_2\}$ no puede ser una base para H. De hecho, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para el *plano* de todos los vectores de la forma $(c_1, c_2, 0)$, pero H es solo una *recta*.

4.4 SISTEMAS DE COORDENADAS

Una razón importante para especificar una base $\mathcal{B}$ para un espacio vectorial V es imponer un "sistema de coordenadas" en V. En esta sección se mostrará que si $\mathcal{B}$ contiene n vectores, entonces el sistema de coordenadas hará que V actúe como $\mathbb{R}^n$. Si V es ya $\mathbb{R}^n$ mismo, entonces $\mathcal{B}$ determinará un sistema de coordenadas que da una nueva perspectiva a V.

La existencia de sistemas de coordenadas se basa en el siguiente resultado fundamental.

TEOREMA 7

Teorema de la representación única

Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Así, para cada $\mathbf{x}$ en V, existe un conjunto único de escalares $c_1,\ldots, c_n$ tal que

$$x = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \tag{1}$$

DEMOSTRACIÓN Puesto que $\mathcal{B}$ genera a V, existen escalares tales que la ecuación (1) es válida. Suponga que $\mathbf{x}$ también tiene la representación

$$\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

para escalares $d_1,\ldots, d_n$. Así, al restar, se tiene

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_n - d_n)\mathbf{b}_n \tag{2}$$

Puesto que $\mathcal{B}$ es linealmente independiente, los pesos en la ecuación (2) deben ser cero. Es decir, $c_j = d_j$ para $1 \leq j \leq n$. ■

DEFINICIÓN

Suponga que $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ es una base para V y que $\mathbf{x}$ está en V. Las **coordenadas de x respecto de la base $\mathcal{B}$** (o las **$\mathcal{B}$-coordenadas de x**) son los pesos $c_1,\ldots, c_n$ tales que $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

Si $c_1,\ldots, c_n$ son las $\mathcal{B}$-coordenadas de $\mathbf{x}$, entonces el vector en $\mathbb{R}^n$

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

es el **vector de coordenadas de x (respecto de $\mathcal{B}$), o el vector de $\mathcal{B}$-coordenadas de x**. El mapeo $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$ es el **mapeo de coordenadas (determinado por $\mathcal{B}$)**.[1]

[1] El concepto de un mapeo de coordenadas supone que la base $\mathcal{B}$ es un conjunto indexado cuyos vectores se listan en un orden fijo asignado con anterioridad. Esta propiedad hace que la definición de $[\,\mathbf{x}\,]_{\mathcal{B}}$ no sea ambigua.

EJEMPLO 1 Considere una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ para $\mathbb{R}^2$, donde $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Supongamos que una $\mathbf{x}$ en $\mathbb{R}^2$ tiene el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$. Determine $\mathbf{x}$.

SOLUCIÓN Las coordenadas $\mathcal{B}$ de $\mathbf{x}$ nos dicen cómo construir $\mathbf{x}$ a partir de los vectores en $\mathcal{B}$. Es decir,

$$\mathbf{x} = (-2)\mathbf{b}_1 + 3\mathbf{b}_2 = (-2)\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$ ■

EJEMPLO 2 Las entradas en el vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ son las coordenadas de $\mathbf{x}$ respecto de la *base estándar* $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, ya que

$$\begin{bmatrix} 1 \\ 6 \end{bmatrix} = 1 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 6 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{e}_1 + 6 \cdot \mathbf{e}_2$$

Si $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$, entonces $[\,\mathbf{x}\,]_{\mathcal{E}} = \mathbf{x}$. ■

Una interpretación gráfica de coordenadas

Un sistema de coordenadas en un conjunto se compone de un mapeo uno a uno de los puntos en el conjunto dentro de $\mathbb{R}^n$. Por ejemplo, el papel común para gráficas proporciona un sistema de coordenadas del plano cuando se seleccionan ejes perpendiculares y una unidad de medida en cada eje. La figura 1 muestra la base estándar $\{\mathbf{e}_1, \mathbf{e}_2\}$, los vectores $\mathbf{b}_1$ $(= \mathbf{e}_1)$ y $\mathbf{b}_2$ del ejemplo 1, y del vector $\mathbf{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$. Las coordenadas 1 y 6 dan la ubicación de $\mathbf{x}$ respecto de la base estándar: 1 unidad en la dirección $\mathbf{e}_1$ y 6 unidades en la dirección $\mathbf{e}_2$.

La figura 2 muestra los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{x}$ de la figura 1. (Desde el punto de vista geométrico, los tres vectores se encuentran en una recta vertical en ambas figuras). Sin embargo, el sistema de coordenadas estándar se borró y se remplazó por una malla especialmente adaptada a la base $\mathcal{B}$ del ejemplo 1. El vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ da la ubicación de $\mathbf{x}$ en este nuevo sistema de coordenadas: -2 unidades en la dirección $\mathbf{b}_1$ y 3 unidades en la dirección $\mathbf{b}_2$.

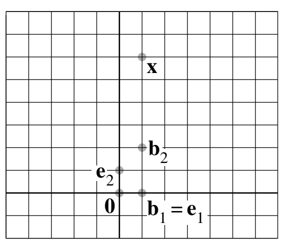


FIGURA 1
Papel cuadriculado estándar.

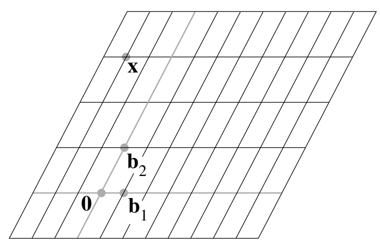
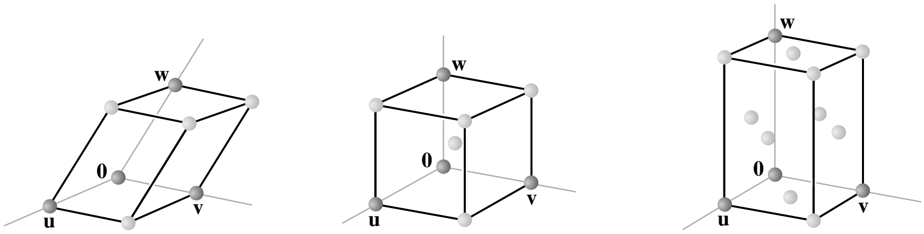


FIGURA 2
Papel para gráficas B.

EJEMPLO 3 En cristalografía, la descripción de una red cristalina se mejora eligiendo una base $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ para $\mathbb{R}^3$ que corresponda a tres aristas adyacentes de una “celda unitaria” del cristal. Se construye una red entera apilando muchas copias de una sola celda. Hay 14 tipos básicos de celdas unitarias; en la figura 3, se muestran tres.[2]

[2] Adaptado de *The Science and Engineering of Materials*, 4a. edición, por Donald R. Askeland (Boston: Prindle, Weber & Schmidt © 2002), p. 36.



a) Monoclínica simple   *b)* Cúbica centrada en el cuerpo   *c)* Ortorrómbica centrada en la cara

FIGURA 3 Ejemplos de celdas unitarias.

Las coordenadas de los átomos dentro del cristal están dadas respecto de la base de la red. Por ejemplo,

$$\begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

identifica el átomo centrado en la cara superior de la celda en la figura 3 *c*). ■

Coordenadas en $\mathbb{R}^n$

Cuando una base $\mathcal{B}$ para $\mathbb{R}^n$ está fija, el vector de coordenadas $\mathcal{B}$ de una $\mathbf{x}$ dada es fácil de determinar, como se muestra en el siguiente ejemplo.

EJEMPLO 4 Sean $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Determine el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ de $\mathbf{x}$ respecto de $\mathcal{B}$.

SOLUCIÓN Las coordenadas $\mathcal{B}$ de c_1, c_2 de $\mathbf{x}$ satisfacen

$$c_1 \underset{\mathbf{b}_1}{\begin{bmatrix} 2 \\ 1 \end{bmatrix}} + c_2 \underset{\mathbf{b}_2}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{\mathbf{x}}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}}$$

o

$$\begin{array}{cc} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \\ \mathbf{b}_1 \quad \mathbf{b}_2 \end{array} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \underset{\mathbf{x}}{\begin{bmatrix} 4 \\ 5 \end{bmatrix}} \tag{3}$$

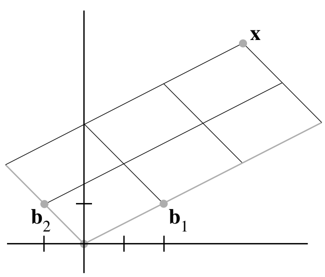


FIGURA 4 El vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$ es (3, 2).

Esta ecuación se puede resolver mediante operaciones de fila en una matriz aumentada o utilizando la inversa de la matriz a la izquierda. En cualquier caso, la solución es $c_1 = 3$, $c_2 = 2$. Por lo tanto, $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2$, y

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$ ■

Véase la figura 4.

La matriz en la ecuación (3) cambia las coordenadas $\mathcal{B}$ de un vector $\mathbf{x}$ en las coordenadas estándar para $\mathbf{x}$. Es posible realizar un cambio análogo de coordenadas en $\mathbb{R}^n$ para una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Sea

$$P_{\mathcal{B}} = \{\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n\}$$

Entonces la ecuación vectorial

$$\mathbf{x} = c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n$$

es equivalente a

$$\boxed{\mathbf{x} = P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}} \tag{4}$$

Se llama $P_{\mathcal{B}}$ a la **matriz de cambio de coordenadas** de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$. Multiplicando por la izquierda a $P_{\mathcal{B}}$ se transforma al vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ en $\mathbf{x}$. La ecuación de cambio de coordenadas (4) es importante y será necesaria en varias secciones de los capítulos 5 y 7.

Puesto que las columnas de $P_{\mathcal{B}}$ forman una base para $\mathbb{R}^n$, $P_{\mathcal{B}}$ es invertible (por el teorema de la matriz invertible). Al multiplicar por la izquierda por $P_{\mathcal{B}}^{-1}$ se convierte a $\mathbf{x}$ en su vector de coordenadas $\mathcal{B}$:

$$P_{\mathcal{B}}^{-1}\mathbf{x} = [\mathbf{x}]_{\mathcal{B}}$$

La correspondencia $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$, producida aquí por $P_{\mathcal{B}}^{-1}$, es el mapeo de coordenadas mencionado anteriormente. Ya que $P_{\mathcal{B}}^{-1}$ es una matriz invertible, el mapeo de coordenadas es una transformación lineal uno a uno de $\mathbb{R}^n$ en $\mathbb{R}^n$, de acuerdo con el teorema de la matriz invertible. (Véase también el teorema 12 de la sección 1.9). Esta propiedad del mapeo de coordenadas también es cierto en un espacio vectorial general que tiene una base, como se verá más adelante.

El mapeo de coordenadas

La elección de una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ de un espacio vectorial V introduce un sistema de coordenadas en V. El mapeo de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ conecta al espacio V, posiblemente desconocido, con el conocido espacio $\mathbb{R}^n$. Véase la figura 5. Los puntos de V ahora se pueden identificar por sus nuevos "nombres".

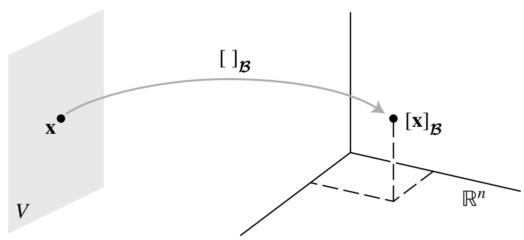


FIGURA 5 El mapeo de coordenadas de V sobre $\mathbb{R}^n$.

TEOREMA 8

Sea $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Así, el mapeo de coordenadas $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ es una transformación lineal uno a uno de V en $\mathbb{R}^n$.

DEMOSTRACIÓN Tome dos vectores típicos en V, por ejemplo,

$$\begin{aligned} \mathbf{u} &= c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \\ \mathbf{w} &= d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n \end{aligned}$$

Luego, utilizando las operaciones de vectores,

$$\mathbf{u} + \mathbf{w} = (c_1 + d_1)\mathbf{b}_1 + \cdots + (c_n + d_n)\mathbf{b}_n$$

De ello se sigue que

$$[\mathbf{u}+\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}$$

Por lo tanto, el mapeo de coordenadas conserva la adición. Si r es un escalar cualquiera, entonces

$$r\mathbf{u} = r(c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n) = (rc_1)\mathbf{b}_1 + \cdots + (rc_n)\mathbf{b}_n$$

De esta forma,

$$[r\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} rc_1 \\ \vdots \\ rc_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = r[\mathbf{u}]_{\mathcal{B}}$$

Así, el mapeo de coordenadas también conserva la multiplicación escalar, y por consiguiente, es una transformación lineal. Véase los ejercicios 23 y 24 para comprobar que el mapeo de coordenadas es uno a uno y mapea V sobre $\mathbb{R}^n$. ■

La linealidad del mapeo de coordenadas se extiende a las combinaciones lineales, al igual que en la sección 1.8. Si $\mathbf{u}_1,\ldots,\mathbf{u}_p$ están en V y si $c_1,\ldots,c_p$ son escalares, entonces

$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_p[\mathbf{u}_p]_{\mathcal{B}} \tag{5}$$

Es decir, (5) nos dice que el vector de $\mathcal{B}$-coordenadas de una combinación lineal de $\mathbf{u}_1,\ldots,\mathbf{u}_p$ es la *misma* combinación lineal de sus vectores de coordenadas.

El mapeo de coordenadas en el teorema 8 es un importante ejemplo de un *isomorfismo* de V en $\mathbb{R}^n$. En general, una transformación lineal uno a uno de un espacio vectorial V en un espacio vectorial W se llama **isomorfismo** de V en W (el término proviene de los vocablos griegos *iso*, que significa "lo mismo", y *morfé*, que significa "forma" o "estructura"). La notación y la terminología para V y W pueden diferir, pero los dos espacios son indistinguibles como espacios vectoriales. *Cada cálculo de espacio vectorial en V se reproduce con exactitud en W, y viceversa*. En particular, cualquier espacio vectorial real con una base de n vectores es indistinguible de $\mathbb{R}^n$. Véase los ejercicios 25 y 26.

EJEMPLO 5 Sea $\mathcal{B}$ la base estándar del espacio $\mathbb{P}_3$ de polinomios; es decir, sea $\mathcal{B} = \{1, t, t^2, t^3\}$. Un elemento típico $\mathbf{p}$ de $\mathbb{P}_3$ tiene la forma

$$\mathbf{p}(t) = a_0 + a_1t + a_2t^2 + a_3t^3$$

Puesto que $\mathbf{p}$ ya se ha desplegado como una combinación lineal de los vectores básicos estándar, concluimos que

$$[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

Así, el mapeo de coordenadas $\mathbf{p} \mapsto [\mathbf{p}]_{\mathcal{B}}$ es un isomorfismo de $\mathbb{P}_3$ sobre $\mathbb{R}^4$. Todas las operaciones de espacio vectorial en $\mathbb{P}_3$ corresponden a operaciones en $\mathbb{R}^4$. ■

Si pensamos en $\mathbb{P}_3$ y $\mathbb{R}^4$ como despliegues en dos pantallas de computadora que se conectan a través del mapeo de coordenadas, entonces cada operación de espacio vectorial en $\mathbb{P}_3$ en una pantalla es duplicada exactamente por una operación vectorial correspondiente en $\mathbb{R}^4$ en la otra pantalla. Los vectores en la pantalla $\mathbb{P}_3$ tienen un aspecto diferente respecto de los que aparecen en la pantalla de $\mathbb{R}^4$, pero "actúan" como vectores exactamente de la misma forma. Véase la figura 6.

FIGURA 6 El espacio $\mathbb{P}_3$ es isomorfo a $\mathbb{R}^4$.

EJEMPLO 6 Utilice vectores de coordenadas para comprobar que los polinomios $1 + 2t^2$, $4 + t + 5t^2$, y $3 + 2t$ son linealmente dependientes en $\mathbb{P}_2$.

SOLUCIÓN El mapeo de coordenadas del ejemplo 5 produce los vectores de coordenadas (1, 0, 2), (4, 1, 5) y (3, 2, 0), respectivamente. Al representar estos vectores como las *columnas* de una matriz A, podemos determinar su independencia mediante reducción por filas de la matriz aumentada de $A\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 5 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las columnas de A son linealmente dependientes, por lo que los polinomios correspondientes son linealmente dependientes. De hecho, es fácil comprobar que la columna 3 de A es la columna 2 multiplicada por 2, menos la columna 1 multiplicada por 5. La relación correspondiente de los polinomios es

$$3 + 2t = 2(4 + t + 5t^2) - 5(1 + 2t^2)$$ ■

El último ejemplo se refiere a un plano en $\mathbb{R}^3$ que es isomorfo a $\mathbb{R}^2$.

EJEMPLO 7 Sea

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}.$$

y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Entonces $\mathcal{B}$ es una base para $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$. Determine si $\mathbf{x}$ se encuentra en H y, si lo está, encuentre el vector de coordenadas de $\mathbf{x}$ respecto de $\mathcal{B}$.

SOLUCIÓN Si $\mathbf{x}$ está en H, entonces la siguiente ecuación vectorial es consistente:

$$c_1 \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 12 \\ 7 \end{bmatrix}$$

Los escalares c_1 y c_2, si existen, son las coordenadas $\mathcal{B}$ de $\mathbf{x}$. Usando operaciones de fila, se obtiene

$$\begin{bmatrix} 3 & -1 & 3 \\ 6 & 0 & 12 \\ 2 & 1 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Así, $c_1 = 2$, $c_2 = 3$ y $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. El sistema de coordenadas en H determinadas por $\mathcal{B}$ se muestra en la figura 7. ■

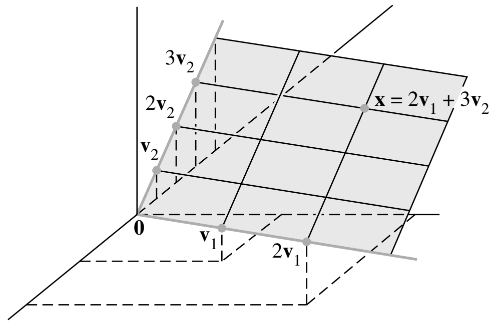


FIGURA 7 Un sistema de coordenadas en un plano H en $\mathbb{R}^3$.

Si se eligiera una base diferente para H, ¿el sistema de coordenadas asociado también haría a H isomorfo a $\mathbb{R}^2$? Sin duda, esto debe ser verdad. Así lo demostraremos en la siguiente sección.

PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

a) Demuestre que el conjunto $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ es una base de $\mathbb{R}^3$.

b) Encuentre la matriz de cambio de las coordenadas de $\mathcal{B}$ a la base estándar.

c) Escriba la ecuación que relaciona $\mathbf{x}$ en $\mathbb{R}^3$ con $[\,\mathbf{x}\,]_{\mathcal{B}}$.

d) Encuentre $[\,\mathbf{x}\,]_{\mathcal{B}}$, para la $\mathbf{x}$ dada anteriormente.

2. El conjunto $\mathcal{B} = \{1 + t, 1 + t^2, t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ en relación con $\mathcal{B}$.

4.4 EJERCICIOS

En los ejercicios 1 a 4, encuentre el vector $\mathbf{x}$ determinado por el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ y la base $\mathcal{B}$.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$

3. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -3 \\ 0 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

4. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} \right\}$, $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$

En los ejercicios 5 a 8, encuentre el vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$ de $\mathbf{x}$ respecto de la base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$.

5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ -6 \end{bmatrix}$

7. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$

En los ejercicios 9 y 10, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a la base estándar en $\mathbb{R}^n$.

9. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \end{bmatrix} \right\}$

10. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \right\}$

En los ejercicios 11 y 12, utilice una matriz inversa para encontrar $[\,\mathbf{x}\,]_{\mathcal{B}}$ para $\mathbf{x}$ y $\mathcal{B}$.

11. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$

12. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

13. El conjunto $\mathcal{B} = \{1 + t^2, t + t^2, 1 + 2t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 1 + 4t + 7t^2$ respecto de $\mathcal{B}$.

14. El conjunto $\mathcal{B} = \{1 - t^2, t - t^2, 2 - t + t^2\}$ es una base para $\mathbb{P}_2$. Encuentre el vector de coordenadas de $\mathbf{p}(t) = 1 + 3t - 6t^2$ respecto de $\mathcal{B}$.

En los ejercicios 15 y 16, marque cada enunciado como verdadero o falso. Justifique sus respuestas. A menos que se indique lo contrario, $\mathcal{B}$ es una base para un espacio vectorial V.

15. *a*) Si $\mathbf{x}$ está en V y si $\mathcal{B}$ contiene n vectores, entonces el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$ está en $\mathbb{R}^n$.

b) Si $P_{\mathcal{B}}$ es la matriz del cambio de coordenadas, entonces $[\,\mathbf{x}\,]_{\mathcal{B}} = P_{\mathcal{B}}\mathbf{x}$, para $\mathbf{x}$ en V.

c) Los espacios vectoriales $\mathbb{P}_3$ y $\mathbb{R}^3$ son isomorfos.

16. *a*) Si $\mathcal{B}$ es la base estándar para $\mathbb{R}^n$, entonces el vector de coordenadas $\mathcal{B}$ de una $\mathbf{x}$ en $\mathbb{R}^n$ es $\mathbf{x}$ misma.

b) La correspondencia $[\,\mathbf{x}\,]_{\mathcal{B}} \mapsto \mathbf{x}$ se llama el mapeo de coordenadas.

c) En algunos casos, un plano en $\mathbb{R}^3$ puede ser isomorfo a $\mathbb{R}^2$.

17. Los vectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ generan a $\mathbb{R}^2$, pero no forman una base. Encuentre formas diferentes de expresar $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

18. Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para un espacio vectorial V. Explique por qué los vectores de coordenadas $\mathcal{B}$ de $\mathbf{b}_1,\ldots, \mathbf{b}_n$ son las columnas $\mathbf{e}_1,\ldots, \mathbf{e}_n$ de la matriz identidad de $n \times n$.

19. Sea S un conjunto finito en un espacio vectorial V con la propiedad de que cada $\mathbf{x}$ en V tiene una representación única como una combinación lineal de elementos de S. Demuestre que S es una base de V.

20. Suponga que $\{\mathbf{v}_1,\ldots, \mathbf{v}_4\}$ es un conjunto linealmente dependiente que genera un espacio vectorial V. Demuestre que cada $\mathbf{w}$ en V se puede expresar en más de una forma como una combinación lineal de $\mathbf{v}_1,\ldots, \mathbf{v}_4$. [*Sugerencia:* Considere $\mathbf{w} = k_1\mathbf{v}_1 + \cdots + k_4\mathbf{v}_4$ un vector arbitrario en V. Utilice la dependencia lineal de $\{\mathbf{v}_1,\ldots, \mathbf{v}_4\}$ para obtener otra representación de $\mathbf{w}$ como una combinación lineal de $\mathbf{v}_1,\ldots, \mathbf{v}_4$].

21. Sea $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Puesto que el mapeo de coordenadas determinado por $\mathcal{B}$ es una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$, este mapeo se debe implementar mediante alguna matriz A de 2×2. Encuéntrela. [*Sugerencia:* La multiplicación por A debería transformar un vector $\mathbf{x}$ en su vector de coordenadas $[\,\mathbf{x}\,]_{\mathcal{B}}$].

22. Sea $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ una base para $\mathbb{R}^n$. Obtenga una descripción de una matriz A de $n \times n$ que implemente el mapeo de coordenadas $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$. (Véase el ejercicio 21).

Los ejercicios 23 a 26 se refieren a un espacio vectorial V, una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, y el mapeo de coordenadas $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$.

23. Demuestre que el mapeo de coordenadas es uno a uno. (*Sugerencia:* Suponga que $[\,\mathbf{u}\,]_{\mathcal{B}} = [\,\mathbf{w}\,]_{\mathcal{B}}$ para algunas $\mathbf{u}$ y $\mathbf{w}$ en V, y demuestre que $\mathbf{u} = \mathbf{w}$).

24. Demuestre que el mapeo de coordenadas es *sobre* $\mathbb{R}^n$. Es decir, dada cualquier $\mathbf{y}$ en $\mathbb{R}^n$, con entradas $y_1,\ldots, y_n$, obtenga $\mathbf{u}$ en V tal que $[\,\mathbf{u}\,]_{\mathcal{B}} = \mathbf{y}$.

25. Demuestre que un subconjunto $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ en V es linealmente independiente si y solo si el conjunto de vectores de coordenadas $\{[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbb{R}^n$. (*Sugerencia:* Como el mapeo de coordenadas es uno a uno, las siguientes ecuaciones tienen las mismas soluciones, $c_1,\ldots, c_p$).

$$c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = \mathbf{0} \qquad \text{El vector cero en } V$$
$$[c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p]_{\mathcal{B}} = [\mathbf{0}]_{\mathcal{B}} \qquad \text{El vector cero en } \mathbb{R}^n$$

26. Dados los vectores $\mathbf{u}_1,\ldots, \mathbf{u}_p$, y $\mathbf{w}$ en V, demuestre que $\mathbf{w}$ es una combinación lineal de $\mathbf{u}_1,\ldots, \mathbf{u}_p$ si y solo si $[\,\mathbf{w}\,]_{\mathcal{B}}$ es una combinación lineal de los vectores de coordenadas $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$.

En los ejercicios 27 a 30, utilice los vectores de coordenadas para probar la independencia lineal de los conjuntos de polinomios. Explique su trabajo.

27. $1 + 2t^3,\ 2 + t - 3t^2,\ -t + 2t^2 - t^3$

28. $1 - 2t^2 - t^3,\ t + 2t^3,\ 1 + t - 2t^2$

29. $(1 - t)^2,\ t - 2t^2 + t^3,\ (1 - t)^3$

30. $(2 - t)^3,\ (3 - t)^2,\ 1 + 6t - 5t^2 + t^3$

31. Utilice los vectores de coordenadas para comprobar si los siguientes conjuntos de polinomios generan a $\mathbb{P}_2$. Justifique sus conclusiones.

a) $1 - 3t + 5t^2, -3 + 5t - 7t^2, -4 + 5t - 6t^2, 1 - t^2$

b) $5t + t^2, 1 - 8t - 2t^2, -3 + 4t + 2t^2, 2 - 3t$

32. Sea $\mathbf{p}_1(t) = 1 + t^2, \mathbf{p}_2(t) = t - 3t^2, \mathbf{p}_3(t) = 1 + t - 3t^2$.

a) Utilice vectores de coordenadas para demostrar que estos polinomios forman una base para $\mathbb{P}_2$.

b) Considere de la base $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ para $\mathbb{P}_2$. Encuentre $\mathbf{q}$ en $\mathbb{P}_2$, considerando que $[\mathbf{q}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$.

En los ejercicios 33 y 34, determine si el conjunto de polinomios forman una base para $\mathbb{P}_3$. Justifique sus conclusiones.

33. [**M**] $3 + 7t, 5 + t - 2t^3, t - 2t^2, 1 + 16t - 6t^2 + 2t^3$

34. [**M**] $5 - 3t + 4t^2 + 2t^3, 9 + t + 8t^2 - 6t^3, 6 - 2t + 5t^2, t^3$

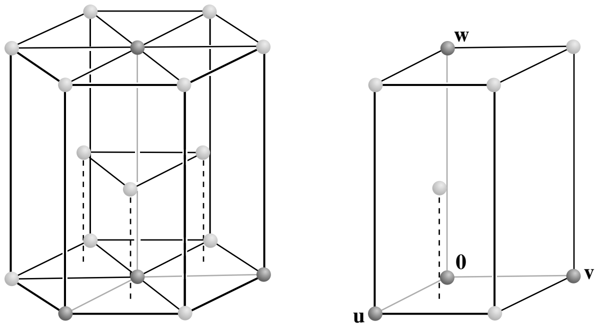
35. [**M**] Sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Demuestre que $\mathbf{x}$ está en H y encuentre el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$, para

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

36. [**M**] Sea $H = \text{Gen}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ y $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Demuestre que $\mathcal{B}$ es una base para H y que $\mathbf{x}$ está en H, y encuentre el vector de coordenadas $\mathcal{B}$ de $\mathbf{x}$, para

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

[**M**] Los ejercicios 37 y 38 se refieren a la red cristalina del titanio, que tiene la estructura hexagonal que se ilustra a la izquierda en la figura adjunta. Los vectores $\begin{bmatrix} 2.6 \\ -1.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4.8 \end{bmatrix}$ en $\mathbb{R}^3$ forman una base para la celda unitaria de la derecha. Los números están en unidades *angstrom* (1 Å = 10^{-8} cm). En las aleaciones de titanio, algunos átomos adicionales pueden estar en la celda unitaria en los sitios *octaédricos* y *tetraédricos* (llamados así por los objetos geométricos que forman los átomos en estos lugares).



La red compacta hexagonal y su celda unitaria.

37. Uno de los sitios octaédricos es $\begin{bmatrix} 1/2 \\ 1/4 \\ 1/6 \end{bmatrix}$, respecto de la base de la red. Determine las coordenadas de este sitio en relación con la base estándar de $\mathbb{R}^3$.

38. Uno de los sitios tetraédricos es $\begin{bmatrix} 1/2 \\ 1/2 \\ 1/3 \end{bmatrix}$. Determine las coordenadas de este sitio en relación con la base estándar de $\mathbb{R}^3$.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *a*) Es evidente que la matriz $P_{\mathcal{B}} = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,]$ es equivalente por filas a la matriz de identidad. Por el teorema de la matriz invertible, $P_{\mathcal{B}}$ es invertible y sus columnas forman una base de $\mathbb{R}^3$.

b) Del inciso *a*), la matriz de cambio de coordenadas es $P_{\mathcal{B}} = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c) $\mathbf{x} = P_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}}$

d) Para resolver la ecuación en *c*), es probable que sea más fácil de reducir por filas una matriz aumentada que calcular $P_{\mathcal{B}}^{-1}$:

$$\underbrace{\begin{bmatrix} 1 & -3 & 3 & -8 \\ 0 & 4 & -6 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix}}_{P_{\mathcal{B}} \quad\quad \mathbf{x}} \sim \underbrace{\begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}}_{I \quad\quad [\,\mathbf{x}\,]_{\mathcal{B}}}$$

De ahí que

$$[\,\mathbf{x}_{\mathcal{B}}\,] = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

2. Las coordenadas de $\mathbf{p}(t) = 6 + 3t - t^2$ respecto de $\mathcal{B}$ satisfacen

$$c_1(1 + t) + c_2(1 + t^2) + c_3(t + t^2) = 6 + 3t - t^2$$

Al igualar los coeficientes de potencias de t, se tiene que

$$\begin{aligned} c_1 + c_2 \quad\;\; &= 6 \\ c_1 \quad\;\; + c_3 &= 3 \\ c_2 + c_3 &= -1 \end{aligned}$$

Al resolver, se encuentra que $c_1 = 5$, $c_2 = 1$, $c_3 = -2$ y $[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

4.5 LA DIMENSIÓN DE UN ESPACIO VECTORIAL

El teorema 8 de la sección 4.4 implica que un espacio vectorial V con una base B que contiene n vectores es isomorfo a $\mathbb{R}^n$. En esta sección veremos que este número n es una propiedad intrínseca (llamada dimensión) del espacio V que no depende de la elección particular de la base. El análisis de la dimensión le ayudará a comprender mejor las propiedades de las bases.

El primer teorema generaliza un resultado bien conocido acerca del espacio vectorial $\mathbb{R}^n$.

TEOREMA 9

Si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, entonces cualquier conjunto en V que contenga más de n vectores debe ser linealmente dependiente.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ un conjunto de V con más de n vectores. Los vectores de coordenadas $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \ldots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$ forman un conjunto linealmente dependiente en $\mathbb{R}^n$, porque hay más vectores (p) que entradas (n) en cada vector. Por lo tanto, existen escalares $c_1,\ldots, c_p$, no todos cero, tales que

$$c_1[\,\mathbf{u}_1\,]_{\mathcal{B}} + \cdots + c_p[\,\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{El vector cero en } \mathbb{R}^n$$

Como el mapeo de coordenadas es una transformación lineal,

$$[\,c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

El vector cero de la derecha muestra los n pesos necesarios para construir el vector $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ de los vectores básicos en $\mathcal{B}$. Es decir, $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = 0 \cdot \mathbf{b}_1 + \cdots + 0 \cdot \mathbf{b}_n = \mathbf{0}$. Puesto que no todas las c_i son cero, $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ es linealmente dependiente.[1] ■

El teorema 9 implica que si un espacio vectorial V tiene una base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$, entonces cada conjunto linealmente independiente ubicado en V no tiene más de n vectores.

[1] El teorema 9 también se aplica a los conjuntos infinitos en V. Se dice que un conjunto infinito es linealmente dependiente si algún subconjunto finito es linealmente dependiente; de lo contrario, el conjunto es linealmente independiente. Si S es un conjunto infinito de V, tome cualquier subconjunto $\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ de S, con $p > n$. La demostración anterior indica que este subconjunto es linealmente dependiente y, por lo tanto, también lo es S.

TEOREMA 10

Si un espacio vectorial V tiene una base de n vectores, entonces toda base de V debe consistir exactamente en n vectores.

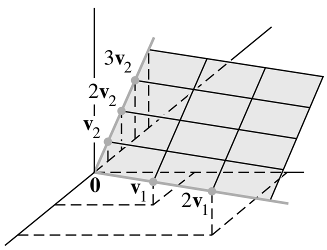
DEMOSTRACIÓN Sea $\mathcal{B}_1$ una base de n vectores y $\mathcal{B}_2$ cualquier otra base (de V). Ya que $\mathcal{B}_1$ es una base y $\mathcal{B}_2$ es linealmente independiente, $\mathcal{B}_2$ no tiene más de n vectores, de acuerdo con el teorema 9. Además, puesto que $\mathcal{B}_2$ es una base y $\mathcal{B}_1$ es linealmente independiente, $\mathcal{B}_2$ tiene al menos n vectores. Así, $\mathcal{B}_2$ se compone de exactamente n vectores. ■

Si un espacio vectorial V distinto de cero es generado por un conjunto finito S, entonces un subconjunto de S es una base para V, de acuerdo con el teorema del conjunto generador. En este caso, el teorema 10 asegura que la siguiente definición tiene sentido.

DEFINICIÓN

Si V es generado por un conjunto finito, entonces se dice que V tiene **dimensión finita**, y la **dimensión** de V, representada como dim V, es el número de vectores en una base para V. La dimensión del espacio vectorial cero $\{\mathbf{0}\}$ se define como cero. Si V no es generado por un conjunto finito, entonces se dice que V tiene **dimensión infinita**.

EJEMPLO 1 La base estándar para $\mathbb{R}^n$ contiene n vectores, por lo que dim $\mathbb{R}^n = n$. La base polinomial estándar $\{1, t, t^2\}$ indica que dim $\mathbb{P}_2 = 3$. En general, dim $\mathbb{P}_n = n + 1$. El espacio $\mathbb{P}$ de todos los polinomios es de dimensión infinita (ejercicio 27). ■



EJEMPLO 2 Sea $H = \text{Gen}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$. Entonces H es el plano estudiado en el ejemplo 7 de la sección 4.4. Una base para H es $\{\mathbf{v}_1, \mathbf{v}_2\}$, ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ no son múltiplos y, por consiguiente, son linealmente independientes. Por lo tanto, dim $H = 2$. ■

EJEMPLO 3 Encuentre la dimensión del subespacio

$$H = \left\{ \begin{bmatrix} a - 3b + 6c \\ 5a + 4d \\ b - 2c - d \\ 5d \end{bmatrix} : a, b, c, d \text{ en } \mathbb{R} \right\}$$

SOLUCIÓN Es fácil ver que H es el conjunto de todas las combinaciones lineales de los vectores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Claramente, $\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{v}_2$ no es un múltiplo de $\mathbf{v}_1$, pero $\mathbf{v}_3$ es un múltiplo de $\mathbf{v}_2$. De acuerdo con el teorema del conjunto generador, podemos descartar a $\mathbf{v}_3$ y aún así tener un conjunto generador H. Finalmente, $\mathbf{v}_4$ no es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Por lo tanto, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ es linealmente independiente (de acuerdo con el teorema 4 de la sección 4.3) y, en consecuencia, es una base para H. Por consiguiente, dim $H = 3$. ■

EJEMPLO 4 Los subespacios de $\mathbb{R}^3$ se pueden clasificar por dimensiones. Véase la figura 1.

Subespacios de dimensión 0. Solo el subespacio cero.

Subespacios de dimensión 1. Cualquier subespacio generado por un solo vector distinto de cero. Tales subespacios son rectas que pasan por el origen.

Subespacios de dimensión 2. Cualquier subespacio generado por dos vectores linealmente independientes. Tales subespacios son planos que pasan por el origen.

Subespacios de dimensión 3. Solo el propio $\mathbb{R}^3$. Cualesquiera tres vectores linealmente independientes en $\mathbb{R}^3$ generan todo $\mathbb{R}^3$, de acuerdo con el teorema de la matriz invertible. ■

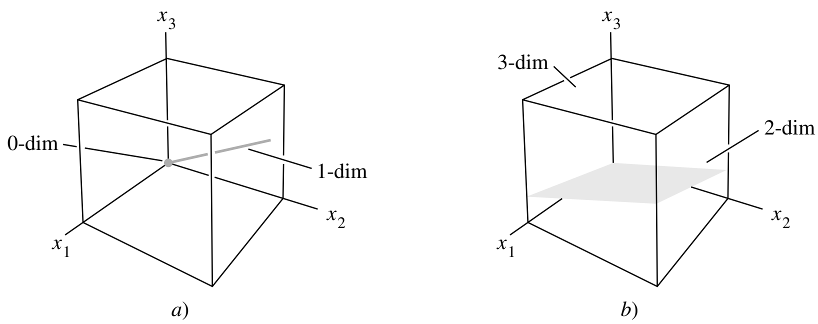


FIGURA 1 Subespacios muestra de $\mathbb{R}^3$.

Subespacios de un espacio de dimensión finita

El siguiente teorema es una contraparte natural del teorema del conjunto generador.

TEOREMA 11

Sea H un subespacio de un espacio vectorial de dimensión finita V. Cualquier conjunto linealmente independiente en H se puede expandir, si es necesario, a una base de H. Además, H tiene dimensión finita y

$$\dim H \leq \dim V$$

DEMOSTRACIÓN Si $H = \{\mathbf{0}\}$, entonces, sin duda, $\dim H = 0 \leq \dim V$. De lo contrario, sea $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ cualquier conjunto linealmente independiente en H. Si S genera a H, entonces S es una base para H. De lo contrario, existe algún $\mathbf{u}_{k+1}$ en H que no está en el generado por S. Pero entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}\}$ será linealmente independiente, ya que ningún vector en el conjunto puede ser una combinación lineal de los vectores que le preceden (de acuerdo con el teorema 4).

En tanto que el nuevo conjunto no genere a H, podemos continuar con este proceso de expansión de S a un conjunto más amplio linealmente independiente en H. Sin embargo, el número de vectores en una expansión linealmente independiente de S nunca podrá superar la dimensión de V, de acuerdo con el teorema 9. Así, finalmente la expansión de S generará a H y, por lo tanto, será una base para H, y $\dim H \leq \dim V$. ■

Cuando se conoce la dimensión de un espacio o subespacio vectorial, la búsqueda de una base se simplifica con el siguiente teorema, el cual dice que si un conjunto tiene el número correcto de elementos, entonces solo se tiene que demostrar ya sea que el conjunto es linealmente independiente o que este genera el espacio. El teorema es de importancia fundamental en muchos problemas de aplicación (por ejemplo, los que implican ecuaciones diferenciales o ecuaciones en diferencias), donde la independencia lineal es mucho más fácil de comprobar que la generación.

TEOREMA 12

El teorema base

Sea V un espacio vectorial de dimensión p, donde $p \geq 1$. Cualquier conjunto linealmente independiente de exactamente p elementos en V es, de forma automática, una base para V. Cualquier conjunto de exactamente p elementos que genera a V es, de manera automática, una base para V.

DEMOSTRACIÓN De acuerdo con el teorema 11, un conjunto linealmente independiente S de p elementos se puede ampliar a una base para V. Sin embargo, esa base debe contener exactamente p elementos, ya que dim $V = p$. Así que S debe ser ya una base para V. Ahora suponga que S tiene p elementos y genera a V. Puesto que V es diferente de cero, el teorema del conjunto generador implica que un subconjunto S' de S es una base de V. Como dim $V = p$, S' debe contener p vectores. Por lo tanto, $S = S'$. ■

Las dimensiones de Nul A y Col A

Como las columnas pivote de una matriz A forman una base para Col A, conocemos la dimensión de Col A tan pronto como se conocen las columnas pivote. Tal vez parezca que la dimensión de Nul A requiere de más trabajo, ya que encontrar una base para Nul A, por lo general, toma más tiempo que una base para Col A. Pero, ¡hay un atajo!

Sea A una matriz de $m \times n$, y supongamos que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene k variables libres. De la sección 4.2, sabemos que el método estándar para encontrar un conjunto generador para Nul A producirá exactamente k vectores linealmente independientes, por ejemplo, $\mathbf{u}_1, \ldots, \mathbf{u}_k$, uno para cada variable libre. Por lo tanto, $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base para Nul A, y el número de variables libres determina el tamaño de la base. Hagamos un resumen de estos hechos para referencia futura.

La dimensión de Nul A es el número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$, y la dimensión de Col A es el número de columnas pivote de A.

EJEMPLO 5 Determine las dimensiones del espacio nulo y el espacio columna de

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

SOLUCIÓN Reduzca por filas la matriz aumentada $[A \quad \mathbf{0}]$ a la forma escalonada:

$$\begin{bmatrix} 1 & -2 & 2 & 3 & -1 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hay tres variables libres: x_2, x_4 y x_5. Por consiguiente, la dimensión de Nul A es 3. Por otra parte, dim Col $A = 2$ porque A tiene dos columnas pivote. ■

PROBLEMAS DE PRÁCTICA

Determine si cada enunciado es verdadero o falso, y dé una razón para cada respuesta. Aquí, V es un espacio vectorial de dimensión finita diferente de cero.

1. Si dim $V = p$, y si S es un subconjunto linealmente dependiente de V, entonces S contiene más que p vectores.
2. Si S genera a V y si T es un subconjunto de V que contiene más vectores que S, entonces T es linealmente dependiente.

4.5 EJERCICIOS

Para cada subespacio en los ejercicios 1 a 8, *a*) encuentre una base para el subespacio, y *b*) indique la dimensión.

1. $\left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ en } \mathbb{R} \right\}$

2. $\left\{ \begin{bmatrix} 2a \\ -4b \\ -2a \end{bmatrix} : a, b \text{ en } \mathbb{R} \right\}$

3. $\left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ en } \mathbb{R} \right\}$

4. $\left\{ \begin{bmatrix} p+2q \\ -p \\ 3p-q \\ p+q \end{bmatrix} : p, q \text{ en } \mathbb{R} \right\}$

5. $\left\{ \begin{bmatrix} p-2q \\ 2p+5r \\ -2q+2r \\ -3p+6r \end{bmatrix} : p, q, r \text{ en } \mathbb{R} \right\}$

6. $\left\{ \begin{bmatrix} 3a-c \\ -b-3c \\ -7a+6b+5c \\ -3a+c \end{bmatrix} : a, b, c \text{ en } \mathbb{R} \right\}$

7. $\{(a, b, c) : a-3b+c=0, b-2c=0, 2b-c=0\}$

8. $\{(a, b, c, d) : a-3b+c=0\}$

9. Encuentre la dimensión del subespacio de todos los vectores en $\mathbb{R}^3$ cuyas entradas primera y tercera sean iguales.

10. Encuentre la dimensión del subespacio H de $\mathbb{R}^2$ generado por
$\begin{bmatrix} 1 \\ -5 \end{bmatrix}, \begin{bmatrix} -2 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 15 \end{bmatrix}$

En los ejercicios 11 y 12, encuentre la dimensión del subespacio generado por los vectores dados.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

12. $\begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \\ 5 \end{bmatrix}$

Determine las dimensiones de Nul A y Col A de las matrices que se muestran en los ejercicios 13 a 18.

13. $A = \begin{bmatrix} 1 & -6 & 9 & 0 & -2 \\ 0 & 1 & 2 & -4 & 5 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 2 & -4 & 3 & -2 & 6 & 0 \\ 0 & 0 & 0 & 1 & 0 & -3 & 7 \\ 0 & 0 & 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

16. $A = \begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix}$

17. $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

18. $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 19 y 20, V es un espacio vectorial. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

19. *a*) El número de columnas pivote de una matriz es igual a la dimensión de su espacio columna.
    b) Un plano en $\mathbb{R}^3$ es un subespacio de dimensión 2 de $\mathbb{R}^3$.
    c) La dimensión del espacio vectorial $\mathbb{P}_4$ es 4.
    d) Si dim $V = n$ y S es un conjunto linealmente independiente en V, entonces S es una base para V.
    e) Si un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ genera un espacio vectorial V de dimensión finita y si T es un conjunto de más de p vectores en V, entonces T es linealmente dependiente.

20. *a*) $\mathbb{R}^2$ es un subespacio de dimensión 2 de $\mathbb{R}^3$.
    b) El número de variables en la ecuación $A\mathbf{x} = \mathbf{0}$ es igual a la dimensión de Nul A.
    c) Un espacio vectorial es de dimensión infinita si es generado por un conjunto infinito.
    d) Si dim $V = n$, y si S genera a V, entonces S es una base de V.
    e) El único subespacio de dimensión 3 de $\mathbb{R}^3$ es el propio $\mathbb{R}^3$.

21. Los primeros cuatro polinomios de Hermite son $1, 2t, -2+4t^2$, y $-12t+8t^3$. Estos polinomios surgen de forma natural en el estudio de ciertas ecuaciones diferenciales importantes en física matemática.[2] Demuestre que los primeros cuatro polinomios de Hermite forman una base de $\mathbb{P}_3$.

22. Los primeros cuatro polinomios de Laguerre son 1, $1-t$, $2-4t+t^2$, y $6-18t+9t^2-t^3$. Demuestre que estos polinomios forman una base de $\mathbb{P}_3$.

23. Sea B la base de $\mathbb{P}_3$ que consta de los polinomios de Hermite en el ejercicio 21, y sea $\mathbf{p}(t) = -1+8t^2+8t^3$. Encuentre el vector de coordenadas de $\mathbf{p}$ respecto de $\mathcal{B}$.

24. Sea $\mathcal{B}$ la base de $\mathbb{P}_2$ que consiste en los tres primeros polinomios de Laguerre listados en el ejercicio 22, y sea $\mathbf{p}(t) = 5+5t-2t^2$. Encuentre el vector de coordenadas de $\mathbf{p}$ respecto de $\mathcal{B}$.

25. Sea S un subconjunto de un espacio vectorial V de dimensión n, y suponga que S contiene menos de n vectores. Explique por qué S no puede generar a V.

26. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Demuestre que $H = V$.

27. Explique por qué el espacio $\mathbb{P}$ de todos los polinomios es un espacio de dimensión infinita.

[2] Véase *Introduction to Functional Analysis*, 2a. edición, por A. E. Taylor y David C. Lay (Nueva York: John Wiley & Sons, 1980), pp. 92-93. También se analizan otros conjuntos de polinomios.

28. Demuestre que el espacio $C(\mathbb{R})$ de todas las funciones continuas definidas en la recta real es un espacio de dimensión infinita.

En los ejercicios 29 y 30, V es un espacio vectorial de dimensión finita diferente de cero, y los vectores mencionados pertenecen a V. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. (Estas preguntas tienen mayor grado de dificultad que las de los ejercicios 19 y 20).

29. *a*) Si existe un conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ que genera a V, entonces $\dim V \leq p$.

b) Si existe un conjunto linealmente independiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V, entonces $\dim V \geq p$.

c) Si $\dim V = p$, entonces existe un conjunto generador de $p + 1$ vectores en V.

30. *a*) Si existe un conjunto linealmente dependiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en V, entonces $\dim V \leq p$.

b) Si cada conjunto de p elementos en V no genera a V, entonces $\dim V > p$.

c) Si $p \geq 2$ y $\dim V = p$, entonces cada conjunto de $p - 1$ vectores distintos de cero es linealmente independiente.

Los ejercicios 31 y 32 se refieren a espacios vectoriales V y W de dimensión finita y a una transformación lineal $T : V \to W$.

31. Sea H un subespacio distinto de cero de V, y sea $T(H)$ el conjunto de imágenes de vectores en H. Entonces, $T(H)$ es un subespacio de W, de acuerdo con el ejercicio 35 en la sección 4.2. Demuestre que $\dim T(H) \leq \dim H$.

32. Sea H un subespacio distinto de cero de V, y suponga que T es un mapeo uno a uno (lineal) de V en W. Demuestre que $\dim T(H) = \dim H$. Si resulta que T es un mapeo uno a uno de V *sobre* W, entonces $\dim V = \dim W$. Espacios vectoriales isomorfos de dimensión finita tienen la misma dimensión.

33. [**M**] De acuerdo con el teorema 11, un conjunto linealmente independiente $\{\mathbf{v}_1,\ldots, \mathbf{v}_k\}$ en $\mathbb{R}^n$ se puede expandir a una base para $\mathbb{R}^n$. Una manera de hacer esto es crear $A = [\mathbf{v}_1 \cdots \mathbf{v}_k \quad \mathbf{e}_1 \cdots \mathbf{e}_n]$, siendo $\mathbf{e}_1,\ldots, \mathbf{e}_n$ las columnas de la matriz identidad; las columnas pivote de A forman una base para $\mathbb{R}^n$.

a) Utilice el método descrito para ampliar los siguientes vectores a una base para $\mathbb{R}^5$:

$$\mathbf{v}_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

b) Explique por qué funciona en general el método: ¿por qué están los vectores originales $\mathbf{v}_1,\ldots, \mathbf{v}_k$ incluidos en la base encontrada para Col A? ¿Por qué es Col $A = \mathbb{R}^n$?

34. [**M**] Sea $\mathcal{B} = \{1, \cos t, \cos^2 t,\ldots, \cos^6 t\}$ y $\mathcal{C} = \{1, \cos t, \cos 2t,\ldots, \cos 6t\}$. Suponga las siguientes identidades trigonométricas (véase el ejercicio 37 de la sección 4.1).

$$\begin{aligned}
\cos 2t &= -1 + 2\cos^2 t \\
\cos 3t &= -3\cos t + 4\cos^3 t \\
\cos 4t &= 1 - 8\cos^2 t + 8\cos^4 t \\
\cos 5t &= 5\cos t - 20\cos^3 t + 16\cos^5 t \\
\cos 6t &= -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t
\end{aligned}$$

Sea H el subespacio de funciones generado por las funciones en $\mathcal{B}$. Entonces $\mathcal{B}$ es una base para H, de acuerdo con el ejercicio 38 de la sección 4.3.

a) Escriba los vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$, y utilícelos para demostrar que $\mathcal{C}$ es un conjunto linealmente independiente en H.

b) Explique por qué $\mathcal{C}$ es una base para H.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Falso. Considere el conjunto $\{\mathbf{0}\}$.
2. Verdadero. De acuerdo con el teorema del conjunto generador, S contiene una base para V; llamémosla base S'. Así, T contendrá más vectores que S'. De acuerdo con el teorema 9, T es linealmente dependiente.

4.6 RANGO

Con la ayuda de los conceptos de espacio vectorial, esta sección ofrece una perspectiva desde *el interior* de una matriz y revela varias relaciones interesantes y útiles, ocultas en sus filas y columnas.

Por ejemplo, imagine que se colocan 2000 números aleatorios en una matriz A de 40×50 y después se determina el número máximo de columnas linealmente independientes de A y el número máximo de columnas linealmente independientes de A^T (filas de A). De manera sorprendente, los dos números son iguales. Como pronto veremos, su valor común es el *rango* de la matriz. Para explicar por qué, necesitamos examinar el subespacio generado por las filas de A.

El espacio fila

Si A es una matriz de $m \times n$, cada fila de A tiene n entradas y, por lo tanto, se puede identificar con un vector en $\mathbb{R}^n$. El conjunto de todas las combinaciones lineales de los vectores fila se denomina **espacio fila** de A y se denota como Fila A. Cada fila tiene n entradas, por lo que Fila A es un subespacio de $\mathbb{R}^n$. Ya que las filas de A se identifican con las columnas de A^T, también podríamos escribir Col A^T en lugar de Fila A.

EJEMPLO 1 Sea

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \quad \text{y} \quad \begin{aligned} \mathbf{r}_1 &= (-2, -5, 8, 0, -17) \\ \mathbf{r}_2 &= (1, 3, -5, 1, 5) \\ \mathbf{r}_3 &= (3, 11, -19, 7, 1) \\ \mathbf{r}_4 &= (1, 7, -13, 5, -3) \end{aligned}$$

El espacio fila de A es el subespacio de $\mathbb{R}^5$ generado por $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Es decir, Fila $A = \text{Gen}\,\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4\}$. Es natural representar vectores fila de forma horizontal; sin embargo, también es posible representarlos como vectores columna si resulta más conveniente. ■

Si supiéramos algo de las relaciones de dependencia lineal entre las filas de la matriz A del ejemplo 1, podríamos usar el teorema del conjunto generador para reducir el tamaño del conjunto generador a una base. Por desgracia, las operaciones de fila en A no nos dan esa información, porque las operaciones de fila cambian las relaciones de dependencia de filas. Pero la reducción de filas de A, ¡sin duda vale la pena, como muestra el siguiente teorema!

TEOREMA 13

> Si dos matrices A y B son equivalentes por filas, entonces sus espacios fila son iguales. Si B está en forma escalonada, las filas de B diferentes de cero forman una base para el espacio fila de A, así como para el de B.

DEMOSTRACIÓN Si B se obtiene a partir de A mediante operaciones de fila, las filas de B son combinaciones lineales de las filas de A. De ello se desprende que cualquier combinación lineal de las filas de B es automáticamente una combinación lineal de las filas de A. Así, el espacio fila de B está contenido en el espacio fila de A. Ya que las operaciones de fila son reversibles, el mismo argumento indica que el espacio fila de A es un subconjunto del espacio fila de B. De manera que los dos espacios fila son iguales. Si B está en forma escalonada, sus filas diferentes de cero son linealmente independientes porque ninguna fila diferente de cero es una combinación lineal de las filas distintas de cero debajo de esta. (Aplique el teorema 4 para las filas diferentes de cero de B en orden inverso, con la primera fila como la última). Así, las filas diferentes de cero de B forman una base del espacio fila (común) de B y A. ■

El resultado principal de esta sección implica los tres espacios: Fila A, Col A y Nul A. El siguiente ejemplo prepara el camino para este resultado y muestra cómo *una* secuencia de operaciones de fila de A conduce a las bases para los tres espacios.

EJEMPLO 2 Encuentre bases para el espacio fila, el espacio columna y el espacio nulo de la matriz

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

SOLUCIÓN Para encontrar las bases para el espacio fila y el espacio columna, reduzca A por filas a una forma escalonada:

$$A \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

De acuerdo con el teorema 13, las tres primeras filas de B forman una base para el espacio fila de A (así como para el espacio fila de B). Así,

$$\text{Base para Fila } A\text{: } \{(1, 3, -5, 1, 5), (0, 1, -2, 2, -7), (0, 0, 0, -4, 20)\}$$

Para el espacio columna, observe a partir de B que los pivotes están en las columnas 1, 2 y 4. Por lo tanto, las columnas 1, 2 y 4 de A (no de B) forman una base para Col A:

$$\text{Base para Col } A\text{: } \left\{ \begin{bmatrix} -2 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 3 \\ 11 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 7 \\ 5 \end{bmatrix} \right\}$$

Considere que cualquier forma escalonada de A proporciona (en sus filas diferentes de cero) una base para Fila A y también identifica las columnas pivote de A para Col A. Sin embargo, para Nul A, se necesita la *forma escalonada reducida*. Otras operaciones de fila sobre B dan como resultado

$$A \sim B \sim C = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La ecuación $A\mathbf{x} = \mathbf{0}$ es equivalente a $C\mathbf{x} = \mathbf{0}$, es decir,

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 - 2x_3 + 3x_5 &= 0 \\ x_4 - 5x_5 &= 0 \end{aligned}$$

Así, $x_1 = -x_3 - x_5$, $x_2 = 2x_3 - 3x_5$, $x_4 = 5x_5$, con x_3 y x_5 como variables libres. Los cálculos usuales (que se analizan en la sección 4.2) demuestran que

$$\text{Base para Nul } A\text{: } \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 0 \\ 5 \\ 1 \end{bmatrix} \right\}$$

Observe que, a diferencia de la base para Col A, las bases para Fila A y Nul A no tienen una relación sencilla con las propias entradas de A.[1] ■

[1] Es posible encontrar una base para Fila A que utiliza filas de A. En primer lugar se determina A^T, y luego se reduce por filas hasta que se encuentren las columnas pivote de A^T. Estas columnas pivote de A^T son filas de A, y forman una base para el espacio fila de A.

Advertencia: A pesar de que las tres primeras filas de B en el ejemplo 2 son linealmente independientes, es erróneo concluir que las tres primeras filas de A son linealmente independientes. (De hecho, la tercera fila de A es la primera fila multiplicada por 2, más la segunda fila multiplicada por 7). Las operaciones de fila pueden cambiar las relaciones de dependencia lineal entre las *filas* de una matriz.

El teorema del rango

WEB

El siguiente teorema describe las relaciones fundamentales entre las dimensiones de Col A, Fila A y Nul A.

DEFINICIÓN

El **rango** de A es la dimensión del espacio columna de A.

Puesto que Fila A es igual que Col A^T, la dimensión del espacio fila de A es el rango de A^T. La dimensión del espacio nulo a veces se llama la **nulidad** de A, aunque no utilizaremos este término.

Tal vez un lector atento ya haya descubierto la totalidad o parte del siguiente teorema, mientras trabajaba con los ejercicios de la sección 4.5 o al leer el ejemplo 2 anterior.

TEOREMA 14

El teorema del rango

Las dimensiones del espacio columna y del espacio fila de una matriz A de $m \times n$ son iguales. Esta dimensión común, el rango de A, también es igual al número de posiciones pivote en A y satisface la ecuación

$$\text{rango } A + \dim \text{Nul } A = n$$

DEMOSTRACIÓN De acuerdo con el teorema 6 de la sección 4.3, rango A es el número de columnas pivote de A. De manera equivalente, rango A es el número de posiciones pivote en una forma escalonada B de A. Además, puesto que B tiene una fila diferente de cero para cada pivote, y como estas filas forman una base para el espacio fila de A, el rango de A también es la dimensión del espacio fila.

A partir de la sección 4.5, la dimensión de Nul A es igual al número de variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$. Dicho de otra manera, la dimensión de Nul A es el número de columnas de A que *no* son columnas pivote. (Es el número de estas columnas, no las columnas mismas, lo que se relaciona con Nul A). Como es evidente,

$$\left\{\begin{matrix}\text{número de}\\ \text{columnas pivote}\end{matrix}\right\} + \left\{\begin{matrix}\text{número de columnas}\\ \text{que no son pivote}\end{matrix}\right\} = \left\{\begin{matrix}\text{número de}\\ \text{columnas}\end{matrix}\right\}$$

Esto demuestra el teorema. ■

Las ideas que sustentan el teorema 14 se pueden distinguir en los cálculos del ejemplo 2. Las tres posiciones pivote en la forma escalonada B determinan las variables básicas e identifican los vectores básicos para Col A y Fila A.

EJEMPLO 3

a) Si A es una matriz de 7×9 con un espacio nulo de dimensión 2, ¿cuál es el rango de A?

b) ¿Podría una matriz de 6×9 tener un espacio nulo de dimensión 2?

SOLUCIÓN

a) Puesto que A tiene 9 columnas, (rango A) + 2 = 9; por lo tanto, rango A = 7.

b) No. Si una matriz de 6×9, llamémosla B, tuviera un espacio nulo de dimensión 2, tendría rango 7, de acuerdo con el teorema del rango. Sin embargo, las columnas de B son vectores en $\mathbb{R}^6$, de manera que la dimensión de Col B no puede ser superior a 6; es decir, el rango de B no puede ser mayor que 6. ■

El siguiente ejemplo proporciona una buena forma de visualizar los subespacios que hemos estudiado. En el capítulo 6 se verá que Fila A y Nul A tienen solo el vector cero en común y, en realidad, son "perpendiculares" entre sí. Este mismo hecho se aplicará a Fila A^T (= Col A) y Nul A^T. Por lo tanto, la figura 1, que acompaña al ejemplo 4, crea una buena imagen mental para el caso general. (El valor de estudiar A^T junto con A se demuestra en el ejercicio 29).

EJEMPLO 4 Sea $A = \begin{bmatrix} 3 & 0 & -1 \\ 3 & 0 & -1 \\ 4 & 0 & 5 \end{bmatrix}$. Se comprueba rápidamente que Nul A es el eje x_2, Fila A es el plano x_1x_3, Col A es el plano cuya ecuación es $x_1 - x_2 = 0$, y Nul A^T es el conjunto de todos los múltiplos de $(1, -1, 0)$. La figura 1 muestra Nul A y Fila A en el dominio de la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$; el rango de este mapeo, Col A, se muestra en una copia separada de $\mathbb{R}^3$, junto con Nul A^T. ■

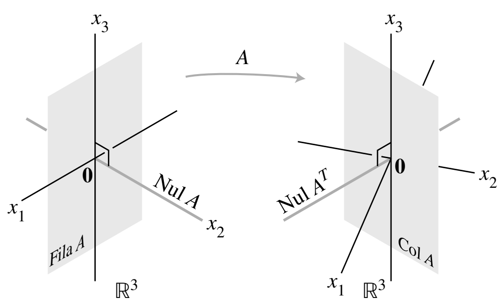


FIGURA 1 Subespacios determinados por una matriz A.

Aplicaciones para sistemas de ecuaciones

El teorema del rango es una poderosa herramienta para el procesamiento de información sobre los sistemas de ecuaciones lineales. El siguiente ejemplo simula la forma como se plantearía un problema de la vida real utilizando ecuaciones lineales, sin mencionar explícitamente términos de álgebra lineal, como matriz, subespacio y dimensión.

EJEMPLO 5 Un científico encontró dos soluciones para un sistema homogéneo de 40 ecuaciones con 42 variables. No son múltiplos las dos soluciones, y todas las demás soluciones se pueden desarrollar sumando múltiplos adecuados de estas dos soluciones. ¿Puede el científico estar *seguro* de que un sistema no homogéneo asociado (con los mismos coeficientes) tiene una solución?

SOLUCIÓN Sí. Sea A una matriz de coeficientes de 40×42 del sistema. La información dada implica que las dos soluciones son linealmente independientes y generan Nul A. Así que dim Nul $A = 2$. De acuerdo con el teorema del rango, dim Col $A = 42 - 2 = 40$. Como $\mathbb{R}^{40}$ es el único subespacio de $\mathbb{R}^{40}$ cuya dimensión es 40, Col A debe ser todo de $\mathbb{R}^{40}$. Esto significa que cada ecuación no homogénea $A\mathbf{x} = \mathbf{b}$ tiene una solución. ■

El rango y el teorema de la matriz invertible

Los distintos conceptos de espacios vectoriales asociados a una matriz proporcionan varios enunciados adicionales al teorema de la matriz invertible. Los nuevos enunciados que se listan a continuación se deducen del teorema de la matriz invertible original de la sección 2.3.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. Cada uno de los siguientes enunciados es equivalente a la afirmación de que A es una matriz invertible.

m) Las columnas de A forman una base de $\mathbb{R}^n$.

n) $\text{Col } A = \mathbb{R}^n$

o) $\dim \text{Col } A = n$

p) $\text{rango } A = n$

q) $\text{Nul } A = \{\mathbf{0}\}$

r) $\dim \text{Nul } A = 0$

DEMOSTRACIÓN El enunciado *m*) es lógicamente equivalente a los enunciados *e*) y *h*) en relación con la independencia lineal y la generación. Los otros cinco enunciados están vinculados a los anteriores del teorema mediante la siguiente cadena de implicaciones casi triviales:

$$g) \Rightarrow n) \Rightarrow o) \Rightarrow p) \Rightarrow r) \Rightarrow q) \Rightarrow d)$$

El enunciado *g*), el cual dice que la ecuación $A\mathbf{x} = \mathbf{b}$ tiene al menos una solución para cada $\mathbf{b}$ en $\mathbb{R}^n$, implica a *n*), porque Col A es precisamente el conjunto de todas las $\mathbf{b}$ tales que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente. Las implicaciones $n) \Rightarrow o) \Rightarrow p)$ se deducen de las definiciones de dimensión y rango. Si el rango de A es n, el número de columnas de A, entonces $\dim \text{Nul } A = 0$, de acuerdo con el teorema del rango, y $\text{Nul } A = \{\mathbf{0}\}$. Por lo tanto, $p) \Rightarrow r) \Rightarrow q)$. Además, *q*) implica que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial, que es el enunciado *d*). Como ya se sabe que los enunciados *d*) y *g*) son equivalentes al enunciado de que A es invertible, la demostración está completa. ■

Nos hemos abstenido de agregar al teorema de la matriz invertible enunciados evidentes acerca del espacio fila de A, ya que el espacio fila es el espacio columna de A^T. Recuerde el enunciado (1) del teorema de la matriz invertible, el cual afirma que A es invertible si y solo si A^T es invertible. Por lo tanto, cada enunciado en el teorema de la matriz invertible también se puede establecer respecto de A^T. Pero hacer esto duplicaría la longitud del teorema y se obtendría una lista ¡de más de 30 enunciados!

NOTA NUMÉRICA

Muchos algoritmos analizados en este libro son útiles para comprender conceptos y efectuar cálculos sencillos a mano. Sin embargo, los algoritmos a menudo son inadecuados para los grandes problemas que surgen en la vida real.

La determinación de un rango es un buen ejemplo. Tal vez parezca fácil reducir una matriz a la forma escalonada y contar los pivotes. Pero, a menos que se realice aritmética exacta en una matriz cuyas entradas se especifiquen con exactitud, las operaciones de fila pueden cambiar el rango aparente de una matriz. Por ejemplo, si el valor de x en la matriz $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ no se almacena exactamente como 7 en una computadora, entonces el rango podría ser 1 o 2, dependiendo de si la computadora trata o no a $x - 7$ como cero.

En las aplicaciones prácticas, el rango efectivo de una matriz A con frecuencia se determina a partir de la descomposición en valores singulares de A, que se analizará en la sección 7.4. Esta descomposición también es una fuente confiable de bases para Col A, Fila A, Nul A y Nul A^T.

WEB

PROBLEMAS DE PRÁCTICA

Las matrices que se muestran a continuación son equivalentes por filas.

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Determine rango A y dim Nul A.
2. Encuentre bases para Col A y Fila A.
3. ¿Cuál es el siguiente paso a realizar con la finalidad de encontrar una base para Nul A?
4. ¿Cuántas columnas pivote están en una forma escalonada por filas de A^T?

4.6 EJERCICIOS

En los ejercicios 1 a 4, suponga que la matriz A es equivalente por filas a B. Sin hacer cálculos, liste rango de A y dim Nul A. Luego, encuentre las bases para Col A, Fila A y Nul A.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 3 & 4 & -1 & 2 \\ 2 & 6 & 6 & 0 & -3 \\ 3 & 9 & 3 & 6 & -3 \\ 3 & 9 & 0 & 9 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 4 & -1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ -2 & -3 & 6 & -3 & 0 & -6 \\ 4 & 9 & -12 & 9 & 3 & 12 \\ -2 & 3 & 6 & 3 & 3 & -6 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ 0 & 3 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & -2 & 0 & 1 & -2 \\ 1 & 2 & -3 & 0 & -2 & -3 \\ 1 & -1 & 0 & 0 & 1 & 6 \\ 1 & -2 & 2 & 1 & -3 & 0 \\ 1 & -2 & 1 & 0 & 2 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 & -2 & 0 & 1 & -2 \\ 0 & 1 & -1 & 0 & -3 & -1 \\ 0 & 0 & 1 & 1 & -13 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

5. Si una matriz A de 4×7 tiene rango 3, determine dim Nul A, dim Fila A y el rango de A^T.

6. Si una matriz A de 7×5 tiene rango 2, determine dim Nul A, dim Fila A y el rango de A^T.

7. Suponga que una matriz A de 4×7 tiene cuatro columnas pivote. ¿Es Col $A = \mathbb{R}^4$? ¿Es Nul $A = \mathbb{R}^3$? Explique sus respuestas.

8. Supongamos que una matriz A de 6×8 tiene cuatro columnas pivote. ¿Cuál es la dim Nul A? ¿Es Col $A = \mathbb{R}^4$? ¿Por qué?

9. Si el espacio nulo de una matriz A de 4×6 es de dimensión 3, ¿cuál es la dimensión del espacio columna de A? ¿Es Col $A = \mathbb{R}^3$? ¿Por qué?

10. Si el espacio nulo de una matriz A de 8×7 es de dimensión 5, ¿cuál es la dimensión del espacio columna de A?

11. Si el espacio nulo de una matriz A de 8×5 es de dimensión 3, ¿cuál es la dimensión del espacio fila de A?

12. Si el espacio nulo de una matriz A de 5×4 es de dimensión 2, ¿cuál es la dimensión del espacio fila de A?

13. Si A es una matriz de 7×5, ¿cuál es el mayor rango posible de A? Si A es una matriz de 5×7, ¿cuál es el mayor rango posible de A? Explique sus respuestas.

14. Si A es una matriz de 5×4, ¿cuál es la mayor dimensión posible del espacio fila de A? Si A es una matriz de 4×5, ¿cuál es la máxima dimensión posible del espacio fila de A? Explique sus respuestas.

15. Si A es una matriz de 3×7, ¿cuál es la menor dimensión posible de Nul A?

16. Si A es una matriz de 7×5, ¿cuál es la menor dimensión posible de Nul A?

En los ejercicios 17 y 18, A es una matriz de $m \times n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) El espacio fila de A es el mismo que el espacio columna de A^T.

b) Si B es cualquier forma escalonada de A, y si B tiene tres filas diferentes de cero, entonces las tres primeras filas de A forman una base para Fila A.

c) Las dimensiones del espacio fila y el espacio columna de A son iguales, incluso si A no es cuadrada.

d) La suma de las dimensiones del espacio fila y el espacio nulo de A es igual al número de filas en A.

e) En una computadora, las operaciones de fila pueden modificar el rango aparente de una matriz.

18. *a*) Si B es cualquier forma escalonada de A, entonces las columnas pivote de B forman una base para el espacio columna de A.

b) Las operaciones de fila preservan las relaciones de dependencia lineal entre las filas de A.

c) La dimensión del espacio nulo de A es el número de columnas de A que *no* son columnas pivote.

d) El espacio fila de A^T es igual que el espacio columna de A.

e) Si A y B son equivalentes por filas, entonces sus espacios fila son iguales.

19. Suponga que todas las soluciones de un sistema homogéneo de cinco ecuaciones lineales con seis incógnitas son múltiplos de una solución diferente de cero. ¿El sistema necesariamente tendrá una solución para cada posible elección de las constantes en el lado derecho de las ecuaciones? Explique su respuesta.

20. Suponga que un sistema no homogéneo de seis ecuaciones lineales con ocho incógnitas tiene una solución, con dos variables libres. ¿Es posible cambiar algunas constantes en los miembros derechos de las ecuaciones para hacer que el nuevo sistema sea inconsistente? Explique su respuesta.

21. Suponga que un sistema no homogéneo de nueve ecuaciones lineales con 10 incógnitas tiene una solución para todas las posibles constantes de los miembros derechos de las ecuaciones. ¿Es posible encontrar dos soluciones diferentes de cero del sistema homogéneo asociado que *no* sean múltiplos una de la otra? Analice.

22. ¿Es posible que todas las soluciones de un sistema homogéneo de 10 ecuaciones lineales con 12 variables sean múltiplos de una solución fija diferente de cero? Analice.

23. Un sistema homogéneo de 12 ecuaciones lineales con ocho incógnitas tiene dos soluciones fijas que no son múltiplos una de la otra, y todas las demás soluciones son combinaciones lineales de estas dos soluciones. ¿Puede describirse el conjunto de todas las soluciones con menos de 12 ecuaciones lineales homogéneas? Si es así, ¿con cuántas? Analice.

24. ¿Es posible que un sistema no homogéneo de siete ecuaciones con seis incógnitas tenga una solución única para algún conjunto de constantes del miembro derecho? ¿Es posible que este sistema tenga una solución única para cada miembro derecho? Explique sus respuestas.

25. Un científico resuelve un sistema no homogéneo de 10 ecuaciones lineales con 12 incógnitas y encuentra que tres de las incógnitas son variables libres. ¿Puede el científico estar seguro de que, si se cambia el lado derecho de las ecuaciones, el nuevo sistema no homogéneo tendrá una solución? Analice.

26. En teoría estadística, un requisito común es que una matriz sea de *rango completo*. Es decir, el rango debe ser tan grande como sea posible. Explique por qué una matriz de $m \times n$ con más filas que columnas tiene rango completo si y solo si sus columnas son linealmente independientes.

Los ejercicios 27 a 29 se refieren a una matriz A de $m \times n$ y a lo que con frecuencia se denomina los *subespacios fundamentales* determinados por A.

27. ¿Cuál de los subespacios Fila A, Col A, Nul A, Fila A^T, Col A^T y Nul A^T están en $\mathbb{R}^m$ y cuáles están en $\mathbb{R}^n$? ¿Cuántos subespacios distintos están en esta lista?

28. Justifique las siguientes igualdades:

a) dim Fila A + dim Nul $A = n$ Número de columnas de A

b) dim Col A + dim Nul $A^T = m$ Número de filas de A

29. Con base en el ejercicio 28, explique por qué la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para toda $\mathbf{b}$ en $\mathbb{R}^m$ si y solo si la ecuación $A^T\mathbf{x} = \mathbf{0}$ tiene solamente la solución trivial.

30. Suponga que A es $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. ¿Qué tiene que ser verdad acerca del rango de $[A \quad \mathbf{b}]$ y el rango A para que la ecuación $A\mathbf{x} = \mathbf{b}$ sea coherente?

Las matrices con rango 1 son importantes en algunos algoritmos de computadora y varios contextos teóricos, incluyendo la descomposición en valores singulares del capítulo 7. Es posible demostrar que una matriz A de $m \times n$ tiene rango 1 si y solo si se trata de un producto externo, es decir, $A = \mathbf{u}\mathbf{v}^T$ por alguna $\mathbf{u}$ en $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$. Los ejercicios 31 a 33 sugieren por qué esta propiedad es verdadera.

31. Compruebe que el rango de $\mathbf{u}\mathbf{v}^T \leq 1$ si $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

32. Sea $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Encuentre $\mathbf{v}$ en $\mathbb{R}^3$ tal que $\begin{bmatrix} 1 & -3 & 4 \\ 2 & -6 & 8 \end{bmatrix} = \mathbf{u}\mathbf{v}^T$.

33. Sea A cualquier matriz de 2×3 tal que rango $A = 1$. Sea $\mathbf{u}$ la primera columna de A, y suponga que $\mathbf{u} \neq \mathbf{0}$. Explique por qué hay un vector $\mathbf{v}$ en $\mathbb{R}^3$ tal que $A = \mathbf{u}\mathbf{v}^T$. ¿Cómo podría modificarse esta construcción si la primera columna de A fuera cero?

34. Sea A una matriz de $m \times n$ de rango $r > 0$, y sea U una forma escalonada de A. Explique por qué existe una matriz invertible E tal que $A = EU$, y utilice esta factorización para escribir A como la suma de r matrices con rango 1. [*Sugerencia*: Véase el teorema 10 de la sección 2.4].

35. [**M**] Sea $A = \begin{bmatrix} 7 & -9 & -4 & 5 & 3 & -3 & -7 \\ -4 & 6 & 7 & -2 & -6 & -5 & 5 \\ 5 & -7 & -6 & 5 & -6 & 2 & 8 \\ -3 & 5 & 8 & -1 & -7 & -4 & 8 \\ 6 & -8 & -5 & 4 & 4 & 9 & 3 \end{bmatrix}$.

a) Construya matrices C y N, cuyas columnas son las bases para Col A y Nul A, respectivamente, y construya una matriz R cuyas *filas* formen una base para Fila A.

b) Construya una matriz M cuyas columnas formen una base para Nul A^T, construya las matrices $S = [R^T \quad N]$ y $T = [C \quad M]$, y explique por qué S y T deberían ser cuadradas. Compruebe que tanto S como T son invertibles.

36. [**M**] Repita el ejercicio 35 para una matriz aleatoria A de 6×7 con valores enteros y rango de 4 o menor. Una manera de construir A es crear una matriz aleatoria J de 6×4 con valores enteros y una matriz aleatoria K de 4×7 con valores enteros, y establecer que $A = JK$. (Véase el ejercicio complementario 12 al final del capítulo; véase también la *Guía de estudio* para los programas de generación de matrices).

37. [**M**] Sea A la matriz del ejercicio 35. Construya una matriz C cuyas columnas sean las columnas pivote de A, y construya una matriz R cuyas filas sean las filas diferentes de cero de la forma escalonada reducida de A. Calcule CR, y analice sus hallazgos.

38. [**M**] Repita el ejercicio 37 para tres matrices aleatorias A de 5×7 con valores enteros y rangos de 5, 4 y 3. Haga una conjetura acerca de cómo se relaciona CR con A para cualquier matriz A. Demuestre su suposición.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. A tiene dos columnas pivote, por lo que rango $A = 2$. Puesto que A tiene 5 columnas en total, dim Nul $A = 5 - 2 = 3$.
2. Las columnas pivote de A son las dos primeras columnas. Por lo tanto, una base para Col A es

$$\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

Las filas diferentes de cero de B forman una base para Fila A, a saber, $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. En este ejemplo en particular, resulta que dos filas cualesquiera de A forman una base para el espacio fila, ya que el espacio fila es de dimensión 2 y ninguna de las filas de A es un múltiplo de otra fila. En general, las filas diferentes de cero de una forma escalonada de A se deberían utilizar como base para Fila A, y no las filas de A.

3. Para Nul A, el siguiente paso es llevar a cabo operaciones de fila en B para obtener la forma escalonada reducida de A.
4. Rango A^T = rango A, de acuerdo con el teorema del rango, ya que Col A^{T} = Fila A. De manera que A^T tiene dos posiciones pivote.

4.7 CAMBIO DE BASE

Cuando se elige una base $\mathcal{B}$ para un espacio vectorial V de dimensión n, el mapeo de coordenadas asociado a $\mathbb{R}^n$ proporciona un sistema de coordenadas para V. Se identifica cada $\mathbf{x}$ en V únicamente por su vector de $\mathcal{B}$-coordenadas, $[\,\mathbf{x}\,]_{\mathcal{B}}$.[1]

En algunas aplicaciones, se describe un problema inicialmente usando una base $\mathcal{B}$, pero la solución del problema se facilita cambiando $\mathcal{B}$ a una nueva base $\mathcal{C}$. (Se darán ejemplos en los capítulos 5 y 7). A cada vector se le asigna un nuevo vector de $\mathcal{C}$-coordenadas. En esta sección se estudia cómo $[\,\mathbf{x}\,]_{\mathcal{C}}$ y $[\,\mathbf{x}\,]_{\mathcal{B}}$ se relacionan para toda $\mathbf{x}$ en V.

Para visualizar el problema, considere los dos sistemas de coordenadas de la figura 1. En la figura 1*a*), $\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2$, mientras que en la figura 1*b*), la misma $\mathbf{x}$ se muestra como $\mathbf{x} = 6\mathbf{c}_1 + 4\mathbf{c}_2$. Es decir,

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Nuestro problema es encontrar la conexión entre los dos vectores de coordenadas. El ejemplo 1 muestra cómo hacer esto, siempre que se conozca cómo se forman $\mathbf{b}_1$ y $\mathbf{b}_2$ a partir de $\mathbf{c}_1$ y $\mathbf{c}_2$.

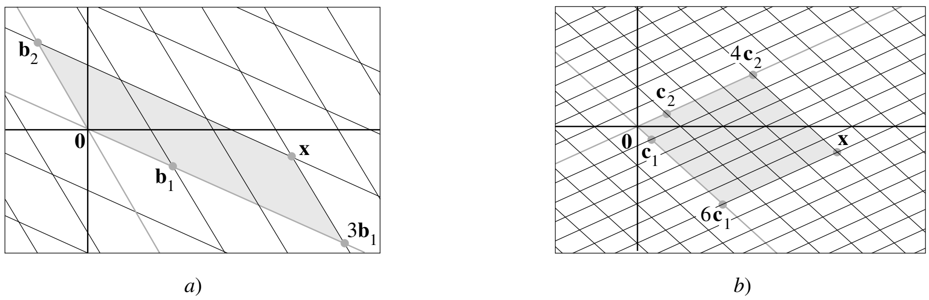


FIGURA 1 Dos sistemas de coordenadas para el mismo espacio vectorial.

EJEMPLO 1 Considere dos bases $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ de un espacio vectorial V, de manera que

$$\mathbf{b}_1 = 4\mathbf{c}_1 + \mathbf{c}_2 \quad \text{y} \quad \mathbf{b}_2 = -6\mathbf{c}_1 + \mathbf{c}_2 \tag{1}$$

Suponga que

$$\mathbf{x} = 3\mathbf{b}_1 + \mathbf{b}_2 \tag{2}$$

Es decir, suponga que $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Encuentre $[\,\mathbf{x}\,]_{\mathcal{C}}$.

SOLUCIÓN Aplique el mapeo de coordenadas determinado por C a $\mathbf{x}$ en la ecuación (2). Puesto que el mapeo de coordenadas es una transformación lineal,

$$\begin{aligned} [\,\mathbf{x}\,]_{\mathcal{C}} &= [\,3\mathbf{b}_1 + \mathbf{b}_2\,]_{\mathcal{C}} \\ &= 3[\,\mathbf{b}_1\,]_{\mathcal{C}} + [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{aligned}$$

Podemos escribir esta ecuación vectorial como una ecuación matricial, utilizando los vectores en la combinación lineal como las columnas de una matriz:

$$[\,\mathbf{x}\,]_{\mathcal{C}} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} \tag{3}$$

[1] Piense en $[\,\mathbf{x}\,]_{\mathcal{B}}$ como un "nombre" de $\mathbf{x}$ que lista los pesos utilizados para construir $\mathbf{x}$ como una combinación lineal de los vectores básicos en $\mathcal{B}$.

Esta fórmula da $[\mathbf{x}]_{\mathcal{C}}$, una vez que se conocen las columnas de la matriz. A partir de (1),

$$[\mathbf{b}_1]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \quad \text{y} \quad [\mathbf{b}_2]_{\mathcal{C}} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$$

Por lo tanto, la ecuación (3) da la solución:

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

Las $\mathcal{C}$-coordenadas de $\mathbf{x}$ coinciden con las de $\mathbf{x}$ en la figura 1. ■

El argumento que se utiliza para deducir la fórmula (3) se puede generalizar para obtener el siguiente resultado. (Véase los ejercicios 15 y 16).

TEOREMA 15

Sean $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ las bases de un espacio vectorial V. Entonces, existe una única matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ de $n \times n$ tal que

$$[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}} \tag{4}$$

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son los vectores de $\mathcal{C}$-coordenadas de los vectores en la base $\mathcal{B}$. Es decir,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{C}} & [\mathbf{b}_2]_{\mathcal{C}} & \cdots & [\mathbf{b}_n]_{\mathcal{C}} \end{bmatrix} \tag{5}$$

La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ en el teorema 15 se denomina **matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$**. La multiplicación por $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ convierte a las $\mathcal{B}$-coordenadas en las $\mathcal{C}$-coordenadas.[2] La figura 2 ilustra la ecuación de cambio de coordenadas (4).

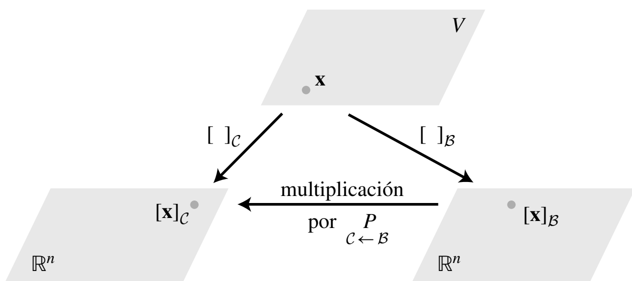


FIGURA 2 Dos sistemas de coordenadas para V.

Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes porque son los vectores de coordenadas del conjunto linealmente independiente $\mathcal{B}$. (Véase el ejercicio 25 de la sección 4.4). Puesto que $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es cuadrada, debe ser invertible, de acuerdo con el teorema de la matriz invertible. Multiplicando por la izquierda ambos lados de la ecuación (4) por $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ se obtiene

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}[\mathbf{x}]_{\mathcal{C}} = [\mathbf{x}]_{\mathcal{B}}$$

---

[2] Para recordar cómo se construye la matriz, piense en $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}}$ como en una combinación lineal de las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. El producto matriz-vector es un vector de $\mathcal{C}$-coordenadas, de modo que las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ también deberían ser los vectores de $\mathcal{C}$-coordenadas.

Por lo tanto, $(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1}$ es la matriz que convierte las $\mathcal{B}$-coordenadas en $\mathcal{C}$-coordenadas. Es decir,

$$(\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P} \tag{6}$$

Cambio de base en $\mathbb{R}^n$

Si $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $\mathcal{E}$ es la *base estándar* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ en $\mathbb{R}^n$, entonces $[\mathbf{b}_1]_\mathcal{E} = \mathbf{b}_1$, y lo mismo para los otros vectores en $\mathcal{B}$. En este caso, $\underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_\mathcal{B}$ que se presentó en la sección 4.4, a saber,

$$P_\mathcal{B} = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n]$$

Para cambiar las coordenadas entre dos bases que no son estándar en $\mathbb{R}^n$, se necesita el teorema 15. El teorema demuestra que para resolver el problema de cambio de base, se necesitan los vectores de coordenadas de la antigua base respecto de la nueva base.

EJEMPLO 2 Sean $\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, y considere las bases de $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

SOLUCIÓN La matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ implica a los vectores de $\mathcal{C}$-coordenadas de $\mathbf{b}_1$ y $\mathbf{b}_2$. Sean $[\,\mathbf{b}_1\,]_\mathcal{C} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ y $[\,\mathbf{b}_2\,]_\mathcal{C} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Así, por definición,

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \quad \text{y} \quad \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2$$

Para resolver ambos sistemas simultáneamente, se aumenta la matriz de coeficientes con $\mathbf{b}_1$ *y* $\mathbf{b}_2$, y se reduce por filas:

$$\left[\begin{array}{cc:cc} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{array}\right] = \left[\begin{array}{rr:rr} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{array}\right] \sim \left[\begin{array}{rr:rr} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{array}\right] \tag{7}$$

Así,

$$[\,\mathbf{b}_1\,]_\mathcal{C} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \quad \text{y} \quad [\,\mathbf{b}_2\,]_\mathcal{C} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

La matriz del cambio de coordenadas deseada es, por consiguiente,

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\,\mathbf{b}_1\,]_\mathcal{C} & [\,\mathbf{b}_2\,]_\mathcal{C} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}$$ ■

Observe que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ del ejemplo 2 ya apareció en la ecuación (7). Esto no es sorprendente, ya que la primera columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ resulta de reducir por filas $[\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1]$ a $[I \,\vdots\, [\mathbf{b}_1]_\mathcal{C}]$, y de manera similar para la segunda columna de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$. Por lo tanto,

$$\boxed{[\,\mathbf{c}_1 \quad \mathbf{c}_2 \,\vdots\, \mathbf{b}_1 \quad \mathbf{b}_2\,] \sim [\,I \,\vdots\, \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}\,]}$$

Un procedimiento análogo funciona para encontrar la matriz de cambio de coordenadas entre dos bases cualesquiera en $\mathbb{R}^n$.

EJEMPLO 3 Sea $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, y considere las bases para $\mathbb{R}^2$ dadas por $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

a) Determine la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

b) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

SOLUCIÓN

a) Considere que $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ se necesita más que $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, y calcule

$$\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \vdots & \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} = \left[\begin{array}{rr:rr} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \sim \left[\begin{array}{rr:rr} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right]$$

Por lo tanto,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

b) De acuerdo con el inciso *a*) y la propiedad (6) anterior (con $\mathcal{B}$ y $\mathcal{C}$ intercambiadas),

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix} \qquad \blacksquare$$

Otra descripción de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ utiliza las matrices de cambio de coordenadas $P_{\mathcal{B}}$ y $P_{\mathcal{C}}$ que convierten a las coordenadas $\mathcal{B}$ y $\mathcal{C}$, respectivamente, en coordenadas estándar. Recuerde que para cada $\mathbf{x}$ en $\mathbb{R}^n$,

$$P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}, \quad P_{\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} = \mathbf{x} \quad \text{y} \quad [\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x}$$

Por lo tanto,

$$[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x} = P_{\mathcal{C}}^{-1}P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$$

En $\mathbb{R}^n$, la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ se puede calcular como $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. En realidad, para matrices más grandes que 2×2, un algoritmo similar al del ejemplo 3 es más rápido que calcular $P_{\mathcal{C}}^{-1}$ y luego $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. Véase el ejercicio 12 en la sección 2.2.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$ y $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2\}$ bases para un espacio vectorial V, y sea P una matriz cuyas columnas son $[\mathbf{f}_1]_{\mathcal{G}}$ y $[\mathbf{f}_2]_{\mathcal{G}}$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{v}$ en V?

   i. $[\mathbf{v}]_{\mathcal{F}} = P[\mathbf{v}]_{\mathcal{G}}$ **ii.** $[\mathbf{v}]_{\mathcal{G}} = P[\mathbf{v}]_{\mathcal{F}}$

2. Sean $\mathcal{B}$ y $\mathcal{C}$ como en el ejemplo 1. Utilice los resultados de ese ejemplo para encontrar la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

4.7 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V y suponga que $\mathbf{b}_1 = 6\mathbf{c}_1 - 2\mathbf{c}_2$ y $\mathbf{b}_2 = 9\mathbf{c}_1 - 4\mathbf{c}_2$.

   a) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

   b) Encuentre $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. Utilice el inciso *a*).

2. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para un espacio vectorial V, y suponga que $\mathbf{b}_1 = -2\mathbf{c}_1 + 4\mathbf{c}_2$ y $\mathbf{b}_2 = 3\mathbf{c}_1 - 6\mathbf{c}_2$.

   a) Encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$.

   b) Determine $[\mathbf{x}]_{\mathcal{C}}$ para $\mathbf{x} = 2\mathbf{b}_1 + 3\mathbf{b}_2$.

3. Sean $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ y $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2\}$ bases para V, y sea P una matriz cuyas columnas son $[\mathbf{u}_1]_{\mathcal{W}}$ y $[\mathbf{u}_2]_{\mathcal{W}}$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{x}$ en V?

i. $[\mathbf{x}]_{\mathcal{U}} = P[\mathbf{x}]_{\mathcal{W}}$ **ii.** $[\mathbf{x}]_{\mathcal{W}} = P[\mathbf{x}]_{\mathcal{U}}$

4. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ bases para V, y sea $P[[\mathbf{d}_1]_{\mathcal{A}}\ [\mathbf{d}_2]_{\mathcal{A}}\ [\mathbf{d}_3]_{\mathcal{A}}]$. ¿Cuál de las siguientes ecuaciones satisface P para toda $\mathbf{x}$ en V?

i. $[\mathbf{x}]_{\mathcal{A}} = P[\mathbf{x}]_{\mathcal{D}}$ **ii.** $[\mathbf{x}]_{\mathcal{D}} = P[\mathbf{x}]_{\mathcal{A}}$

5. Sean $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{a}_1 = 4\mathbf{b}_1 - \mathbf{b}_2$, $\mathbf{a}_2 = -\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, y $\mathbf{a}_3 = \mathbf{b}_2 - 2\mathbf{b}_3$.

a) Encuentre la matriz de cambio de coordenadas de $\mathcal{A}$ a $\mathcal{B}$.

b) Determine $[\mathbf{x}]_{\mathcal{B}}$ para $\mathbf{x} = 3\mathbf{a}_1 + 4\mathbf{a}_2 + \mathbf{a}_3$.

6. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ y $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ bases para un espacio vectorial V, y suponga que $\mathbf{f}_1 = 2\mathbf{d}_1 - \mathbf{d}_2 + \mathbf{d}_3$, $\mathbf{f}_2 = 3\mathbf{d}_2 + \mathbf{d}_3$, y $\mathbf{f}_3 = -3\mathbf{d}_1 + 2\mathbf{d}_3$.

a) Encuentre la matriz de cambio de las coordenadas de $\mathcal{F}$ a $\mathcal{D}$.

b) Determine $[\mathbf{x}]_{\mathcal{D}}$ para $\mathbf{x} = \mathbf{f}_1 - 2\mathbf{f}_2 + 2\mathbf{f}_3$.

En los ejercicios 7 a 10, sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ bases para $\mathbb{R}^2$. En cada ejercicio, encuentre la matriz de cambio de coordenadas de $\mathcal{B}$ a $\mathcal{C}$ y la matriz de cambio de coordenadas de $\mathcal{C}$ a $\mathcal{B}$.

7. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -7 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

9. $\mathbf{b}_1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

10. $\mathbf{b}_1 = \begin{bmatrix} 6 \\ -12 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$

En los ejercicios 11 y 12, $\mathcal{B}$ y $\mathcal{C}$ son bases para un espacio vectorial V. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) Las columnas de la matriz de cambio de coordenadas $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son vectores de $\mathcal{B}$-coordenadas de los vectores en $\mathcal{C}$.

b) Si $V = \mathbb{R}^n$, y $\mathcal{C}$ es la base *estándar* para V, entonces $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es igual a la matriz de cambio de coordenadas $P_{\mathcal{B}}$ que se presentó en la sección 4.4.

12. *a*) Las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ son linealmente independientes.

b) Si $V = \mathbb{R}^2$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, entonces la reducción por filas de $[\mathbf{c}_1\ \ \mathbf{c}_2\ \ \mathbf{b}_1\ \ \mathbf{b}_2]$ a $[I\ \ P]$ produce una matriz P que satisface $[\mathbf{x}]_{\mathcal{B}} = P[\mathbf{x}]_{\mathcal{C}}$ para toda $\mathbf{x}$ en V.

13. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 2t + t^2, 3 - 5t + 4t^2, 2t + 3t^2\}$ a la base estándar de $\mathcal{C} = \{1, t, t^2\}$. Luego, encuentre el vector de $\mathcal{B}$-coordenadas para $-1 + 2t$.

14. En $\mathbb{P}_2$, encuentre la matriz de cambio de coordenadas de la base $\mathcal{B} = \{1 - 3t^2, 2 + t - 5t^2, 1 + 2t\}$ a la base estándar. Luego, escriba t^2 como una combinación lineal de los polinomios en $\mathcal{B}$.

Los ejercicios 15 y 16 brindan una demostración del teorema 15. Complete la justificación para cada paso.

15. Dada $\mathbf{v}$ en V, existen escalares $x_1,\ldots, x_n$, tales que

$$\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$$

porque *a*) ______ . Aplique el mapeo de coordenadas determinado por la base $\mathcal{C}$, y obtenga

$$[\mathbf{v}]_{\mathcal{C}} = x_1[\mathbf{b}_1]_{\mathcal{C}} + x_2[\mathbf{b}_2]_{\mathcal{C}} + \cdots + x_n[\mathbf{b}_n]_{\mathcal{C}}$$

porque *b*) ______ . Esta ecuación se puede escribir en la forma

$$[\,\mathbf{v}\,]_{\mathcal{C}} = \begin{bmatrix} [\,\mathbf{b}_1\,]_{\mathcal{C}} & [\,\mathbf{b}_2\,]_{\mathcal{C}} & \cdots & [\,\mathbf{b}_n\,]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{8}$$

de acuerdo con la definición de *c*) ______ . Esto indica que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ que se muestra en la ecuación (5) satisface $[\mathbf{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{v}]_{\mathcal{B}}$ para toda $\mathbf{v}$ en V, porque el vector del lado derecho en (8) es *d*) ______ .

16. Suponga que Q es cualquier matriz de manera que

$$[\mathbf{v}]_{\mathcal{C}} = Q[\mathbf{v}]_{\mathcal{B}} \quad \text{para toda } \mathbf{v} \text{ en } V \tag{9}$$

Establezca $\mathbf{v} = \mathbf{b}_1$ en la ecuación (9). Después (9) indica que $[\mathbf{b}_1]_{\mathcal{C}}$ es la primera columna de Q porque *a*) ______ . De manera similar, para $k = 2,\ldots, n$, la k-ésima columna de Q es *b*) _____ porque *c*) _____ . Esto indica que la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ definida por (5) en el teorema 15 es la única matriz que satisface la condición (4).

17. **[M]** Sea $\mathcal{B} = \{\mathbf{x}_0,\ldots, \mathbf{x}_6\}$ y $\mathcal{C} = \{\mathbf{y}_0,\ldots, \mathbf{y}_6\}$, donde $\mathbf{x}_k$ es la función $\cos^k t$ y $\mathbf{y}_k$ es la función $\cos kt$. El ejercicio 34 de la sección 4.5 mostró que tanto B como C son bases para el espacio vectorial $H = \text{Gen}\,\{\mathbf{x}_0,\ldots, \mathbf{x}_6\}$.

a) Sea $P = \begin{bmatrix} [\,\mathbf{y}_0\,]_{\mathcal{B}} & \cdots & [\,\mathbf{y}_6\,]_{\mathcal{B}} \end{bmatrix}$, y calcule P^{-1}.

b) Explique por qué las columnas de P^{-1} son los vectores de $\mathcal{C}$-coordenadas de $\mathbf{x}_0,\ldots, \mathbf{x}_6$. Después, utilice estos vectores de coordenadas para escribir identidades trigonométricas que expresen potencias de $\cos t$ en términos de las funciones en $\mathcal{C}$.

18. **[M]** (*Se requiere cálculo*)[3] Recuerde de sus clases de cálculo que las integrales como

$$\int (5\cos^3 t - 6\cos^4 t + 5\cos^5 t - 12\cos^6 t)\, dt \tag{10}$$

son tediosas de calcular. (El método habitual es aplicar integración por partes varias veces y usar la fórmula de la mitad del ángulo). Utilice la matriz P o P^{-1} del ejercicio 17 para transformar (10), y después calcule la integral.

[3] La idea de los ejercicios 17 y 18 y cinco ejercicios relacionados en las secciones anteriores provienen de un documento de Jack W. Rogers, Jr., de la Universidad de Auburn, presentado en una reunión de la International Linear Algebra Society, en agosto de 1995. Véase la sección "Applications of Linear Algebra in Calculus", *American Mathematical Monthly* **104** (1), 1997.

19. [**M**] Sea

$$P = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -5 & 0 \\ 4 & 6 & 1 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$$

a) Encuentre una base $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ para $\mathbb{R}^3$ tal que P es la matriz de cambio de coordenadas de $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ a la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. [*Sugerencia:* Pregúntese qué representan las columnas de $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$].

b) Encuentre una base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ para $\mathbb{R}^3$ tal que P es la matriz de cambio de coordenadas de $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

20. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ y bases para un espacio vectorial de dimensión 2.

a) Escriba una ecuación que relacione las matrices $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, $\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}$ y $\underset{\mathcal{D}\leftarrow\mathcal{B}}{P}$. Justifique su resultado.

b) [**M**] Utilice un programa de matrices ya sea para ayudarle a encontrar la ecuación o para comprobar la ecuación que escribió. Trabaje con tres bases de $\mathbb{R}^2$. (Véase los ejercicios 7 a 10).

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Como las columnas de P son vectores de coordenadas $\mathcal{G}$, un vector de la forma $P\mathbf{x}$ debe ser un vector de coordenadas $\mathcal{G}$. Por lo tanto, P satisface la ecuación **ii**.
2. Los vectores de coordenadas que se encuentran en el ejemplo 1 indican que

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \big[\, [\mathbf{b}_1]_{\mathcal{C}} \quad [\mathbf{b}_2]_{\mathcal{C}} \,\big] = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Por lo tanto,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} .1 & .6 \\ -.1 & .4 \end{bmatrix}$$

4.8 APLICACIONES A LAS ECUACIONES EN DIFERENCIAS

Ahora que se dispone de poderosas computadoras, cada vez más problemas científicos y de ingeniería se analizan utilizando datos discretos, o digitales, en lugar de datos continuos. Las ecuaciones en diferencias con frecuencia son la herramienta adecuada para analizar esos datos. Incluso cuando una ecuación diferencial se utiliza para modelar un proceso continuo, a menudo se obtiene una solución numérica a partir de una ecuación en diferencias relacionada.

En esta sección se ponen de relieve algunas propiedades fundamentales de las ecuaciones en diferencias lineales que se explican mejor utilizando el álgebra lineal.

Señales discretas de tiempo

El espacio vectorial $\mathbb{S}$ de las señales discretas de tiempo se presentó en la sección 4.1. Una **señal** en $\mathbb{S}$ es una función definida solo con números enteros y se visualiza como una secuencia de números, por ejemplo, $\{y_k\}$. La figura 1 muestra tres señales típicas cuyos términos generales son $(.7)^k$, 1^k y $(-1)^k$, respectivamente.

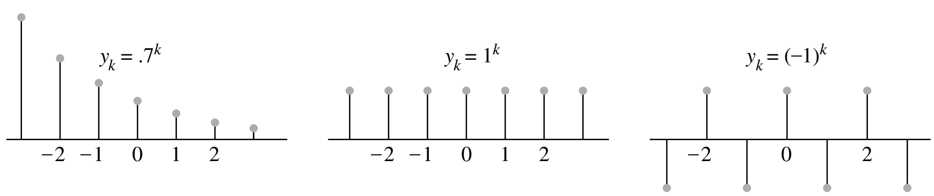


FIGURA 1 Tres señales en $\mathbb{S}$.

Las señales digitales, desde luego, surgen en ingeniería de sistemas eléctricos y de control, pero secuencias de datos discretos también se generan en biología, física, economía, demografía y muchas otras áreas, donde se mide un proceso, o se *muestrea*, en intervalos discretos de tiempo. Cuando se inicia un proceso en un momento específico, a veces es conveniente representar una señal como una secuencia de la forma $(y_0, y_1, y_2, \ldots)$. Se supone que los términos y_k para $k < 0$ son cero o simplemente se omiten.

EJEMPLO 1 El sonido cristalino de un reproductor de discos compactos proviene de música de la que se han tomado muestras a una velocidad de 44,100 veces por segundo. Véase la figura 2. En cada medición, la amplitud de la señal de música se registra como un número, por ejemplo, y_k. La música original está compuesta de muchos sonidos diferentes de diversas frecuencias; sin embargo, la secuencia $\{y_k\}$ contiene suficiente información para reproducir todas las frecuencias en el sonido hasta aproximadamente 20,000 ciclos por segundo, más allá de lo que el oído humano puede percibir. ∎

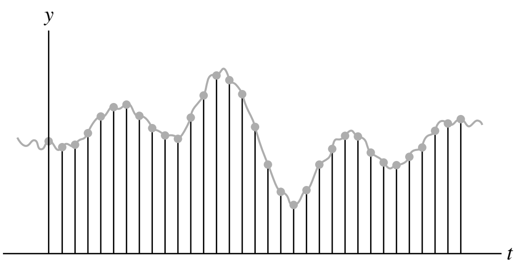


FIGURA 2 Datos muestreados a partir de una señal de música.

Independencia lineal en el espacio $\mathbb{S}$ de las señales

Para simplificar la notación, consideramos un conjunto de solo tres señales en $\mathbb{S}$, por ejemplo, $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Son linealmente independientes precisamente cuando la ecuación

$$c_1 u_k + c_2 v_k + c_3 w_k = 0 \quad \text{para toda } k \tag{1}$$

implica que $c_1 = c_2 = c_3 = 0$. La frase "para toda k" significa para todos los enteros: positivos, negativos y cero. También se podría considerar que las señales comienzan con $k = 0$, por ejemplo; en tal caso, "para toda k" significaría para todos los enteros mayores que o iguales a cero ($k \geq 0$).

Suponga que c_1, c_2, c_3 satisfacen la ecuación (1). Entonces la ecuación (1) es válida para cualesquiera tres valores consecutivos de k, por ejemplo, k, $k + 1$ y $k + 2$. Así, la ecuación (1) implica que

$$c_1 u_{k+1} + c_2 v_{k+1} + c_3 w_{k+1} = 0 \quad \text{para toda } k$$

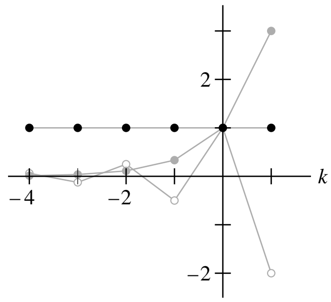
y

$$c_1 u_{k+2} + c_2 v_{k+2} + c_3 w_{k+2} = 0 \quad \text{para toda } k$$

Por lo tanto c_1, c_2, c_3 satisfacen

$$\begin{bmatrix} u_k & v_k & w_k \\ u_{k+1} & v_{k+1} & w_{k+1} \\ u_{k+2} & v_{k+2} & w_{k+2} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{para toda } k \tag{2}$$

La matriz de coeficientes de este sistema se llama la **matriz de Casorati** de las señales, y el determinante de la matriz se denomina **casoratiano** de $\{u_k\}$, $\{v_k\}$ y $\{w_k\}$. Si la matriz de Casorati es invertible, para al menos un valor de k, entonces la ecuación (2) implicará que $c_1 = c_2 = c_3 = 0$, lo que demuestra que las tres señales son linealmente independientes.



Las señales 1^k, $(-2)^k$ y 3^k.

EJEMPLO 2 Compruebe que las señales 1^k, $(-2)^k$ y 3^k son linealmente independientes.

SOLUCIÓN La matriz de Casorati es

$$\begin{bmatrix} 1^k & (-2)^k & 3^k \\ 1^{k+1} & (-2)^{k+1} & 3^{k+1} \\ 1^{k+2} & (-2)^{k+2} & 3^{k+2} \end{bmatrix}$$

Con las operaciones de fila se puede demostrar muy fácilmente que esta matriz siempre es invertible. Sin embargo, es más rápido sustituir k por un valor, por ejemplo, $k = 0$ y reducir por filas la matriz numérica:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 3 \\ 1 & 4 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 3 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & 2 \\ 0 & 0 & 10 \end{bmatrix}$$

La matriz de Casorati es invertible para $k = 0$. Por lo tanto, 1^k, $(-2)^k$ y 3^k son linealmente independientes. ■

Si una matriz de Casorati no es invertible, las señales asociadas que se están sometiendo a prueba pueden o no ser linealmente dependientes. (Véase el ejercicio 33). Sin embargo, es posible demostrar que si todas las señales son las soluciones de la *misma* ecuación en diferencias homogénea (que se describe a continuación), entonces la matriz de Casorati es invertible para toda k y las señales son linealmente independientes, o bien, la matriz de Casorati no es invertible para toda k y las señales son linealmente dependientes. En la *Guía de estudio* se presenta una excelente demostración utilizando transformaciones lineales.

Ecuaciones lineales en diferencias

Considerando los escalares $a_0,\dots, a_n$, con a_0 y a_n diferentes de cero, y dada una señal $\{z_k\}$, la ecuación

$$a_0 y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{3}$$

se llama una **ecuación lineal en diferencias** (o **relación de recurrencia lineal**) **de orden n**. Para simplificar, a_0 con frecuencia se considera igual a 1. Si $\{z_k\}$ es la secuencia cero, la ecuación es **homogénea**; de lo contrario, la ecuación es **no homogénea**.

EJEMPLO 3 En el procesamiento de señal digital, una ecuación en diferencias tal como la ecuación (3) describe un **filtro lineal**, y $a_0,\dots, a_n$ se denominan los **coeficientes de filtro**. Si $\{y_k\}$ se trata como la entrada y $\{z_k\}$ como la salida, entonces las soluciones de la ecuación homogénea asociada son las señales que se filtran *hacia fuera* y se transforman en señal cero. Vamos a alimentar dos señales diferentes en el filtro

$$.35y_{k+2} + .5y_{k+1} + .35y_k = z_k$$

Aquí, .35 es una abreviatura de $\sqrt{2}/4$. La primera señal se crea mediante el muestreo de la señal continua $y = \cos(\pi t/4)$ para valores enteros de t, como en la figura 3*a*). La señal discreta es

$$\{y_k\} = \{\dots, \cos(0), \cos(\pi/4), \cos(2\pi/4), \cos(3\pi/4), \dots\}$$

Para simplificar, se escribe $\pm.7$ en lugar de $\pm\sqrt{2}/2$, de manera que

$$\{y_k\} = \{\dots, \underset{\substack{\uparrow \\ k=0}}{1}, .7, 0, -.7, -1, -.7, 0, .7, 1, .7, 0, \dots\}$$

La tabla 1 muestra el cálculo de la secuencia de salida $\{z_k\}$, donde .35(.7) es una abreviatura de $(\sqrt{2}/4)(\sqrt{2}/2) = .25$. La salida es $\{y_k\}$, corrida un término.

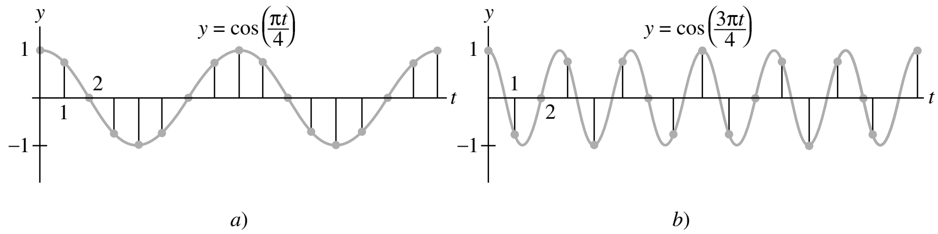


FIGURA 3 Señales discretas con diferentes frecuencias.

TABLA 1 Cálculo de la salida de un filtro

k	y_k	y_{k+1}	y_{k+2}	$.35y_k$	$+ .5y_{k+1}$	$+ .35y_{k+2}$	$=$	z_k
0	1	.7	0	.35(1)	+ .5(.7)	+ .35(0)	=	.7
1	.7	0	−.7	.35(.7)	+ .5(0)	+ .35(−.7)	=	0
2	0	−.7	−1	.35(0)	+ .5(−.7)	+ .35(−1)	=	−.7
3	−.7	−1	−.7	.35(−.7)	+ .5(−1)	+ .35(−.7)	=	−1
4	−1	−.7	0	.35(−1)	+ .5(−.7)	+ .35(0)	=	−.7
5	−.7	0	.7	.35(−.7)	+ .5(0)	+ .35(.7)	=	0
⋮	⋮							⋮

Una señal de entrada diferente se produce a partir de la mayor frecuencia de la señal $y = \cos(3\pi t/4)$, que se muestra en la figura 3*b*). El muestreo a la misma tasa de antes produce una nueva secuencia de entrada:

$$\{w_k\} = \{\ldots, \underset{\substack{\uparrow \\ k=0}}{1}, -.7,\ 0,\ .7,\ -1,\ .7,\ 0,\ -.7,\ 1,\ -.7,\ 0, \ldots\}$$

Cuando se alimenta al filtro con $\{\omega_k\}$, la salida es la secuencia cero. El filtro, conocido como *filtro pasa bajos*, hace que $\{y_k\}$ pase, pero detiene a las señales de mayor frecuencia $\{\omega_k\}$. ■

En muchas aplicaciones, se especifica una secuencia $\{z_k\}$ para el lado derecho de una ecuación en diferencias (3), y una $\{y_k\}$ que satisface la ecuación (3) se considera una **solución** de la ecuación. El siguiente ejemplo ilustra cómo encontrar soluciones para una ecuación homogénea.

EJEMPLO 4 Soluciones de una ecuación homogénea en diferencias a menudo tienen la forma $y_k = r^k$ para alguna r. Encuentre algunas soluciones de la ecuación

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k \tag{4}$$

SOLUCIÓN Se sustituye y^k por r^k en la ecuación y se factoriza el lado izquierdo:

$$r^{k+3} - 2r^{k+2} - 5r^{k+1} + 6r^k = 0 \tag{5}$$

$$r^k(r^3 - 2r^2 - 5r + 6) = 0$$

$$r^k(r-1)(r+2)(r-3) = 0 \tag{6}$$

Como (5) es equivalente a (6), r^k satisface la ecuación en diferencias (4) si y solo si r satisface (6). Así, 1^k, $(-2)^k$ y 3^k son todas soluciones de (4). Por ejemplo, para comprobar que 3^k es una solución de (4), calcule

$$\begin{aligned} &3^{k+3} - 2 \cdot 3^{k+2} - 5 \cdot 3^{k+1} + 6 \cdot 3^k \\ &= 3^k(27 - 18 - 15 + 6) = 0 \quad \text{para toda } k \end{aligned}$$

■

En general, una señal distinta de cero r^k satisface la ecuación homogénea en diferencias

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

si y solo si r es una raíz de la **ecuación auxiliar**

$$r^n + a_1 r^{n-1} + \cdots + a_{n-1} r + a_n = 0$$

No vamos a considerar el caso en que r es una raíz repetida de la ecuación auxiliar. Cuando la ecuación auxiliar tiene una *raíz compleja*, la ecuación en diferencias tiene soluciones de la forma $s^k \cos k\omega$ y s^k sen $k\omega$, para las constantes s y ω. Esto sucedió en el ejemplo 3.

Conjuntos de soluciones de ecuaciones lineales en diferencias

A partir de $a_1,\ldots, a_n$, considere el mapeo $T : \mathbb{S} \to \mathbb{S}$ que transforma una señal $\{y_k\}$ en una señal $\{\omega_k\}$ dada por

$$w_k = y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k$$

Es fácil comprobar que T es una transformación *lineal*. Esto implica que el conjunto solución de la ecuación homogénea

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

es el núcleo de T (el conjunto de las señales que T mapea en la señal cero) y, por lo tanto, el conjunto solución es un *subespacio* de $\mathbb{S}$. Cualquier combinación lineal de soluciones es de nuevo una solución.

El siguiente teorema, un resultado sencillo, pero básico, conducirá a mayor información acerca de los conjuntos solución de ecuaciones en diferencias.

TEOREMA 16

Si $a_n \neq 0$ y si $\{z_k\}$ está dada, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{7}$$

tiene una única solución siempre que se especifican $y_0,\ldots, y_{n-1}$.

DEMOSTRACIÓN Si $y_0,\ldots, y_{n-1}$ se especifican, utilice la ecuación (7) para *definir*

$$y_n = z_0 - [\, a_1 y_{n-1} + \cdots + a_{n-1} y_1 + a_n y_0 \,]$$

Y ahora que $y_1,\ldots, y_n$ se especifican, utilice (7) para definir y_{n+1}. En general, utilice la relación de recurrencia

$$y_{n+k} = z_k - [\, a_1 y_{k+n-1} + \cdots + a_n y_k \,] \tag{8}$$

para definir y_{n+k} para $k \geq 0$. Para definir y_k para $k < 0$, utilice la relación de recurrencia

$$y_k = \frac{1}{a_n} z_k - \frac{1}{a_n} [\, y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} \,] \tag{9}$$

Esto produce una señal que satisface la ecuación (7). A la inversa, cualquier señal que satisface la ecuación (7) para toda k sin duda satisface las ecuaciones (8) y (9), por lo que la solución de la ecuación (7) es única. ■

TEOREMA 17

El conjunto H de todas las soluciones de la ecuación lineal homogénea en diferencias de n-ésimo orden

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k \tag{10}$$

es un espacio vectorial de dimensión n.

DEMOSTRACIÓN Como se indicó antes, H es un subespacio de $\mathbb{S}$ porque H es el núcleo de una transformación lineal. Para $\{y_k\}$ en H, sea $F\{y_k\}$ el vector en $\mathbb{R}^n$ dado por $(y_0, y_1, \ldots, y_{n-1})$. Es posible comprobar fácilmente que $F : H \to \mathbb{R}^n$ es una transformación lineal. Dado cualquier vector $(y_0, y_1, \ldots, y_{n-1})$ en $\mathbb{R}^n$, el teorema 16 dice que hay una única señal $\{y_k\}$ en H tal que $F\{y_k\} = (y_0, y_1, \ldots, y_{n-1})$. Esto significa que F es una transformación lineal uno a uno de H sobre $\mathbb{R}^n$, es decir, F es un isomorfismo. Así, $\dim H = \dim \mathbb{R}^n = n$. (Véase el ejercicio 32 en la sección 4.5). ■

EJEMPLO 5 Encuentre una base para el conjunto de todas las soluciones a la ecuación en diferencias

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

SOLUCIÓN Nuestro trabajo en álgebra lineal ¡realmente está rindiendo frutos ahora! Sabemos a partir de los ejemplos 2 y 4 que 1^k, $(-2)^k$ y 3^k son soluciones linealmente independientes. En general, puede ser difícil comprobar directamente que un conjunto de las señales *genera* el espacio solución. Pero eso no es problema en este caso, debido a dos importantes teoremas: el teorema 17, que demuestra que el espacio solución es exactamente de dimensión 3, y el teorema de la base de la sección 4.5, el cual afirma que un conjunto linealmente independiente de n vectores en un espacio de dimensión n es automáticamente una base. Así, 1^k, $(-2)^k$ y 3^k forman una base para el espacio solución. ■

La solución estándar para describir la “solución general” de la ecuación en diferencias (10) es mostrar una base para el subespacio de todas las soluciones. Dicha base se suele denominar **conjunto fundamental de soluciones** de (10). En la práctica, si se encuentran n señales linealmente independientes que satisfagan (10), de manera automática generarán el espacio de soluciones de dimensión n, como se explica en el ejemplo 5.

Ecuaciones no homogéneas

La solución general de la ecuación no homogénea en diferencias

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = z_k \quad \text{para toda } k \tag{11}$$

se puede escribir como una solución particular de (11) más una combinación lineal arbitraria de un conjunto fundamental de soluciones de la ecuación homogénea correspondiente (10). Este hecho es análogo al resultado de la sección 1.5 que muestra que los conjuntos solución de $A\mathbf{x} = \mathbf{b}$ y $A\mathbf{x} = \mathbf{0}$ son paralelos. Ambos resultados tienen la misma explicación: el mapeo $\mathbf{x} \mapsto A\mathbf{x}$ es lineal, y el mapeo que transforma la señal $\{y_k\}$ en la señal $\{z_k\}$ en (11) es lineal. Véase el ejercicio 35.

EJEMPLO 6 Compruebe que la señal $y_k = k^2$ satisface la ecuación en diferencias

$$y_{k+2} - 4y_{k+1} + 3y_k = -4k \quad \text{para toda } k \tag{12}$$

Luego, encuentre una descripción de todas las soluciones de esta ecuación.

SOLUCIÓN Sustituya y_k por k^2 en el lado izquierdo de (12):

$$\begin{aligned}(k+2)^2 - 4(k+1)^2 + 3k^2 & \\ &= (k^2 + 4k + 4) - 4(k^2 + 2k + 1) + 3k^2 \\ &= -4k\end{aligned}$$

Así, k^2 es de hecho una solución de (12). El siguiente paso es resolver la ecuación homogénea

$$y_{k+2} - 4y_{k+1} + 3y_k = 0 \tag{13}$$

La ecuación auxiliar es

$$r^2 - 4r + 3 = (r-1)(r-3) = 0$$

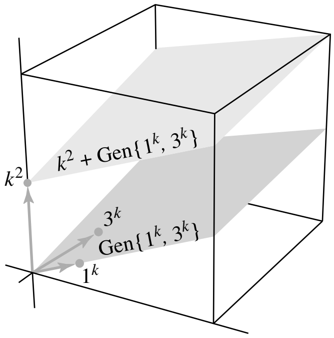


FIGURA 4
Conjuntos solución de las ecuaciones en diferencias (12) y (13).

Las raíces son $r = 1, 3$. Así que dos soluciones de la ecuación homogénea en diferencias son 1^k y 3^k. Evidentemente, no son múltiplos una de la otra, por lo que son señales linealmente independientes. De acuerdo con el teorema 17, el espacio solución es de dimensión 2, de manera que 3^k y 1^k forman una base para el conjunto de soluciones de la ecuación (13). Al trasladar ese conjunto mediante una solución particular de la ecuación no homogénea (12), se obtiene la solución general de (12):

$$k^2 + c_1 1^k + c_2 3^k \quad \text{o} \quad k^2 + c_1 + c_2 3^k$$

La figura 4 muestra una visualización geométrica de los dos conjuntos solución. Cada punto de la figura corresponde a una señal en $\mathbb{S}$. ■

Reducción a sistemas de ecuaciones de primer orden

Una forma moderna de estudiar una ecuación en diferencias homogénea de n-ésimo orden es remplazarla por un sistema equivalente de ecuaciones en diferencias de primer orden, escrito en la forma

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k$$

donde los vectores $\mathbf{x}_k$ están en $\mathbb{R}^n$ y A es una matriz de $n \times n$.

Un ejemplo sencillo de tal ecuación en diferencias (con valor vectorial) se estudió ya en la sección 1.10. Otros ejemplos se tratarán en las secciones 4.9 y 5.6.

EJEMPLO 7 Escriba la siguiente ecuación en diferencias como un sistema de primer orden:

$$y_{k+3} - 2y_{k+2} - 5y_{k+1} + 6y_k = 0 \quad \text{para toda } k$$

SOLUCIÓN Para cada k, establezca

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

La ecuación en diferencias dice que $y_{k+3} = -6y_k + 5y_{k+1} + 2y_{k+2}$, por lo que

$$\mathbf{x}_{k+1} = \begin{bmatrix} y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix} = \begin{bmatrix} 0 & + & y_{k+1} & + 0 \\ 0 & + 0 & + & y_{k+2} \\ -6y_k & + 5y_{k+1} & + 2y_{k+2} & \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \end{bmatrix}$$

Es decir,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{para toda } k, \quad \text{donde} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 5 & 2 \end{bmatrix}$$ ■

En general, la ecuación

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \quad \text{para toda } k$$

se puede rescribir como $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para toda k, donde

$$\mathbf{x}_k = \begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+n-1} \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix}$$

Lecturas adicionales

Hamming, R. W., *Digital Filters*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1989), pp. 1-37.

Kelly, W. G. y A. C. Peterson, *Difference Equations*, 2a. ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2a. ed. (Nueva York: Van Nostrand Reinhold, 1990), pp. 88-141.

Oppenheim, A. V. y A. S. Willsky, *Signals and Systems*, 2a. ed. (Upper Saddle River, NJ: Prentice-Hall, 1997), pp. 1-14, 21-30, 38-43.

PROBLEMA DE PRÁCTICA

Es posible demostrar que las señales 2^k, $3^k \operatorname{sen} \frac{k\pi}{2}$, y $3^k \cos \frac{k\pi}{2}$ son soluciones de

$$y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$$

Demuestre que estas señales forman una base para el conjunto de todas las soluciones de la ecuación en diferencias.

4.8 EJERCICIOS

Compruebe que las señales de los ejercicios 1 y 2 son soluciones de la ecuación en diferencias correspondiente.

1. $2^k, (-4)^k;\ y_{k+2} + 2y_{k+1} - 8y_k = 0$

2. $5^k, (-5)^k;\ y_{k+2} - 25y_k = 0$

Demuestre que las señales en los ejercicios 3 a 6 forman una base para el conjunto solución de la ecuación en diferencias correspondiente.

3. Las señales y la ecuación en el ejercicio 1

4. Las señales y la ecuación en el ejercicio 2

5. $(-2)^k, k(-2)^k;\ y_{k+2} + 4y_{k+1} + 4y_k = 0$

6. $4^k \cos\left(\frac{k\pi}{2}\right), 4^k \operatorname{sen}\left(\frac{k\pi}{2}\right);\ y_{k+2} + 16y_k = 0$

En los ejercicios 7 a 12, suponga que las señales listadas son soluciones de la ecuación en diferencias dada. ¿Las señales forman una base para el espacio solución de la ecuación? Justifique sus respuestas usando los teoremas adecuados.

7. $1^k, 2^k, (-2)^k;\ y_{k+3} - y_{k+2} - 4y_{k+1} + 4y_k = 0$

8. $(-1)^k, 2^k, 3^k;\ y_{k+3} - 4y_{k+2} + 1y_{k+1} + 6y_k = 0$

9. $2^k, 5^k \cos\left(\frac{k\pi}{2}\right), 5^k \operatorname{sen}\left(\frac{k\pi}{2}\right);$

$y_{k+3} - 2y_{k+2} + 25y_{k+1} - 50y_k = 0$

10. $(-2)^k, k(-2)^k, 3^k;\ y_{k+3} + y_{k+2} - 8y_{k+1} - 12y_k = 0$

11. $(-1)^k, 2^k;\ y_{k+3} - 3y_{k+2} + 4y_k = 0$

12. $3^k, (-2)^k;\ y_{k+4} - 13y_{k+2} + 36y_k = 0$

En los ejercicios 13 a 16, encuentre una base para el espacio solución de la ecuación en diferencias. Demuestre que las soluciones que encuentre generan al conjunto solución.

13. $y_{k+2} - y_{k+1} + \frac{2}{9}y_k = 0$ **14.** $y_{k+2} - 5y_{k+1} + 6y_k = 0$

15. $6y_{k+2} + y_{k+1} - 2y_k = 0$ **16.** $y_{k+2} - 25y_k = 0$

Los ejercicios 17 y 18 se refieren a un modelo sencillo de la economía nacional descrito por la ecuación en diferencias

$$Y_{k+2} - a(1+b)Y_{k+1} + abY_k = 1 \tag{14}$$

Aquí Y_k es el ingreso nacional total durante el año k, a es una constante menor que 1, que se llama *propensión marginal al consumo*, y b es una *constante de ajuste* positiva que describe cómo los cambios en el gasto de consumo afectan la tasa anual de inversión privada.[1]

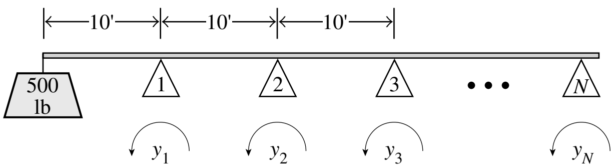
17. Determine la solución general de la ecuación (14) cuando $a = .9$ y $b = \frac{4}{9}$. ¿Qué pasa con Y_k conforme k aumenta? [*Sugerencia:* En primer lugar, encuentre una solución particular de la forma $Y_k = T$, donde T es una constante, llamada el nivel de equilibrio del ingreso nacional].

18. Determine la solución general de la ecuación (14), cuando $a = .9$ y $b = .5$.

[1] Por ejemplo, véase *Discrete Dynamical Systems*, de James T. Sandefur (Oxford: Clarendon Press, 1990), pp. 267-276. El *modelo original acelerador-multiplicador* se atribuye al economista P. A. Samuelson.

Una viga ligera en voladizo está soportada en los puntos N espaciados por una distancia de 10 ft, y se coloca un peso de 500 lb en el extremo de la viga, a 10 ft del primer soporte, como se muestra en la figura. Sea y_k el momento flector en el k-ésimo soporte. Luego, $y_1 = 5000$ ft-lb. Suponga que la viga está rígidamente sujeta al N-ésimo soporte y el momento flector ahí es cero. En medio, los momentos satisfacen la *ecuación de tres momentos*

$$y_{k+2} + 4y_{k+1} + y_k = 0 \quad \text{para } k = 1, 2, \ldots, N-2 \tag{15}$$



Momentos flectores en una viga en voladizo.

19. Determine la solución general de la ecuación en diferencias (15). Justifique su respuesta.

20. Encuentre la solución particular de la ecuación (15) que satisface las *condiciones frontera* $y_1 = 5000$ y $y_N = 0$. (La respuesta implica a N).

21. Cuando se produce una señal a partir de una secuencia de mediciones efectuadas en un proceso (una reacción química, un flujo de calor a través de un tubo, un brazo robot en movimiento, etcétera), la señal, por lo general, contiene *ruido* aleatorio producido por errores de medición. Un método estándar de procesamiento previo de los datos para reducir el ruido es suavizar o filtrar los datos. Un sencillo filtro es un *promedio móvil* que sustituye cada y_k por su promedio con los dos valores adyacentes:

$$\tfrac{1}{3}y_{k+1} + \tfrac{1}{3}y_k + \tfrac{1}{3}y_{k-1} = z_k \quad \text{para } k = 1, 2, \ldots$$

Suponga que una señal y_k, para $k = 0, \ldots, 14$, es

9, 5, 7, 3, 2, 4, 6, 5, 7, 6, 8, 10, 9, 5, 7

Utilice el filtro para calcular $z_1, \ldots, z_{13}$. Trace una gráfica de línea discontinua que se superponga a la señal original y a la señal suavizada.

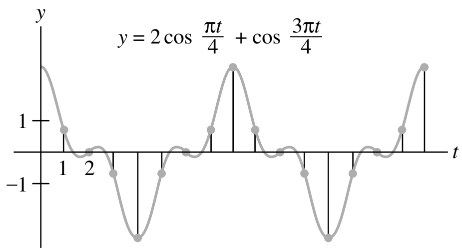
22. Sea $\{y_k\}$ la secuencia producida por la toma de muestras de señal continua $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$ en $t = 0, 1, 2, \ldots$, como se indica en la figura. Los valores de y_k, empezando con $k = 0$, son

$3, .7, 0, -.7, -3, -.7, 0, .7, 3, .7, 0, \ldots$

donde .7 es una abreviatura de $\sqrt{2}/2$.

a) Calcule la señal de salida $\{z_k\}$ cuando $\{y_k\}$ alimenta al filtro del ejemplo 3.

b) Explique cómo y por qué la salida en el inciso *a*) está relacionada con los cálculos del ejemplo 3.



Datos de la muestra de $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$.

Los ejercicios 23 y 24 se refieren a una ecuación en diferencias de la forma $y_{k+1} - ay_k = b$, para constantes adecuadas a y b.

23. Un préstamo de \$10,000 tiene una tasa de interés del 1% mensual y requiere un pago mensual de \$450. El préstamo se realiza en el mes $k = 0$, y el primer pago se realiza un mes después, en $k = 1$. Para $k = 0, 1, 2, \ldots$, sea y_k el saldo insoluto del préstamo justo después del k-ésimo pago mensual. Así,

$$\underset{\text{Nuevo saldo}}{y_1} = \underset{\text{Saldo adeudado}}{10{,}000} + \underset{\text{Intereses agregados}}{(.01)10{,}000} - \underset{\text{Pago}}{450}$$

a) Escriba una ecuación en diferencias satisfecha por $\{y_k\}$.

b) **[M]** Cree una tabla que muestre k, y el saldo y_k al mes k. Liste el programa o los tecleos (las instrucciones) para crear la tabla.

c) **[M]** ¿Cuál es el valor de k cuando se efectúa el último pago? ¿De cuánto será el último pago? ¿Cuánto dinero en total pagó el deudor?

24. En el tiempo $k = 0$, con una inversión inicial de \$1000 se abre una cuenta de ahorros que paga el 6% de interés anual con capitalización mensual. (La tasa de interés mensual es de .005). Cada mes después de la inversión inicial, se agregan \$200 a la cuenta. Para $k = 0, 1, 2, \ldots$, sea y_k la cantidad que hay en la cuenta al momento k, justo después de que se hace un depósito.

a) Escriba una ecuación en diferencias que satisfaga $\{y_k\}$.

b) **[M]** Cree una tabla que muestre k, y la cantidad total en la cuenta de ahorros en el mes k, para $k = 0$ a 60. Liste su programa o los tecleos (las instrucciones) que utilizó para crear la tabla.

c) **[M]** ¿Cuánto habrá en la cuenta después de dos años (es decir, 24 meses), cuatro años, y cinco años? ¿Qué parte del total a los cinco años corresponderá a los intereses?

En los ejercicios 25 a 28, demuestre que la señal dada es una solución de la ecuación en diferencias. Luego, determine la solución general de esa ecuación en diferencias.

25. $y_k = k^2;\ y_{k+2} + 3y_{k+1} - 4y_k = 7 + 10k$

26. $y_k = 1 + k;\ y_{k+2} - 6y_{k+1} + 5y_k = -4$

27. $y_k = k - 2;\ y_{k+2} - 4y_k = 8 - 3k$

28. $y_k = 1 + 2k;\ y_{k+2} - 25y_k = -48k - 20$

Escriba las ecuaciones en diferencias en los ejercicios 29 y 30 como sistemas de primer orden, $\mathbf{x}_{k+1} = A\mathbf{x}_k$, para toda k.

29. $y_{k+4} + 3y_{k+3} - 8y_{k+2} + 6y_{k+1} - 2y_k = 0$

30. $y_{k+3} - 5y_{k+2} + 8y_k = 0$

31. ¿La siguiente ecuación en diferencias es de orden 3? Explique su respuesta.

$$y_{k+3} + 5y_{k+2} + 6y_{k+1} = 0$$

32. ¿Cuál es el orden de la siguiente ecuación en diferencias? Explique su respuesta.

$$y_{k+3} + a_1 y_{k+2} + a_2 y_{k+1} + a_3 y_k = 0$$

33. Sea $y_k = k^2$ y $z_k = 2k|k|$. ¿Las señales $\{y_k\}$ y $\{z_k\}$ son linealmente independientes? Evalúe la matriz de Casorati $C(k)$ asociada para $k = 0$, $k = -1$ y $k = -2$, y analice sus resultados.

34. Sean f, g y h, funciones linealmente independientes definidas para todos los números reales, y construya tres señales por muestreo de los valores de las funciones de los números enteros:

$$u_k = f(k), \qquad v_k = g(k), \qquad w_k = h(k)$$

¿Es necesario que las señales sean linealmente independientes en $\mathbb{S}$? Analícelo.

35. Sean a y b números diferentes de cero. Demuestre que el mapeo T definido por $T\{y_k\} = \{w_k\}$, donde

$$w_k = y_{k+2} + a y_{k+1} + b y_k$$

es una transformación lineal de $\mathbb{S}$ en $\mathbb{S}$.

36. Sea V un espacio vectorial, y sea $T : V \to V$ una transformación lineal. Dada $\mathbf{z}$ en V, suponga que $\mathbf{x}_p$ en V satisface $T(\mathbf{x}_p) = \mathbf{z}$, y sea $\mathbf{u}$ cualquier vector en el núcleo de T. Demuestre que $\mathbf{u} + \mathbf{x}_p$ satisface la ecuación no homogénea $T(\mathbf{x}) = \mathbf{z}$.

37. Sea $\mathbb{S}_0$ el espacio vectorial de todas las secuencias de la forma $(y_0, y_1, y_2, \ldots)$, y defina las transformaciones lineales T y D de $\mathbb{S}_0$ dentro de $\mathbb{S}_0$ por

$$T(y_0, y_1, y_2, \ldots) = (y_1, y_2, y_3, \ldots)$$
$$D(y_0, y_1, y_2, \ldots) = (0, y_0, y_1, y_2, \ldots)$$

Demuestre que $TD = I$ (la transformación identidad en $\mathbb{S}_0$) y que, sin embargo, $DT \neq I$.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Examine la matriz de Casorati:

$$C(k) = \begin{bmatrix} 2^k & 3^k \operatorname{sen}\frac{k\pi}{2} & 3^k \cos\frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} \operatorname{sen}\frac{(k+1)\pi}{2} & 3^{k+1} \cos\frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} \operatorname{sen}\frac{(k+2)\pi}{2} & 3^{k+2} \cos\frac{(k+2)\pi}{2} \end{bmatrix}$$

Sea $k = 0$ y reduzca por filas la matriz para comprobar que tiene tres posiciones pivote y que, por lo tanto, es invertible:

$$C(0) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix}$$

La matriz de Casorati es invertible en $k = 0$, por lo que las señales son linealmente independientes. Puesto que hay tres señales, y el espacio solución H de la ecuación en diferencias es de dimensión 3 (teorema 17), las señales forman una base para H, de acuerdo con el teorema de la base.

4.9 APLICACIONES A CADENAS DE MARKOV

Las cadenas de Markov descritas en esta sección se utilizan como modelos matemáticos de una amplia variedad de situaciones en biología, negocios, química, ingeniería, física y otros campos. En cada caso, el modelo se utiliza para describir un experimento o una medición que se realiza muchas veces de la misma manera, cuando el resultado de cada ensayo del experimento será uno de varios posibles resultados especificados, y depende solo del ensayo inmediato anterior.

Por ejemplo, si la población de una ciudad y sus suburbios se midieran cada año, entonces un vector tal como

$$\mathbf{x}_0 = \begin{bmatrix} .60 \\ .40 \end{bmatrix} \tag{1}$$

podría indicar que el 60% de la población vive en la ciudad y el 40% en los suburbios. Los decimales en $\mathbf{x}_0$ suman 1, ya que dan cuenta de toda la población de la región. Los porcentajes son más convenientes para nuestros propósitos en este momento que los totales de población.

Un vector con entradas no negativas que suman 1 se llama **vector de probabilidad**. Una **matriz estocástica** es una matriz cuadrada cuyas columnas son vectores de probabilidad. Una **cadena de Markov** es una secuencia de vectores de probabilidad de $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots,$ junto con una matriz estocástica P, tal que

$$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, \quad \ldots$$

Así, la cadena de Markov se describe mediante la ecuación en diferencias de primer orden

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para } k = 0, 1, 2, \ldots$$

Cuando una cadena de Markov de vectores en $\mathbb{R}^n$ describe un sistema o una serie de experimentos, las entradas en $\mathbf{x}_k$ listan, respectivamente, las probabilidades de que el sistema esté en cada uno de los n posibles estados, o las probabilidades de que el resultado del experimento sea uno de los n posibles resultados. Por esta razón, $\mathbf{x}_k$ con frecuencia se llama un **vector de estado**.

EJEMPLO 1 En la sección 1.10 se examinó un modelo para los movimientos de población entre una ciudad y sus suburbios. Véase la figura 1. La migración anual entre estas dos partes de la zona metropolitana se regía por la *matriz de migración* M:

$$\begin{array}{cc} & \begin{array}{cc} \multicolumn{2}{c}{\text{De:}} \\ \text{Ciudad} & \text{Suburbios} \end{array} \quad \text{De:} \\ M = & \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \quad \begin{array}{l} \text{Ciudad} \\ \text{Suburbios} \end{array} \end{array}$$

Es decir, cada año el 5% de la población de la ciudad se muda a los suburbios, y el 3% de la población de los suburbios se muda a la ciudad. Las columnas de M son vectores de probabilidad, por lo que M es una matriz estocástica. Suponga la población de la región en el año 2000 era de 600,000 en la ciudad y de 400,000 en los suburbios. Entonces, la distribución inicial de la población en la región está dada por $\mathbf{x}_0$ en la ecuación (1) anterior. ¿Cuál es la distribución de la población en el año 2001? ¿Y en 2002?

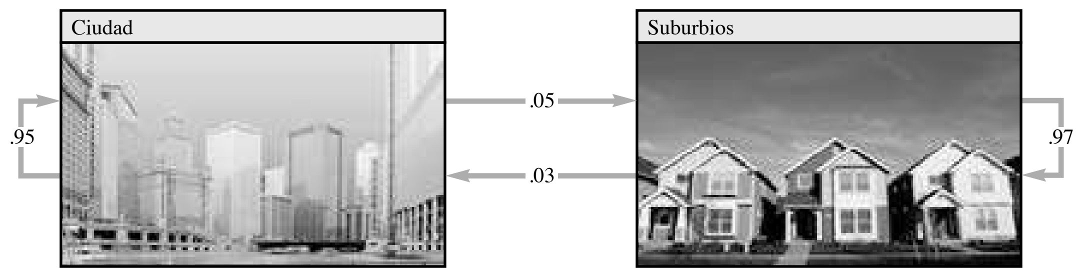


FIGURA 1 Migración porcentual anual entre la ciudad y los suburbios.

SOLUCIÓN En el ejemplo 3 de la sección 1.10, vimos que después de un año, el vector de población $\begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$ cambió a

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

Si dividimos ambos lados de esta ecuación entre la población total de un millón y consideramos el hecho de que $kM\mathbf{x} = M(k\mathbf{x})$, se encuentra que

$$\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .600 \\ .400 \end{bmatrix} = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$$

El vector $\mathbf{x}_1 = \begin{bmatrix} .582 \\ .418 \end{bmatrix}$ da la distribución de la población en 2001. Es decir, el 58.2% de la región vivía en la ciudad y el 41.8% vivía en los suburbios. Del mismo modo, la distribución de la población en el año 2002 se describe por un vector $\mathbf{x}_2$, donde

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .582 \\ .418 \end{bmatrix} = \begin{bmatrix} .565 \\ .435 \end{bmatrix} \qquad \blacksquare$$

EJEMPLO 2 Suponga que el resultado de la votación durante las elecciones de un congreso en un distrito de votación está representado por un vector $\mathbf{x}$ en $\mathbb{R}^3$:

$$\mathbf{x} = \begin{bmatrix} \text{\% de votación por los demócratas (D)} \\ \text{\% de votación por los republicanos (R)} \\ \text{\% de votación por los liberales (L)} \end{bmatrix}$$

Suponga que se registra el resultado de las elecciones al congreso cada dos años mediante un vector de este tipo y que el resultado de una elección solo depende de los resultados de la elección anterior. Entonces, la secuencia de los vectores que describen los votos cada dos años puede ser una cadena de Markov. Como un ejemplo de una matriz estocástica P para esta cadena, tomamos

$$\begin{array}{cc} & \begin{array}{ccc} & \text{De:} & \\ \text{D} & \text{R} & \text{L} \end{array} \\ P = & \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \end{array} \begin{array}{c} \text{A:} \\ \text{D} \\ \text{R} \\ \text{L} \end{array}$$

Las entradas en la primera columna, denotada como D, describen qué van a hacer en las próximas elecciones las personas que votaron por los demócratas en una elección. Aquí supusimos que el 70% votará de nuevo por los demócratas en las próximas elecciones, el 20% votará por los republicanos, y el 10% dará su voto a los liberales. Similares interpretaciones valen para las otras columnas de P. Un diagrama de esta matriz se presenta en la figura 2.

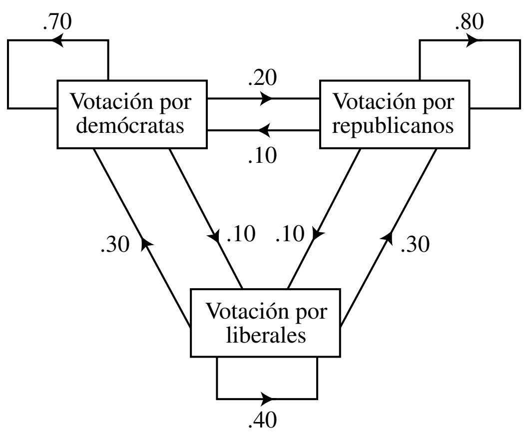


FIGURA 2 Cambios en la votación de una elección a la siguiente.

Si los porcentajes de "transición" se mantienen constantes a lo largo de muchos años de una elección a otra, entonces la secuencia de los vectores que dan los resultados de la votación constituye una cadena de Markov. Suponga que el resultado de una elección está dado por

$$\mathbf{x}_0 = \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix}$$

Determine el resultado probable de la siguiente elección y el resultado probable de la elección posterior.

SOLUCIÓN El resultado de la siguiente elección se describe mediante el vector de estado $\mathbf{x}_1$ y el de la elección posterior por $\mathbf{x}_2$, donde

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .55 \\ .40 \\ .05 \end{bmatrix} = \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} \quad \begin{matrix} \text{El 44\% votará por D} \\ \text{El 44.5\% votará por R} \\ \text{El 11.5\% votará por L} \end{matrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .70 & .10 & .30 \\ .20 & .80 & .30 \\ .10 & .10 & .40 \end{bmatrix} \begin{bmatrix} .440 \\ .445 \\ .115 \end{bmatrix} = \begin{bmatrix} .3870 \\ .4785 \\ .1345 \end{bmatrix} \quad \begin{matrix} \text{El 38.7\% votará por D} \\ \text{El 47.8\% votará por R} \\ \text{El 13.5\% votará por L} \end{matrix}$$

Para entender por qué $\mathbf{x}_1$, efectivamente, da el resultado de las próximas elecciones, suponga que 1000 personas votaron en la "primera" elección, con 550 votos a favor de D, 400 votos a favor de R y 50 votos a favor de L. (Véase los porcentajes en $\mathbf{x}_0$). En la siguiente elección, el 70% de los 550 votarán de nuevo por D, el 10% de los 400 cambiará su voto de R a D, y el 30% de los 50 cambiará de L a D. De esta forma, el total de votos para D será

$$.70(550) + .10(400) + .30(50) = 385 + 40 + 15 = 440 \tag{2}$$

Así, el 44% de la próxima votación será a favor del candidato demócrata. El cálculo en (2) es, en esencia, el mismo que se utilizó para calcular la primera entrada en $\mathbf{x}_1$. Se podrían hacer cálculos análogos para las otras entradas en $\mathbf{x}_1$, para las entradas en $\mathbf{x}_2$, y así sucesivamente. ■

Predicciones del futuro lejano

El aspecto más interesante de las cadenas de Markov es el estudio del comportamiento a largo plazo de una cadena. Por ejemplo, ¿qué se puede decir en el ejemplo 2 de la votación después de que han pasado muchas elecciones (suponiendo que la matriz estocástica dada continúe describiendo los porcentajes de transición de una elección a la siguiente)? O, ¿qué sucede con la distribución de la población en el ejemplo 1 "a largo plazo"? Antes de responder estas preguntas, veamos un ejemplo numérico.

EJEMPLO 3 Sea $P = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Considere un sistema cuyo estado está descrito por la cadena de Markov $\mathbf{x}_{k+1} = P\mathbf{x}_k$, para $k = 0, 1, \ldots$ ¿Qué sucede con el sistema a medida que transcurre el tiempo? Para averiguarlo, calcule los vectores de estado de $\mathbf{x}_1, \ldots, \mathbf{x}_{15}$.

SOLUCIÓN

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix}$$

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .5 \\ .3 \\ .2 \end{bmatrix} = \begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix}$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .37 \\ .45 \\ .18 \end{bmatrix} = \begin{bmatrix} .329 \\ .525 \\ .146 \end{bmatrix}$$

Los resultados de los cálculos posteriores se muestran a continuación, con las entradas redondeadas a cuatro o cinco cifras significativas.

$$\mathbf{x}_4 = \begin{bmatrix} .3133 \\ .5625 \\ .1242 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} .3064 \\ .5813 \\ .1123 \end{bmatrix}, \quad \mathbf{x}_6 = \begin{bmatrix} .3032 \\ .5906 \\ .1062 \end{bmatrix}, \quad \mathbf{x}_7 = \begin{bmatrix} .3016 \\ .5953 \\ .1031 \end{bmatrix}$$

$$\mathbf{x}_8 = \begin{bmatrix} .3008 \\ .5977 \\ .1016 \end{bmatrix}, \quad \mathbf{x}_9 = \begin{bmatrix} .3004 \\ .5988 \\ .1008 \end{bmatrix}, \quad \mathbf{x}_{10} = \begin{bmatrix} .3002 \\ .5994 \\ .1004 \end{bmatrix}, \quad \mathbf{x}_{11} = \begin{bmatrix} .3001 \\ .5997 \\ .1002 \end{bmatrix}$$

$$\mathbf{x}_{12} = \begin{bmatrix} .30005 \\ .59985 \\ .10010 \end{bmatrix}, \quad \mathbf{x}_{13} = \begin{bmatrix} .30002 \\ .59993 \\ .10005 \end{bmatrix}, \quad \mathbf{x}_{14} = \begin{bmatrix} .30001 \\ .59996 \\ .10002 \end{bmatrix}, \quad \mathbf{x}_{15} = \begin{bmatrix} .30001 \\ .59998 \\ .10001 \end{bmatrix}$$

Estos vectores parecen acercarse a $\mathbf{q} = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$. Las probabilidades son difíciles de cambiar de un valor de k al siguiente. Observe que el siguiente cálculo es exacto (sin errores de redondeo):

$$P\mathbf{q} = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}\begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix} = \begin{bmatrix} .15 + .12 + .03 \\ .09 + .48 + .03 \\ .06 + 0 + .04 \end{bmatrix} = \begin{bmatrix} .30 \\ .60 \\ .10 \end{bmatrix} = \mathbf{q}$$

Cuando el sistema está en estado $\mathbf{q}$, no hay ningún cambio en el sistema de una medición a la siguiente. ■

Vectores de estado estable

Si P es una matriz estocástica, entonces un **vector de estado estable** (o **vector de equilibrio**) de P es un vector de probabilidad $\mathbf{q}$ tal que

$$P\mathbf{q} = \mathbf{q}$$

Es posible demostrar que toda matriz estocástica tiene un vector de estado estable. En el ejemplo 3, $\mathbf{q}$ es un vector de estado estable de P.

EJEMPLO 4 El vector de probabilidad $\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$ es un vector de estado estable para la matriz de migración de población M en el ejemplo 1, porque

$$M\mathbf{q} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .375 \\ .625 \end{bmatrix} = \begin{bmatrix} .35625 + .01875 \\ .01875 + .60625 \end{bmatrix} = \begin{bmatrix} .375 \\ .625 \end{bmatrix} = \mathbf{q}$$ ■

Si el total de la población de la región metropolitana en el ejemplo 1 es de 1 millón, entonces **q** del ejemplo 4 correspondería a tener 375,000 personas en la ciudad y 625,000 en los suburbios. Al final de un año, la *emigración* de la ciudad sería (.05)(375,000) = 18,750 personas, y la *inmigración* a la ciudad proveniente de los suburbios sería (.03)(625,000) = 18,750 personas. Como resultado, la población de la ciudad seguiría siendo la misma. Asimismo, la población de los suburbios se mantendría estable.

El siguiente ejemplo muestra cómo *encontrar* un vector de estado estable.

EJEMPLO 5 Sea $P = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$. Encuentre un vector de estado estable para P.

SOLUCIÓN En primer lugar, se resuelve la ecuación $P\mathbf{x} = \mathbf{x}$.

$$\begin{aligned} P\mathbf{x} - \mathbf{x} &= \mathbf{0} \\ P\mathbf{x} - I\mathbf{x} &= \mathbf{0} \quad \text{Recuerde de la sección 1.4 que } I\mathbf{x} = \mathbf{x}. \\ (P - I)\mathbf{x} &= \mathbf{0} \end{aligned}$$

Para la P anterior,

$$P - I = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -.4 & .3 \\ .4 & -.3 \end{bmatrix}$$

Para encontrar todas las soluciones de $(P - I)\mathbf{x} = 0$, se reduce por filas la matriz aumentada:

$$\begin{bmatrix} -.4 & .3 & 0 \\ .4 & -.3 & 0 \end{bmatrix} \sim \begin{bmatrix} -.4 & .3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3/4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Entonces $x_1 = \frac{3}{4}x_2$ y x_2 es libre. La solución general es $x_2 \begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$.

A continuación, se elige una base sencilla para el espacio solución. Una elección evidente es $\begin{bmatrix} 3/4 \\ 1 \end{bmatrix}$, pero una mejor opción, sin fracciones, es $\mathbf{w} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ (correspondiente a $x_2 = 4$).

Por último, encuentre un vector de probabilidad en el conjunto de todas las soluciones de $P\mathbf{x} = \mathbf{x}$. Este proceso es fácil, ya que cada solución es un múltiplo de la solución **w** anterior. Divida **w** entre la suma de sus entradas para obtener

$$\mathbf{q} = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$$

Para comprobar, calcule

$$P\mathbf{q} = \begin{bmatrix} 6/10 & 3/10 \\ 4/10 & 7/10 \end{bmatrix}\begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix} = \begin{bmatrix} 18/70 + 12/70 \\ 12/70 + 28/70 \end{bmatrix} = \begin{bmatrix} 30/70 \\ 40/70 \end{bmatrix} = \mathbf{q} \qquad \blacksquare$$

El siguiente teorema demuestra que lo que sucedió en el ejemplo 3 es característico de muchas matrices estocásticas. Decimos que una matriz estocástica es **regular** si alguna potencia de la matriz P^k contiene solo entradas estrictamente positivas. Para P en el ejemplo 3,

$$P^2 = \begin{bmatrix} .37 & .26 & .33 \\ .45 & .70 & .45 \\ .18 & .04 & .22 \end{bmatrix}$$

Ya que cada entrada en P^2 es estrictamente positiva, P es una matriz estocástica regular.

Además, se dice que una secuencia de vectores $\{\mathbf{x}_k\colon k = 1, 2, \ldots\}$ **converge** a un vector **q** conforme $k \to \infty$ si las entradas en $\mathbf{x}_k$ se pueden acercar tanto como se desee a las entradas correspondientes en **q** al hacer a k suficientemente grande.

TEOREMA 18

Si P es una matriz de $n \times n$ estocástica regular, entonces P tiene un único vector de estado estable $\mathbf{q}$. Además, si $\mathbf{x}_0$ es cualquier estado inicial y una $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$, entonces la cadena de Markov $\{\mathbf{x}_k\}$ converge a $\mathbf{q}$ conforme $k \to \infty$.

Este teorema se demuestra en los libros que analizan cadenas de Markov. La parte sorprendente del teorema es que el estado inicial no tiene ningún efecto sobre el comportamiento a largo plazo de la cadena de Markov. Se verá más adelante (en la sección 5.2) por qué esto es cierto para varias matrices estocásticas estudiadas aquí.

EJEMPLO 6 En el ejemplo 2, ¿qué porcentaje de los electores tienden a votar por el candidato republicano en alguna elección muchos años después a partir de ahora, suponiendo que el resultado de las elecciones forman una cadena de Markov?

SOLUCIÓN Para los cálculos a mano, el enfoque *equivocado* es elegir algún vector inicial $\mathbf{x}_0$ y calcular $\mathbf{x}_1, \ldots, \mathbf{x}_k$ para algún valor grande de k. No hay manera de saber cuántos vectores habrá que calcular, y no se puede estar seguro de los valores límite de las entradas en $\mathbf{x}_k$.

El enfoque correcto es calcular el vector de estado estable y, luego, recurrir al teorema 18. Dado P como en el ejemplo 2, se forma $P - I$ restando 1 de cada entrada diagonal en P. Después se reduce por filas la matriz aumentada:

$$[\,(P - I) \quad \mathbf{0}\,] = \begin{bmatrix} -.3 & .1 & .3 & 0 \\ .2 & -.2 & .3 & 0 \\ .1 & .1 & -.6 & 0 \end{bmatrix}$$

Recuerde, de trabajos anteriores con decimales, que el cálculo se simplifica al multiplicar cada fila por 10.[1]

$$\begin{bmatrix} -3 & 1 & 3 & 0 \\ 2 & -2 & 3 & 0 \\ 1 & 1 & -6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -9/4 & 0 \\ 0 & 1 & -15/4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general de $(P - I)\mathbf{x} = \mathbf{0}$ es $x_1 = \frac{9}{4}x_3$, $x_2 = \frac{15}{4}x_3$ y x_3 es libre. Eligiendo $x_3 = 4$, se obtiene una base para el espacio solución cuyas entradas son enteros, y a partir de ello se encuentra fácilmente el vector de estado estable cuyas entradas suman 1:

$$\mathbf{w} = \begin{bmatrix} 9 \\ 15 \\ 4 \end{bmatrix} \quad \text{y} \quad \mathbf{q} = \begin{bmatrix} 9/28 \\ 15/28 \\ 4/28 \end{bmatrix} \approx \begin{bmatrix} .32 \\ .54 \\ .14 \end{bmatrix}$$

Las entradas en $\mathbf{q}$ describen la distribución de los votos en una elección que se celebrará dentro de muchos años (suponiendo que la matriz estocástica continúa describiendo los cambios de una elección a otra). Así, finalmente, alrededor del 54% de los votos serán para el candidato republicano. ■

[1] *Advertencia:* No solo multiplique P por 10. En vez de ello, multiplique la matriz aumentada para la ecuación $(P - I)\mathbf{x} = 0$ por 10.

NOTA NUMÉRICA

Quizás haya notado que si $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, \ldots$ entonces

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0,$$

y, en general,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para } k = 0, 1, \ldots$$

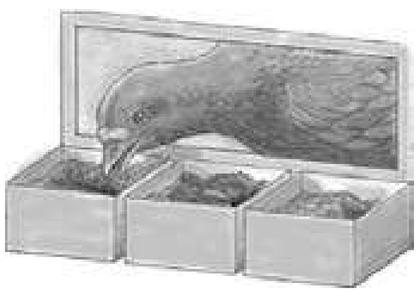
Para calcular un vector específico, tal como $\mathbf{x}_3$, se necesitan menos operaciones aritméticas para calcular $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$, más que P^3 y $P^3\mathbf{x}_0$. Sin embargo, si P es pequeña, por ejemplo, 30×30, el tiempo de cálculo de la máquina es insignificante con ambos métodos, y tal vez se prefiera un comando para calcular $P^3\mathbf{x}_0$, ya que requiere teclear menos.

PROBLEMAS DE PRÁCTICA

1. Suponga que los habitantes de una región metropolitana se desplazan de acuerdo con las probabilidades de la matriz de migración M del ejemplo 1, y que se elige un residente "al azar". Entonces, un vector de estado para un año determinado se puede interpretar como un indicador de las probabilidades de que la persona sea residente de la ciudad o de los suburbios en ese momento.
   a) Suponga que la persona elegida es un residente de la ciudad ahora, de manera que $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. ¿Cuál es la probabilidad de que esa persona viva en los suburbios el año próximo?
   b) ¿Cuál es la probabilidad de que esa persona viva en los suburbios en dos años?
2. Sea $P = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}$ y $\mathbf{q} = \begin{bmatrix} .3 \\ .7 \end{bmatrix}$. ¿Es $\mathbf{q}$ un vector de estado estable para P?
3. ¿Qué porcentaje de la población en el ejemplo 1 vivirá en los suburbios después de muchos años?

4.9 EJERCICIOS

1. Un pueblo remoto y pequeño recibe transmisiones de radio de dos estaciones radiofónicas: una de noticias y una emisora de música. De los oyentes que sintonizan la emisora de noticias, el 70% seguirá escuchando las noticias después del corte de estación que ocurre cada media hora, mientras que el 30% sintonizará la estación de música durante el corte de estación. De los oyentes que sintonizan la estación de música, el 60% cambiará a la estación de noticias en el corte de estación, mientras que el 40% permanecerá escuchando la música. Suponga que todos los habitantes están escuchando las noticias a las 8:15 A.M.
   a) Dé la matriz estocástica que describe cómo los oyentes de radio tienden a cambiar las estaciones en cada corte de estación. Rotule las filas y las columnas.
   b) Dé el vector de estado inicial.
   c) ¿Qué porcentaje de los oyentes estará escuchando la emisora de música a las 9:25 A.M. (después de los cortes de estación de las 8:30 y de las 9:00 A.M.)?
2. Un animal de laboratorio puede comer cualquiera de tres alimentos cada día. Los registros de laboratorio indican que si el animal elige un alimento en un ensayo, elegirá el mismo alimento en el siguiente ensayo con una probabilidad del 60%, y elegirá cualquiera de los otros alimentos en el siguiente ensayo con iguales probabilidades del 20%.
   a) ¿Cuál es la matriz estocástica de esta situación?
   b) Si el animal elige el alimento #1 en un ensayo inicial, ¿cuál es la probabilidad de que elija el alimento #2 en el segundo ensayo después del inicial?



3. En un día cualquiera, un estudiante está sano o enfermo. De los estudiantes que están sanos hoy, el 95% estarán sanos mañana.

De los estudiantes que están enfermos hoy, el 55% seguirán enfermos mañana.

a) ¿Cuál es la matriz estocástica de esta situación?

b) Suponga que el 20% de los alumnos están enfermos el lunes. ¿Qué fracción o porcentaje de los estudiantes es probable que estén enfermos el martes? ¿Y el miércoles?

c) Si un estudiante está sano hoy, ¿cuál es la probabilidad de que esté sano dentro de dos días?

4. El clima en Columbus es bueno, regular o malo en un día determinado. Si el clima es bueno hoy, hay un 40% de probabilidad de que sea bueno mañana, un 30% de probabilidad de que sea regular, y un 30% de que sea malo. Si el clima es regular hoy, existe un 50% de probabilidad de que sea bueno mañana, y un 20% de probabilidad de que sea regular. Por último, si el clima es malo hoy, existe un 30% de probabilidad de que sea bueno mañana y un 40% de que sea regular.

a) ¿Cuál es la matriz estocástica de esta situación?

b) Suponga que hay una probabilidad del 50% de buen clima hoy, y una probabilidad del 50% de clima regular. ¿Cuáles son las probabilidades de que el clima sea malo mañana?

c) Suponga que, de acuerdo con los pronósticos para el lunes, hay un 60% de probabilidad de que el clima sea regular y un 40% de que sea malo. ¿Cuáles son las probabilidades de tener buen clima el miércoles?

En los ejercicios 5 a 8, encuentre el vector de estado estable.

5. $\begin{bmatrix} .1 & .5 \\ .9 & .5 \end{bmatrix}$

6. $\begin{bmatrix} .4 & .8 \\ .6 & .2 \end{bmatrix}$

7. $\begin{bmatrix} .7 & .1 & .1 \\ .2 & .8 & .2 \\ .1 & .1 & .7 \end{bmatrix}$

8. $\begin{bmatrix} .4 & .5 & .8 \\ 0 & .5 & .1 \\ .6 & 0 & .1 \end{bmatrix}$

9. Determine si $P = \begin{bmatrix} .2 & 1 \\ .8 & 0 \end{bmatrix}$ es una matriz estocástica regular.

10. Determine si $P = \begin{bmatrix} 1 & .3 \\ 0 & .7 \end{bmatrix}$ es una matriz estocástica regular.

11. *a*) Encuentre el vector de estado estable para la cadena de Markov en el ejercicio 1.

b) En algún momento al final del día, ¿qué fracción de los oyentes escuchará las noticias?

12. Consulte el ejercicio 2. ¿Qué alimento prefiere el animal después de muchos ensayos?

13. *a*) Encuentre el vector de estado estable para la cadena de Markov en el ejercicio 3.

b) ¿Cuál es la probabilidad de que, después de muchos días, un estudiante en particular esté enfermo? ¿Importa si esa persona está enferma hoy?

14. Consulte el ejercicio 4. En el largo plazo, ¿qué tan probable es que el clima en Columbus sea bueno en un día determinado?

15. [**M**] La Unidad de Investigación Demográfica del Departamento de Finanzas del Estado de California proporcionó los siguientes datos para la matriz de migración, la cual describe el movimiento de la población de Estados Unidos durante 1989. En ese año, alrededor del 11.7% de la población total vivía en California. ¿Qué porcentaje de la población total con el tiempo vivirá en California si las probabilidades de migración mencionadas se mantuvieran constantes durante muchos años?

$$\begin{array}{cc} \multicolumn{2}{c}{\text{De:}} & \\ \text{CA} & \text{Resto de E.U.A.} & \text{A:} \\ \left[\begin{array}{cc} .9821 & .0029 \\ .0179 & .9971 \end{array}\right] & & \begin{array}{l} \text{California} \\ \text{Resto de E.U.A.} \end{array} \end{array}$$

16. [**M**] En Detroit, Hertz Rent A Car cuenta con una flota de cerca de 2000 automóviles. El patrón de puntos de alquiler y devolución de las unidades está descrito por las fracciones en la siguiente tabla. En un día típico, ¿cuántos autos estarán listos para rentarse en la sucursal ubicada en el centro de la ciudad?

$$\begin{array}{ccc} \multicolumn{3}{c}{\text{Autos rentados en:}} & \\ \text{Aeropuerto de la cd.} & \text{Centro} & \text{Apto. fuera de cd.} & \text{Devueltos en:} \\ \left[\begin{array}{ccc} .90 & .01 & .09 \\ .01 & .90 & .01 \\ .09 & .09 & .90 \end{array}\right] & & & \begin{array}{l} \text{Aeropuerto de la cd.} \\ \text{Centro} \\ \text{Apto. fuera de cd.} \end{array} \end{array}$$

17. Sea P una matriz estocástica de $n \times n$. El siguiente argumento indica que la ecuación $P\mathbf{x} = \mathbf{x}$ tiene una solución no trivial. (De hecho, una solución de estado estable existe con entradas no negativas. En algunos libros de texto avanzados se da una demostración). Justifique cada una de las siguientes afirmaciones. (Mencione un teorema cuando sea pertinente).

a) Si todas las otras filas de $P - I$ se suman a la fila de abajo, el resultado es una fila de ceros.

b) Las filas de $P - I$ son linealmente dependientes.

c) La dimensión del espacio fila de $P - I$ es menor que n.

d) $P - I$ tiene un espacio nulo no trivial.

18. Demuestre que toda matriz estocástica de 2×2 tiene al menos un vector de estado estable. Cualquier matriz de este tipo se puede representar en la forma $P = \begin{bmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{bmatrix}$, donde α y β son constantes entre 0 y 1. (Hay dos vectores de estado estable, linealmente independientes, si $\alpha = \beta = 0$. De lo contrario, solo hay uno).

19. Sea S una matriz fila de $1 \times n$ con un 1 en cada columna, $S = [1 \ \ 1 \ \cdots \ 1]$

a) Explique por qué un vector $\mathbf{x}$ en $\mathbb{R}^n$ es un vector de probabilidad si y solo si sus entradas son no negativas y $S\mathbf{x} = 1$. (Una matriz de 1×1 como el producto $S\mathbf{x}$ generalmente se representa sin los símbolos de corchete de la matriz).

b) Sea P una matriz estocástica de $n \times n$. Explique por qué $SP = S$.

c) Sea P una matriz estocástica de $n \times n$, y sea $\mathbf{x}$ un vector de probabilidad. Demuestre que $P\mathbf{x}$ es también un vector de probabilidad.

20. Con base en el ejercicio 19, demuestre que si P es una matriz estocástica de $n \times n$, entonces también lo es P^2.

21. [**M**] Examine las potencias de una matriz estocástica regular.

a) Calcule P^k para $k = 2, 3, 4, 5$, cuando

$$P = \begin{bmatrix} .3355 & .3682 & .3067 & .0389 \\ .2663 & .2723 & .3277 & .5451 \\ .1935 & .1502 & .1589 & .2395 \\ .2047 & .2093 & .2067 & .1765 \end{bmatrix}$$

Presente los cálculos con cuatro decimales. ¿Qué ocurre a las columnas de P^k conforme k aumenta? Calcule el vector de estado estable para P.

b) Calcule Q^k para $k = 10, 20,\ldots, 80$, cuando

$$Q = \begin{bmatrix} .97 & .05 & .10 \\ 0 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix}$$

(La estabilidad de Q^k con cuatro decimales puede requerir $k = 116$ o más). Calcule el vector de estado estable para Q. Haga una conjetura sobre lo que podría ser cierto para cualquier matriz estocástica regular.

c) Con base en el teorema 18, explique lo que encontró en los incisos *a*) y *b*).

22. [**M**] Compare dos métodos para encontrar el vector de estado estable **q** de una matriz estocástica regular P: **1.** calculando **q** como en el ejemplo 5, o **2.** calculando P^k para un valor grande de k y utilizando una de las columnas de P^k como una aproximación para **q**. [La *Guía de estudio* describe un programa de *base nula* que casi automatiza el método 1].

Experimente con las matrices aleatorias estocásticas más grandes que su programa de matrices le permita, y utilice $k = 100$ o algún otro valor grande. Para cada método, describa el tiempo que *usted* necesita para teclear y ejecutar su programa. (Algunas versiones de MATLAB tienen los comandos `flops` y `tic` …`toc` que registran el número de operaciones de punto flotante y el tiempo total transcurrido que utiliza MATLAB). Compare las ventajas de cada método y determine cuál prefiere.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. *a*) Como el 5% de los residentes de la ciudad se mudará a los suburbios en menos de un año, hay una probabilidad del 5% de elegir a esa persona. Sin profundizar en el conocimiento acerca del individuo, se dice que hay un 5% de probabilidad de que este se mude a los suburbios. Este hecho se encuentra en la segunda entrada del vector de estado $\mathbf{x}_1$, donde

$$\mathbf{x}_1 = M\mathbf{x}_0 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .95 \\ .05 \end{bmatrix}$$

b) La probabilidad de que la persona viva en los suburbios después de dos años es del 9.6%, porque

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .95 \\ .05 \end{bmatrix} = \begin{bmatrix} .904 \\ .096 \end{bmatrix}$$

2. El vector de estado estable satisface $P\mathbf{x} = \mathbf{x}$. Como

$$P\mathbf{q} = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}\begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .32 \\ .68 \end{bmatrix} \neq \mathbf{q}$$

llegamos a la conclusión de que **q** *no* es el vector de estado estable para P.

3. M en el ejemplo 1 es una matriz estocástica regular, ya que sus entradas son estrictamente positivas. Así que podemos utilizar el teorema 18. Ya conocemos el vector de estado estable del ejemplo 4. Por lo tanto, los vectores de distribución de la población $\mathbf{x}_k$ convergen en

$$\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$$

WEB

Al final, el 62.5% de la población vivirá en los suburbios.

CAPÍTULO 4 EJERCICIOS COMPLEMENTARIOS

1. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. (Si es verdadero, cite hechos o teoremas pertinentes. Si es falso, explique por qué o dé un contraejemplo que muestre por qué el enunciado no es cierto en todos los casos). En los incisos *a*) a *f*), $\mathbf{v}_1,\ldots, \mathbf{v}_p$ son vectores en un espacio vectorial V con dimensión finita diferente de cero, y $S = \{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$.

a) El conjunto de todas las combinaciones lineales de $\mathbf{v}_1,\ldots, \mathbf{v}_p$ es un espacio vectorial.

b) Si $\{\mathbf{v}_1,\ldots, \mathbf{v}_{p-1}\}$ genera a V, entonces S genera a V.

c) Si $\{\mathbf{v}_1,\ldots, \mathbf{v}_{p-1}\}$ es linealmente independiente, también S lo es.

d) Si S es linealmente independiente, entonces S es una base para V.

e) Si Gen $S = V$, entonces algún subconjunto de S es una base para V.

f) Si dim $V = p$ y Gen $S = V$, entonces S no puede ser linealmente dependiente.

g) Un plano en $\mathbb{R}^3$ es un subespacio de dimensión 2.

h) Las columnas de una matriz que no son pivote siempre son linealmente dependientes.

i) Las operaciones de fila en una matriz A pueden cambiar las relaciones de dependencia lineal entre las filas de A.

j) Las operaciones de fila en una matriz pueden cambiar el espacio nulo.

k) El rango de una matriz es igual al número de filas diferentes de cero.

l) Si una matriz A de $m \times n$ es equivalente por filas a una matriz escalonada U, y si U tiene k filas diferentes de cero, entonces la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$ es $m - k$.

m) Si B se obtiene a partir de una matriz A con varias operaciones elementales de fila, entonces rango B = rango A.

n) Las filas diferentes de cero de una matriz A forman una base para Fila A.

o) Si las matrices A y B tienen la misma forma escalonada reducida, entonces Fila A = Fila B.

p) Si H es un subespacio de $\mathbb{R}^3$, entonces hay una matriz A de 3×3 tal que H = Col A.

q) Si A es $m \times n$ y rango $A = m$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno.

r) Si A es de $m \times n$ y la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ es sobre, entonces rango $A = m$.

s) Una matriz de cambio de coordenadas siempre es invertible.

t) Si $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$ y $\mathcal{C} = \{\mathbf{c}_1,\ldots, \mathbf{c}_n\}$ son bases para un espacio vectorial V, entonces la j-ésima columna de la matriz de cambio de coordenadas de la matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ es el vector de coordenadas $[\mathbf{c}_j]_{\mathcal{B}}$.

2. Encuentre una base para el conjunto de todos los vectores de la forma

$$\begin{bmatrix} a - 2b + 5c \\ 2a + 5b - 8c \\ -a - 4b + 7c \\ 3a + b + c \end{bmatrix}. \text{ (Tenga cuidado).}$$

3. Sean $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ y W = Gen $\{\mathbf{u}_1, \mathbf{u}_2\}$. Encuentre una descripción *implícita* de W, es decir, encuentre un conjunto de una o más ecuaciones homogéneas que caracterizan a los puntos de W. [*Sugerencia:* Pregúntese cuándo está $\mathbf{b}$ en W].

4. Explique lo que está mal en el siguiente análisis: Sea $\mathbf{f}(t) = 3 + t$ y $\mathbf{g}(t) = 3t + t^2$, y note que $\mathbf{g}(t) = t\mathbf{f}(t)$. Entonces, $\{\mathbf{f}, \mathbf{g}\}$ es linealmente dependiente, porque $\mathbf{g}$ es un múltiplo de $\mathbf{f}$.

5. Considere los polinomios $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, $\mathbf{p}_3(t) = 4$, $\mathbf{p}_4(t) = t + t^2$, y $\mathbf{p}_5(t) = 1 + 2t + t^2$, y sea H el subespacio de $\mathbb{P}_5$ generado por el conjunto $S = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$. Utilice el método descrito en la demostración del teorema del conjunto generador (sección 4.3) con la finalidad de obtener una base para H. (Explique cómo seleccionar los miembros adecuados de S).

6. Suponga que $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ son polinomios específicos que generan un subespacio H de dimensión 2 de $\mathbb{P}_5$. Describa cómo se puede encontrar una base para H mediante el examen de los cuatro polinomios y sin hacer casi ningún cálculo.

7. ¿Qué tendría que saber acerca del conjunto solución de un sistema homogéneo de 18 ecuaciones lineales con 20 variables con la finalidad de determinar que cada ecuación no homogénea asociada tiene una solución? Analice.

8. Sea H un subespacio de dimensión n de un espacio vectorial V de dimensión n. Explique por qué $H = V$.

9. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal.

a) ¿Cuál es la dimensión del rango de T si T es un mapeo uno a uno? Explique su respuesta.

b) ¿Cuál es la dimensión del núcleo de T (véase la sección 4.2) si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^m$? Explique su respuesta.

10. Sea S un subconjunto maximal linealmente independiente de un espacio vectorial V. Es decir, S tiene la propiedad de que si un vector que no está en S se adjunta a S, entonces el nuevo conjunto ya no será linealmente independiente. Demuestre que S debe ser una base para V. [*Sugerencia:* Pregúntese qué pasa si S es linealmente independiente, pero no una base de V].

11. Sea S un conjunto generador mínimo finito de un espacio vectorial V. Es decir, S tiene la propiedad de que si un vector se elimina de S, entonces el nuevo conjunto ya no genera a V. Demuestre que S debe ser una base para V.

Los ejercicios 12 a 17 desarrollan las propiedades del rango que a veces son necesarias en las aplicaciones. Suponga que la matriz A es de $m \times n$.

12. A partir de los incisos *a*) y *b*) demuestre que el rango AB no puede exceder al rango de A o al rango de B. (En general, el rango de un producto de matrices no puede exceder el rango de cualquier otro factor en el producto).

a) Demuestre que si B es de $n \times p$, entonces rango $AB \le$ rango A. [*Sugerencia:* Explique por qué cada vector en el espacio columna de AB está en el espacio columna de A].

b) Demuestre que si B es $n \times p$, entonces rango $AB \le$ rango B. [*Sugerencia:* Tome en cuenta el inciso *a*) para estudiar el rango de $(AB)^T$].

13. Demuestre que si P es una matriz invertible de $m \times m$, entonces el rango PA = rango A. [*Sugerencia:* Aplique el ejercicio 12 a PA y $P^{-1}(PA)$].

14. Demuestre que si Q es invertible, entonces rango AQ = rango A. [*Sugerencia:* Tome en cuenta el ejercicio 13 para estudiar el rango de $(AQ)^T$].

15. Sea A una matriz de $m \times n$, y B una matriz de $n \times p$ tales que $AB = 0$. Demuestre que rango A + rango $B \le n$. [*Sugerencia:* Considere que uno de los cuatro subespacios Nul A, Col A, Nul B y Col B se encuentra en uno de los otros tres subespacios].

16. Si A es una matriz de $m \times n$ de rango r, entonces la *factorización de rango* de A es una ecuación de la forma $A = CR$, donde C es una matriz de $m \times r$ de rango r, y R es una matriz de $r \times n$

de rango r. Tal factorización siempre existe (ejercicio 38 en la sección 4.6). Dadas dos matrices A y B de $m \times n$, utilice factorizaciones de rango de A y B para demostrar que

rango $(A + B) \leq$ rango A + rango B

[*Sugerencia:* Escriba $A + B$ como el producto de dos matrices particionadas].

17. Una **submatriz** de una matriz A es cualquier matriz que resulta de la eliminación de algunas filas y/o columnas (o de ninguna) de A. Es posible demostrar que A tiene rango r si y solo si A contiene una submatriz invertible de $r \times r$ y ninguna submatriz cuadrada mayor es invertible. Demuestre parte de este enunciado explicando: *a*) por qué una matriz A de $m \times n$ y rango r tiene una submatriz A_1 de $m \times r$ y rango r, y *b*) por qué A_1 tiene una submatriz invertible A_2 de $r \times r$.

El concepto de rango desempeña un papel importante en el diseño de sistemas de control de ingeniería, como el sistema del transbordador espacial mencionado en el ejemplo introductorio de este capítulo. Un *modelo de espacios de estado* de un sistema de control incluye una ecuación en diferencias de la forma

$$\mathbf{x}_{k+1} = A\mathbf{x} + B\mathbf{u}_k \quad \text{para } k = 0, 1, \ldots \tag{1}$$

donde A es de $n \times n$, B es de $n \times m$, $\{\mathbf{x}_k\}$ es una secuencia de "vectores de estado" en $\mathbb{R}^n$ que describe el estado del sistema en tiempos discretos, y $\{\boldsymbol{u}_k\}$ es una secuencia de *control* o de *entrada*. Se dice que el par (A, B) es **controlable** si

$$\text{rango } [B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B] = n \tag{2}$$

La matriz que aparece en la ecuación (2) se llama **matriz de controlabilidad** del sistema. Si (A, B) es controlable, entonces el sistema se puede controlar, o conducir del estado **0** a cualquier estado especificado $\mathbf{v}$ (en $\mathbb{R}^n$) en un máximo de n pasos, simplemente eligiendo una secuencia de control adecuada en $\mathbb{R}^m$. Este hecho se ilustra en el ejercicio 18 para $n = 4$ y $m = 2$. Para un análisis adicional de la capacidad de control, visite el sitio Web de este libro (estudio de caso del capítulo 4).

WEB

18. Suponga que A es una matriz de 4×4 y B es una matriz de 4×2, y sea que $\mathbf{u}_0, \ldots, \mathbf{u}_3$ representen una secuencia de vectores de entrada en $\mathbb{R}^2$.

a) Establezca $\mathbf{x}_0 = \mathbf{0}$, calcule $\mathbf{x}_1, \ldots, \mathbf{x}_4$ a partir de la ecuación (1), y escriba una fórmula para $\mathbf{x}_4$ que implique a la matriz de controlabilidad M que aparece en la ecuación (2). (*Nota:* La matriz M se construye como una matriz particionada; su tamaño total aquí es 4×8).

b) Suponga que (A, B) es controlable y $\mathbf{v}$ es cualquier vector en $\mathbb{R}^4$. Explique por qué existe una secuencia de control $\mathbf{u}_0, \ldots, \mathbf{u}_3$ en $\mathbb{R}^2$ tal que $\mathbf{x}_4 = \mathbf{v}$.

Determine si los pares de matrices de los ejercicios 19 a 22 son controlables.

19. $A = \begin{bmatrix} .9 & 1 & 0 \\ 0 & -.9 & 0 \\ 0 & 0 & .5 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

20. $A = \begin{bmatrix} .8 & -.3 & 0 \\ .2 & .5 & 1 \\ 0 & 0 & -.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

21. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -4.2 & -4.8 & -3.6 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

22. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -13 & -12.2 & -1.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

5

Valores propios y vectores propios

EJEMPLO INTRODUCTORIO

Sistemas dinámicos y búhos manchados

En 1990 el búho manchado norteño se convirtió en el centro de una controversia nacional sobre el uso y el mal uso de los majestuosos bosques en el Pacífico Noroccidental de Estados Unidos. Los ambientalistas convencieron al gobierno federal de que el búho estaría en riesgo de extinción si continuaba la tala en los antiguos bosques (con árboles de más de 200 años de edad), donde el búho prefiere vivir. La industria maderera, anticipando la pérdida de 30,000 a 100,000 puestos de trabajo como resultado de las nuevas restricciones gubernamentales sobre la tala, argumentó que el búho no debería estar clasificado como "especie amenazada" y citó diferentes reportes científicos publicados para apoyar su caso.[1]

En medio del fuego cruzado entre los grupos antagónicos, los ecologistas matemáticos intensificaron su afán de comprender la dinámica poblacional del búho. El ciclo de vida de un búho manchado se divide naturalmente en tres etapas: juvenil (hasta un año de edad), subadulto (de 1 a 2 años) y adulto (más de 2 años). Los búhos se aparean de por vida en las etapas de subadulto y adulto, y empiezan a criar cuando son adultos. Viven alrededor de 20 años y cada pareja requiere aproximadamente 1000 hectáreas (4 millas cuadradas) como su propio territorio. Un tiempo fundamental en el ciclo de vida es cuando los búhos juveniles dejan el nido. Para sobrevivir y convertirse en un subadulto, un búho juvenil debe tener éxito para encontrar un nuevo territorio para vivir (y en general, una pareja).

Un primer paso en el estudio de la dinámica poblacional consiste en modelar la población a intervalos anuales, en tiempos que se denotan con $k = 0, 1, 2,...$ En general, se supone que hay una relación 1:1 de machos a hembras en cada etapa del ciclo de vida y se cuentan solamente las hembras. La población en cada año k se describe con un vector $\boldsymbol{x}_k = (j_k, s_k, a_k)$, donde j_k, s_k y a_k son los números de hembras en las etapas juvenil, subadulto y adulto, respectivamente.

Utilizando datos reales de estudios demográficos, R. Lamberson y sus colaboradores consideraron el siguiente *modelo matricial por etapas*:[2]

$$\begin{bmatrix} j_{k+1} \\ s_{k+1} \\ a_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \begin{bmatrix} j_k \\ s_k \\ a_k \end{bmatrix}$$

Aquí, el número de nuevas hembras juveniles en el año $k + 1$ es .33 veces el número de hembras adultas en el año k (según la tasa promedio de nacimientos por pareja de búhos). Asimismo, 18% de los búhos juveniles sobreviven para convertirse en subadultos, en tanto que 71% de los subadultos y 94% de los adultos sobreviven para contarse como adultos.

El modelo matricial por etapas es una ecuación en diferencias de la forma $\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k$. A esta ecuación con frecuencia se le llama **sistema dinámico** (o **sistema**

[1] "The Great Spotted Owl War", *Reader's Digest*, noviembre de 1992, pp. 91-95.

[2] R. H. Lamberson, R. McKelvey, B. R. Noon y C. Voss, "A Dynamic Analysis of the Viability of the Northern Spotted Owl in a Fragmented Forest Environment", *Conservation Biology* **6** (1992), pp. 505-512. También, una comunicación privada del profesor Lamberson, 1993.

dinámico lineal discreto), ya que describe los cambios en un sistema conforme transcurre el tiempo.

La tasa de sobrevivencia del 18% de búhos juveniles en la matriz por etapas de Lamberson es la entrada más afectada por la cantidad de bosque antiguo disponible. En realidad, el 60% de los búhos juveniles normalmente sobreviven para dejar el nido, pero en la región de Willow Creek, California, estudiada por Lamberson y sus colaboradores, tan solo al 30% de los búhos juveniles que dejaron el nido les fue posible encontrar un nuevo territorio para vivir. Los restantes perecieron durante la búsqueda.

Un motivo significativo para el fracaso de los búhos en encontrar nuevas áreas de distribución es el aumento en la fragmentación de las áreas con árboles antiguos, debido a la tala total de áreas dispersas en terrenos con árboles antiguos. Cuando un búho sale de la región protectora del bosque y cruza un área talada, aumenta drásticamente el riesgo de sufrir un ataque por parte de los depredadores. En la sección 5.6 se mostrará que el modelo descrito predice la eventual extinción de búhos, pero que si el 50% de los búhos juveniles que sobreviven al dejar el nido también encuentran nuevas áreas de distribución, entonces se salvaría su población.

WEB

El objetivo de este capítulo es descomponer la acción de una transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ en elementos que sean de fácil visualización. Excepto por un breve paréntesis en la sección 5.4, en este capítulo todas las matrices son cuadradas. Las principales aplicaciones descritas aquí son para sistemas dinámicos discretos, incluyendo el ya mencionado asunto de los búhos manchados. Sin embargo, los conceptos básicos, vectores propios y valores propios, son útiles en matemáticas puras y aplicadas, y se presentan en situaciones más generales que las que aquí se consideran. Los valores propios también sirven para estudiar ecuaciones diferenciales y sistemas dinámicos *continuos*, brindan información crítica en diseños de ingeniería, y se originan de manera natural en campos de la física y la química.

5.1 VECTORES PROPIOS Y VALORES PROPIOS

No obstante que una transformación $\mathbf{x} \mapsto A\mathbf{x}$ puede mover vectores en diversas direcciones, a menudo sucede que hay vectores especiales sobre los cuales la acción de A es bastante simple.

EJEMPLO 1 Sean $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. En la figura 1 se muestran las imágenes de $\mathbf{u}$ y $\mathbf{v}$ bajo la multiplicación por A. En efecto, $A\mathbf{v}$ es precisamente $2\mathbf{v}$. Así, A tan solo "estira" o dilata a $\mathbf{v}$. ■

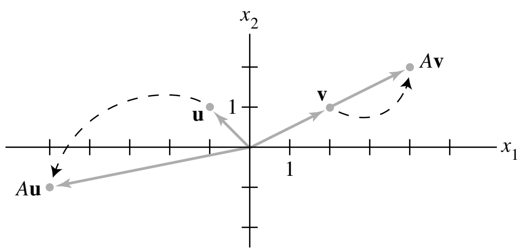


FIGURA 1 Efectos de la multiplicación por A.

Como otro ejemplo, los lectores de la sección 4.9 recordarán que si A es una matriz estocástica, entonces el vector de estado estable $\mathbf{q}$ para A satisface la ecuación $A\mathbf{x} = \mathbf{x}$. Es decir, $A\mathbf{q} = 1 \cdot \mathbf{q}$.

Esta sección estudia ecuaciones tales como

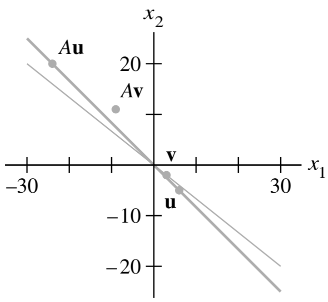
$$A\mathbf{x} = 2\mathbf{x} \quad \text{o bien} \quad A\mathbf{x} = -4\mathbf{x}$$

donde los vectores especiales son transformados por A en múltiplos escalares de sí mismos.

DEFINICIÓN

Un **vector propio** de una matriz A de $n \times n$ es un vector distinto de cero $\mathbf{x}$ tal que $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ. Un escalar λ es un **valor propio** de A si existe una solución no trivial $\mathbf{x}$ de $A\mathbf{x} = \lambda\mathbf{x}$; tal $\mathbf{x}$ es el *vector propio correspondiente a* λ.[1]

Es fácil determinar si un vector dado es un vector propio de una matriz. También es sencillo decidir si cierto escalar es un valor propio.



$A\mathbf{u} = -4\mathbf{u}$, pero $A\mathbf{v} \neq \lambda\mathbf{v}$.

EJEMPLO 2 Sean $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. ¿Son $\mathbf{u}$ y $\mathbf{v}$ vectores propios de A?

SOLUCIÓN

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Entonces, $\mathbf{u}$ es un vector propio correspondiente a un valor propio (-4), pero $\mathbf{v}$ no es un vector propio de A porque $A\mathbf{v}$ no es múltiplo de $\mathbf{v}$. ■

EJEMPLO 3 Demuestre que 7 es un valor propio de la matriz A del ejemplo 2, y determine los vectores propios correspondientes.

SOLUCIÓN El escalar 7 es un valor propio de A, si y solo si la ecuación

$$A\mathbf{x} = 7\mathbf{x} \tag{1}$$

tiene una solución no trivial. No obstante, la ecuación (1) es equivalente a $A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$, o bien,

$$(A - 7I)\mathbf{x} = \mathbf{0} \tag{2}$$

Para resolver esta ecuación homogénea, forme la matriz

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

Como es evidente, las columnas de $A - 7I$ son linealmente dependientes, de manera que la ecuación (2) tiene soluciones no triviales. Por lo tanto, 7 *es* un valor propio de A. Para encontrar los vectores propios correspondientes, utilice operaciones de fila:

$$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La solución general tiene la forma $x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Cada vector de esta forma con $x_2 \neq 0$ es un vector propio correspondiente a $\lambda = 7$. ■

[1] Advierta que, por definición, un vector propio debe ser *distinto* de cero, pero un valor propio puede ser cero. Después del ejemplo 5, se analizará el caso donde el número 0 es un valor propio.

Advertencia: No obstante que en el ejemplo 3 se utilizó reducción por filas para encontrar *vectores* propios, este método no se puede emplear para determinar *valores* propios. En general, una forma escalonada de una matriz A *no* presenta los valores propios de A.

Evidentemente, la equivalencia de las ecuaciones (1) y (2) es válida para cualquier λ en vez de $\lambda = 7$. Entonces, λ es un valor propio de la matriz A de $n \times n$ si y solo si la ecuación

$$(A - \lambda I)\mathbf{x} = \mathbf{0} \tag{3}$$

tiene una solución no trivial. El conjunto de *todas* las soluciones de (3) es justo el espacio nulo de la matriz de $A - \lambda I$. Así que este conjunto es un *subespacio* de $\mathbb{R}^n$ y se denomina el **espacio propio** de A correspondiente a λ. El espacio propio consiste en el vector cero y de todos los vectores propios correspondientes a λ.

El ejemplo 3 demuestra que para la matriz A del ejemplo 2, el espacio propio correspondiente a $\lambda = 7$ consiste en *todos* los múltiplos de (1, 1), que es la recta que pasa tanto por (1, 1) como por el origen. Con el ejemplo 2, se comprueba que el espacio propio asociado con $\lambda = -4$ es la recta que pasa por $(6, -5)$. En la figura 2, se presentan esos espacios propios, junto con los vectores propios (1, 1) y $(3/2, -5/4)$ y la acción geométrica de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ sobre cada espacio propio.

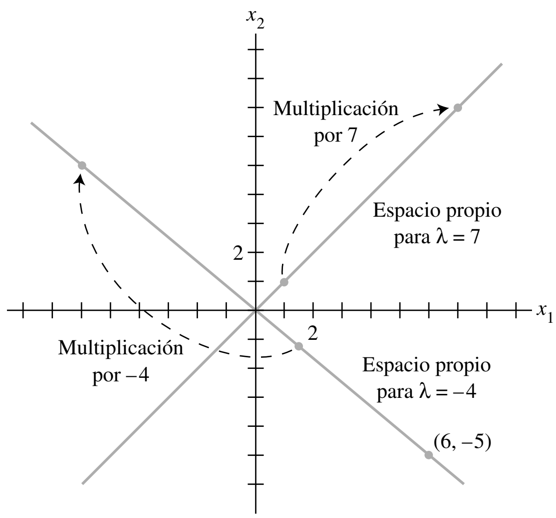


FIGURA 2 Espacios propios para $\lambda = -4$ y $\lambda = 7$.

EJEMPLO 4 Sea $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. Un valor propio de A es 2. Determine una base para el espacio propio correspondiente.

SOLUCIÓN Forme

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

y reduzca por filas la matriz aumentada para $(A - 2I)\mathbf{x} = \mathbf{0}$:

$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

En este punto, resulta claro que 2 es un valor propio de A porque la ecuación $(A - 2I)\mathbf{x} = \mathbf{0}$ tiene variables libres. La solución general es:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ y } x_3 \text{ son libres}$$

El espacio propio, que se muestra en la figura 3, es un subespacio bidimensional de $\mathbb{R}^3$. Una base es

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

■

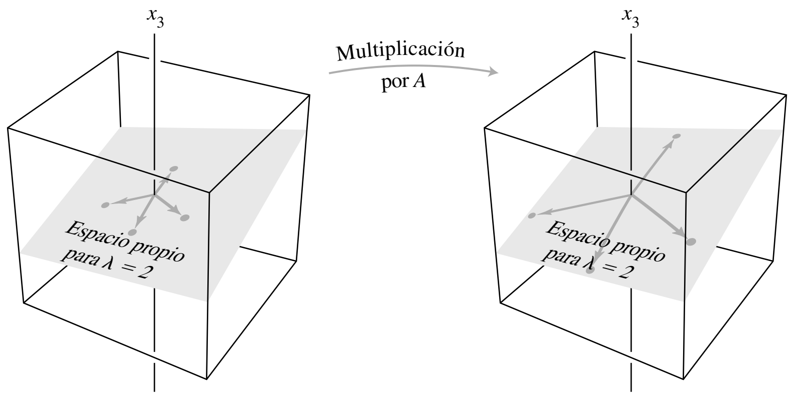


FIGURA 3 A actúa como una dilatación sobre el espacio propio.

NOTA NUMÉRICA

El ejemplo 4 es un buen método para el cálculo manual de vectores propios en casos simples donde se conoce un valor propio. En general, se recomienda el uso de un programa matricial y la reducción por filas para encontrar un espacio propio (para un valor propio dado), aunque no es totalmente confiable. Ocasionalmente, el error por redondeo lleva a una forma escalonada reducida con el número de pivotes equivocado. Los mejores programas computacionales calculan simultáneamente aproximaciones para los valores propios y los vectores propios, con el grado de exactitud que se requiera, en matrices que no son muy grandes. El tamaño de las matrices que se logra analizar se incrementa cada año conforme mejora la potencia computacional y la eficiencia del software.

El siguiente teorema describe uno de los pocos casos especiales donde los valores propios se pueden determinar con precisión. En la sección 5.2 también se analizará el cálculo de valores propios.

TEOREMA 1 Los valores propios de una matriz triangular son las entradas sobre su diagonal principal.

DEMOSTRACIÓN Por sencillez, considere el caso 3 × 3. Si A es triangular superior, entonces, $A - \lambda I$ tiene la forma

$$\begin{aligned} A - \lambda I &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} - \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \\ &= \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{bmatrix} \end{aligned}$$

El escalar λ es un valor propio de A si y solo si la ecuación $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial, es decir, si y solo si la ecuación tiene una variable libre. Gracias a las entradas cero en $A - \lambda I$, es fácil ver que $(A - \lambda I)\mathbf{x} = \mathbf{0}$ tiene una variable libre si y solo si al menos una de las entradas sobre la diagonal de $A - \lambda I$ es cero. Esto ocurre si y solo si λ es igual a una de las entradas a_{11}, a_{22}, a_{33} en A. Véase el ejercicio 28 para el caso donde A es triangular inferior. ■

EJEMPLO 5 Sean $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. Los valores propios de A son 3, 0 y 2. Los valores propios de B son 4 y 1. ■

¿Qué significa para una matriz A tener un valor propio de 0, tal como en el ejemplo 5? Esto ocurre si y solo si la ecuación

$$A\mathbf{x} = 0\mathbf{x} \tag{4}$$

tiene una solución no trivial. Pero la ecuación (4) es equivalente a $A\mathbf{x} = \mathbf{0}$, que tiene una solución no trivial si y solo si A no es invertible. Entonces, 0 *es un valor propio de A si y solo si A no es invertible.* En la sección 5.2 este hecho se agregará al teorema de la matriz invertible.

El siguiente teorema importante se necesitará más adelante. Su demostración muestra un cálculo típico con vectores propios.

TEOREMA 2

Si $\mathbf{v}_1,\ldots, \mathbf{v}_r$ son vectores propios que corresponden a distintos valores propios $\lambda_1,\ldots, \lambda_r$ de una matriz A de $n \times n$, entonces el conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_r\}$ es linealmente independiente.

DEMOSTRACIÓN Suponga que $\{\mathbf{v}_1,\ldots, \mathbf{v}_r\}$ es linealmente dependiente. Puesto que $\mathbf{v}_1$ es diferente de cero, el teorema 7 de la sección 1.7 indica que uno de los vectores en el conjunto es una combinación lineal de los vectores precedentes. Sea p el índice mínimo tal que $\mathbf{v}_{p+1}$ sea una combinación lineal de los vectores precedentes (linealmente independientes). Entonces, existen escalares $c_1,\ldots, c_p$ tales que

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{v}_{p+1} \tag{5}$$

Multiplicando por A ambos miembros de la ecuación (5) y considerando el hecho de que $A\mathbf{v}_k = \lambda_k\mathbf{v}_k$ para cada k,

$$\begin{aligned} c_1A\mathbf{v}_1 + \cdots + c_pA\mathbf{v}_p &= A\mathbf{v}_{p+1} \\ c_1\lambda_1\mathbf{v}_1 + \cdots + c_p\lambda_p\mathbf{v}_p &= \lambda_{p+1}\mathbf{v}_{p+1} \end{aligned} \tag{6}$$

Multiplicando por λ_{p+1} ambos lados de la ecuación (5) y restando el resultado de (6),

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \cdots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0} \tag{7}$$

Puesto que $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es linealmente independiente, todos los pesos en la ecuación (7) son iguales a cero. Sin embargo, ninguno de los factores $\lambda_i - \lambda_{p+1}$ son cero, ya que los valores propios son distintos. Por lo tanto, $c_i = 0$ para $i = 1,\ldots, p$. No obstante, la ecuación (5) indica que $\mathbf{v}_{p+1} = \mathbf{0}$, lo cual es imposible. En consecuencia, $\{\mathbf{v}_1,\ldots, \mathbf{v}_r\}$ no puede ser linealmente dependiente y, por consiguiente, debe ser linealmente independiente. ■

Vectores propios y ecuaciones en diferencias

Esta sección concluye con la demostración cómo construir soluciones de la ecuación en diferencias de primer orden analizada en el ejemplo introductorio del capítulo:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad (k = 0, 1, 2, \ldots) \tag{8}$$

Si A es una matriz de $n \times n$, entonces la ecuación (8) es una descripción *recursiva* de una secuencia $\{\mathbf{x}_k\}$ en $\mathbb{R}^n$. Una **solución** de la ecuación (8) es una descripción explícita de $\{\mathbf{x}_k\}$, cuya fórmula para cada $\mathbf{x}_k$ no depende directamente de A ni de los términos precedentes en la secuencia, excepto del término inicial $\mathbf{x}_0$.

La forma más sencilla de construir una solución de la ecuación (8) es tomar un vector propio $\mathbf{x}_0$ y su correspondiente valor propio λ, y hacer

$$\mathbf{x}_k = \lambda^k \mathbf{x}_0 \quad (k = 1, 2, \ldots) \tag{9}$$

Esta secuencia es una solución, ya que

$$A\mathbf{x}_k = A(\lambda^k \mathbf{x}_0) = \lambda^k (A\mathbf{x}_0) = \lambda^k (\lambda \mathbf{x}_0) = \lambda^{k+1}\mathbf{x}_0 = \mathbf{x}_{k+1}$$

Combinaciones lineales de soluciones en la forma de la ecuación (9) ¡también son soluciones! Véase el ejercicio 33.

PROBLEMAS DE PRÁCTICA

1. ¿Es 5 un valor propio de $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?
2. Si $\mathbf{x}$ es un vector propio de A correspondiente a λ, ¿qué ocurre con $A^3\mathbf{x}$?
3. Suponga que $\mathbf{b}_1$ y $\mathbf{b}_2$ son vectores propios correspondientes a distintos valores propios λ_1 y λ_2, respectivamente, y también que $\mathbf{b}_3$ y $\mathbf{b}_4$ son vectores propios linealmente independientes asociados a un tercer valor propio λ_3 distinto. ¿Necesariamente se deduce que $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ es un conjunto linealmente independiente? [*Sugerencia:* Considere la ecuación $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$].

5.1 EJERCICIOS

1. ¿Es $\lambda = 2$ un valor propio de $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? ¿Por qué?
2. ¿Es $\lambda = -3$ un valor propio de $\begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$? ¿Por qué?
3. ¿Es $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 1 & -1 \\ 6 & -4 \end{bmatrix}$? Si lo es, encuentre el valor propio.
4. ¿Es $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 5 & 2 \\ 3 & 6 \end{bmatrix}$? Si así es, determine el valor propio.
5. ¿Es $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$ un vector propio de $\begin{bmatrix} -4 & 3 & 3 \\ 2 & -3 & -2 \\ -1 & 0 & -2 \end{bmatrix}$? En caso afirmativo, determine el valor propio.
6. ¿Es $\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$ un vector propio de $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 2 & 7 \\ 5 & 6 & 4 \end{bmatrix}$? Si lo es, encuentre el valor propio.
7. ¿Es $\lambda = 4$ un valor propio de $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? Si es así, determine un vector propio correspondiente.
8. ¿Es $\lambda = 1$ un valor propio de $\begin{bmatrix} 4 & -2 & 3 \\ 0 & -1 & 3 \\ -1 & 2 & -2 \end{bmatrix}$? Si es así, determine un vector propio correspondiente.

En los ejercicios 9 a 16, determine una base para el espacio propio asociado con cada valor propio indicado.

9. $A = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}$, $\lambda = 1, 3$

10. $A = \begin{bmatrix} -4 & 2 \\ 3 & 1 \end{bmatrix}, \lambda = -5$

11. $A = \begin{bmatrix} 1 & -3 \\ -4 & 5 \end{bmatrix}, \lambda = -1, 7$

12. $A = \begin{bmatrix} 4 & 1 \\ 3 & 6 \end{bmatrix}, \lambda = 3, 7$

13. $A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \lambda = 1, 2, 3$

14. $A = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 0 & 3 \\ 2 & -2 & 5 \end{bmatrix}, \lambda = 3$

15. $A = \begin{bmatrix} -4 & 1 & 1 \\ 2 & -3 & 2 \\ 3 & 3 & -2 \end{bmatrix}, \lambda = -5$

16. $A = \begin{bmatrix} 5 & 0 & -1 & 0 \\ 1 & 3 & 0 & 0 \\ 2 & -1 & 3 & 0 \\ 4 & -2 & -2 & 4 \end{bmatrix}, \lambda = 4$

En los ejercicios 17 y 18, determine los valores propios de las matrices.

17. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & -2 \end{bmatrix}$ **18.** $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{bmatrix}$

19. Para $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, encuentre un valor propio, sin realizar cálculos. Justifique su respuesta.

20. Sin hacer cálculos, obtenga un valor propio y dos vectores propios linealmente independientes de $A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$. Justifique su respuesta.

En los ejercicios 21 y 22, A es una matriz de $n \times n$. Marque cada enunciado como verdadero o falso. Justifique su respuesta.

21. *a*) Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$, entonces λ es un valor propio de A.

b) Una matriz A no es invertible si y solo si 0 es un valor propio de A.

c) Un número c es un valor propio de A si y solo si la ecuación $(A - cI)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial.

d) Quizá sea difícil obtener un vector propio de A, pero es sencillo comprobar si un vector dado es, de hecho, un vector propio.

e) Para determinar los valores propios de A, se reduce A a una forma escalonada.

22. *a*) Si $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ, entonces $\mathbf{x}$ es un vector propio de A.

b) Si $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios linealmente independientes, entonces corresponden a distintos valores propios.

c) Un vector de estado estable para una matriz estocástica es realmente un vector propio.

d) Los valores propios de una matriz están sobre su diagonal principal.

e) Un espacio propio de A es el espacio nulo de cierta matriz.

23. Explique por qué una matriz de 2×2 puede tener, cuando mucho, dos valores propios distintos. También indique por qué una matriz de $n \times n$ puede tener, cuando mucho, n valores propios diferentes.

24. Construya un ejemplo de una matriz de 2×2 con tan solo un valor propio distinto.

25. Sea λ un valor propio de una matriz A invertible. Demuestre que λ^{-1} es un valor propio de A^{-1}. [*Sugerencia:* Suponga que $\mathbf{x}$ distinto de cero satisface $A\mathbf{x} = \lambda\mathbf{x}$].

26. Demuestre que si A^2 es la matriz cero, entonces el único valor propio de A es 0.

27. Demuestre que λ es un valor propio de A si y solo si λ es un valor propio de A^T. [*Sugerencia:* Determine cómo se relaciona $A - \lambda I$ con $A^T - \lambda I$].

28. Utilice el ejercicio 27 para completar la demostración del teorema 1 para el caso en que A es triangular inferior.

29. Considere una matriz A de $n \times n$ con la propiedad de que la suma de las filas sea igual al mismo número s. Demuestre que s es un valor propio de A. [*Sugerencia:* Encuentre un vector propio].

30. Considere una matriz A de $n \times n$ tal que las sumas de columnas sean iguales al mismo número s. Demuestre que s es un valor propio de A. [*Sugerencia:* Utilice los ejercicios 27 y 29].

En los ejercicios 31 y 32, sea A la matriz de la transformación lineal T. Sin escribir A, encuentre un valor propio de A y describa el espacio propio.

31. T es la transformación sobre $\mathbb{R}^2$ que refleja los puntos con respecto a una recta que pasa por el origen.

32. T es la transformación en $\mathbb{R}^3$ que gira los puntos alrededor de alguna recta que pasa por el origen.

33. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores propios de una matriz A, con valores propios correspondientes λ y μ, y sean c_1 y c_2 escalares. Defina:

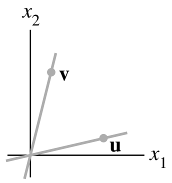
$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \ldots)$$

a) ¿Qué es $\mathbf{x}_{k+1}$ por definición?

b) Calcule $A\mathbf{x}_k$ de la fórmula para $\mathbf{x}_k$ y demuestre que $A\mathbf{x}_k = \mathbf{x}_{k+1}$. Este cálculo probará que la secuencia $\{\mathbf{x}_k\}$ que se acaba de definir satisface la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2,\ldots$).

34. Describa cómo intentaría construir una solución de una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0,1,2,\ldots$), si le dieran $\mathbf{x}_0$ inicial y resultara que este no fuera un vector propio de A. [*Sugerencia:* ¿Cómo podría relacionar $\mathbf{x}_0$ con los vectores propios de A?]

35. Sean $\mathbf{u}$ y $\mathbf{v}$ los vectores que se muestran en la figura, y suponga que $\mathbf{u}$ y $\mathbf{v}$ son vectores propios de una matriz A de 2×2 correspondientes a los valores propios 2 y 3, respectivamente. Sea $T : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación lineal dada por $T(\mathbf{x}) = A\mathbf{x}$ para cada $\mathbf{x}$ en $\mathbb{R}^2$, y sea $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Haga una copia de la figura, y sobre el mismo sistema de coordenadas grafique cuidadosamente los vectores $T(\mathbf{u})$, $T(\mathbf{v})$ y $T(\mathbf{w})$.



36. Repita el ejercicio 35, suponiendo que $\mathbf{u}$ y $\mathbf{v}$ son vectores propios de A correspondientes a los valores propios -1 y 3, respectivamente.

[**M**] En los ejercicios 37 a 40, utilice un programa matricial para determinar los valores propios de la matriz. Después, aplique el método del ejemplo 4 con una rutina de reducción por filas para producir una base para cada espacio propio.

37. $\begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix}$

38. $\begin{bmatrix} 5 & -2 & 2 & -4 \\ 7 & -4 & 2 & -4 \\ 4 & -4 & 2 & 0 \\ 3 & -1 & 1 & -3 \end{bmatrix}$

39. $\begin{bmatrix} 12 & -90 & 30 & 30 & 30 \\ 8 & -49 & 15 & 15 & 15 \\ 16 & -52 & 12 & 0 & 20 \\ 0 & -30 & 10 & 22 & 10 \\ 8 & -41 & 15 & 15 & 7 \end{bmatrix}$

40. $\begin{bmatrix} -23 & 57 & -9 & -15 & -59 \\ -10 & 12 & -10 & 2 & -22 \\ 11 & 5 & -3 & -19 & -15 \\ -27 & 31 & -27 & 25 & -37 \\ -5 & -15 & -5 & 1 & 31 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El número 5 es un valor propio de A si y solo si la ecuación $(A - 5I)\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Forme

$$A - 5I = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 1 \\ 3 & -5 & 5 \\ 2 & 2 & 1 \end{bmatrix}$$

y reduzca por filas la matriz aumentada:

$$\begin{bmatrix} 1 & -3 & 1 & 0 \\ 3 & -5 & 5 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 8 & -1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 1 & 0 \\ 0 & 4 & 2 & 0 \\ 0 & 0 & -5 & 0 \end{bmatrix}$$

En este punto, es evidente que el sistema homogéneo no tiene variables libres. Entonces, $A - 5I$ es una matriz invertible, lo cual significa que 5 *no* es un valor propio de A.

2. Si $\mathbf{x}$ es un vector propio de A correspondiente a λ, entonces $A\mathbf{x} = \lambda\mathbf{x}$ y

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Otra vez, $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. El patrón general, $A^k\mathbf{x} = \lambda^k\mathbf{x}$, se demuestra por inducción.

3. Sí. Suponga que $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 = \mathbf{0}$. Puesto que cualquier combinación lineal de vectores propios del mismo valor propio es otra vez un vector propio para ese valor propio, $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ es un vector propio para λ_3. De acuerdo con el teorema 2, los vectores $\mathbf{b}_1$, $\mathbf{b}_2$ y $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ son linealmente independientes, de manera que

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$$

implica que $c_1 = c_2 = 0$. Pero, c_3 y c_4 también deben ser cero porque $\mathbf{b}_3$ y $\mathbf{b}_4$ son linealmente independientes. Por lo tanto, todos los coeficientes en la ecuación original deben ser iguales a cero, y los vectores $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ y $\mathbf{b}_4$ son linealmente independientes.

5.2 LA ECUACIÓN CARACTERÍSTICA

Información útil acerca de los valores propios de una matriz cuadrada A está codificada en una ecuación escalar especial llamada ecuación característica de A. Un simple ejemplo nos llevará al caso general.

EJEMPLO 1 Determine los valores propios de $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

SOLUCIÓN Deben obtenerse todos los escalares λ tales que la ecuación matricial

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

tenga una solución no trivial. De acuerdo con el teorema de la matriz invertible de la sección 2.3, este problema es equivalente a encontrar todas las λ tales que la matriz de $A - \lambda I$ *no* sea invertible, donde

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

De acuerdo con el teorema 4 de la sección 2.2, esta matriz no es invertible precisamente cuando su determinante es cero. De manera que los valores propios de A son las soluciones de la ecuación

$$\det(A - \lambda I) = \det \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$$

Recuerde que

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

Entonces,

$$\begin{aligned} \det(A - \lambda I) &= (2-\lambda)(-6-\lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &= \lambda^2 + 4\lambda - 21 \\ &= (\lambda - 3)(\lambda + 7) \end{aligned}$$

Si $\det(A - \lambda I) = 0$, entonces $\lambda = 3$, o bien, $\lambda = -7$. Por lo tanto, los valores propios de A son 3 y -7. ■

El determinante en el ejemplo 1 transformó la ecuación matricial $(A - \lambda I)\mathbf{x} = \mathbf{0}$, que implica *dos* incógnitas (λ y $\mathbf{x}$), en la ecuación escalar $\lambda^2 + 4\lambda - 21 = 0$, que tan solo implica una incógnita. La misma idea funciona para matrices de $n \times n$. Sin embargo, antes de pasar a matrices más grandes, se resumen las propiedades necesarias de los determinantes para estudiar valores propios.

Determinantes

Sean A una matriz de $n \times n$, y U una forma escalonada de A obtenida mediante remplazos e intercambios de fila (sin escalamiento), y sea r el número de tales intercambios de fila. Entonces, el **determinante** de A, que se escribe $\det A$, es $(-1)^r$ veces el producto de las entradas diagonales $u_{11}, \dots, u_{nn}$ en U. Si A es invertible, todas la entradas $u_{11}, \dots, u_{nn}$ son *pivotes* (porque $A \sim I_n$ y las u_{ii} no se escalaron a 1). Es decir, al menos u_{nn} es cero y el producto $u_{11} \cdots u_{nn}$ es cero. Así,[1]

$$\det A = \begin{cases} (-1)^r \cdot \begin{pmatrix} \text{producto de} \\ \text{pivotes en } U \end{pmatrix}, & \text{cuando } A \text{ es invertible} \\ 0, & \text{cuando } A \text{ no es invertible} \end{cases} \qquad (1)$$

[1] La fórmula (1) se dedujo en la sección 3.2. Los lectores que no hayan estudiado el capítulo 3 pueden emplear esta fórmula como la definición de det A. Es un hecho notable y no trivial que cualquier forma escalonada U obtenida de A sin escalar dé el mismo valor para det A.

EJEMPLO 2 Calcule det A para $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

SOLUCIÓN La siguiente reducción por filas utiliza un intercambio de filas:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & -6 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -1 \end{bmatrix} = U_1$$

De manera que det $A = (-1)^1(1)(-2)(-1) = -2$. La siguiente reducción por filas alternativa evita el intercambio de filas y produce una forma escalonada diferente. El último paso suma $-1/3$ por la fila 2 a la fila 3:

$$A \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & -2 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & 0 \\ 0 & -6 & -1 \\ 0 & 0 & 1/3 \end{bmatrix} = U_2$$

Esta vez, det A es $(-1)^0(1)(-6)(1/3) = -2$, igual que antes. ■

La fórmula (1) para el determinante muestra que A es invertible si y solo si det A es distinto de cero. Este hecho, y la caracterización de invertibilidad establecida en la sección 5.1, se agregan al teorema de la matriz invertible.

TEOREMA

El teorema de la matriz invertible (continuación)

Sea A una matriz de $n \times n$. Entonces, A es invertible si y solo si:

s) El número 0 *no es* un valor propio de A.

t) El determinante de A *no es* cero.

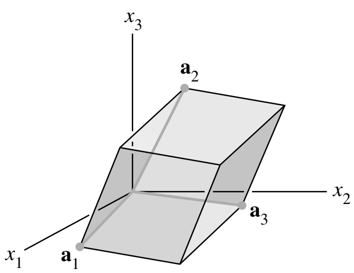


FIGURA 1

Cuando A es una matriz de 3×3, $|\det A|$ resulta ser el *volumen* del paralelepípedo definido por las columnas $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ de A, como se muestra en la figura 1. (Para más detalles, véase la sección 3.3). Este volumen es *distinto de cero* si y solo si los vectores $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ son linealmente independientes, en cuyo caso la matriz A es invertible. (Si los vectores son distintos de cero y linealmente dependientes, entonces se encuentran sobre un plano o sobre una recta).

El siguiente teorema lista los resultados, que se necesitarán, de las secciones 3.1 y 3.2. Aquí se incluye el inciso *a*) como una referencia conveniente.

TEOREMA 3

Propiedades de los determinantes

Sean A y B matrices de $n \times n$.

a) A es invertible si y solo si det $A \neq 0$.

b) $\det(AB) = (\det A)(\det B)$.

c) $\det A^T = \det A$.

d) Si A es triangular, entonces det A es el producto de las entradas sobre la diagonal principal de A.

e) La operación de remplazo de filas sobre A no cambia el determinante. Un intercambio de filas cambia el signo del determinante. Un escalamiento de filas también escala al determinante por el mismo factor de escala.

La ecuación característica

El teorema 3*a*) demuestra cómo determinar cuando una matriz de la forma $A - \lambda I$ *no* es invertible. La ecuación escalar $\det(A - \lambda I) = 0$ es la **ecuación característica** de A, y el argumento en el ejemplo 1 justifica el siguiente hecho.

> Un escalar λ es un valor propio de una matriz A de $n \times n$ si y solo si λ satisface la ecuación característica
>
> $$\det(A - \lambda I) = 0$$

EJEMPLO 3 Encuentre la ecuación característica de

$$A = \begin{bmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

SOLUCIÓN Forme $A - \lambda I$ y utilice el teorema 3*d*):

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix} \\ &= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda) \end{aligned}$$

La ecuación característica es

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0$$

o bien,

$$(\lambda - 5)^2(\lambda - 3)(\lambda - 1) = 0$$

Desarrollando el producto, se escribe

$$\lambda^4 - 14\lambda^3 + 68\lambda^2 - 130\lambda + 75 = 0$$ ■

En los ejemplos 1 y 3, $\det(A - \lambda I)$ es un polinomio en λ. Se puede demostrar que si A es una matriz de $n \times n$, entonces $\det(A - \lambda I)$ es un polinomio de grado n llamado **polinomio característico** de A.

El valor propio 5, en el ejemplo 3, tiene *multiplicidad* 2 porque $(\lambda - 5)$ ocurre dos veces como factor del polinomio característico. En general, la **multiplicidad** (**algebraica**) de un valor propio λ es su multiplicidad como una raíz de la ecuación característica.

EJEMPLO 4 El polinomio característico de una matriz de 6×6 es $\lambda^6 - 4\lambda^5 - 12\lambda^4$. Determine los valores propios y sus multiplicidades.

SOLUCIÓN Factorice el polinomio

$$\lambda^6 - 4\lambda^5 - 12\lambda^4 = \lambda^4(\lambda^2 - 4\lambda - 12) = \lambda^4(\lambda - 6)(\lambda + 2)$$

Los valores propios son 0 (multiplicidad 4), 6 (multiplicidad 1) y -2 (multiplicidad 1). ■

Se podrían listar los valores propios del ejemplo 4 como 0, 0, 0, 0, 6 y -2, de modo que los valores propios están repetidos de acuerdo con sus multiplicidades.

Puesto que la ecuación característica para una matriz de $n \times n$ implica un polinomio de grado n, la ecuación tiene exactamente n raíces, contando las multiplicidades, si se permiten raíces complejas. En la sección 5.5 se analizarán dichas raíces complejas, llamadas *valores propios complejos*. Hasta entonces, solo se considerarán valores propios reales, y los escalares continuarán siendo números reales.

La ecuación característica es importante para fines teóricos. Sin embargo, en el trabajo práctico, los valores propios de una matriz más grande que 2×2 deberían encontrarse usando una computadora, a menos que la matriz sea triangular o tenga otras propiedades especiales. No obstante que un polinomio característico 3×3 es fácil de calcular a mano, quizá sea difícil factorizarlo (a menos que la matriz sea cuidadosamente seleccionada). Véase las "Notas numéricas" del final de esta sección.

Similitud

El siguiente teorema ilustra un uso del polinomio característico, y brinda un fundamento para varios métodos iterativos que *aproximan* valores propios. Si A y B son matrices de $n \times n$, entonces A **es similar a** B si hay una matriz invertible P tal que $P^{-1}AP = B$, o bien, de manera equivalente, $A = PBP^{-1}$. Escribiendo Q para P^{-1}, se tiene $Q^{-1}BQ = A$. Así, B es también similar a A, y se dice simplemente que A y B **son similares**. Cambiar A por $P^{-1}AP$ es una **transformación de similitud**.

TEOREMA 4 Si las matrices A y B de $n \times n$ son similares, entonces tienen el mismo polinomio característico y, por lo tanto, los mismos valores propios (con las mismas multiplicidades).

DEMOSTRACIÓN Si $B = P^{-1}AP$, entonces,

$$B - \lambda I = P^{-1}AP - \lambda P^{-1}P = P^{-1}(AP - \lambda P) = P^{-1}(A - \lambda I)P$$

Utilizando la propiedad multiplicativa *b*) del teorema 3, se calcula

$$\begin{aligned} \det(B - \lambda I) &= \det[P^{-1}(A - \lambda I)P] \\ &= \det(P^{-1}) \cdot \det(A - \lambda I) \cdot \det(P) \end{aligned} \tag{2}$$

Puesto que $\det(P^{-1}) \cdot \det(P) = \det(P^{-1}P) = \det I = 1$, de la ecuación (2) se observa que $\det(B - \lambda I) = \det(A - \lambda I)$. ■

ADVERTENCIAS:

1. Las matrices

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$

no son similares a pesar de tener los mismos valores propios.

2. Similitud no es lo mismo que equivalencia por filas. (Si A es equivalente por filas a B, entonces $B = EA$ para alguna matriz E invertible). En general, las operaciones con filas sobre una matriz cambian sus valores propios.

Aplicación a sistemas dinámicos

Los valores propios y vectores propios tienen la clave para la evolución discreta de un sistema dinámico, como se mencionó en la introducción al capítulo.

EJEMPLO 5 Sea $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analice el comportamiento a largo plazo del sistema dinámico definido por $\mathbf{x}_{k+1} = A\mathbf{x}_k$ ($k = 0, 1, 2, \ldots$), con $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

SOLUCIÓN El primer paso consiste en determinar los valores propios de A y una base para cada espacio propio. La ecuación característica de A es

$$\begin{aligned} 0 &= \det \begin{bmatrix} .95 - \lambda & .03 \\ .05 & .97 - \lambda \end{bmatrix} = (.95 - \lambda)(.97 - \lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

Por la fórmula cuadrática,

$$\begin{aligned} \lambda &= \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} \\ &= \frac{1.92 \pm .08}{2} = 1 \quad \text{o bien} \quad .92 \end{aligned}$$

Es rápido comprobar que los vectores propios correspondientes a $\lambda = 1$ y $\lambda = 0.92$ son múltiplos de

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectivamente.

El siguiente paso es escribir $\mathbf{x}_0$ dada, en términos de $\mathbf{v}_1$ y $\mathbf{v}_2$. Esto se puede hacer porque $\{\mathbf{v}_1, \mathbf{v}_2\}$ evidentemente es una base para $\mathbb{R}^2$. (¿Por qué?) De manera que existen pesos c_1 y c_2 tales que

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

De hecho,

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= [\,\mathbf{v}_1 \quad \mathbf{v}_2\,]^{-1}\mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1} \begin{bmatrix} .60 \\ .40 \end{bmatrix} \\ &= \frac{1}{-8}\begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix}\begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned} \tag{4}$$

Cada $\mathbf{x}_k$ es fácil de calcular, ya que $\mathbf{v}_1$ y $\mathbf{v}_2$ en la ecuación (3) son vectores propios de A, con $A\mathbf{v}_1 = \mathbf{v}_1$ y $A\mathbf{v}_2 = .92\mathbf{v}_2$:

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1 A\mathbf{v}_1 + c_2 A\mathbf{v}_2 && \text{Utilizando linealidad de } \mathbf{x} \mapsto A\mathbf{x}, \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 && \mathbf{v}_1 \text{ y } \mathbf{v}_2 \text{ son vectores propios.} \\ \mathbf{x}_2 = A\mathbf{x}_1 &= c_1 A\mathbf{v}_1 + c_2(.92)A\mathbf{v}_2 \\ &= c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2 \end{aligned}$$

y así sucesivamente. En general,

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \ldots)$$

Utilizando c_1 y c_2 de la ecuación (4),

$$\mathbf{x}_k = .125\begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \ldots) \tag{5}$$

Esta fórmula explícita para $\mathbf{x}_k$ da la solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Conforme $k \to \infty$, $(.92)^k$ tiende a cero y $\mathbf{x}_k$ tiende a $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$. ■

Los cálculos del ejemplo 5 tienen una aplicación interesante a las cadenas de Markov analizadas en la sección 4.9. Quienes hayan leído esa sección reconocerán que la matriz A del ejemplo 5 anterior es la misma que la matriz M de migración de la sección 4.9, que $\mathbf{x}_0$ es la distribución de población inicial entre la ciudad y los suburbios, y que $\mathbf{x}_k$ representa la distribución de población después de k años.

El teorema 18 de la sección 4.9 estableció que para una matriz tal como A, la secuencia $\mathbf{x}_k$ tiende a un vector de estado estable. Ahora se sabe *por qué* las $\mathbf{x}_k$ se comportan de esa forma, al menos para la matriz de migración. El vector de estado estable es $.125\mathbf{v}_1$, un múltiplo del vector propio $\mathbf{v}_1$, y la fórmula (5) para $\mathbf{x}_k$ muestra precisamente por qué $\mathbf{x}_k \to .125\mathbf{v}_1$.

NOTAS NUMÉRICAS

1. El software computacional como Mathematica y Maple utilizan cálculos simbólicos para encontrar el polinomio característico de una matriz de tamaño moderado. Pero no hay una fórmula o un algoritmo finito para resolver la ecuación característica de una matriz general de $n \times n$ para $n \geq 5$.
2. Los mejores métodos numéricos para encontrar valores propios evitan totalmente el polinomio característico. En efecto, MATLAB determina el polinomio característico de una matriz A calculando primero los valores propios $\lambda_1,\ldots,\lambda_n$ de A y, luego, desarrollando el producto $(\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$.
3. Varios algoritmos comunes para estimar los valores propios de una matriz A se basan en el teorema 4. En los ejercicios se analiza el poderoso *algoritmo QR*. Otra técnica, llamada *método de Jacobi*, funciona cuando $A = A^T$ y calcula una secuencia de matrices de la forma

$$A_1 = A \quad \text{y} \quad A_{k+1} = P_k^{-1} A_k P_k \quad (k = 1, 2, \ldots)$$

Cada matriz en la secuencia es similar a A y así tiene los mismos valores propios que A. Las entradas no diagonales de A_{k+1} tienden a cero conforme aumenta k, y las entradas diagonales se van aproximando a los valores propios de A.

4. En la sección 5.8 se analizan otros métodos para calcular valores propios.

PROBLEMA DE PRÁCTICA

Encuentre la ecuación característica y los valores propios de $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.2 EJERCICIOS

En los ejercicios 1 a 8, encuentre el polinomio característico y los valores propios reales de las matrices dadas.

1. $\begin{bmatrix} 2 & 7 \\ 7 & 2 \end{bmatrix}$ **2.** $\begin{bmatrix} -4 & -1 \\ 6 & 1 \end{bmatrix}$

3. $\begin{bmatrix} -4 & 2 \\ 6 & 7 \end{bmatrix}$ **4.** $\begin{bmatrix} 8 & 2 \\ 3 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}$ **6.** $\begin{bmatrix} 9 & -2 \\ 2 & 5 \end{bmatrix}$

7. $\begin{bmatrix} 5 & 3 \\ -4 & 4 \end{bmatrix}$ **8.** $\begin{bmatrix} -4 & 3 \\ 2 & 1 \end{bmatrix}$

Los ejercicios 9 a 14 requieren las técnicas de la sección 3.1. Determine el polinomio característico de cada matriz, utilizando un desarrollo por cofactores o la fórmula especial para determinantes de 3×3 descrita antes de los ejercicios 15 a 18 de la sección 3.1. [*Nota:* Mediante operaciones de filas no es sencillo encontrar el polinomio característico de una matriz de 3×3, porque está implicada la variable λ].

9. $\begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \\ 1 & 0 & 2 \end{bmatrix}$ **10.** $\begin{bmatrix} 3 & 1 & 1 \\ 0 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 0 & 0 \\ 2 & 1 & 4 \\ 1 & 0 & 4 \end{bmatrix}$ **12.** $\begin{bmatrix} -1 & 0 & 2 \\ 3 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 4 & 0 & -1 \\ -1 & 0 & 4 \\ 0 & 2 & 3 \end{bmatrix}$

Para las matrices de los ejercicios 15 a 17, liste los valores propios reales, repetidos de acuerdo con su multiplicidad.

15. $\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ **16.** $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 0 & 3 & 6 & 0 \\ 2 & 3 & 3 & -5 \end{bmatrix}$

17. $\begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ -5 & 1 & 0 & 0 & 0 \\ 3 & 8 & 0 & 0 & 0 \\ 0 & -7 & 2 & 1 & 0 \\ -4 & 1 & 9 & -2 & 3 \end{bmatrix}$

18. Se puede demostrar que la multiplicidad algebraica de un valor propio λ es siempre mayor o igual que la dimensión del espacio propio correspondiente a λ. Determine h en la matriz A de abajo, tal que el espacio propio para $\lambda = 4$ sea bidimensional:

$$A = \begin{bmatrix} 4 & 2 & 3 & 3 \\ 0 & 2 & h & 3 \\ 0 & 0 & 4 & 14 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

19. Sea A una matriz de $n \times n$ y suponga que A tiene n valores propios reales, $\lambda_1, \ldots, \lambda_n$, repetidos de acuerdo con sus multiplicidades, tal que

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda)$$

Explique por qué $\det A$ es el producto de los n valores propios de A. (Este resultado es válido para cualquier matriz cuadrada cuando se consideran valores propios complejos).

20. Utilice una propiedad de determinantes para demostrar que A y A^T tienen el mismo polinomio característico.

En los ejercicios 21 y 22, A y B son matrices de $n \times n$. Indique si cada enunciado es verdadero o falso. Justifique cada respuesta.

21. *a*) El determinante de A es el producto de las entradas diagonales en A.

b) Una operación de filas elemental sobre A no cambia el valor del determinante.

c) $(\det A)(\det B) = \det AB$

d) Si $\lambda + 5$ es un factor del polinomio característico de A, entonces 5 es un valor propio de A.

22. *a*) Si A es 3×3, con columnas $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, entonces $\det A$ es igual al volumen del paralelepípedo formado por $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

b) $\det A^T = (-1) \det A$.

c) La multiplicidad de una raíz r de la ecuación característica de A se llama multiplicidad algebraica de r como un valor propio de A.

d) Una operación de remplazo por filas sobre A no altera sus valores propios.

Un método ampliamente utilizado para estimar los valores propios de una matriz general A es el *algoritmo QR*. En condiciones adecuadas, este algoritmo produce una secuencia de matrices, todas similares a A, casi triangulares superiores, con entradas diagonales que aproximan los valores propios de A. La idea principal es factorizar A (u otra matriz similar a A) en la forma $A = Q_1R_1$, donde $Q_1{}^T = Q_1{}^{-1}$ y R_1 es triangular superior. Los factores se intercambian para formar $A_1 = R_1Q_1$, que se factoriza de nuevo como $A_1 = Q_2R_2$; entonces, se forma $A_2 = R_2Q_2$, y así sucesivamente. La similitud de $A, A_1, \ldots$ se deduce del resultado más general del ejercicio 23.

23. Demuestre que si $A = QR$ con Q invertible, entonces A es similar a $A_1 = RQ$.

24. Demuestre que si A y B son similares, entonces $\det A = \det B$.

25. Sean $A = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 5 de la sección 4.9].

a) Encuentre una base para $\mathbb{R}^2$ que consiste en $\mathbf{v}_1$ y otro vector propio $\mathbf{v}_2$ de A.

b) Compruebe que $\mathbf{x}_0$ se escribe en la forma $\mathbf{x}_0 = \mathbf{v}_1 + c\mathbf{v}_2$.

c) Para $k = 1,2,\ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$, y escriba una fórmula para $\mathbf{x}_k$. Después, demuestre que $\mathbf{x}_k \to \mathbf{v}_1$ conforme k aumenta.

26. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Utilice la fórmula (1) para un determinante (dada antes del ejemplo 2) para demostrar que $\det A = ad - bc$. Considere dos casos: $a \neq 0$ y $a = 0$.

27. Sean $A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$,

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

a) Demuestre que $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ son vectores propios de A. [*Nota:* A es la matriz estocástica estudiada en el ejemplo 3 de la sección 4.9].

b) Sea $\mathbf{x}_0$ cualquier vector en $\mathbb{R}^3$ con entradas no negativas cuya suma sea 1. (En la sección 4.9, a $\mathbf{x}_0$ se llamó vector de probabilidad). Explique por qué existen constantes c_1, c_2, c_3 tales que $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Calcule $\mathbf{w}^T\mathbf{x}_0$, y deduzca que $c_1 = 1$.

c) Para $k = 1, 2,\ldots$, defina $\mathbf{x}_k = A^k\mathbf{x}_0$, con $\mathbf{x}_0$ como en el inciso *b*). Demuestre que $\mathbf{x}_k \to \mathbf{v}_1$ cuando k aumenta.

28. [**M**] Construya una matriz A de 4×4, valuada en los enteros, y compruebe que A y A^T tienen el mismo polinomio característico (los mismos valores propios con las mismas multiplicidades). ¿A y A^T tienen los mismos vectores propios? Realice el mismo análisis con una matriz de 5×5. Escriba las matrices y sus conclusiones.

29. [**M**] Construya una matriz A de 4×4 valuada en los enteros.
   a) Reduzca A a una forma escalonada U sin emplear escalamiento por filas, y utilice U en la fórmula (1) (que está antes del ejemplo 2) para calcular det A. (Si A resulta ser singular, entonces inicie con una nueva matriz aleatoria).
   b) Determine los valores propios de A y el producto de esos valores propios (tan exactamente como sea posible).
   c) Escriba la matriz A y, con cuatro lugares decimales, liste los pivotes en U y los valores propios de A. Calcule det A con su programa matricial, y compárelo con los productos encontrados en los incisos *a*) y *b*).

30. [**M**] Sea $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. Para cada valor de a en el conjunto {32, 31.9, 31.8, 32.1, 32.2}, calcule el polinomio característico de A y los valores propios. En cada caso, trace una gráfica del polinomio característico $p(t) = \det(A - tI)$ para $0 \leq t \leq 3$. Si es posible, elabore todas las gráficas sobre un mismo sistema de coordenadas. Describa cómo las gráficas revelan los cambios en los valores propios conforme a cambia.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

La ecuación característica es

$$\begin{aligned} 0 = \det(A - \lambda I) &= \det \begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix} \\ &= (1-\lambda)(2-\lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18 \end{aligned}$$

De la fórmula cuadrática,

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} = \frac{3 \pm \sqrt{-63}}{2}$$

Es evidente que la ecuación característica no tiene soluciones reales, así que A no tiene valores propios reales. La matriz A está actuando sobre el espacio vectorial real $\mathbb{R}^2$, y ahí no existe un vector $\mathbf{v}$ distinto de cero, tal que $A\mathbf{v} = \lambda\mathbf{v}$ para algún escalar λ.

5.3 DIAGONALIZACIÓN

En muchos casos, la información valor propio-vector propio contenida en una matriz A se presenta en una factorización útil de la forma $A = PDP^{-1}$ donde D es una matriz diagonal. En esta sección, la factorización permite calcular rápidamente las potencias A^k para valores grandes de k: una idea fundamental en varias aplicaciones del álgebra lineal. Posteriormente, en las secciones 5.6 y 5.7, se utilizará la factorización para analizar (y *desacoplar*) sistemas dinámicos.

El siguiente ejemplo muestra que las potencias de una matriz diagonal son fáciles de calcular.

EJEMPLO 1 Si $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, entonces, $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$

y

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

En general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{para } k \geq 1$$

■

Si $A = PDP^{-1}$ para alguna P invertible con D diagonal, entonces A^k es también fácil de calcular, como lo demuestra el siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una fórmula para A^k, dado que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

SOLUCIÓN La fórmula estándar para la inversa de una matriz de 2×2 da

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Entonces, por asociatividad de la multiplicación matricial,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Nuevamente,

$$A^3 = (PDP^{-1})A^2 = (PD\underbrace{P^{-1})P}_{I}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

En general, para $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$ ■

Una matriz cuadrada A es **diagonalizable** si A es similar a una matriz diagonal, es decir, si $A = PDP^{-1}$ para alguna matriz P invertible y alguna matriz D diagonal. El siguiente teorema da una caracterización de matrices diagonalizables e indica cómo construir una factorización conveniente.

TEOREMA 5

Teorema de diagonalización

Una matriz A de $n \times n$ es diagonalizable, si y solo si A tiene n vectores propios linealmente independientes.

En efecto, $A = PDP^{-1}$, con D como matriz diagonal, si y solo si las columnas de P son n vectores propios linealmente independientes de A. En este caso, las entradas diagonales de D son valores propios de A que corresponden, respectivamente, a los vectores propios en P.

En otras palabras, A es diagonalizable si y solo si hay suficientes vectores propios para formar una base de $\mathbb{R}^n$. Tal base es una **base de vectores propios** de $\mathbb{R}^n$.

DEMOSTRACIÓN Primero, observe que si P es cualquier matriz de $n \times n$ con columnas $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si D es cualquier matriz diagonal con entradas diagonales $\lambda_1, \ldots, \lambda_n$, entonces,

$$AP = A[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,] = [\,A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n\,] \tag{1}$$

mientras que

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\,\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n\,] \tag{2}$$

Ahora suponga que A es diagonalizable y $A = PDP^{-1}$. Por lo tanto, esta relación se multiplica por la derecha por P y se obtiene $AP = PD$. En este caso, las ecuaciones (1) y (2) implican que

$$[\,A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n\,] = [\,\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n\,] \tag{3}$$

Igualando columnas,

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \quad \ldots, \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n \tag{4}$$

Puesto que P es invertible, sus columnas $\mathbf{v}_1,\ldots, \mathbf{v}_n$ deben ser linealmente independientes. Asimismo, como esas columnas no son cero, entonces las ecuaciones en (4) muestran que $\lambda_1,\ldots, \lambda_n$ son valores propios y $\mathbf{v}_1,\ldots, \mathbf{v}_n$ son los vectores propios correspondientes. Este argumento demuestra las partes "solo si" del primer y segundo enunciados, junto con el tercer enunciado, del teorema.

Por último, dados cualesquiera n vectores propios $\mathbf{v}_1,\ldots, \mathbf{v}_n$, úselos para construir las columnas de P y utilice los valores propios correspondientes $\lambda_1,\ldots, \lambda_n$ para formar D. Por las ecuaciones (1) a (3), $AP = PD$. Esto es válido sin condiciones sobre los vectores propios. Si, en efecto, los vectores propios son linealmente independientes, entonces P es invertible (según el teorema de la matriz invertible), y $AP = PD$ implica que $A = PDP^{-1}$. ■

Diagonalización de matrices

EJEMPLO 3 Diagonalice la siguiente matriz, si es posible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Es decir, encuentre una matriz P invertible y una matriz D diagonal tales que $A = PDP^{-1}$.

SOLUCIÓN Hay cuatro pasos para implementar la descripción en el teorema 5.

Paso 1. Determine los valores propios de A. Como se mencionó en la sección 5.2, la mecánica de este paso es adecuada para una computadora, cuando la matriz es mayor que 2×2. Para eliminar distracciones innecesarias, el texto generalmente dará información adicional de utilidad para este paso. En el presente caso, la ecuación característica implica un polinomio cúbico que se factoriza como:

$$\begin{aligned} 0 = \det(A - \lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$.

Paso 2. Encuentre tres vectores propios de A linealmente independientes. Se necesitan *tres* vectores porque A es una matriz de 3×3. Este es el paso crítico. Si falla, entonces el teorema 5 indica que A no se puede diagonalizar. El método de la sección 5.1 produce una base para cada espacio propio:

$$\text{Base para } \lambda = 1: \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Puede comprobar que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto linealmente independiente.

Paso 3. Construya P con los vectores del paso 2. No es importante el orden de los vectores. Utilizando el orden seleccionado en el paso 2, forme

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Paso 4. Construya D con los valores propios correspondientes. En este paso, resulta esencial que el orden de los valores propios coincida con el orden elegido para las columnas de P. Utilice dos veces el valor propio $\lambda = -2$, uno para cada vector propio correspondiente a $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Es buena idea comprobar que P y D realmente funcionen. Para evitar el cálculo de P^{-1}, simplemente compruebe que $AP = PD$. Esto es equivalente a $A = PDP^{-1}$ cuando P es invertible. (Sin embargo, ¡asegúrese que P sea invertible!) Calcule

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$ ■

EJEMPLO 4 Si es posible, diagonalice la siguiente matriz.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

SOLUCIÓN La ecuación característica de A resulta ser exactamente la misma que en el ejemplo 3:

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

Los valores propios son $\lambda = 1$ y $\lambda = -2$. No obstante, es fácil comprobar que cada espacio propio es tan solo unidimensional:

$$\text{Base para } \lambda = 1: \qquad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -2: \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

No existen más valores propios, en tanto que cada vector propio de A es un múltiplo de $\mathbf{v}_1$ o de $\mathbf{v}_2$. Por consiguiente, es imposible construir una base de $\mathbb{R}^3$ utilizando vectores propios de A. De acuerdo con el teorema 5, A *no es* diagonalizable. ■

El siguiente teorema proporciona una condición *suficiente* para que una matriz sea diagonalizable.

TEOREMA 6 Una matriz de $n \times n$ con n valores propios distintos es diagonalizable.

DEMOSTRACIÓN Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ los vectores propios correspondientes a los n valores propios diferentes de una matriz A. Entonces, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente, de acuerdo con el teorema 2 de la sección 5.1. Por lo tanto, A es diagonalizable, según el teorema 5. ■

No es *necesario* que una matriz de $n \times n$ tenga n valores propios distintos para ser diagonalizable. La matriz de 3×3 del ejemplo 3 es diagonalizable aun cuando tenga solamente dos valores propios diferentes.

EJEMPLO 5 Determine si la siguiente matriz es diagonalizable.

$$A = \begin{bmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{bmatrix}$$

SOLUCIÓN ¡Esto es fácil! Como la matriz es triangular, sus valores propios son evidentemente 5, 0 y -2. Al ser A una matriz de 3×3 con tres valores propios distintos, entonces A es diagonalizable. ■

Matrices cuyos valores propios no son diferentes

Si una matriz A de $n \times n$ tiene n valores propios distintos, con los vectores propios correspondientes $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y si $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, entonces P es automáticamente invertible porque sus columnas son linealmente independientes, según el teorema 2. Cuando A es diagonalizable, pero tiene menos de n valores propios diferentes, sigue siendo posible construir P en una forma que la hace automáticamente invertible, como lo demuestra el siguiente teorema.[1]

TEOREMA 7

Sea A una matriz de $n \times n$ cuyos distintos valores propios son $\lambda_1, \ldots, \lambda_p$.

a) Para $1 \leq k \leq p$, la dimensión del espacio propio para λ_k es menor o igual que la multiplicidad del valor propio λ_k.

b) La matriz A es diagonalizable si y solo si la suma de las dimensiones de los espacios propios es igual a n, y esto ocurre si y solo si: **i.** se factoriza completamente el polinomio característico en factores lineales y **ii.** la dimensión del espacio propio para cada λ_k es igual a la multiplicidad de λ_k.

c) Si A es diagonalizable y $\mathcal{B}_k$ es una base para el espacio propio asociado con λ_k para cada k, entonces la colección de vectores total en los conjuntos $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forma una base de vectores propios para $\mathbb{R}^n$.

EJEMPLO 6 Diagonalice la siguiente matriz, si es posible.

$$A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}$$

[1] La demostración del teorema 7 es larga, pero no difícil. Por ejemplo, véase S. Friedberg, A. Insel, y L. Spence, *Linear Algebra*, 4a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 2002), sección 5.2.

SOLUCIÓN Puesto que A es una matriz triangular, los valores propios son 5 y -3, cada uno con multiplicidad 2. Utilizando el método de la sección 5.1, se encuentra una base para cada espacio propio.

$$\text{Base para } \lambda = 5: \qquad \mathbf{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Base para } \lambda = -3: \qquad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

El conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es linealmente independiente, según el teorema 7. Así, la matriz $P = [\mathbf{v}_1 \cdots \mathbf{v}_4]$ es invertible, en tanto que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \qquad \blacksquare$$

PROBLEMAS DE PRÁCTICA

1. Calcule A^8, donde $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.
2. Sean $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suponga que le han dicho que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A. Utilice esta información para diagonalizar A.
3. Sea A una matriz de 4×4 con valores propios 5, 3 y -2, y suponga que se sabe que el espacio propio para $\lambda = 3$ es bidimensional. ¿Se dispone de suficiente información para determinar si A es diagonalizable?

5.3 EJERCICIOS

En los ejercicios 1 y 2, sea $A = PDP^{-1}$ y calcule A^4.

1. $P = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

2. $P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

En los ejercicios 3 y 4, use la factorización $A = PDP^{-1}$ para calcular A^k, donde k representa un entero positivo arbitrario.

3. $\begin{bmatrix} a & 0 \\ 2(a-b) & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

4. $\begin{bmatrix} 1 & -6 \\ 2 & -6 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -3 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ -2 & 3 \end{bmatrix}$

En los ejercicios 5 y 6, la matriz A se factoriza en la forma PDP^{-1}. Utilice el teorema de diagonalización para encontrar los valores propios de A y una base para cada espacio propio.

5. $A = \begin{bmatrix} 2 & -1 & -1 \\ 1 & 4 & 1 \\ -1 & -1 & 2 \end{bmatrix}$
$$= \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

6. $A = \begin{bmatrix} 3 & 0 & 0 \\ -3 & 4 & 9 \\ 0 & 0 & 3 \end{bmatrix}$
$$= \begin{bmatrix} 3 & 0 & -1 \\ 0 & 1 & -3 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ -3 & 1 & 9 \\ -1 & 0 & 3 \end{bmatrix}$$

En los ejercicios 7 a 20, diagonalice las matrices, si es posible. Para los ejercicios 11 a 16 y 18, se incluyen los valores propios reales debajo de la matriz.

7. $\begin{bmatrix} 1 & 0 \\ 6 & -1 \end{bmatrix}$

8. $\begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$

9. $\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$

11. $\begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 3 & 2 \end{bmatrix}$ $\lambda = -1,\ 5$

12. $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ $\lambda = 2,\ 5$

13. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$ $\lambda = 1,\ 5$

14. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ $\lambda = 2,\ 3$

15. $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ $\lambda = 0,\ 1$

16. $\begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -2 \\ 1 & 3 & 1 \end{bmatrix}$ $\lambda = 0$

17. $\begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix}$ $\lambda = -2, -1, 0$

19. $\begin{bmatrix} 5 & -3 & 0 & 9 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

20. $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$

En los ejercicios 21 y 22, A, B, P y D son matrices de $n \times n$. Indique si cada enunciado es verdadero o falso. Justifique cada respuesta. (Antes de intentar resolver estos ejercicios, estudie con cuidado los teoremas 5 y 6, y los ejemplos de esta sección).

21. *a*) A es diagonalizable si $A = PDP^{-1}$ para alguna matriz D y alguna matriz P invertible.

b) Si $\mathbb{R}^n$ tiene una base de vectores propios de A, entonces A es diagonalizable.

c) A es diagonalizable si y solo si A tiene n valores propios, contando multiplicidades.

d) Si A es diagonalizable, entonces A es invertible.

22. *a*) A es diagonalizable si A tiene n vectores propios.

b) Si A es diagonalizable, entonces A tiene n valores propios distintos.

c) Si $AP = PD$, con D diagonal, entonces las columnas distintas de cero de P deben ser vectores propios de A.

d) Si A es invertible, entonces A es diagonalizable.

23. A es una matriz de 5×5 con dos valores propios. Un espacio propio es tridimensional y el otro espacio propio es bidimensional. ¿A es diagonalizable? ¿Por qué?

24. A es una matriz de 3×3 con dos valores propios. Cada espacio propio es unidimensional. ¿A es diagonalizable? ¿Por qué?

25. A es una matriz de 4×4 con tres valores propios. Un espacio propio es unidimensional, y uno de los otros espacios propios es bidimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

26. A es una matriz de 7×7 con tres valores propios. Un espacio propio es bidimensional, y uno de los otros espacios propios es tridimensional. ¿Es posible que A *no* sea diagonalizable? Justifique su respuesta.

27. Demuestre que si A es diagonalizable e invertible, entonces también lo es A^{-1}.

28. Demuestre que si A tiene n vectores propios linealmente independientes, entonces ocurre lo mismo con A^T. [*Sugerencia:* Use el teorema de diagonalización].

29. Una factorización $A = PDP^{-1}$ no es única. Demuestre esto para la matriz A del ejemplo 2. Con $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, utilice la información del ejemplo 2 para encontrar una matriz P_1 tal que $A = P_1 D_1 P_1{}^{-1}$.

30. Con A y D como en el ejemplo 2, encuentre una P_2 invertible, distinta de la P del ejemplo 2, tal que $A = P_2 D P_2{}^{-1}$.

31. Construya una matriz de 2×2 distinta de cero que sea invertible, pero no diagonalizable.

32. Construya una matriz de 2×2 no diagonal que sea diagonalizable, pero no invertible.

[**M**] En los ejercicios 33 a 36, diagonalice las matrices. Utilice la instrucción de valores propios de su programa matricial para obtener los valores propios y, luego, calcule bases para los espacios propios como en la sección 5.1.

33. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

34. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$

35. $\begin{bmatrix} 13 & -12 & 9 & -15 & 9 \\ 6 & -5 & 9 & -15 & 9 \\ 6 & -12 & -5 & 6 & 9 \\ 6 & -12 & 9 & -8 & 9 \\ -6 & 12 & 12 & -6 & -2 \end{bmatrix}$

36. $\begin{bmatrix} 24 & -6 & 2 & 6 & 2 \\ 72 & 51 & 9 & -99 & 9 \\ 0 & -63 & 15 & 63 & 63 \\ 72 & 15 & 9 & -63 & 9 \\ 0 & 63 & 21 & -63 & -27 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. det $(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. Los valores propios son 2 y 1, y los vectores propios correspondientes son $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Ahora, forme

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \qquad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Como $A = PDP^{-1}$,

$$\begin{aligned} A^8 = PD^8P^{-1} &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

2. Calcule $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1$, y

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

De manera que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios para los valores propios 1 y 3, respectivamente. Entonces,

$$A = PDP^{-1} \quad \text{donde,} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

3. Sí, A es diagonalizable. Existe una base $\{\mathbf{v}_1, \mathbf{v}_2\}$ para el espacio propio asociado a $\lambda = 3$. Asimismo, habrá al menos un vector propio para $\lambda = 5$ y uno para $\lambda = -2$, denotados mediante $\mathbf{v}_3$ y $\mathbf{v}_4$. Entonces, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es linealmente independiente de acuerdo con el teorema 2 y el problema de práctica 3 de la sección 5.1. No pueden existir más vectores propios que sean linealmente independientes de $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, porque los vectores están en $\mathbb{R}^4$. Por consiguiente, los espacios propios para $\lambda = 5$ y $\lambda = -2$ son ambos unidimensionales. De acuerdo con el teorema 7*b*) se tiene que A es diagonalizable.

5.4 VECTORES PROPIOS Y TRANSFORMACIONES LINEALES

El objetivo de esta sección es entender la factorización matricial $A = PDP^{-1}$ como un enunciado sobre transformaciones lineales. Se verá que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es esencialmente lo mismo que el muy simple mapeo $\mathbf{u} \mapsto D\mathbf{u}$, cuando se estudia desde un enfoque adecuado. Una interpretación similar se aplicará a D y A, aun cuando D no sea una matriz diagonal.

De la sección 1.9, recuerde que cualquier transformación lineal T de $\mathbb{R}^n$ a $\mathbb{R}^m$ se implementa mediante la multiplicación por la izquierda por una matriz A, llamada *matriz estándar* de T. Ahora se necesita el mismo tipo de representación para cualquier transformación lineal entre dos espacios vectoriales de dimensión finita.

La matriz de una transformación lineal

Sean V un espacio vectorial n-dimensional, W un espacio vectorial m-dimensional y T cualquier transformación lineal de V a W. Para asociar una matriz con T, se eligen bases (ordenadas) $\mathcal{B}$ y $\mathcal{C}$ para V y W, respectivamente.

Dada cualquier $\mathbf{x}$ en V, el vector de coordenadas $[\mathbf{x}]_{\mathcal{B}}$ está en $\mathbb{R}^n$ y el vector de coordenadas de su imagen, $[T(\mathbf{x})]_{\mathcal{C}}$, está en $\mathbb{R}^m$, como se indica en la figura 1.

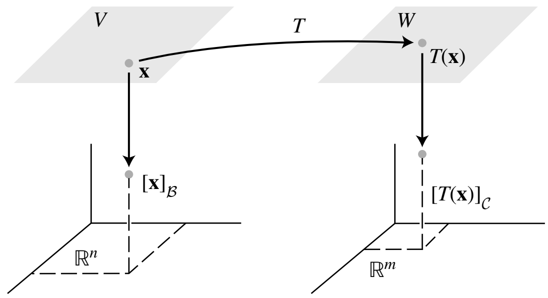


FIGURA 1 Una transformación lineal de V a W.

Es fácil encontrar la conexión entre $[\mathbf{x}]_{\mathcal{B}}$ y $[T(\mathbf{x})]_{\mathcal{C}}$. Sea $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ la base $\mathcal{B}$ para V. Si $\mathbf{x} = r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n$, entonces,

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

y

$$T(\mathbf{x}) = T(r_1\mathbf{b}_1 + \cdots + r_n\mathbf{b}_n) = r_1T(\mathbf{b}_1) + \cdots + r_nT(\mathbf{b}_n) \tag{1}$$

porque T es lineal. Ahora, como el mapeo coordenado de W a $\mathbb{R}^m$ es lineal (teorema 8 de la sección 4.4), la ecuación (1) conduce a

$$[\,T(\mathbf{x})\,]_{\mathcal{C}} = r_1[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} + \cdots + r_n[\,T(\mathbf{b}_n)\,]_{\mathcal{C}} \tag{2}$$

Puesto que los vectores de coordenadas C están en $\mathbb{R}^m$, así la ecuación vectorial (2) se escribe como una ecuación matricial; a saber,

$$[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}} \tag{3}$$

donde

$$M = \begin{bmatrix} [T(\mathbf{b}_1)]_{\mathcal{C}} & [T(\mathbf{b}_2)]_{\mathcal{C}} & \cdots & [T(\mathbf{b}_n)]_{\mathcal{C}} \end{bmatrix} \tag{4}$$

$$\begin{tikzcd} \mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\ {[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplicación}", "\text{por } M"'] & {[T(\mathbf{x})]_{\mathcal{C}}} \end{tikzcd}$$

FIGURA 2

La matriz M es una representación matricial de T, llamada **matriz para T respecto a las bases $\mathcal{B}$ y $\mathcal{C}$**. Véase la figura 2.

La ecuación (3) indica que, en lo concerniente a los vectores de coordenadas, la acción de T sobre $\mathbf{x}$ se puede considerar una multiplicación por la izquierda por M.

EJEMPLO 1 Suponga que $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ es una base para V y $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ es una base para W. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{b}_1) = 3\mathbf{c}_1 - 2\mathbf{c}_2 + 5\mathbf{c}_3 \quad \text{y} \quad T(\mathbf{b}_2) = 4\mathbf{c}_1 + 7\mathbf{c}_2 - \mathbf{c}_3$$

Determine la matriz M para T respecto a $\mathcal{B}$ y $\mathcal{C}$.

SOLUCIÓN Los vectores de coordenadas $\mathcal{C}$ de las *imágenes* de $\mathbf{b}_1$ y $\mathbf{b}_2$ son

$$[\,T(\mathbf{b}_1)\,]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \quad \text{y} \quad [\,T(\mathbf{b}_2)\,]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}$$

Por lo tanto,

$$M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}$$ ■

Si $\mathcal{B}$ y $\mathcal{C}$ son bases para el mismo espacio V y si T es la transformación identidad $T(\mathbf{x}) = \mathbf{x}$ para $\mathbf{x}$ en V, entonces la matriz M en la ecuación (4) es justo una matriz de cambio de coordenadas (véase la sección 4.7).

Transformaciones lineales de V en V

$$\begin{tikzcd} \mathbf{x} \arrow[r, "T"] \arrow[d] & T(\mathbf{x}) \arrow[d] \\ {[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplicación}", "\text{por } [T]_{\mathcal{B}}"'] & {[T(\mathbf{x})]_{\mathcal{B}}} \end{tikzcd}$$

FIGURA 3

En el caso común donde W es igual a V y la base $\mathcal{C}$ coincide con $\mathcal{B}$, la matriz M en (4) se denomina **matriz para T respecto a $\mathcal{B}$**, o simplemente **$\mathcal{B}$-matriz para T**, y se denota con $[T]_{\mathcal{B}}$. Véase la figura 3.

La $\mathcal{B}$-matriz para $T : V \to V$ satisface

$$[\,T(\mathbf{x})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}} \quad \text{para toda } \mathbf{x} \text{ en } V \tag{5}$$

EJEMPLO 2 El mapeo $T : \mathbb{P}_2 \to \mathbb{P}_2$ definido por

$$T(a_0 + a_1 t + a_2 t^2) = a_1 + 2a_2 t$$

es una transformación lineal. (Los estudiantes de cálculo reconocerán T como el operador derivada).

a) Determine la $\mathcal{B}$-matriz para T, cuando $\mathcal{B}$ es la base $\{1, t, t^2\}$.

b) Compruebe que $[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}}$ para toda $\mathbf{p}$ en $\mathbb{P}_2$.

SOLUCIÓN

a) Determine las imágenes de los vectores básicos:

$T(1) = 0$ El polinomio cero

$T(t) = 1$ El polinomio cuyo valor siempre es 1

$T(t^2) = 2t$

Después, escriba los vectores de $\mathcal{B}$-coordenadas de $T(1)$, $T(t)$ y $T(t^2)$ (que en este ejemplo se localizan por inspección) y júntelos en la $\mathcal{B}$-matriz para T:

$$[\,T(1)\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [\,T(t)\,]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [\,T(t^2)\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

b) Para un general $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$,

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,a_1 + 2a_2 t\,]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2a_2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [\,T\,]_{\mathcal{B}} [\,\mathbf{p}\,]_{\mathcal{B}}$$

Véase la figura 4. ■

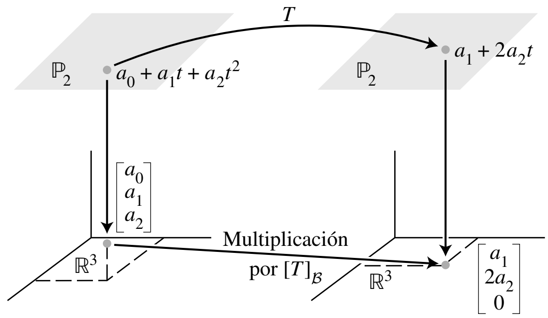


FIGURA 4 Representación matricial de una transformación lineal.

WEB

Transformaciones lineales sobre $\mathbb{R}^n$

En un problema aplicado que implica a $\mathbb{R}^n$, en general una transformación lineal T aparece primero como una transformación matricial, $\mathbf{x} \mapsto A\mathbf{x}$. Si A es diagonalizable, entonces hay una base $\mathcal{B}$ para $\mathbb{R}^n$ que consiste en vectores propios de A. A continuación, el teorema 8 demuestra que, en este caso, la $\mathcal{B}$-matriz para T es diagonal. Diagonalizar A equivale a encontrar una representación matricial diagonal $\mathbf{x} \mapsto A\mathbf{x}$.

TEOREMA 8 **Representación matricial diagonal**

Suponga que $A = PDP^{-1}$, donde D es una matriz de $n \times n$ diagonal. Si $\mathcal{B}$ es la base para $\mathbb{R}^n$ formada con las columnas de P, entonces D es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

DEMOSTRACIÓN Denote las columnas de P con $\mathbf{b}_1, \ldots, \mathbf{b}_n$, de manera que $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y $P = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$. En este caso, P es la matriz de cambio de coordenadas $P_{\mathcal{B}}$ analizada en la sección 4.4, donde

$$P[\,\mathbf{x}\,]_{\mathcal{B}} = \mathbf{x} \quad \text{y} \quad [\,\mathbf{x}\,]_{\mathcal{B}} = P^{-1}\mathbf{x}$$

Si $T(\mathbf{x}) = A\mathbf{x}$ para $\mathbf{x}$ en $\mathbb{R}^n$, entonces,

$$\begin{aligned}
[\,T\,]_{\mathcal{B}} &= \big[\,[\,T(\mathbf{b}_1)\,]_{\mathcal{B}} \quad \cdots \quad [\,T(\mathbf{b}_n)\,]_{\mathcal{B}}\,\big] && \text{Definición de } [T]_{\mathcal{B}} \\
&= \big[\,[\,A\mathbf{b}_1\,]_{\mathcal{B}} \quad \cdots \quad [\,A\mathbf{b}_n\,]_{\mathcal{B}}\,\big] && \text{Porque } T(\mathbf{x}) = A\mathbf{x} \\
&= [\,P^{-1}A\mathbf{b}_1 \quad \cdots \quad P^{-1}A\mathbf{b}_n\,] && \text{Cambio de coordenadas} \\
&= P^{-1}A[\,\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n\,] && \text{Multiplicación matricial} \\
&= P^{-1}AP && (6)
\end{aligned}$$

Puesto que $A = PDP^{-1}$, entonces se tiene $[\,T\,]_{\mathcal{B}} = P^{-1}AP = D$. ■

EJEMPLO 3 Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ por $T(\mathbf{x}) = A\mathbf{x}$, donde $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Encuentre una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que la $\mathcal{B}$-matriz para T es una matriz diagonal.

SOLUCIÓN Del ejemplo 2 de la sección 5.3, se conoce que $A = PDP^{-1}$, donde

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

Las columnas de P, denotadas mediante $\mathbf{b}_1$ y $\mathbf{b}_2$, son vectores propios de A. Según el teorema 8, D es la $\mathcal{B}$-matriz para T cuando $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Los mapeos $\mathbf{x} \mapsto A\mathbf{x}$ y $\mathbf{u} \mapsto D\mathbf{u}$ describen la misma transformación lineal respecto a diferentes bases. ■

Similitud de representaciones matriciales

La demostración del teorema 8 no utilizó la información de que D era diagonal. Por lo tanto, si A es similar a una matriz C, con $A = PCP^{-1}$, entonces C es la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ cuando la base $\mathcal{B}$ está formada por las columnas de P. La figura 5 ilustra la factorización $A = PCP^{-1}$.

$$\begin{tikzcd}
\mathbf{x} \arrow[r, "\text{Multiplicación}", "\text{por } A"'] \arrow[d, "\text{Multiplicación por } P^{-1}"'] & A\mathbf{x} \\
{[\mathbf{x}]_{\mathcal{B}}} \arrow[r, "\text{Multiplicación}", "\text{por } C"'] & {[A\mathbf{x}]_{\mathcal{B}}} \arrow[u, "\text{Multiplicación por } P"']
\end{tikzcd}$$

FIGURA 5 Similitud de dos representaciones matriciales: $A = PCP^{-1}$.

Inversamente, si $T : \mathbb{R}^n \to \mathbb{R}^n$ está definida por $T(\mathbf{x}) = A\mathbf{x}$, y si $\mathcal{B}$ es cualquier base para $\mathbb{R}^n$, entonces la $\mathcal{B}$-matriz para T es similar a A. En efecto, los cálculos en la demostración del teorema 8 indican que si P es la matriz cuyas columnas provienen de los vectores en $\mathcal{B}$, entonces, $[T]_{\mathcal{B}} = P^{-1}AP$. Así, el conjunto de todas las matrices similares a una matriz A coincide con el conjunto de todas las representaciones matriciales de la transformación $\mathbf{x} \mapsto A\mathbf{x}$.

EJEMPLO 4 Sean $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. El polinomio característico de A es $(\lambda + 2)^2$, sin embargo, el espacio propio para el valor propio -2 es unidimensional; de manera que A no es diagonalizable. Sin embargo, la base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ tiene la propiedad de que la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es una matriz triangular llamada *forma de Jordan* de A.[1] Determine esta $\mathcal{B}$-matriz.

SOLUCIÓN Si $P = [\mathbf{b}_1 \quad \mathbf{b}_2]$, entonces la $\mathcal{B}$-matriz es $P^{-1}AP$. Calcule

$$AP = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix}$$

$$P^{-1}AP = \begin{bmatrix} -1 & 2 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} -6 & -1 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix}$$

Observe que el valor propio de A está sobre la diagonal. ■

[1] Cada matriz cuadrada A es similar a una matriz en forma de Jordan. La base utilizada para generar una forma de Jordan consiste en vectores propios y los llamados “vectores propios generalizados” de A. Véase el capítulo 9 de *Applied Linear Algebra*, 3a. ed. (Englewood Cliffs, NJ: Prentice-Hall, 1988), de B. Noble y J. W. Daniel.

NOTA NUMÉRICA

Una forma eficiente de calcular una $\mathcal{B}$-matriz de $P^{-1}AP$ es determinando AP y, luego, reduciendo por filas la matriz aumentada $[P \quad AP]$ a $[I \quad P^{-1}AP]$. Es innecesario un cálculo por separado de P^{-1}. Véase el ejercicio 15 de la sección 2.2.

PROBLEMAS DE PRÁCTICA

1. Determine $T(a_0 + a_1t + a_2t^2)$, si T es la transformación lineal de $\mathbb{P}_2$ a $\mathbb{P}_2$ cuya matriz respecto a $\mathcal{B} = \{1, t, t^2\}$ es

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

2. Sean A, B y C matrices de $n \times n$. En el texto se ha demostrado que si A es similar a B, entonces B es similar a A. Esta propiedad, junto con los enunciados que se muestran debajo, indica que "similar a" es una *relación de equivalencia*. (La equivalencia por fila es otro ejemplo de una relación de equivalencia). Compruebe los incisos *a*) y *b*).

a) A es similar a A.

b) Si A es similar a B, y B es similar a C, entonces A es similar a C.

5.4 EJERCICIOS

1. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ y $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ las bases de los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{b}_1) = 3\mathbf{d}_1 - 5\mathbf{d}_2, \quad T(\mathbf{b}_2) = -\mathbf{d}_1 + 6\mathbf{d}_2, \quad T(\mathbf{b}_3) = 4\mathbf{d}_2$$

Determine la matriz para T respecto a $\mathcal{B}$ y $\mathcal{D}$.

2. Sean $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ las bases para los espacios vectoriales V y W, respectivamente. Sea $T : V \to W$ una transformación lineal con la propiedad

$$T(\mathbf{d}_1) = 3\mathbf{b}_1 - 3\mathbf{b}_2, \quad T(\mathbf{d}_2) = -2\mathbf{b}_1 + 5\mathbf{b}_2$$

Determine la matriz para T respecto a $\mathcal{D}$ y $\mathcal{B}$.

3. Sean $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base estándar para $\mathbb{R}^3$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V, y $T : \mathbb{R}^3 \to V$ una transformación lineal con la propiedad

$$T(x_1, x_2, x_3) = (2x_3 - x_2)\mathbf{b}_1 - (2x_2)\mathbf{b}_2 + (x_1 + 3x_3)\mathbf{b}_3$$

a) Calcule $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$ y $T(\mathbf{e}_3)$.

b) Obtenga $[T(\mathbf{e}_1)]_{\mathcal{B}}$, $[T(\mathbf{e}_2)]_{\mathcal{B}}$ y $[T(\mathbf{e}_3)]_{\mathcal{B}}$.

c) Determine la matriz para T respecto a $\mathcal{E}$ y $\mathcal{B}$.

4. Sean $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V y $T : V \to \mathbb{R}^2$ una transformación lineal con la propiedad

$$T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3) = \begin{bmatrix} 2x_1 - 3x_2 + x_3 \\ -2x_1 + 5x_3 \end{bmatrix}$$

Encuentre la matriz para T respecto a $\mathcal{B}$ y la base estándar para $\mathbb{R}^2$.

5. Sea $T : \mathbb{P}_2 \to \mathbb{P}_3$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $(t + 3)\mathbf{p}(t)$.

a) Encuentre la imagen de $\mathbf{p}(t) = 3 - 2t + t^2$.

b) Demuestre que T es una transformación lineal.

c) Obtenga la matriz para T respecto a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3\}$.

6. Sea $T : \mathbb{P}_2 \to \mathbb{P}_4$ la transformación que mapea un polinomio $\mathbf{p}(t)$ en el polinomio $\mathbf{p}(t) + 2t^2\mathbf{p}(t)$.

a) Determine la imagen de $\mathbf{p}(t) = 3 - 2t + t^2$.

b) Demuestre que T es una transformación lineal.

c) Encuentre la matriz para T respecto a las bases $\{1, t, t^2\}$ y $\{1, t, t^2, t^3, t^4\}$.

7. Suponga que el mapeo $T : \mathbb{P}_2 \to \mathbb{P}_2$ definido por

$$T(a_0 + a_1t + a_2t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$$

es lineal. Obtenga la representación matricial de T respecto a la base $\mathcal{B} = \{1, t, t^2\}$.

8. Sea $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ una base para un espacio vectorial V. Encuentre $T(4\mathbf{b}_1 - 3\mathbf{b}_2)$ cuando T es una transformación lineal de V a V cuya matriz respecto a $\mathcal{B}$ es

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$$

9. Defina $T : \mathbb{P}_2 \to \mathbb{R}^3$ por $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$.

a) Encuentre la imagen bajo T de $\mathbf{p}(t) = 5 + 3t$.

b) Pruebe que T es una transformación lineal.

c) Obtenga la matriz para T respecto a la base $\{1, t, t^2\}$ para $\mathbb{P}_2$ y la base estándar para $\mathbb{R}^3$.

10. Defina $T : \mathbb{P}_3 \to \mathbb{R}^4$ por $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-2) \\ \mathbf{p}(3) \\ \mathbf{p}(1) \\ \mathbf{p}(0) \end{bmatrix}$.

a) Demuestre que T es una transformación lineal.

b) Encuentre la matriz para T respecto a la base $\{1, t, t^2, t^3\}$ para $\mathbb{P}_3$ y la base estándar para $\mathbb{R}^4$.

En los ejercicios 11 y 12, determine la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ donde $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

11. $A = \begin{bmatrix} -4 & -1 \\ 6 & 1 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

12. $A = \begin{bmatrix} -6 & -2 \\ 4 & 0 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

En los ejercicios 13 a 16, defina $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ por $T(\mathbf{x}) = A\mathbf{x}$. Determina una base $\mathcal{B}$ para $\mathbb{R}^2$ con la propiedad de que $[T]_{\mathcal{B}}$ es diagonal.

13. $A = \begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$

14. $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$

16. $A = \begin{bmatrix} 4 & -2 \\ -1 & 5 \end{bmatrix}$

17. Sean $A = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$ y $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ para $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Defina $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que $T(\mathbf{x}) = A\mathbf{x}$.

a) Compruebe que $\mathbf{b}_1$ es un vector propio de A, pero que A no es diagonalizable.

b) Encuentre la $\mathcal{B}$-matriz para T.

18. Defina $T : \mathbb{R}^3 \to \mathbb{R}^3$ mediante $T(\mathbf{x}) = A\mathbf{x}$, donde A es una matriz de 3×3 con valores propios de 5, 5 y -2. ¿Existe una base $\mathcal{B}$ para $\mathbb{R}^3$ tal que la $\mathcal{B}$-matriz para T sea una matriz diagonal? Analice.

En los ejercicios 19 a 24, compruebe los enunciados. Las matrices son cuadradas.

19. Si A es invertible y similar a B, entonces B es invertible y A^{-1} es similar a B^{-1}. [*Sugerencia:* $P^{-1}AP = B$ para alguna P invertible. Explique por qué B es invertible. Luego, determine una Q invertible tal que $Q^{-1}A^{-1}Q = B^{-1}$].

20. Si A es similar a B, entonces A^2 es similar a B^2.

21. Si B es similar a A y C es similar a A, entonces B es similar a C.

22. Si A es diagonalizable y B es similar a A, entonces B también es diagonalizable.

23. Si $B = P^{-1}AP$ y $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ, entonces $P^{-1}\mathbf{x}$ es un vector propio de B también asociado con λ.

24. Si A y B son similares, entonces tienen el mismo rango (dimensión del espacio imagen). [*Sugerencia:* Consulte los ejercicios complementarios 13 y 14 del capítulo 4].

25. La *traza* de una matriz cuadrada A es la suma de las entradas diagonales en A y se denota con tr A. Se puede comprobar que $\operatorname{tr}(FG) = \operatorname{tr}(GF)$ para cualesquiera dos matrices F y G de $n \times n$. Demuestre que si A y B son similares, entonces $\operatorname{tr} A = \operatorname{tr} B$.

26. Se puede demostrar que la traza de una matriz A es igual a la suma de los valores propios de A. Compruebe este enunciado para el caso cuando A es diagonalizable.

27. Sean $V = \mathbb{R}^n$ con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, y $W = \mathbb{R}^n$ con la base estándar que se denota mediante $\mathcal{E}$; y considere la transformación identidad $I : \mathbb{R}^n \to \mathbb{R}^n$, donde $I(\mathbf{x}) = \mathbf{x}$. Obtenga la matriz para I respecto a $\mathcal{B}$ y $\mathcal{E}$. ¿Cómo se llamó a esta matriz en la sección 4.4?

28. Sean V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ y W el mismo espacio V con una base $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$; I es la transformación identidad $I : V \to W$. Encuentre la matriz para I respecto a $\mathcal{B}$ y $\mathcal{C}$. ¿Cómo se llamó a esta matriz en la sección 4.7?

29. Sea V un espacio vectorial con una base $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Determine la $\mathcal{B}$-matriz para la transformación identidad $I : V \to V$.

[M] En los ejercicios 30 y 31, determine la $\mathcal{B}$-matriz para la transformación $\mathbf{x} \mapsto A\mathbf{x}$ donde $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

30. $A = \begin{bmatrix} 6 & -2 & -2 \\ 3 & 1 & -2 \\ 2 & -2 & 2 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$

31. $A = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. **[M]** Sea T la transformación cuya matriz estándar se presenta a continuación. Determine una base para $\mathbb{R}^4$ con la propiedad de que $[T]_{\mathcal{B}}$ sea diagonal.

$$A = \begin{bmatrix} -6 & 4 & 0 & 9 \\ -3 & 0 & 1 & 6 \\ -1 & -2 & 1 & 0 \\ -4 & 4 & 0 & 7 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sea $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$, y calcule

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

Así, $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. *a*) $A = (I)^{-1}AI$, entonces A es similar a A.

b) Por hipótesis, existen matrices invertibles P y Q con la propiedad de que $B = P^{-1}AP$ y $C = Q^{-1}BQ$. Sustituya la fórmula para B en la expresión para C, y utilice un resultado acerca de la inversa de un producto:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

Esta ecuación tiene la forma adecuada para demostrar que A es similar a C.

5.5 VALORES PROPIOS COMPLEJOS

Puesto que la ecuación característica de una matriz de $n \times n$ implica un polinomio de grado n, la ecuación siempre tiene exactamente n raíces, contando las multiplicidades, *siempre y cuando se admitan raíces complejas.* Esta sección demuestra que si la ecuación característica de una matriz real A tiene algunas raíces complejas, entonces esas raíces aportan información importante acerca de A. La clave es dejar que A actúe sobre el espacio $\mathbb{C}^n$ de n-adas de números complejos.[1]

Nuestro interés en $\mathbb{C}^n$ no se origina por el deseo de "generalizar" los resultados de los capítulos anteriores, no obstante que ello abriría nuevas aplicaciones significativas en álgebra lineal.[2] Más bien, este estudio de valores propios complejos es esencial para revelar información "oculta" sobre ciertas matrices con entradas reales que se presentan en gran variedad de problemas de la vida cotidiana. Tales problemas incluyen muchos sistemas dinámicos reales que implican movimiento periódico, vibración o algún tipo de rotación en el espacio.

La teoría del valor propio-vector propio matricial ya desarrollada para $\mathbb{R}^n$ se aplica bien a $\mathbb{C}^n$. Así que un escalar complejo λ satisface $\det(A - \lambda I) = 0$ si y solo si existe un vector distinto de cero $\mathbf{x}$ en $\mathbb{C}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$. Decimos que λ es un **valor propio (complejo)** y $\mathbf{x}$ su correspondiente **vector propio (complejo)**.

EJEMPLO 1 Si $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ on $\mathbb{R}^2$ gira el plano en sentido antihorario, un ángulo de 90º. La acción de A es periódica, ya que, después de cuatro de tales aplicaciones, un vector regresa a su posición original. Evidentemente, ningún vector distinto de cero se mapea sobre un múltiplo de sí mismo, de manera que A no tiene vectores propios en $\mathbb{R}^2$ y, por lo tanto, carece de vectores propios reales. En efecto, la ecuación característica de A es

$$\lambda^2 + 1 = 0$$

[1] Consulte el apéndice B para un breve análisis de números complejos. El álgebra matricial y los conceptos sobre espacios vectoriales reales se podrían ampliar al caso de entradas y escalares complejos. En particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, para A de $m \times n$, con entradas complejas, $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{C}^n$, y c, d en $\mathbb{C}$.

[2] Con frecuencia, en un segundo curso en álgebra lineal se analizan dichos temas, que son de mucha importancia en ingeniería eléctrica.

Las únicas raíces son complejas: $\lambda = i$ y $\lambda = -i$. Sin embargo, si se permite que A actúe sobre $\mathbb{C}^2$, entonces,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} i \\ 1 \end{bmatrix} = i\begin{bmatrix} 1 \\ -i \end{bmatrix}$$
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} -i \\ 1 \end{bmatrix} = -i\begin{bmatrix} 1 \\ i \end{bmatrix}$$

Así, i y $-i$ son valores propios, con $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$ como sus vectores propios correspondientes. (En el ejemplo 2 se analiza un método para *determinar* vectores propios complejos). ■

El principal interés de esta sección será la matriz del siguiente ejemplo.

EJEMPLO 2 Sea $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$. Determine los valores propios de A, y encuentre una base para cada espacio propio.

SOLUCIÓN La ecuación característica de A es

$$\begin{aligned} 0 = \det\begin{bmatrix} .5 - \lambda & -.6 \\ .75 & 1.1 - \lambda \end{bmatrix} &= (.5 - \lambda)(1.1 - \lambda) - (-.6)(.75) \\ &= \lambda^2 - 1.6\lambda + 1 \end{aligned}$$

De la fórmula cuadrática, $\lambda = \frac{1}{2}[1.6 \pm \sqrt{(-1.6)^2 - 4}] = .8 \pm .6i$. Para el valor propio $\lambda = 0.8 \pm .6i$, construya

$$\begin{aligned} A - (.8 - .6i)I &= \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix} - \begin{bmatrix} .8 - .6i & 0 \\ 0 & .8 - .6i \end{bmatrix} \\ &= \begin{bmatrix} -.3 + .6i & -.6 \\ .75 & .3 + .6i \end{bmatrix} \end{aligned} \tag{1}$$

La reducción por filas de la matriz aumentada usual es muy engorrosa a mano debido a su aritmética compleja. Sin embargo, a continuación se presenta una agradable observación que realmente simplifica el asunto: Puesto que $.8 - .6i$ es un valor propio, el sistema

$$\begin{aligned} (-.3 + .6i)x_1 - \quad\quad .6x_2 &= 0 \\ .75x_1 + (.3 + .6i)x_2 &= 0 \end{aligned} \tag{2}$$

tiene una solución no trivial (con x_1 y x_2 posiblemente números complejos). Por lo tanto, *ambas ecuaciones en la ecuación (2) determinan la relación entre x_1 y x_2, y cualquier ecuación sirve para expresar una variable en términos de la otra.*[3]

La segunda ecuación en (2) conduce a

$$\begin{aligned} .75x_1 &= (-.3 - .6i)x_2 \\ x_1 &= (-.4 - .8i)x_2 \end{aligned}$$

Se elige $x_2 = 5$ para así eliminar los decimales, y se obtiene $x_1 = -2 - 4i$. Una base para el espacio propio correspondiente a $\lambda = .8 - .6i$ es

$$\mathbf{v}_1 = \begin{bmatrix} -2 - 4i \\ 5 \end{bmatrix}$$

[3] Otra forma de ver esto es considerando que la matriz en la ecuación (1) no es invertible, de manera que sus filas son linealmente dependientes (como vectores en $\mathbb{C}^2$) y, por consiguiente, una fila es múltiplo (complejo) de la otra.

Cálculos semejantes para $\lambda = .8 + .6i$ producen el vector propio

$$\mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix}$$

Como una comprobación, calcule

$$A\mathbf{v}_2 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \begin{bmatrix} -4+2i \\ 4+3i \end{bmatrix} = (.8+.6i)\mathbf{v}_2 \qquad \blacksquare$$

Sorprendentemente, la matriz A del ejemplo 2 determina una transformación $\mathbf{x} \mapsto A\mathbf{x}$ que en esencia es una rotación. Este hecho se hace evidente cuando se grafican los puntos adecuados.

EJEMPLO 3 Una manera de ver cómo la multiplicación por la matriz A del ejemplo 2 afecta los puntos consiste en graficar un punto inicial arbitrario, por ejemplo, $\mathbf{x}_0 = (2, 0)$, y después graficar imágenes sucesivas de este punto con multiplicaciones repetidas por A. Es decir, grafique

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} \\ \mathbf{x}_2 &= A\mathbf{x}_1 = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} 1.0 \\ 1.5 \end{bmatrix} = \begin{bmatrix} -.4 \\ 2.4 \end{bmatrix} \\ \mathbf{x}_3 &= A\mathbf{x}_2, \ldots \end{aligned}$$

La figura 1 muestra $\mathbf{x}_0, \ldots, \mathbf{x}_8$ como los puntos más grandes. Los puntos más pequeños son las ubicaciones de $\mathbf{x}_9, \ldots, \mathbf{x}_{100}$. La secuencia se encuentra sobre una órbita elíptica. $\blacksquare$

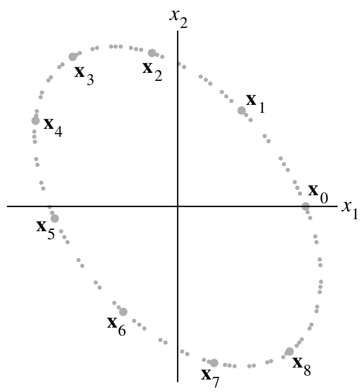


FIGURA 1 Iteración de un punto $\mathbf{x}_0$ por la acción de una matriz con un valor propio complejo.

Desde luego, la figura 1 no explica *por qué* ocurre la rotación. El secreto de la rotación está oculto en las partes real e imaginaria de un vector propio complejo.

Partes real e imaginaria de vectores

El complejo conjugado de un vector complejo $\mathbf{x}$ en $\mathbb{C}^n$ es el vector $\bar{\mathbf{x}}$ en $\mathbb{C}^n$ cuyas entradas están los complejos conjugados de las entradas en $\mathbf{x}$. Las **partes real** e **imaginaria** de un vector complejo $\mathbf{x}$ son los vectores Re $\mathbf{x}$ e Im $\mathbf{x}$ en $\mathbb{R}^n$ formados con las partes real e imaginaria de las entradas de $\mathbf{x}$.

EJEMPLO 4 Si $\mathbf{x} = \begin{bmatrix} 3-i \\ i \\ 2+5i \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} + i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$; entonces

$$\operatorname{Re}\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \quad \operatorname{Im}\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} \quad \text{y} \quad \bar{\mathbf{x}} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} - i \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 3+i \\ -i \\ 2-5i \end{bmatrix} \qquad \blacksquare$$

Si B es una matriz de $m \times n$ con posibles entradas complejas, entonces $\overline{B}$ denota la matriz cuyas entradas son los complejos conjugados de las entradas en B. Las propiedades de conjugados de números complejos son válidas también en álgebra matricial compleja:

$$\overline{r\mathbf{x}} = \bar{r}\,\bar{\mathbf{x}}, \quad \overline{B\mathbf{x}} = \overline{B}\,\bar{\mathbf{x}}, \quad \overline{BC} = \overline{B}\,\overline{C} \quad \text{y} \quad \overline{rB} = \bar{r}\,\overline{B}$$

Valores propios y vectores propios de una matriz real que actúa sobre $\mathbb{C}^n$

Sea A una matriz de $n \times n$ cuyas entradas son reales. Entonces, $\overline{A\mathbf{x}} = \overline{A}\bar{\mathbf{x}} = A\bar{\mathbf{x}}$. Si λ es un valor propio de A y $\mathbf{x}$ su vector propio correspondiente en $\mathbb{C}^n$, entonces,

$$A\bar{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$$

Por lo tanto, $\bar{\lambda}$ también es un valor propio de A, con $\bar{\mathbf{x}}$ su vector propio correspondiente. Esto demuestra que *cuando A es real, sus valores propios complejos se presentan en pares conjugados*. (Aquí y en todas partes, se utiliza el término *valor propio complejo* para referirse a un valor propio $\lambda = a + bi$, con $b \neq 0$).

EJEMPLO 5 Los valores propios de la matriz real del ejemplo 2 son complejos conjugados; a saber, $.8 - .6i$ y $.8 + .6i$. Los vectores propios correspondientes que se encontraron en el ejemplo 2 también son conjugados:

$$\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -2+4i \\ 5 \end{bmatrix} = \bar{\mathbf{v}}_1 \qquad \blacksquare$$

El siguiente ejemplo proporciona el “bloque básico” para todas las matrices reales de 2×2 con valores propios complejos.

EJEMPLO 6 Si $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, donde a y b son reales, pero no ambos cero, entonces los valores propios de C son $\lambda = a \pm bi$. (Véase el problema de práctica al final de esta sección). También, si $r = |\lambda| = \sqrt{a^2 + b^2}$, entonces,

$$C = r\begin{bmatrix} a/r & -b/r \\ b/r & a/r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{bmatrix}$$

donde φ es el ángulo entre el eje x positivo y el rayo que pasa por el origen y por (a, b). Véase la figura 2 y el apéndice B. El ángulo φ se llama el *argumento* de $\lambda = a + bi$. Por consiguiente, la transformación $\mathbf{x} \mapsto C\mathbf{x}$ se puede ver como la composición de una rotación por un ángulo φ y un escalamiento por $|\lambda|$ (véase la figura 3). $\blacksquare$

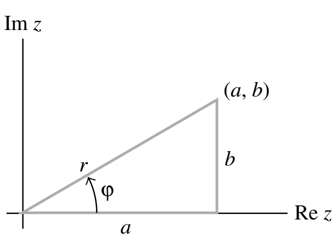


FIGURA 2

Finalmente, se está listo para descubrir la rotación que está oculta dentro de una matriz real que tiene un valor propio complejo.

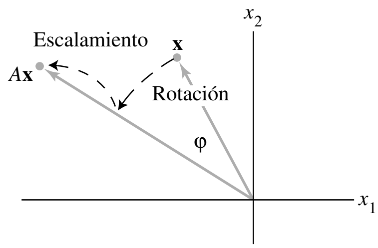


FIGURA 3 Una rotación seguida por un escalamiento.

EJEMPLO 7 Sea $A = \begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}$, $\lambda = .8 - .6i$ y $\mathbf{v}_1 = \begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$, como en el ejemplo 2. También, sea P la matriz real de 2×2

$$P = \begin{bmatrix} \operatorname{Re}\mathbf{v}_1 & \operatorname{Im}\mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix}$$

y sea

$$C = P^{-1}AP = \frac{1}{20}\begin{bmatrix} 0 & 4 \\ -5 & -2 \end{bmatrix}\begin{bmatrix} .5 & -.6 \\ .75 & 1.1 \end{bmatrix}\begin{bmatrix} -2 & -4 \\ 5 & 0 \end{bmatrix} = \begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}$$

Por el ejemplo 6, C es una rotación pura porque $|\lambda|^2 = (.8)^2 + (.6)^2 = 1$. De $C = P^{-1}AP$, se obtiene

$$A = PCP^{-1} = P\begin{bmatrix} .8 & -.6 \\ .6 & 8 \end{bmatrix}P^{-1}$$

¡Aquí está la rotación "dentro" de A! La matriz P proporciona un cambio de variable, por ejemplo, $\mathbf{x} = P\mathbf{u}$. La acción de A equivale a un cambio de variable de $\mathbf{x}$ a $\mathbf{u}$, seguido de una rotación y, después, de regreso a la variable original. Véase la figura 4. La rotación produce una elipse, como en la figura 1, en vez de un círculo, ya que el sistema de coordenadas determinado por las columnas de P no es rectangular ni tiene longitudes unitarias idénticas sobre los dos ejes. ■

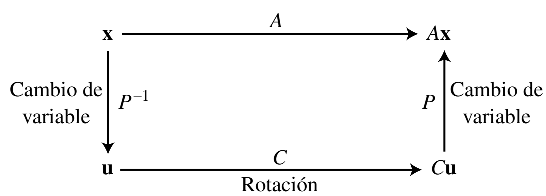


FIGURA 4 Rotación debida a un valor propio complejo.

El siguiente teorema demuestra que los cálculos del ejemplo 7 se adaptan a cualquier matriz real de 2×2 que tenga un valor propio complejo λ. La demostración utiliza el hecho de que si las entradas de A son reales, entonces $A(\operatorname{Re}\mathbf{x}) = \operatorname{Re} A\mathbf{x}$ y $A(\operatorname{Im}\mathbf{x}) = \operatorname{Im} A\mathbf{x}$, y si $\mathbf{x}$ es un vector propio para un valor propio complejo, entonces $\operatorname{Re}\mathbf{x}$ e $\operatorname{Im}\mathbf{x}$ son linealmente independientes en $\mathbb{R}^2$. (Véase los ejercicios 25 y 26). Se omiten los detalles.

TEOREMA 9

Sea A una matriz real de 2×2 con un valor propio complejo $\lambda = a - bi$ ($b \neq 0$) y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$. Entonces,

$$A = PCP^{-1} \quad \text{donde} \quad P = [\operatorname{Re}\mathbf{v} \quad \operatorname{Im}\mathbf{v}] \quad \text{y} \quad C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

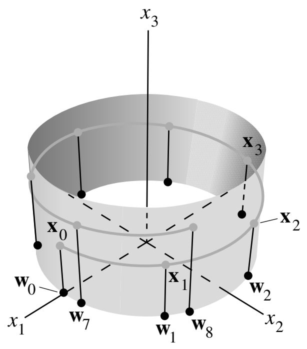


FIGURA 5
Iteraciones de dos puntos bajo la acción de una matriz de 3×3 con un valor propio complejo.

El fenómeno que se ilustra en el ejemplo 7 persiste en dimensiones más altas. Por ejemplo, si A es una matriz de 3×3 con un valor propio complejo, entonces existe un plano en $\mathbb{R}^3$ sobre el cual A actúa como una rotación (posiblemente combinada con escalamiento). Cada vector en ese plano se gira en otro punto sobre el mismo plano. Se dice que el plano es **invariante** bajo A.

EJEMPLO 8 La matriz $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ tiene valores propios $.8 \pm .6i$ y 1.07. Cualquier vector $\mathbf{w}_0$ en el plano x_1x_2 (con la tercera coordenada cero) es girado por A en otro punto del plano. Cualquier vector $\mathbf{x}_0$ que no esté en el plano tiene su coordenada x_3 multiplicada por 1.07. En la figura 5 se presentan las iteraciones de los puntos $\mathbf{w}_0 = (2, 0, 0)$ y $\mathbf{x}_0 = (2, 0, 1)$ con multiplicación por A. ■

PROBLEMA DE PRÁCTICA

Demuestre que si a y b son reales, entonces los valores propios de $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ son $a \pm bi$, con los vectores propios correspondientes $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ y $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

5.5 EJERCICIOS

En los ejercicios 1 a 6, cada matriz actúa sobre $\mathbb{C}^2$. Determine los valores propios y una base para cada espacio propio en $\mathbb{C}^2$.

1. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **2.** $\begin{bmatrix} 3 & -3 \\ 3 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 5 & 1 \\ -8 & 1 \end{bmatrix}$ **4.** $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 1 \\ -2 & 5 \end{bmatrix}$ **6.** $\begin{bmatrix} 7 & -5 \\ 1 & 3 \end{bmatrix}$

En los ejercicios 7 a 12, utilice el ejemplo 6 para listar los valores propios de A. En cada caso, la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es la composición de una rotación y de un escalamiento. Determine el ángulo φ de la rotación, donde $-\pi < \varphi \leq \pi$, así como el factor de escala de r.

7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$ **8.** $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$

9. $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} 0 & .5 \\ -.5 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$ **12.** $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

En los ejercicios 13 a 20, encuentre una matriz P invertible y una matriz C de la forma $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ tal que la matriz dada tenga la forma $A = PCP^{-1}$.

13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$ **16.** $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$

17. $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$ **18.** $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$

19. $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$ **20.** $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

21. En el ejemplo 2, despeje x_2 en la primera ecuación en (2) en términos de x_1, para de ahí producir el vector propio $\mathbf{y} = \begin{bmatrix} 2 \\ -1 + 2i \end{bmatrix}$ para la matriz A. Demuestre que este $\mathbf{y}$ es un múltiplo (complejo) del vector $\mathbf{v}_1$ utilizado en el ejemplo 2.

22. Sean A una matriz de $n \times n$ compleja (o real), y $\mathbf{x}$ en $\mathbb{C}^n$ un vector propio correspondiente a un valor propio λ en $\mathbb{C}$. Demuestre que para cada escalar complejo μ distinto de cero, el vector $\mu\mathbf{x}$ es un vector propio de A.

El capítulo 7 se enfocará en matrices con la propiedad $A^T = A$. Los ejercicios 23 y 24 demuestran que cada valor propio de dichas matrices es necesariamente real.

23. Sean A una matriz real de $n \times n$ tal que $A^T = A$, $\mathbf{x}$ cualquier vector en $\mathbb{C}^n$ y $q = \bar{\mathbf{x}}^T A\mathbf{x}$. Las igualdades que se presentan a continuación demuestran que q es un número real comprobando que $\bar{q} = q$. Dé una razón para cada paso.

$$\bar{q} = \overline{\bar{\mathbf{x}}^T A\mathbf{x}} \underset{a)}{=} \mathbf{x}^T \overline{A\mathbf{x}} \underset{b)}{=} \mathbf{x}^T A\bar{\mathbf{x}} \underset{c)}{=} (\mathbf{x}^T A\bar{\mathbf{x}})^T \underset{d)}{=} \bar{\mathbf{x}}^T A^T \mathbf{x} \underset{e)}{=} q$$

24. Sea A una matriz real de $n \times n$ con la propiedad $A^T = A$. Demuestre que si $A\mathbf{x} = \lambda\mathbf{x}$ para algún vector $\mathbf{x}$ distinto de cero en $\mathbb{C}^n$, entonces, en efecto, λ es real y la parte real de $\mathbf{x}$ es un vector propio de A. [*Sugerencia:* Calcule $\bar{\mathbf{x}}^T A\mathbf{x}$, y utilice el ejercicio 23. Asimismo, examine las partes real e imaginaria de $A\mathbf{x}$].

25. Sean A una matriz real de $n \times n$, y $\mathbf{x}$ un vector en $\mathbb{C}^n$. Demuestre que $\text{Re}(A\mathbf{x}) = A(\text{Re}\,\mathbf{x})$ y $\text{Im}(A\mathbf{x}) = A(\text{Im}\,\mathbf{x})$.

26. Sea A una matriz real de 2×2 con un valor propio complejo $\lambda = a - bi$ ($b \neq 0$) y un vector propio asociado $\mathbf{v}$ en $\mathbb{C}^2$.

a) Demuestre que $A(\text{Re}\ \mathbf{v}) = a\ \text{Re}\ \mathbf{v} + b\ \text{Im}\ \mathbf{v}$ y $A(\text{Im}\ \mathbf{v}) = -b\ \text{Re}\ \mathbf{v} + a\ \text{Im}\ \mathbf{v}$. [*Sugerencia:* Escriba $\mathbf{v} = \text{Re}\ \mathbf{v} + i\ \text{Im}\ \mathbf{v}$, y calcule $A\mathbf{v}$].

b) Compruebe que si P y C están dadas como en el teorema 9, entonces $AP = PC$.

[**M**] En los ejercicios 27 y 28, encuentre una factorización de la matriz A dada en la forma $A = PCP^{-1}$, donde C sea una matriz diagonal a bloques de 2×2 como se indica en el ejemplo 6. (Para cada par conjugado de valores propios, utilice las partes real e imaginaria de un vector propio en $\mathbb{C}^4$ para crear dos columnas de P).

27. $A = \begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$

28. $A = \begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Recuerde que es fácil comprobar si un vector es un vector propio. No hay necesidad de examinar la ecuación característica. Calcule

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a + bi \\ b - ai \end{bmatrix} = (a + bi)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Así $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ es un vector propio correspondiente a $\lambda = a + bi$. Del análisis en esta sección, resulta que $\begin{bmatrix} 1 \\ i \end{bmatrix}$ debe ser un vector propio asociado con $\overline{\lambda} = a - bi$.

5.6 SISTEMAS DINÁMICOS DISCRETOS

Los valores propios y vectores propios ofrecen la clave para entender el comportamiento a largo plazo, o *evolución*, de un sistema dinámico descrito por una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Dicha ecuación se utilizó en la sección 1.10 para modelar el movimiento de población; en la sección 4.9, varias cadenas de Markov; y en el ejemplo introductorio de este capítulo, la población de búhos manchados. Los vectores $\mathbf{x}_k$ dan información acerca del sistema conforme transcurre el tiempo (que se denota con k). En el caso de los búhos manchados, por ejemplo, $\mathbf{x}_k$ listó el número de búhos en tres clases de edades al tiempo k.

En esta sección las aplicaciones se enfocan sobre problemas ecológicos, ya que son más fáciles de plantear y de explicar que, por ejemplo, problemas de física o de ingeniería. Sin embargo, los sistemas dinámicos se originan en muchos campos científicos. En cursos estándar de licenciatura de sistemas de control se estudian diversos aspectos de sistemas dinámicos. En tales cursos el método moderno de diseño del *espacio de estados* se basa significativamente en álgebra de matrices.[1] La respuesta de *estado estable* de un sistema de control es el equivalente en ingeniería de lo que aquí se denomina "comportamiento a largo plazo" del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

[1] Véase G. F. Franklin, J. D. Powell y A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 5a. ed. (Upper Saddle River, NJ: Prentice-Hall, 2006). Este texto tiene una excelente introducción a modelos dinámicos (cap.2). Los capítulos 7 y 8 tratan el diseño del espacio de estados.

Hasta el ejemplo 6, se supone que A es diagonalizable con n vectores propios linealmente independientes, $\mathbf{v}_1, \ldots, \mathbf{v}_n$, y sus valores propios correspondientes $\lambda_1, \ldots, \lambda_n$. Por conveniencia, se supone que los vectores propios están arreglados de tal manera que $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$. Puesto que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base para $\mathbb{R}^n$, entonces cualquier vector inicial $\mathbf{x}_0$ se puede escribir unívocamente en la forma

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \tag{1}$$

Esta *descomposición de vectores propios* de $\mathbf{x}_0$ determina lo que sucede a la secuencia $\{\mathbf{x}_k\}$. El siguiente cálculo generaliza el caso simple examinado en el ejemplo 5 de la sección 5.2. Como los $\mathbf{v}_i$ son vectores propios

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + \cdots + c_nA\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_n\lambda_n\mathbf{v}_n \end{aligned}$$

En general,

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \qquad (k = 0, 1, 2, \ldots) \tag{2}$$

Los siguientes ejemplos ilustran qué sucede en (2) conforme $k \to \infty$.

Un sistema depredador-presa

En lo profundo de los bosques de secuoyas en California, las ratas de bosque de patas oscuras suministran el 80% de la dieta de los búhos manchados, el principal depredador de esos roedores nocturnos. El ejemplo 1 utiliza un sistema dinámico lineal para modelar el sistema físico de los búhos y las ratas. (Hay que admitir que aunque el modelo no es realista en varios aspectos, puede dar el punto de partida para el estudio de modelos no lineales más complicados utilizados por científicos ambientales).

EJEMPLO 1 Denote las poblaciones de búhos y ratas de bosque al tiempo k por el vector $\mathbf{x}_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$, donde k es el tiempo en meses, O_k es el número de búhos en la región bajo estudio y R_k es el número de ratas (medido en miles). Suponga que

$$\begin{aligned} O_{k+1} &= (.5)O_k + (.4)R_k \\ R_{k+1} &= -p \cdot O_k + (1.1)R_k \end{aligned} \tag{3}$$

donde p es un parámetro positivo que se debe especificar. $(.5)O_k$ en la primera ecuación indica que si no hay ratas para alimentarse, entonces cada mes sobreviviría tan solo la mitad de los búhos; mientras que $(1.1)R_k$ en la segunda ecuación implica que sin los búhos como depredadores, la población de ratas crecería 10% cada mes. Si las ratas son abundantes, $(.4)R_k$ hará que crezca la población de búhos, en tanto que el término negativo $-p \cdot O_k$ mide las muertes de ratas por la depredación de los búhos. (En efecto, $1000p$ es el número promedio de ratas devoradas por un búho en un mes). Determine la evolución de este sistema, cuando el parámetro de depredación p es .104.

SOLUCIÓN Cuando $p = .104$, los valores propios de la matriz de coeficientes A para las ecuaciones (3) resultan ser $\lambda_1 = 1.02$ y $\lambda_2 = .58$. Los vectores propios correspondientes son

$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

Un vector inicial $\mathbf{x}_0$ se escribe como $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Entonces, para $k \geq 0$,

$$\begin{aligned} \mathbf{x}_k &= c_1(1.02)^k\mathbf{v}_1 + c_2(.58)^k\mathbf{v}_2 \\ &= c_1(1.02)^k\begin{bmatrix} 10 \\ 13 \end{bmatrix} + c_2(.58)^k\begin{bmatrix} 5 \\ 1 \end{bmatrix} \end{aligned}$$

Conforme $k \to \infty$, entonces $(.58)^k$ rápidamente se aproxima a cero. Suponga que $c_1 > 0$. Por lo tanto, para toda k suficientemente grande, $\mathbf{x}_k$ es aproximadamente igual a $c_1(1.02)^k\mathbf{v}_1$, y se escribe

$$\mathbf{x}_k \approx c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \tag{4}$$

La aproximación en (4) mejora conforme k se incrementa, y así para k grandes,

$$\mathbf{x}_{k+1} \approx c_1(1.02)^{k+1} \begin{bmatrix} 10 \\ 13 \end{bmatrix} = (1.02)c_1(1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \approx 1.02\mathbf{x}_k \tag{5}$$

La aproximación en la ecuación (5) indica que a final de cuentas ambas entradas de $\mathbf{x}_k$ (los números de búhos y ratas) crecerán por un factor de casi 1.02 cada mes, un 2% de tasa de crecimiento mensual. Por (4), $\mathbf{x}_k$ es aproximadamente un múltiplo de (10, 13), de manera que las entradas en $\mathbf{x}_k$ están casi en la misma razón de 10 a 13. Es decir, por cada 10 búhos hay cerca de 13 mil ratas. ■

El ejemplo 1 muestra dos hechos generales acerca de un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ donde A es $n \times n$, sus valores propios satisfacen $|\lambda_1| \geq 1$ y $1 > |\lambda_1|$ para $j = 2, \ldots, n$ y $\mathbf{v}_1$ es un vector propio correspondiente a λ_1. Si $\mathbf{x}_0$ está dado por la ecuación (1), con $c_1 \neq 0$, entonces para todas las k suficientemente grandes,

$$\mathbf{x}_{k+1} \approx \lambda_1 \mathbf{x}_k \tag{6}$$

y

$$\mathbf{x}_k \approx c_1(\lambda_1)^k \mathbf{v}_1 \tag{7}$$

Las aproximaciones en (6) y (7) se pueden volver tan cercanas como se desee tomando una k muy grande. Por (6), $\mathbf{x}_k$ finalmente crecerá casi por un factor de λ_1 cada vez, de manera que λ_1 determina la eventual tasa de crecimiento del sistema. Asimismo, por (7), la razón de cualesquiera dos entradas en $\mathbf{x}_k$ (para k grande) es casi la misma que la razón de las entradas correspondientes en $\mathbf{v}_1$. En el ejemplo 5 de la sección 5.2, se ilustra el caso donde $\lambda_1 = 1$.

Descripción gráfica de soluciones

Cuando A es una matriz de 2×2, los cálculos algebraicos se complementan con una descripción geométrica de la evolución del sistema. La ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ se puede considerar una descripción de qué sucede al punto inicial $\mathbf{x}_0$ en $\mathbb{R}^2$, conforme se transforma repetidamente por el mapeo $\mathbf{x} \mapsto A\mathbf{x}$. La gráfica de $\mathbf{x}_0, \mathbf{x}_1, \ldots$ es la **trayectoria** del sistema dinámico.

EJEMPLO 2 Grafique varias trayectorias del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, cuando

$$A = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

SOLUCIÓN Los valores propios de A son .8 y .64, con vectores propios $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces,

$$\mathbf{x}_k = c_1(.8)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Desde luego, $\mathbf{x}_k$ tiende a $\mathbf{0}$ porque $(.8)^k$ y $(.64)^k$ se aproximan a 0 conforme $k \to \infty$. No obstante, resulta interesante *la manera* en que $\mathbf{x}_k$ va hacia $\mathbf{0}$. La figura 1 (en la página 304) muestra los primeros términos escasos de varias trayectorias que empiezan en puntos sobre la frontera de la caja con vértices en $(\pm 3, \pm 3)$. Los puntos sobre cada trayectoria están conectados por una curva delgada, para que así la trayectoria sea más fácil de observar. ■

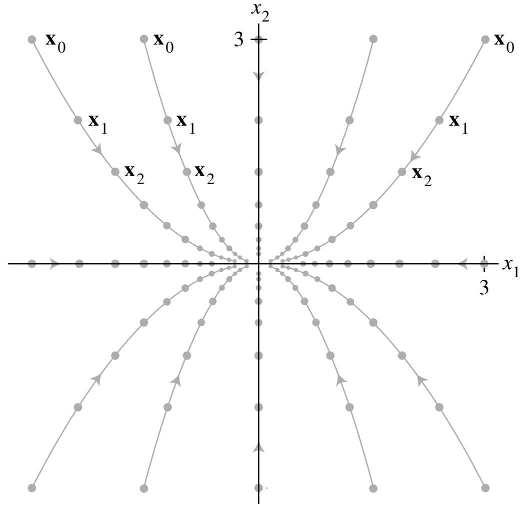


FIGURA 1 El origen como un atractor.

En el ejemplo 2, al origen se le denomina **atractor** del sistema dinámico porque todas las trayectorias tienden hacia **0**. Esto ocurre siempre que ambos valores propios sean menores que 1 en magnitud. La dirección de la atracción más grande está a lo largo de la recta que pasa por **0** y el vector propio $\mathbf{v}_2$ para el valor propio de menor magnitud.

En el siguiente ejemplo, ambos valores propios de A son mayores que 1 en magnitud, y **0** es un **repulsor** del sistema dinámico. Todas las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$ excepto la solución cero (constante), no están acotadas y tienden a alejarse del origen.[2]

EJEMPLO 3 Grafique varias soluciones típicas de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

SOLUCIÓN Los valores propios de A son 1.44 y 1.2. Si $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces,

$$\mathbf{x}_k = c_1(1.44)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Ambos términos aumentaron en tamaño, pero el primer término lo hizo más rápido. Así, la dirección de la mayor repulsión es la recta que pasa por **0** y por el vector propio para el valor propio de magnitud más grande. La figura 2 muestra varias trayectorias que inician en puntos muy cercanos a **0**. ∎

En el siguiente ejemplo, **0** es un **punto silla** porque el origen atrae soluciones de algunas direcciones y las repele en otras direcciones. Esto ocurre siempre que un valor propio sea más grande que 1 en magnitud y el otro sea menor que 1 en magnitud. La dirección de mayor atracción está determinada por un vector propio del valor propio de menor magnitud. La dirección de la mayor repulsión se determina mediante un vector propio del valor propio de magnitud más grande.

---

[2] El origen es el único posible atractor o repulsor en un sistema dinámico *lineal*, aunque podrían existir múltiples atractores y repulsores en un sistema dinámico más general para el cual el mapeo $\mathbf{x}_k \rightarrow \mathbf{x}_{k+1}$ no es lineal. En tal sistema, los atractores y los repulsores están definidos en términos de los valores propios de una matriz especial (con entradas variables) llamada *matriz jacobiana* del sistema.

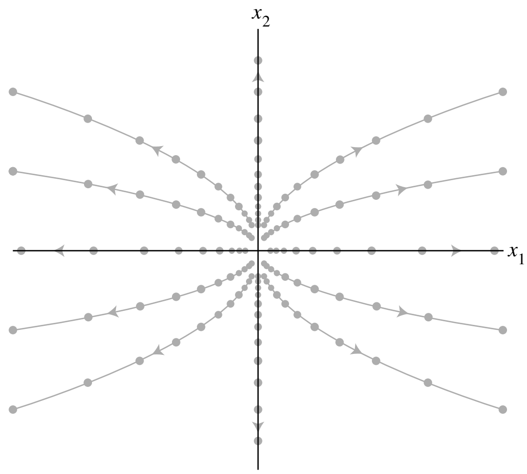


FIGURA 2 El origen como un repulsor.

EJEMPLO 4 Grafique varias soluciones típicas de la ecuación $\mathbf{y}_{k+1} = D\mathbf{y}_k$, donde

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

(Aquí se usaron D y $\mathbf{y}$ en vez de A y $\mathbf{x}$ porque este ejemplo se utilizará más adelante). Demuestre que una solución $\{\mathbf{y}_k\}$ no está acotada, cuando su punto inicial no está sobre el eje x_2.

SOLUCIÓN Los valores propios de D son 2 y 0.5. Si $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, entonces,

$$\mathbf{y}_k = c_1 2^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.5)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{8}$$

Si $\mathbf{y}_0$ está sobre el eje x_2, entonces $c_1 = 0$ y $\mathbf{y}_k \to \mathbf{0}$ conforme $k \to \infty$. Pero si $\mathbf{y}_0$ no se encuentra sobre el eje x_2, entonces el primer término en la suma para $\mathbf{y}_k$ se vuelve arbitrariamente grande y, por lo tanto, $\{\mathbf{y}_k\}$ no está acotada. La figura 3 presenta diez trayectorias que inician cerca del eje x_2 o sobre este. ■

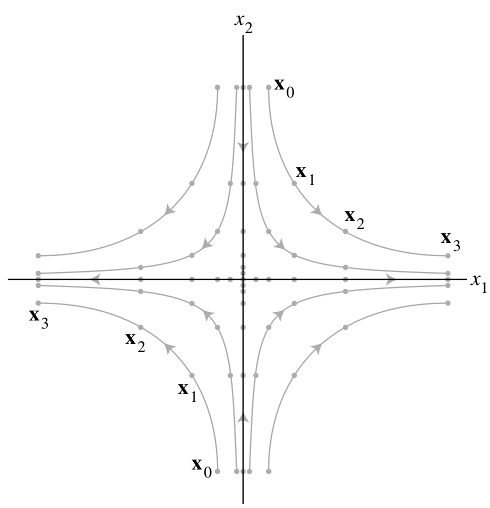


FIGURA 3 El origen como un punto silla.

Cambio de variable

Los tres ejemplos anteriores implicaron matrices diagonales. Para tratar el caso no diagonal, regrese por un momento al caso $n \times n$ donde los vectores propios de A forman una base $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ para $\mathbb{R}^n$. Sean $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, y D la matriz diagonal con los valores propios correspondientes sobre la diagonal. Dada una secuencia $\{\mathbf{x}_k\}$ que comprueba $\mathbf{x}_{k+1} = A\mathbf{x}_k$, defina una nueva secuencia $\{\mathbf{y}_k\}$ mediante

$$\mathbf{y}_k = P^{-1}\mathbf{x}_k \quad \text{o, de manera equivalente,} \quad \mathbf{x}_k = P\mathbf{y}_k$$

Sustituyendo esas relaciones en la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ y utilizando que $A = PDP^{-1}$, se tiene que

$$P\mathbf{y}_{k+1} = AP\mathbf{y}_k = (PDP^{-1})P\mathbf{y}_k = PD\mathbf{y}_k$$

Se multiplican ambos lados por la izquierda por P^{-1} y se obtiene

$$\mathbf{y}_{k+1} = D\mathbf{y}_k$$

Si se escribe $\mathbf{y}_k$ como $\mathbf{y}(k)$ y se denotan las entradas en $\mathbf{y}(k)$ por $y_1(k), \ldots, y_n(k)$, entonces,

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ \vdots \\ y_n(k+1) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_n(k) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}_k$ a $\mathbf{y}_k$ ha *desacoplado* el sistema de ecuaciones en diferencias. La evolución de $y_1(k)$, por ejemplo, no se ve afectada por lo que suceda a $y_2(k), \ldots, y_n(k)$, porque $y_1(k + 1) = \lambda_1 \cdot y_1(k)$ para cada k.

La ecuación $\mathbf{x}_k = P\mathbf{y}_k$ indica que $\mathbf{y}_k$ es el vector de coordenadas de $\mathbf{x}_k$ con respecto a la base de vectores propios $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Se puede desacoplar el sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$ efectuando cálculos en el nuevo sistema de coordenadas de vectores propios. Cuando $n = 2$, esto equivale a emplear papel para graficar con ejes en las direcciones de los dos vectores propios.

EJEMPLO 5 Demuestre que el origen es un punto silla para las soluciones de $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 1.25 & -.75 \\ -.75 & 1.25 \end{bmatrix}$$

Encuentre las direcciones de mayor atracción y de mayor repulsión.

SOLUCIÓN Utilizando técnicas estándar, se determina que los valores propios de A son 2 y .5, con los vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectivamente. Como $|2| > 1$ y $|.5| < 1$, entonces el origen es un punto silla del sistema dinámico. Si $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$,

$$\mathbf{x}_k = c_1 2^k \mathbf{v}_1 + c_2(.5)^k \mathbf{v}_2 \tag{9}$$

Esta ecuación se ve exactamente igual que la ecuación (8) del ejemplo 4, con $\mathbf{v}_1$ y $\mathbf{v}_2$ en lugar de la base estándar.

Sobre el papel para graficar, dibuje ejes a través de $\mathbf{0}$ y de los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$. Véase la figura 4. El movimiento a lo largo de los ejes corresponde al movimiento sobre los ejes estándar de la figura 3. En la figura 4, la dirección de mayor *repulsión* es la recta que pasa por $\mathbf{0}$ y el vector propio $\mathbf{v}_1$ cuyo valor propio es mayor que 1 en magnitud. Si $\mathbf{x}_0$ está sobre esta recta, c_2 de la ecuación (9) es cero y $\mathbf{x}_k$ se aleja rápidamente de $\mathbf{0}$. La dirección de mayor *atracción* está determinada por el vector propio $\mathbf{v}_2$ cuyo valor propio es menor que 1 en magnitud.

En la figura 4 se ilustra un número de trayectorias. Cuando esta gráfica se observa en términos de los ejes de vectores propios, el esquema "parece" en esencia igual que el esquema de la figura 3. ■

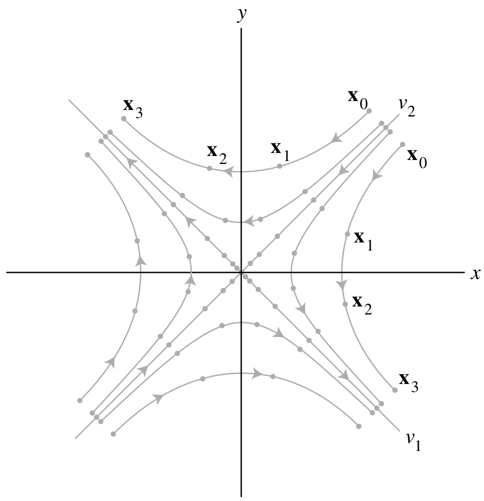


FIGURA 4 El origen como un punto silla.

Valores propios complejos

Cuando una matriz real A de 2×2 tiene valores propios complejos, entonces A no es diagonalizable (cuando actúa sobre $\mathbb{R}^2$), pero el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ es fácil de describir. El ejemplo 3 de la sección 5.5 ilustró el caso donde los valores propios tienen valor absoluto 1. La iteración de un punto $\mathbf{x}_0$ giró alrededor del origen con una trayectoria elíptica.

Si A tiene dos valores propios complejos cuyo valor absoluto es mayor que 1, entonces $\mathbf{0}$ es un repulsor y la iteración de $\mathbf{x}_0$ girará en espiral hacia fuera en torno al origen. Si los valores absolutos de los valores propios complejos son menores que 1, entonces el origen es un atractor y la iteración de $\mathbf{x}_0$ da una espiral hacia adentro acercándose al origen, como en el siguiente ejemplo.

EJEMPLO 6 Se puede comprobar que la matriz

$$A = \begin{bmatrix} .8 & .5 \\ -.1 & 1.0 \end{bmatrix}$$

tiene valores propios $.9 \pm .2i$, con vectores propios $\begin{bmatrix} 1 \mp 2i \\ 1 \end{bmatrix}$. La figura 5 (en la página 308) muestra tres trayectorias del sistema $\mathbf{x}_{k+1} = A\mathbf{x}_k$, con vectores iniciales $\begin{bmatrix} 0 \\ 2.5 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ -2.5 \end{bmatrix}$. ■

Supervivencia de los búhos manchados

Recuerde del ejemplo introductorio de este capítulo que la población de búhos manchados en el área Willow Creek, California, se modeló con un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ donde las entradas en $\mathbf{x}_k = (j_k, s_k, a_k)$ listaban los números de hembras (al tiempo k) en las etapas de vida juvenil, subadulto y adulto, respectivamente, y A es la matriz de etapas

$$A = \begin{bmatrix} 0 & 0 & .33 \\ .18 & 0 & 0 \\ 0 & .71 & .94 \end{bmatrix} \tag{10}$$

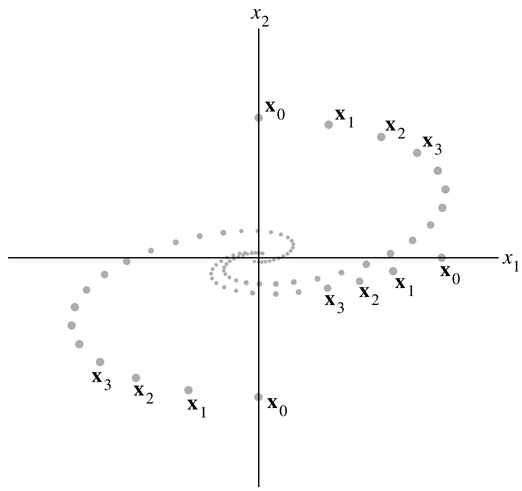


FIGURA 5 Rotación asociada con valores propios complejos.

MATLAB muestra que los valores propios de A son aproximadamente $\lambda_1 = .98$, $\lambda_2 = -.02 + .21i$ y $\lambda_3 = -.02 - .21i$. Observe que los tres valores propios son menores que 1 en magnitud, ya que $|\lambda_2|^2 = |\lambda_3|^2 = (-.02)^2 + (.21)^2 = 0.445$.

Por el momento, deje que A actúe sobre el espacio vectorial complejo $\mathbb{C}^3$. Puesto que A tiene tres valores propios distintos, entonces los tres vectores propios correspondientes son linealmente independientes y forman una base para $\mathbb{C}^3$. Denote los vectores propios con $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Así, la solución general de $\mathbf{x}_{k+1} = A\mathbf{x}_k$ (utilizando vectores en $\mathbb{C}^3$) tiene la forma

$$\mathbf{x}_k = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + c_3(\lambda_3)^k\mathbf{v}_3 \tag{11}$$

Si $\mathbf{x}_0$ es un vector inicial real, entonces $\mathbf{x}_1 = A\mathbf{x}_0$ es real porque A es real. Asimismo, la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ indica que cada $\mathbf{x}_k$ en el miembro izquierdo de la ecuación (11) es real, aun cuando se expresa como una suma de vectores complejos. Sin embargo, cada término en el miembro derecho de la ecuación (11) se aproxima al vector cero, ya que todos los valores propios son menores que 1 en magnitud. Por lo tanto, la secuencia real $\mathbf{x}_k$ también tiende al vector cero. Tristemente, este modelo predice que, a final de cuentas, perecerán todos los búhos.

¿Hay alguna esperanza para los búhos manchados? Recuerde del ejemplo introductorio que la entrada de 18% en la matriz A de la ecuación (10) es porque, no obstante que el 60% de los búhos juveniles sobreviven lo suficiente para dejar el nido y buscar nuevas áreas de distribución, tan solo el 30% de ese grupo sobrevive a dicha búsqueda y encuentra un nuevo hogar. La supervivencia a la mencionada búsqueda está significativamente influida por el número de áreas deforestadas en el bosque, lo cual vuelve más difícil y peligrosa esa búsqueda.

Algunas poblaciones de búhos habitan en zonas con pocas o ninguna área deforestada. Quizás un porcentaje aún mayor de búhos juveniles sobrevivan ahí y encuentren nuevos territorios. De hecho, el problema del búho manchado es más complicado que como se ha expuesto, no obstante que el ejemplo final da un final feliz a la historia.

EJEMPLO 7 Suponga que la tasa de supervivencia a la búsqueda, de los búhos juveniles, sea de 50%, de manera que la entrada (2, 1) de la matriz por etapas (de estados) A en (10) es .3 en vez de .18. ¿Qué predice el modelo de la matriz por etapas acerca de la población de búhos manchados?

SOLUCIÓN Ahora los valores propios de A resultan ser aproximadamente $\lambda_1 = 1.01$, $\lambda_2 = -.03 + .26i$ y $\lambda_3 = -.03 - .26i$. Un vector propio para λ_1 es aproximadamente $\mathbf{v}_1 = (10, 3, 31)$. Sean $\mathbf{v}_2$ y $\mathbf{v}_3$ vectores propios (complejos) para λ_2 y λ_3. En este caso, la ecuación (11) se convierte en

$$\mathbf{x}_k = c_1(1.01)^k\mathbf{v}_1 + c_2(-.03 + .26i)^k\mathbf{v}_2 + c_3(-.03 - .26i)^k\mathbf{v}_3$$

Cuando $k \to \infty$, el segundo de los dos vectores tiende a cero. Así, $\mathbf{x}_k$ se vuelve cada vez más semejante al vector (real) $c_1(1.01)^k\mathbf{v}_1$. Las aproximaciones aplicadas en (6) y (7) provienen del ejemplo 1. Además, se puede demostrar que la constante c_1 en la descomposición inicial de $\mathbf{x}_0$ es positiva cuando las entradas en $\mathbf{x}_0$ son no negativas. Entonces, la población de búhos crecerá lentamente, con una tasa de crecimiento asintótica de 1.01. El vector propio $\mathbf{v}_1$ describe la distribución final de los búhos en sus etapas de vida: por cada 31 adultos, habrá 10 juveniles y 3 subadultos. ■

Lecturas adicionales

Franklin G. F., J. D. Powell y M. L. Workman. *Digital Control of Dynamic Systems*, 3a. ed., Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems-Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Phlip. *Management of a Buffalo Herd*, UMAP módulo 207. Lexington, MA: COMAP, 1980.

PROBLEMAS DE PRÁCTICA

1. La matriz A que se muestra a continuación tiene los valores propios 1, $\frac{2}{3}$ y $\frac{1}{3}$ con los vectores propios correspondientes $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$:

$$A = \frac{1}{9}\begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

Encuentre la solución general de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 11 \\ -2 \end{bmatrix}$.

2. ¿Qué ocurre en la secuencia $\{\mathbf{x}_k\}$ del problema de práctica 1, conforme $k \to \infty$?

5.6 EJERCICIOS

1. Sea A una matriz de 2×2 con valores propios 3 y $1/3$, y vectores propios correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Sea $\{\mathbf{x}_k\}$ una solución de la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $\mathbf{x}_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$.

a) Calcule $\mathbf{x}_1 = A\mathbf{x}_0$. [*Sugerencia:* No se necesita conocer A].

b) Determine una fórmula para $\mathbf{x}_k$ que implique k y los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$.

2. Suponga que los valores propios de A de 3×3 son 3, $4/5$ y $3/5$, con sus vectores propios correspondientes $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$ y $\begin{bmatrix} -3 \\ -3 \\ 7 \end{bmatrix}$. Sea $\mathbf{x}_0 = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$. Encuentre la solución de la ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $\mathbf{x}_0$ especificado, y describa qué sucede cuando $k \to \infty$.

En los ejercicios 3 a 6, suponga que cualquier vector inicial $\mathbf{x}_0$ tiene una descomposición de vectores propios tal que el coeficiente c_1, en la ecuación (1) de esta sección, es positivo.[3]

3. Determine la evolución del sistema dinámico del ejemplo 1, cuando el parámetro de depredación p es .2 en la ecuación (3). (Dé una fórmula para $\mathbf{x}_k$). ¿Crece o disminuye la población de búhos? ¿Qué se puede decir sobre la población de ratas de bosque?

4. Determine la evolución del sistema dinámico del ejemplo 1, cuando el parámetro de depredación p es .125. (Dé una fórmula para $\mathbf{x}_k$). Conforme el tiempo transcurre, ¿qué ocurre a los tamaños de las poblaciones de búhos y de ratas bosque? El sistema tiende hacia lo que algunas veces se denomina equilibrio inestable. ¿Qué piensa que pueda ocurrirle al sistema si algún aspecto del modelo (como las tasas de nacimiento o de depredación) cambiara ligeramente?

5. En los antiguos bosques de abetos en Douglas, los búhos manchados cenan básicamente ardillas voladoras. Suponga que la matriz depredador-presa para esas dos poblaciones es $A = \begin{bmatrix} .4 & .3 \\ -p & 1.2 \end{bmatrix}$. Demuestre que si el parámetro de depredación p es .325, entonces ambas poblaciones crecen. Estime la tasa de crecimiento a largo plazo y la proporción final entre búhos y ardillas voladoras.

6. Demuestre que si el parámetro de depredación p del ejercicio 5 es .5, entonces a final de cuentas desaparecerán tanto los búhos como las ardillas. Encuentre un valor de p para el cual ambas poblaciones tienden a mantener niveles constantes. En este caso, ¿cuál es la proporción entre las dos poblaciones?

7. Sea que A tiene las propiedades descritas en el ejercicio 1.

a) ¿El origen es un atractor, un repulsor, o un punto silla del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$?

b) Encuentre las direcciones de mayor atracción y/o repulsión para este sistema dinámico.

c) Haga una descripción gráfica del sistema, que indique las direcciones de mayor atracción o repulsión. Incluya un esquema de varias trayectorias típicas (sin calcular puntos específicos).

8. Determine la naturaleza del origen (atractor, repulsor o punto silla) para el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ si A tiene las propiedades descritas en el ejercicio 2. Obtenga las direcciones de mayor atracción o repulsión.

En los ejercicios 9 a 14, clasifique el origen como atractor, repulsor o punto silla del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Encuentre las direcciones de mayor atracción y/o repulsión.

9. $A = \begin{bmatrix} 1.7 & -.3 \\ -1.2 & .8 \end{bmatrix}$

10. $A = \begin{bmatrix} .3 & .4 \\ -.3 & 1.1 \end{bmatrix}$

[3] Una de las limitaciones del modelo del ejemplo 1 es que ahí siempre existen vectores $\mathbf{x}_0$ de población inicial con entradas positivas, tales que el coeficiente c_1 es negativo. La aproximación (7) aún es válida, pero las entradas en $\mathbf{x}_k$ finalmente se vuelven negativas.

11. $A = \begin{bmatrix} .4 & .5 \\ -.4 & 1.3 \end{bmatrix}$

12. $A = \begin{bmatrix} .5 & .6 \\ -.3 & 1.4 \end{bmatrix}$

13. $A = \begin{bmatrix} .8 & .3 \\ -.4 & 1.5 \end{bmatrix}$

14. $A = \begin{bmatrix} 1.7 & .6 \\ -.4 & .7 \end{bmatrix}$

15. Sea $A = \begin{bmatrix} .4 & 0 & .2 \\ .3 & .8 & .3 \\ .3 & .2 & .5 \end{bmatrix}$. El vector $\mathbf{v}_1 = \begin{bmatrix} .1 \\ .6 \\ .3 \end{bmatrix}$ es un vector propio de A, y dos valores propios son .5 y .2. Construya la solución del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que satisface $\mathbf{x}_0 = (0, .3, .7)$. ¿Qué sucede a $\mathbf{x}_k$ conforme $k \to \infty$?

16. [**M**] Obtenga la solución general del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ cuando A es la matriz estocástica para el modelo Rent A Car de Hertz del ejercicio 16 de la sección 4.9.

17. Construya un modelo matricial por etapas para una especie animal que tiene dos etapas de vida: juvenil (hasta 1 año de edad) y adulto. Suponga que las hembras adultas cada año dan a luz a un promedio de 1.6 hembras juveniles. Cada año, el 30% de las hembras juveniles sobreviven para convertirse en adultos y el 80% de los adultos sobreviven. Para $k \geq 0$, sea $\mathbf{x}_k = (j_k, a_k)$, donde las entradas en $\mathbf{x}_k$ son los números de hembras juveniles y hembras adultas en el año k.

a) Construya la matriz por etapas A tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b) Demuestre que la población está creciendo, calcule la tasa de crecimiento final de la población y dé la razón final entre adultos y juveniles.

c) [**M**] Suponga que inicialmente hay 15 hembras juveniles y 10 hembras adultas en la población. Elabore cuatro gráficas que muestren cómo cambia la población durante 8 años: *a*) el número de juveniles, *b*) el número de adultas, *c*) la población total y *d*) la razón entre adultas y juveniles (cada año). ¿Cuándo parece estabilizarse la razón en *d*)? Incluya una lista del programa o las instrucciones empleadas para realizar las gráficas en *c*) y *d*).

18. Una manada de búfalos americanos (bisontes) se podría modelar con una matriz por etapas, similar a la empleada para los búhos manchados. Las hembras se pueden dividir en terneras (hasta 1 año de edad), becerras (de 1 a 2 años) y adultas. Suponga que un promedio de 42 hembras terneras nacen cada año por cada 100 adultas. (Tan solo las adultas tienen descendencia). Cada año sobreviven cerca del 60% de terneras, 75% de becerras y 95% de adultas. Para $k \geq 0$, sea $\mathbf{x}_k = (c_k, y_k, a_k)$, donde las entradas en $\mathbf{x}_k$ son los números de hembras en cada etapa de vida en el año k.

a) Construya la matriz por etapas A para la manada de búfalos, tal que $\mathbf{x}_{k+1} = A\mathbf{x}_k$ para $k \geq 0$.

b) [**M**] Demuestre que la manada de búfalos está creciendo, determine la tasa de crecimiento esperada después de muchos años, y dé los números esperados de terneras y becerras por cada 100 adultas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. El primer paso consiste en escribir $\mathbf{x}_0$ como una combinación lineal de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. La reducción por filas de $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{x}_0]$ genera los pesos $c_1 = 2$, $c_2 = 1$ y $c_3 = 3$, tal que

$$\mathbf{x}_0 = 2\mathbf{v}_1 + 1\mathbf{v}_2 + 3\mathbf{v}_3$$

Como los valores propios son 1, $\frac{2}{3}$ y $\frac{1}{3}$, la solución general es

$$\begin{aligned} \mathbf{x}_k &= 2 \cdot 1^k \mathbf{v}_1 + 1 \cdot \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3 \cdot \left(\frac{1}{3}\right)^k \mathbf{v}_3 \\ &= 2\begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + \left(\frac{2}{3}\right)^k \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + 3 \cdot \left(\frac{1}{3}\right)^k \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \end{aligned} \tag{12}$$

2. Conforme $k \to \infty$, el segundo y tercer términos en (12) tienden al vector cero, y

$$\mathbf{x}_k = 2\mathbf{v}_1 + \left(\frac{2}{3}\right)^k \mathbf{v}_2 + 3\left(\frac{1}{3}\right)^k \mathbf{v}_3 \to 2\mathbf{v}_1 = \begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$$

5.7 APLICACIONES A ECUACIONES DIFERENCIALES

Esta sección describe análogos continuos de las ecuaciones en diferencias estudiadas en la sección 5.6. En muchos problemas aplicados, algunas cantidades varían continuamente en el tiempo, y se relacionan mediante un sistema de ecuaciones diferenciales:

$$\begin{aligned} x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n \end{aligned}$$

Aquí $x_1, \ldots, x_n$ son funciones diferenciables de t, con derivadas $x_1', \ldots, x_n'$, y las a_{ij} son constantes. El aspecto esencial de este sistema es que es *lineal*. Para observarlo, escriba el sistema como una ecuación diferencial matricial

$$\mathbf{x}'(t) = A\mathbf{x}(t) \tag{1}$$

donde

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Una **solución** de la ecuación (1) es una función valuada en vectores que satisface la ecuación (1) para toda t en algún intervalo de números reales, tales como $t \geq 0$.

La ecuación (1) es *lineal* porque la derivación de funciones y la multiplicación de vectores por una matriz son transformaciones lineales. Así, cuando $\mathbf{u}$ y $\mathbf{v}$ son soluciones de $\mathbf{x}' = A\mathbf{x}$, entonces $c\mathbf{u} + d\mathbf{v}$ también es una solución, porque

$$\begin{aligned} (c\mathbf{u} + d\mathbf{v})' &= c\mathbf{u}' + d\mathbf{v}' \\ &= cA\mathbf{u} + dA\mathbf{v} = A(c\mathbf{u} + d\mathbf{v}) \end{aligned}$$

(A esta propiedad los ingenieros la denominan *superposición* de soluciones). Asimismo, la función idénticamente cero es una solución (trivial) de (1). En la terminología del capítulo 4, el conjunto de todas las soluciones de (1) es un *subespacio* del conjunto de todas las funciones continuas con valores en $\mathbb{R}^n$.

Los libros estándar de ecuaciones diferenciales demuestran que siempre existe el **conjunto fundamental de soluciones** de (1). Si A es $n \times n$, entonces hay n funciones linealmente independientes en un conjunto fundamental, y cada solución de (1) es una combinación lineal única de tales n funciones. Es decir, un conjunto fundamental de soluciones es una *base* para el conjunto de todas las soluciones de (1), y el conjunto solución es un espacio vectorial n-dimensional de funciones. Si se especifica un vector $\mathbf{x}_0$, entonces el **problema con valores iniciales** consiste en construir la función (única) $\mathbf{x}$ tal que $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$.

Cuando A es una matriz diagonal, las soluciones de (1) se generan mediante cálculo elemental. Por ejemplo, considere

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & -5 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{2}$$

es decir,

$$\begin{aligned} x_1'(t) &= 3x_1(t) \\ x_2'(t) &= -5x_2(t) \end{aligned} \tag{3}$$

Se dice que el sistema (2) está *desacoplado* porque cada derivada de una función tan solo depende de la propia función, y no de alguna combinación o “acoplamiento” de $x_1(t)$ y $x_2(t)$. De cálculo, las soluciones de (3) son $x_1(t) = c_1e^{3t}$ y $x_2(t) = c_2e^{-5t}$, para las constantes cualesquiera c_1 y c_2. Cada solución de la ecuación (2) se escribe en la forma

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} c_1e^{3t} \\ c_2e^{-5t} \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix}e^{3t} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}e^{-5t}$$

Este ejemplo sugiere que para la ecuación general $\mathbf{x}' = A\mathbf{x}$, una solución podría ser una combinación lineal de funciones con la forma

$$\mathbf{x}(t) = \mathbf{v}e^{\lambda t} \tag{4}$$

para algún escalar λ y algún vector fijo $\mathbf{v}$ distinto de cero. [Si $\mathbf{v} = \mathbf{0}$, la función $\mathbf{x}(t)$ es idénticamente cero y por lo tanto satisface $\mathbf{x}' = A\mathbf{x}$]. Observe que

$$\begin{aligned} \mathbf{x}'(t) &= \lambda\mathbf{v}e^{\lambda t} && \text{Por Cálculo, ya que } \mathbf{v} \text{ es un vector constante} \\ A\mathbf{x}(t) &= A\mathbf{v}e^{\lambda t} && \text{Multiplicando ambos lados de (4) por } A \end{aligned}$$

Como $e^{\lambda t}$ nunca es cero, $\mathbf{x}'(t)$ será igual a $A\mathbf{x}(t)$ si y solo si $\lambda\mathbf{v} = A\mathbf{v}$, es decir, si y solo si λ es un valor propio de A y $\mathbf{v}$ es un vector propio correspondiente. Así, cada par valor propio-vector propio ofrece una solución (4) de $\mathbf{x}' = A\mathbf{x}$. Algunas veces dichas soluciones se conocen como *funciones propias* de la ecuación diferencial. Las funciones propias dan la clave para resolver sistemas de ecuaciones diferenciales.

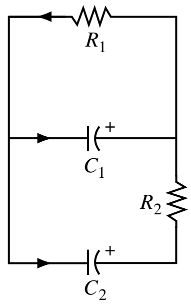


FIGURA 1

EJEMPLO 1 El circuito de la figura 1 se describe con la ecuación diferencial

$$\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} -(1/R_1 + 1/R_2)/C_1 & 1/(R_2C_1) \\ 1/(R_2C_2) & -1/(R_2C_2) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

donde $x_1(t)$ y $x_2(t)$ son los voltajes en los dos capacitores al tiempo t. Suponga que el resistor R_1 es de 1 ohm, $R_2 = 2$ ohms, capacitor $C_1 = 1$ farad, y $C_2 = .5$ farad; con cargas iniciales de 5 volts en el capacitor C_1 y de 4 volts en C_2. Encuentre las fórmulas para $x_1(t)$ y $x_2(t)$ que describan cómo los voltajes cambian en el tiempo.

SOLUCIÓN Sea que A denote la matriz mostrada arriba y sea $\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$. Para los datos dados, $A = \begin{bmatrix} -1.5 & .5 \\ 1 & -1 \end{bmatrix}$ y $\mathbf{x}(0) = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. Los valores propios de A son $\lambda_1 = -.5$ y $\lambda_2 = -2$, con los vectores propios correspondientes

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Ambas funciones propias $\mathbf{x}_1(t) = \mathbf{v}_1 e^{\lambda_1 t}$ y $\mathbf{x}_2(t) = \mathbf{v}_2 e^{\lambda_2 t}$ satisfacen $\mathbf{x}' = A\mathbf{x}$, y lo mismo ocurre con cualquier combinación lineal de $\mathbf{x}_1$ y $\mathbf{x}_2$. Se hace

$$\mathbf{x}(t) = c_1 \mathbf{v}_1 e^{\lambda_1 t} + c_2 \mathbf{v}_2 e^{\lambda_2 t} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

y observe que $\mathbf{x}(0) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. De manera evidente, puesto que $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes y por consiguiente generan $\mathbb{R}^2$, c_1 y c_2 se determinan tal que $\mathbf{x}(0)$ sea igual a $\mathbf{x}_0$. En efecto, la ecuación

$$c_1 \underset{\substack{\uparrow \\ \mathbf{v}_1}}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} + c_2 \underset{\substack{\uparrow \\ \mathbf{v}_2}}{\begin{bmatrix} -1 \\ 1 \end{bmatrix}} = \underset{\substack{\uparrow \\ \mathbf{x}_0}}{\begin{bmatrix} 5 \\ 4 \end{bmatrix}}$$

conduce fácilmente a $c_1 = 3$ y $c_2 = -2$. Por lo tanto, la solución deseada de la ecuación diferencial $\mathbf{x}' = A\mathbf{x}$ es

$$\mathbf{x}(t) = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-2t}$$

o bien,

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 3e^{-.5t} + 2e^{-2t} \\ 6e^{-.5t} - 2e^{-2t} \end{bmatrix}$$

La figura 2 muestra la gráfica, o *trayectoria*, de $\mathbf{x}(t)$, para $t \geq 0$, junto con trayectorias para algunos otros puntos iniciales. Las trayectorias de las dos funciones propias $\mathbf{x}_1$ y $\mathbf{x}_2$ están en los espacios propios de A.

Las funciones $\mathbf{x}_1$ y $\mathbf{x}_2$ decaen a cero cuando $t \to \infty$, pero los valores de $\mathbf{x}_2$ decrecen más rápido porque su exponente es más negativo. Las entradas en el vector propio correspondiente $\mathbf{v}_2$ muestran que los voltajes en los capacitores decaerán a cero tan rápido como sea posible, si los voltajes iniciales son iguales en magnitud pero opuestos en signo. ■

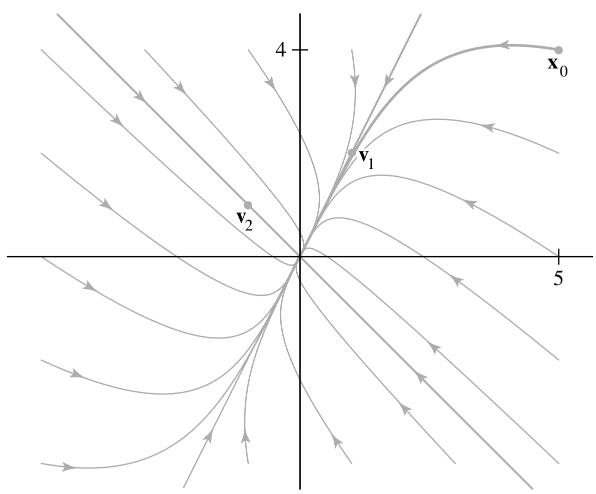


FIGURA 2 El origen como un atractor.

En la figura 2, el origen es un **atractor** o **sumidero** del sistema dinámico porque todas las trayectorias caen al origen. La dirección de mayor atracción está sobre la trayectoria de la función propia $\mathbf{x}_2$ (a lo largo de la recta que pasa por $\mathbf{0}$ y $\mathbf{v}_2$) correspondiente al valor propio más negativo, $\lambda = -2$. Las trayectorias que inician en puntos fuera de esta recta se convierten en asíntotas a la recta que pasa por $\mathbf{0}$ y $\mathbf{v}_1$, ya que sus componentes en la dirección $\mathbf{v}_2$ decaen muy rápidamente.

Si los valores propios del ejemplo 1 fueran positivos en vez de negativos, las trayectorias correspondientes serían similares en forma, aunque las trayectorias estarían *alejándose* del origen. En tal caso, el origen es un **repulsor**, o **fuente**, del sistema dinámico, y la dirección de mayor repulsión es la recta que contiene la trayectoria de la función propia asociada con el valor propio más positivo.

EJEMPLO 2 Suponga que una partícula se mueve en un campo de fuerza plano y que su vector de posición $\mathbf{x}$ satisface $\mathbf{x}' = A\mathbf{x}$ y $\mathbf{x}(0) = \mathbf{x}_0$, donde

$$A = \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix}, \qquad \mathbf{x}_0 = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Resuelva este problema de valor inicial para $t \geq 0$, y bosqueje la trayectoria de la partícula.

SOLUCIÓN Los valores propios de A resultan ser $\lambda_1 = 6$ y $\lambda_2 = -1$, con los vectores propios correspondientes $\mathbf{v}_1 = (-5, 2)$ y $\mathbf{v}_2 = (1, 1)$. Para cualesquiera constantes c_1 y c_2, la función

$$\mathbf{x}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} = c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

es una solución de $\mathbf{x}' = A\mathbf{x}$. Se requiere que c_1 y c_2 satisfagan $\mathbf{x}(0) = \mathbf{x}_0$, es decir,

$$c_1\begin{bmatrix} -5 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} -5 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 2.9 \\ 2.6 \end{bmatrix}$$

Los cálculos demuestran que $c_1 = -3/70$ y $c_2 = 188/70$, y así la función deseada es

$$\mathbf{x}(t) = \frac{-3}{70}\begin{bmatrix} -5 \\ 2 \end{bmatrix} e^{6t} + \frac{188}{70}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-t}$$

En la figura 3 se ilustran las trayectorias de $\mathbf{x}$ y otras soluciones. ■

En la figura 3, el origen es un **punto silla** del sistema dinámico porque algunas trayectorias primero se acercan al origen y, después, cambian de dirección y se alejan de él. Un punto silla se presenta cuando la matriz tiene valores propios tanto positivos como negativos. La dirección de mayor repulsión es la recta que pasa por $\mathbf{v}_1$ y $\mathbf{0}$, asociada con el valor propio positivo. La dirección de mayor atracción es la recta que pasa por $\mathbf{v}_2$ y $\mathbf{0}$, correspondiente al valor propio negativo.

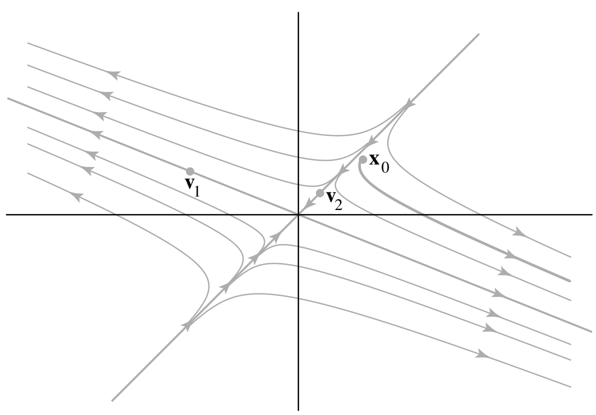


FIGURA 3 El origen como un punto silla.

Desacoplamiento de un sistema dinámico

El siguiente análisis muestra que el método de los ejemplos 1 y 2 produce un conjunto fundamental de soluciones para cualquier sistema dinámico descrito por $\mathbf{x}' = A\mathbf{x}$ cuando A es de $n \times n$ y tiene n vectores propios linealmente independientes, es decir, cuando A es diagonalizable. Suponga que las funciones propias para A son

$$\mathbf{v}_1 e^{\lambda_1 t}, \quad \ldots, \mathbf{v}_n e^{\lambda_n t}$$

con $\mathbf{v}_1,\ldots, \mathbf{v}_n$ vectores propios linealmente independientes. Sean $P = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ y D la matriz diagonal con entradas $\lambda_1,\ldots, \lambda_n$ tal que $A = PDP^{-1}$. Ahora se hace un *cambio de variable*, definiendo una nueva función $\mathbf{y}$ por

$$\mathbf{y}(t) = P^{-1}\mathbf{x}(t) \quad \text{o, de manera equivalente,} \quad \mathbf{x}(t) = P\mathbf{y}(t)$$

La ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ indica que $\mathbf{y}(t)$ es el vector de coordenadas de $\mathbf{x}(t)$ respecto a la base de vectores propios. La sustitución de $P\mathbf{y}$ para $\mathbf{x}$ en la ecuación $\mathbf{x}' = A\mathbf{x}$ da

$$\frac{d}{dt}(P\mathbf{y}) = A(P\mathbf{y}) = (PDP^{-1})P\mathbf{y} = PD\mathbf{y} \tag{5}$$

Como P es una matriz constante, el lado izquierdo de (5) es $P\mathbf{y}'$. Ambos lados de (5) se multiplican por la izquierda por P^{-1} y se obtiene $\mathbf{y}' = D\mathbf{y}$, o bien,

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}$$

El cambio de variable de $\mathbf{x}$ a $\mathbf{y}$ *desacopló* el sistema de ecuaciones diferenciales, porque la derivada de cada función escalar y_k tan solo depende de y_k. (Revise el cambio de variables análogo de la sección 5.6). Como $y_1' = \lambda_1 y_1$, se tiene que $y_1(t) = c_1 e^{\lambda_1 t}$, con fórmulas similares para $y_2,\ldots, y_n$. Así,

$$\mathbf{y}(t) = \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}, \quad \text{donde} \quad \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{y}(0) = P^{-1}\mathbf{x}(0) = P^{-1}\mathbf{x}_0$$

Para obtener la solución general $\mathbf{x}$ del sistema original, calcule

$$\begin{aligned} \mathbf{x}(t) &= P\mathbf{y}(t) = [\,\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n\,]\,\mathbf{y}(t) \\ &= c_1\mathbf{v}_1 e^{\lambda_1 t} + \cdots + c_n \mathbf{v}_n e^{\lambda_n t} \end{aligned}$$

Este es el desarrollo de funciones propias construidas como en el ejemplo 1.

Valores propios complejos

En el siguiente ejemplo, una matriz real A tiene un par de valores propios complejos λ y $\bar{\lambda}$, con vectores propios complejos asociados $\mathbf{v}$ y $\bar{\mathbf{v}}$. (Recuerde de la sección 5.5 que para una matriz real, los valores propios complejos y sus vectores propios correspondientes se presentan en pares conjugados). Entonces, dos soluciones de $\mathbf{x}' = A\mathbf{x}$ son

$$\mathbf{x}_1(t) = \mathbf{v}e^{\lambda t} \quad \text{y} \quad \mathbf{x}_2(t) = \bar{\mathbf{v}}e^{\bar{\lambda} t} \tag{6}$$

Utilizando una representación en serie de potencias de la función exponencial compleja se demuestra que $\mathbf{x}_2(t) = \overline{\mathbf{x}_1(t)}$. No obstante que las funciones propias complejas $\mathbf{x}_1$ y $\mathbf{x}_2$ son convenientes para algunos cálculos (sobre todo en ingeniería eléctrica), las funciones reales son más adecuadas para diversos fines. Afortunadamente, las partes real e imaginaria de $\mathbf{x}_1$ son soluciones (reales) de $\mathbf{x}' = A\mathbf{x}$, porque son combinaciones lineales de las soluciones de (6):

$$\operatorname{Re}(\mathbf{v}e^{\lambda t}) = \frac{1}{2}[\,\mathbf{x}_1(t) + \overline{\mathbf{x}_1(t)}\,], \qquad \operatorname{Im}(\mathbf{v}e^{\lambda t}) = \frac{1}{2i}[\,\mathbf{x}_1(t) - \overline{\mathbf{x}_1(t)}\,]$$

Para entender la naturaleza de $\operatorname{Re}(\mathbf{v}e^{\lambda t})$, recuerde que en cálculo para cualquier número x, la función exponencial e^x se determina mediante la serie de potencias:

$$e^x = 1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n + \cdots$$

la cual se utiliza para definir $e^{\lambda t}$ cuando λ es complejo:

$$e^{\lambda t} = 1 + (\lambda t) + \frac{1}{2!}(\lambda t)^2 + \cdots + \frac{1}{n!}(\lambda t)^n + \cdots$$

Escribiendo $\lambda = a + bi$ (con a y b reales), y utilizando series de potencias similares para las funciones coseno y seno, se obtiene que

$$e^{(a+bi)t} = e^{at} \cdot e^{ibt} = e^{at}(\cos bt + i \operatorname{sen} bt) \tag{7}$$

Por lo tanto,

$$\begin{aligned}\mathbf{v}e^{\lambda t} &= (\operatorname{Re}\mathbf{v} + i \operatorname{Im}\mathbf{v}) \cdot e^{at}(\cos bt + i \operatorname{sen} bt) \\ &= [\,(\operatorname{Re}\mathbf{v})\cos bt - (\operatorname{Im}\mathbf{v})\operatorname{sen} bt\,]e^{at} \\ &\quad + i[\,(\operatorname{Re}\mathbf{v})\operatorname{sen} bt + (\operatorname{Im}\mathbf{v})\cos bt\,]e^{at}\end{aligned}$$

Entonces, dos soluciones reales de $\mathbf{x}' = A\mathbf{x}$ son

$$\begin{aligned}\mathbf{y}_1(t) &= \operatorname{Re}\mathbf{x}_1(t) = [(\operatorname{Re}\mathbf{v})\cos bt - (\operatorname{Im}\mathbf{v})\operatorname{sen} bt]e^{at} \\ \mathbf{y}_2(t) &= \operatorname{Im}\mathbf{x}_1(t) = [(\operatorname{Re}\mathbf{v})\operatorname{sen} bt + (\operatorname{Im}\mathbf{v})\cos bt]e^{at}\end{aligned}$$

Se puede demostrar que $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes (cuando $b \neq 0$).[1]

EJEMPLO 3 El circuito de la figura 4 se describe con la ecuación

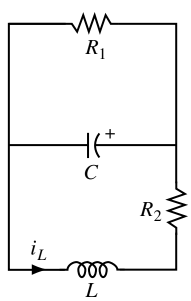


FIGURA 4

$$\begin{bmatrix} i_L' \\ v_C' \end{bmatrix} = \begin{bmatrix} -R_2/L & -1/L \\ 1/C & -1/(R_1C) \end{bmatrix}\begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente que pasa en el inductor L y v_C es la caída de voltaje en el capacitor C. Suponga que $R_1 = 5$ ohms, $R_2 = .8$ ohms, $C = .1$ farad y $L = .4$ henry. Encuentre las fórmulas para i_L y v_C, si la corriente inicial en el inductor es 3 amperes y el voltaje inicial en el capacitor es 3 volts.

SOLUCIÓN Para los datos dados, $A = \begin{bmatrix} -2 & -2.5 \\ 10 & -2 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$. El método analizado en la sección 5.5 produce el valor propio $\lambda = -2 + 5i$ y el vector propio correspondiente $\mathbf{v}_1 = \begin{bmatrix} i \\ 2 \end{bmatrix}$. Las soluciones complejas de $\mathbf{x}' = A\mathbf{x}$ son combinaciones lineales complejas de

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix}e^{(-2+5i)t} \quad \text{y} \quad \mathbf{x}_2(t) = \begin{bmatrix} -i \\ 2 \end{bmatrix}e^{(-2-5i)t}$$

---

[1] Puesto que $\mathbf{x}_2(t)$ es el complejo conjugado de $\mathbf{x}_1(t)$, las partes real e imaginaria de $\mathbf{x}_2(t)$ son $\mathbf{y}_1(t)$ y $-\mathbf{y}_2(t)$, respectivamente. Entonces, se utiliza ya sea $\mathbf{x}_1(t)$ o $\mathbf{x}_2(t)$, pero no ambas, para producir dos soluciones reales linealmente independientes de $\mathbf{x}' = A\mathbf{x}$.

Ahora, utilice la ecuación (7) para escribir

$$\mathbf{x}_1(t) = \begin{bmatrix} i \\ 2 \end{bmatrix} e^{-2t}(\cos 5t + i \operatorname{sen} 5t)$$

Las partes real e imaginaria de $\mathbf{x}_1$ dan soluciones reales:

$$\mathbf{y}_1(t) = \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t}, \qquad \mathbf{y}_2(t) = \begin{bmatrix} \cos 5t \\ 2 \operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Como $\mathbf{y}_1$ y $\mathbf{y}_2$ son funciones linealmente independientes, forman una base para el espacio vectorial bidimensional real de soluciones de $\mathbf{x}' = A\mathbf{x}$. Así, la solución general es

$$\mathbf{x}(t) = c_1 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \cos 5t \\ 2 \operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

Para satisfacer $\mathbf{x}(0) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, se necesita $c_1 \begin{bmatrix} 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, lo que conduce a $c_1 = 1.5$ y $c_2 = 3$. Por lo tanto,

$$\mathbf{x}(t) = 1.5 \begin{bmatrix} -\operatorname{sen} 5t \\ 2\cos 5t \end{bmatrix} e^{-2t} + 3 \begin{bmatrix} \cos 5t \\ 2 \operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

o bien,

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -1.5 \operatorname{sen} 5t + 3\cos 5t \\ 3\cos 5t + 6 \operatorname{sen} 5t \end{bmatrix} e^{-2t}$$

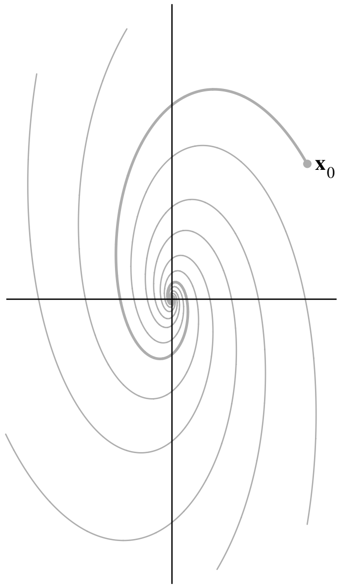


FIGURA 5
El origen como un punto espiral.

Véase la figura 5. ■

En la figura 5, el origen es un **punto espiral** del sistema dinámico. La rotación es causada por las funciones seno y coseno provenientes de un valor propio complejo. Las trayectorias en espiral entran, ya que el factor e^{-2t} tiende a cero. Recuerde que -2 es la parte real del valor propio del ejemplo 3. Cuando A tiene un valor propio complejo con parte real positiva, entonces las trayectorias en espiral son hacia fuera. Si la parte real del valor propio es cero, las trayectorias forman elipses alrededor del origen.

PROBLEMAS DE PRÁCTICA

A es una matriz real de 3×3 con valores propios $-.5$, $.2 + .3i$, y $.2 - .3i$, y vectores propios correspondientes

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 = \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix}$$

1. ¿A es diagonalizable como $A = PDP^{-1}$, utilizando matrices complejas?
2. Escriba la solución general de $\mathbf{x}' = A\mathbf{x}$ utilizando funciones propias complejas y, después, encuentre la solución general real.
3. Describa las formas de las trayectorias típicas.

5.7 EJERCICIOS

1. Una partícula que se mueve bajo un campo de fuerza plano tiene un vector de posición $\mathbf{x}$ que satisface $\mathbf{x}' = A\mathbf{x}$. La matriz A de 2×2 tiene valores propios 4 y 2, con sus vectores propios asociados $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Encuentre la posición de la partícula al tiempo t, suponiendo que $\mathbf{x}(0) = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$.

2. Sea A una matriz de 2×2 con valores propios -3 y -1 con vectores propios asociados $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Sea $\mathbf{x}(t)$ la posición de una partícula al tiempo t. Resuelva el problema de valor inicial $\mathbf{x}' = A\mathbf{x}$, $\mathbf{x}(0) = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

En los ejercicios 3 a 6, resuelva el problema de valor inicial $\mathbf{x}'(t) = A\mathbf{x}(t)$ para $t \geq 0$, con $\mathbf{x}(0) = (3, 2)$. Clasifique la naturaleza del origen como un atractor, repulsor o punto silla del sistema dinámico descrito por $\mathbf{x}' = A\mathbf{x}$. Encuentre las direcciones de mayor atracción y/o repulsión. Cuando el origen sea un punto silla, dibuje trayectorias típicas.

3. $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$ **4.** $A = \begin{bmatrix} -2 & -5 \\ 1 & 4 \end{bmatrix}$

5. $A = \begin{bmatrix} 7 & -1 \\ 3 & 3 \end{bmatrix}$ **6.** $A = \begin{bmatrix} 1 & -2 \\ 3 & -4 \end{bmatrix}$

En los ejercicios 7 y 8, haga un cambio de variable que desacople la ecuación $\mathbf{x}' = A\mathbf{x}$. Escriba la ecuación $\mathbf{x}(t) = P\mathbf{y}(t)$ y muestre el cálculo que conduce al sistema desacoplado $\mathbf{y}' = D\mathbf{y}$, especificando P y D.

7. A como en el ejercicio 5. **8.** A como en el ejercicio 6.

En los ejercicios 9 a 18, construya la solución general de $\mathbf{x}' = A\mathbf{x}$ que implique funciones propias complejas y, después, determine la solución general real. Describa las formas de trayectorias típicas.

9. $A = \begin{bmatrix} -3 & 2 \\ -1 & -1 \end{bmatrix}$ **10.** $A = \begin{bmatrix} 3 & 1 \\ -2 & 1 \end{bmatrix}$

11. $A = \begin{bmatrix} -3 & -9 \\ 2 & 3 \end{bmatrix}$ **12.** $A = \begin{bmatrix} -7 & 10 \\ -4 & 5 \end{bmatrix}$

13. $A = \begin{bmatrix} 4 & -3 \\ 6 & -2 \end{bmatrix}$ **14.** $A = \begin{bmatrix} -2 & 1 \\ -8 & 2 \end{bmatrix}$

15. [**M**] $A = \begin{bmatrix} -8 & -12 & -6 \\ 2 & 1 & 2 \\ 7 & 12 & 5 \end{bmatrix}$

16. [**M**] $A = \begin{bmatrix} -6 & -11 & 16 \\ 2 & 5 & -4 \\ -4 & -5 & 10 \end{bmatrix}$

17. [**M**] $A = \begin{bmatrix} 30 & 64 & 23 \\ -11 & -23 & -9 \\ 6 & 15 & 4 \end{bmatrix}$

18. [**M**] $A = \begin{bmatrix} 53 & -30 & -2 \\ 90 & -52 & -3 \\ 20 & -10 & 2 \end{bmatrix}$

19. [**M**] Encuentre fórmulas para los voltajes v_1 y v_2 (como funciones del tiempo t) en el circuito del ejemplo 1, suponiendo que $R_1 = 1/5$ ohm, $R_2 = 1/3$ ohm, $C_1 = 4$ farads, $C_2 = 3$ farads, y el voltaje inicial en cada capacitor es 4 volts.

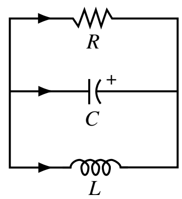
20. [**M**] Obtenga fórmulas para los voltajes v_1 y v_2 en el circuito del ejemplo 1, suponiendo que $R_1 = 1/15$ ohm, $R_2 = 1/3$ ohm, $C_1 = 9$ farads, $C_2 = 2$ farads y el voltaje inicial sobre cada capacitor es 3 volts.

21. [**M**] Encuentre fórmulas para la corriente i_L y el voltaje v_C para el circuito del ejemplo 3, suponiendo que $R_1 = 1$ ohm, $R_2 = .125$ ohm, $C = .2$ farad, $L = .125$ henry, la corriente inicial es 0 amp y el voltaje inicial es 15 volts.

22. [**M**] El circuito en la figura se describe por la ecuación

$$\begin{bmatrix} i'_L \\ v'_C \end{bmatrix} = \begin{bmatrix} 0 & 1/L \\ -1/C & -1/(RC) \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix}$$

donde i_L es la corriente en el inductor L y v_C es la caída de voltaje en el capacitor C. Obtenga fórmulas para i_L y v_C cuando $R = .5$ ohm, $C = 2.5$ farads, $L = .5$ henry, la corriente inicial es 0 amp y el voltaje inicial de 12 volts.



SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Sí, la matriz de 3×3 es diagonalizable ya que tiene tres valores propios distintos. El teorema 2 de la sección 5.1 y el teorema 5 de la sección 5.3 son válidos cuando se utilizan escalares complejos. (Las pruebas son esencialmente las mismas que para los escalares reales).
2. La solución general tiene la forma

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{(.2+.3i)t} + c_3 \begin{bmatrix} 1-2i \\ -4i \\ 2 \end{bmatrix} e^{(.2-.3i)t}$$

Aquí los escalares c_1, c_2, c_3 pueden ser números complejos cualesquiera. El primer término en $\mathbf{x}(t)$ es real. Es posible obtener dos soluciones reales adicionales utilizando las partes real e imaginaria del segundo término en $\mathbf{x}(t)$:

$$\begin{bmatrix} 1+2i \\ 4i \\ 2 \end{bmatrix} e^{.2t}(\cos .3\text{t} + i \operatorname{sen} .3t)$$

La solución general real tiene la siguiente forma, con escalares *reales* c_1, c_2, c_3:

$$c_1\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} e^{-.5t} + c_2 \begin{bmatrix} \cos .3t - 2 \operatorname{sen} .3t \\ -4 \operatorname{sen} .3t \\ 2\cos .3t \end{bmatrix} e^{.2t} + c_3 \begin{bmatrix} \operatorname{sen} .3t + 2\cos .3t \\ 4\cos .3t \\ 2 \operatorname{sen} .3t \end{bmatrix} e^{.2t}$$

3. Cualquier solución con $c_2 = c_3 = 0$ es atraída hacia el origen debido al factor exponencial negativo. Otras soluciones tienen componentes que crecen sin cota, en tanto que las trayectorias son espirales hacia afuera.

Sea cuidadoso para no confundir este problema con el de la sección 5.6. Ahí la condición para atracción hacia **0** fue que un valor propio era menor que 1 en magnitud, para que $|\lambda|^k \to 0$. Aquí la parte real del valor propio debe ser negativa, para que $e^{\lambda t} \to 0$.

5.8 ESTIMACIONES ITERATIVAS PARA VALORES PROPIOS

En las aplicaciones científicas del álgebra lineal, los valores propios rara vez se conocen con precisión. Por fortuna, una aproximación numérica cercana es bastante satisfactoria. En efecto, algunas aplicaciones tan solo requieren una aproximación burda al valor propio más grande. El primer algoritmo que se describe abajo sirve muy bien para este caso. Además, proporciona un fundamento para un método más poderoso que también aportaría estimaciones rápidas para los valores propios restantes.

El método de potencias

El método de potencias se aplica a una matriz A de $n \times n$, con un **valor propio estrictamente dominante** λ_1, lo cual significa que λ_1 debe ser, en valor absoluto, mucho más grande que todos los demás valores propios. En tal caso, el método de potencias produce una secuencia escalar que se aproxima a λ_1 y una secuencia vectorial que se aproxima al vector propio correspondiente. El antecedente para el método es la descomposición de vectores propios que se utilizó al inicio de la sección 5.6.

Por sencillez, suponga que A es diagonalizable y que $\mathbb{R}^n$ tiene una base de vectores propios $\mathbf{v}_1, \ldots, \mathbf{v}_n$, colocados de manera que los valores propios correspondientes $\lambda_1, \ldots, \lambda_n$ disminuyan de tamaño, con el primer valor propio estrictamente dominante. Es decir,

$$|\lambda_1| \underset{\text{estrictamente mayor}}{>} |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \tag{1}$$

Como se vio en la ecuación (2) de la sección 5.6, si $\mathbf{x}$ en $\mathbb{R}^n$ se escribe en la forma $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, entonces,

$$A^k\mathbf{x} = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n \quad (k = 1, 2, \ldots)$$

Suponga que $c_1 \neq 0$. Por lo tanto, dividiendo entre $(\lambda_1)^k$,

$$\frac{1}{(\lambda_1)^k}A^k\mathbf{x} = c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n \quad (k = 1, 2, \ldots) \tag{2}$$

De la desigualdad (1), las fracciones $\lambda_2/\lambda_1, \ldots, \lambda_n/\lambda_1$ son menores que 1 en magnitud y por ende sus potencias tienden a cero. Así,

$$(\lambda_1)^{-k}A^k\mathbf{x} \to c_1\mathbf{v}_1 \quad \text{como } k \to \infty \tag{3}$$

De manera que para k grande, un escalar múltiplo de $A^k\mathbf{x}$ casi determina la misma *dirección* como el vector propio $c_1\mathbf{v}_1$. Puesto que un múltiplo escalar positivo no cambia la dirección de un vector, entonces $A^k\mathbf{x}$ casi apunta en la misma dirección que $\mathbf{v}_1$ o $-\mathbf{v}_1$, si $c_1 \neq 0$.

EJEMPLO 1 Sean $A = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -.5 \\ 1 \end{bmatrix}$. Entonces, A tiene los valores propios 2 y 1, y el espacio propio para $\lambda_1 = 2$ es la línea que pasa por $\mathbf{0}$ y $\mathbf{v}_1$. Para $k = 0, \ldots, 8$, calcule $A^k\mathbf{x}$ y construya la línea por $\mathbf{0}$ y $A^k\mathbf{x}$. ¿Qué sucede conforme k se incrementa?

SOLUCIÓN Los primeros tres cálculos son

$$A\mathbf{x} = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.5 \\ 1 \end{bmatrix} = \begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$$

$$A^2\mathbf{x} = A(A\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix} = \begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$$

$$A^3\mathbf{x} = A(A^2\mathbf{x}) = \begin{bmatrix} 1.8 & .8 \\ .2 & 1.2 \end{bmatrix}\begin{bmatrix} .7 \\ 1.3 \end{bmatrix} = \begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$$

Cálculos semejantes permiten elaborar la tabla 1.

TABLA 1 Iteración de un vector

k	0	1	2	3	4	5	6	7	8
$A^k\mathbf{x}$	$\begin{bmatrix} -.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -.1 \\ 1.1 \end{bmatrix}$	$\begin{bmatrix} .7 \\ 1.3 \end{bmatrix}$	$\begin{bmatrix} 2.3 \\ 1.7 \end{bmatrix}$	$\begin{bmatrix} 5.5 \\ 2.5 \end{bmatrix}$	$\begin{bmatrix} 11.9 \\ 4.1 \end{bmatrix}$	$\begin{bmatrix} 24.7 \\ 7.3 \end{bmatrix}$	$\begin{bmatrix} 50.3 \\ 13.7 \end{bmatrix}$	$\begin{bmatrix} 101.5 \\ 26.5 \end{bmatrix}$

En la figura 1 se ilustran los vectores $\mathbf{x}, A\mathbf{x}, \ldots, A^4\mathbf{x}$. Los otros crecen demasiado y se dificulta dibujarlos. Sin embargo, se han trazado segmentos de recta para mostrar las direcciones de esos vectores. En efecto, las direcciones de los vectores es lo que realmente se desea observar, no los vectores mismos. Las rectas parecen aproximarse a la recta que representa el espacio propio generado por $\mathbf{v}_1$. Más precisamente, el ángulo entre la recta (subespacio) determinada por $A^k\mathbf{x}$ y la recta (espacio propio) definida por $\mathbf{v}_1$ tiende a cero cuando $k \to \infty$. ■

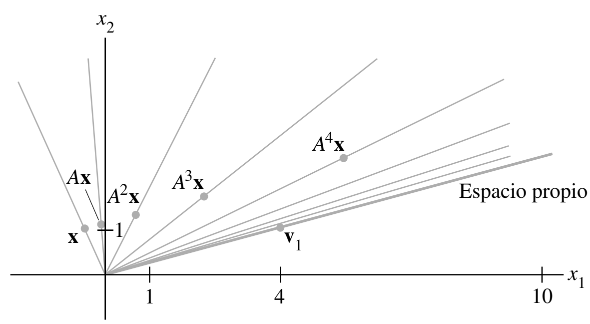


FIGURA 1 Direcciones determinadas por $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

Se escalan los vectores $(\lambda_1)^{-k}A^k\mathbf{x}$ en la ecuación (3) para hacerlos converger a $c_1\mathbf{v}_1$, si $c_1 \neq 0$. No se puede escalar $A^k\mathbf{x}$ en esta forma porque se desconoce λ_1. Sin embargo, sí es posible escalar cada $A^k\mathbf{x}$ para hacer que su entrada sea 1. Entonces, resulta que la secuencia $\{\boldsymbol{x}_k\}$ convergerá a un múltiplo de $\mathbf{v}_1$ cuya mayor entrada es 1. La figura 2 muestra la secuencia escalada para el ejemplo 1. El valor propio λ_1 también se estima a partir de la secuencia $\{\mathbf{x}_k\}$. Cuando $\mathbf{x}_k$ se acerca a un vector propio para λ_1, el vector $A\mathbf{x}_k$ está cerca de $\lambda_1\mathbf{x}_k$, con cada entrada en $A\mathbf{x}_k$ aproximadamente igual a λ_1 veces la entrada correspondiente en $\mathbf{x}_k$. Como la entrada más grande en $\mathbf{x}_k$ es 1, por lo tanto, la mayor entrada en $A\mathbf{x}_k$ es casi λ_1. (Se omiten pruebas cuidadosas de estos enunciados).

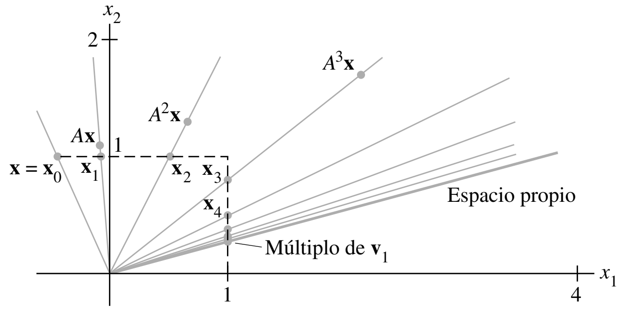


FIGURA 2 Múltiplos escalados de $\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \ldots, A^7\mathbf{x}$.

MÉTODO DE POTENCIAS PARA ESTIMAR UN VALOR PROPIO ESTRICTAMENTE DOMINANTE

1. Seleccione un vector inicial $\mathbf{x}_0$ cuya entrada más grande sea 1.
2. Para $k = 0, 1, \ldots,$
   a) Calcule $A\mathbf{x}_k$.
   b) Sea μ_k una entrada en $A\mathbf{x}_k$ cuyo valor absoluto es tan grande como sea posible.
   c) Determine $\mathbf{x}_{k+1} = (1/\mu_k)A\mathbf{x}_k$.
3. Para casi todas las elecciones de $\mathbf{x}_0$, la secuencia $\{\mu_k\}$ se aproxima al valor propio dominante, y la secuencia $\{\mathbf{x}_k\}$ tiende a un vector propio correspondiente.

EJEMPLO 2 Aplique el método de potencias a $A = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}$ con $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Pare cuando $k = 5$, y estime el valor propio dominante y un vector propio correspondiente de A.

SOLUCIÓN Los cálculos en este ejemplo y en el siguiente fueron realizados con MATLAB, que calcula con 16 dígitos de exactitud, no obstante que aquí tan solo se muestran pocos decimales. Para comenzar, calcule $A\mathbf{x}_0$ e identifique la entrada más grande μ_0 en $A\mathbf{x}_0$:

$$A\mathbf{x}_0 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad \mu_0 = 5$$

Escale $A\mathbf{x}_0$ por $1/\mu_0$ para obtener $\mathbf{x}_1$, calcule $A\mathbf{x}_1$ e identifique la mayor entrada en $A\mathbf{x}_1$:

$$\mathbf{x}_1 = \frac{1}{\mu_0}A\mathbf{x}_0 = \frac{1}{5}\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ .4 \end{bmatrix}$$

$$A\mathbf{x}_1 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .4 \end{bmatrix} = \begin{bmatrix} 8 \\ 1.8 \end{bmatrix}, \quad \mu_1 = 8$$

Escale $A\mathbf{x}_1$ por $1/\mu_1$ para obtener $\mathbf{x}_2$, determine $A\mathbf{x}_2$ e identifique la mayor entrada en $A\mathbf{x}_2$:

$$\mathbf{x}_2 = \frac{1}{\mu_1}A\mathbf{x}_1 = \frac{1}{8}\begin{bmatrix} 8 \\ 1.8 \end{bmatrix} = \begin{bmatrix} 1 \\ .225 \end{bmatrix}$$

$$A\mathbf{x}_2 = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .225 \end{bmatrix} = \begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}, \quad \mu_2 = 7.125$$

Escale $A\mathbf{x}_2$ por $1/\mu_2$ para obtener $\mathbf{x}_3$, y así sucesivamente. Los resultados de los cálculos con MATLAB para las primeras cinco iteraciones se dan en la tabla 2.

TABLA 2 Método de potencias para el ejemplo 2

k	**0**	**1**	**2**	**3**	**4**	**5**
$\mathbf{x}_k$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .4 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .225 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2035 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .2005 \end{bmatrix}$	$\begin{bmatrix} 1 \\ .20007 \end{bmatrix}$
$A\mathbf{x}_k$	$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 8 \\ 1.8 \end{bmatrix}$	$\begin{bmatrix} 7.125 \\ 1.450 \end{bmatrix}$	$\begin{bmatrix} 7.0175 \\ 1.4070 \end{bmatrix}$	$\begin{bmatrix} 7.0025 \\ 1.4010 \end{bmatrix}$	$\begin{bmatrix} 7.00036 \\ 1.40014 \end{bmatrix}$
μ_k	5	8	7.125	7.0175	7.0025	7.00036

La evidencia de la tabla 2 sugiere significativamente que $\{\mathbf{x}_k\}$ se aproxima a (1, .2) y que $\{\mu_k\}$ se aproxima a 7. Si es así, entonces (1, .2) es un vector propio y 7 es el valor propio dominante. Esto se comprueba fácilmente calculando

$$A\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ .2 \end{bmatrix} = \begin{bmatrix} 7 \\ 1.4 \end{bmatrix} = 7\begin{bmatrix} 1 \\ .2 \end{bmatrix} \qquad \blacksquare$$

La secuencia $\{\mu_k\}$ del ejemplo 2 tuvo convergencia rápida a $\lambda_1 = 7$ porque el segundo valor propio de A fue mucho más pequeño. (En efecto, $\lambda_2 = 1$). En general, la rapidez de convergencia depende de la razón $|\lambda_2/\lambda_1|$, porque en la ecuación (2) el vector $c_2(\lambda_2/\lambda_1)^k\mathbf{v}_2$ es la principal fuente de error, cuando se utiliza una versión escalada de $A^k\mathbf{x}$ como un estimador de $c_1\mathbf{v}_1$. (Las otras fracciones λ_j/λ_1 son mucho más pequeñas). Si $|\lambda_2/\lambda_1|$ es cercana a 1, entonces $\{\mu_k\}$ y $\{\mathbf{x}_k\}$ pueden converger muy lentamente y se preferirían otros métodos de aproximación.

Con el método de potencias, hay una leve posibilidad de que el vector inicial elegido $\mathbf{x}$ no tenga componente en la dirección $\mathbf{v}_1$ (cuando $c_1 = 0$). Pero el error de redondeo computacional durante los cálculos de $\mathbf{x}_k$ podría crear un vector con al menos una pequeña componente en la dirección de $\mathbf{v}_1$. Si eso ocurre, entonces $\mathbf{x}_k$ empezará a converger a un múltiplo de $\mathbf{v}_1$.

Método de potencias inverso

Este método brinda una aproximación para *cualquier* valor propio, si se tiene una buena estimación inicial α del valor propio λ. En tal caso, sea $B = (A - \alpha I)^{-1}$ y se le aplica el método de potencias a B. Se puede demostrar que si los valores propios de A son $\lambda_1, \ldots, \lambda_n$, entonces los valores propios de B son

$$\frac{1}{\lambda_1 - \alpha}, \quad \frac{1}{\lambda_2 - \alpha}, \quad \ldots, \quad \frac{1}{\lambda_n - \alpha}$$

y los vectores propios correspondientes coinciden con los de A. (Véase los ejercicios 15 y 16).

Suponga, por ejemplo, que α está más cerca de λ_2 que de los otros valores propios de A. Entonces, $1/(\lambda_2 - \alpha)$ será un valor propio estrictamente dominante de B. Si α está realmente cerca de λ_2, entonces $1/(\lambda_2 - \alpha)$ será *mucho* mayor que los demás valores propios de B, y el método de potencias inverso produce una aproximación rápida a λ_2 para casi todas las elecciones de $\mathbf{x}_0$. El siguiente algoritmo presenta los detalles.

MÉTODO DE POTENCIAS INVERSO PARA ESTIMAR UN VALOR PROPIO λ DE A

1. Seleccione una estimación inicial λ lo suficientemente cerca a λ.
2. Elija un vector inicial $\mathbf{x}_0$ cuya entrada más grande sea 1.
3. Para $k = 0, 1, \ldots,$
   a) Resuelva $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$.
   b) Sea μ_k una entrada en $\mathbf{y}_k$ cuyo valor absoluto es tan grande como sea posible.
   c) Calcule $\nu_k = \alpha + (1/\mu_k)$.
   d) Determine $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$.
4. Para casi todas las elecciones de $\mathbf{x}_0$, la secuencia $\{\nu_k\}$ se aproxima al valor propio λ de A, y la secuencia $\{\mathbf{x}_k\}$ tiende a un vector propio correspondiente.

Observe que B, o bien $(A - \alpha I)^{-1}$, no se encuentra en el algoritmo. En vez de calcular $(A - \alpha I)^{-1}\mathbf{x}_k$ para obtener el siguiente vector en la secuencia, es mucho mejor *resolver* la ecuación $(A - \alpha I)\mathbf{y}_k = \mathbf{x}_k$ para $\mathbf{y}_k$ (y después escalar $\mathbf{y}_k$ para producir $\mathbf{x}_{k+1}$). Como esta ecuación para $\mathbf{y}_k$ se debe resolver para cada k, entonces una factorización LU de $A - \alpha I$ acelerará el proceso.

EJEMPLO 3 En algunas aplicaciones no es poco común necesitar conocer el valor propio más pequeño de una matriz A y tener a mano estimaciones burdas de los valores propios. Suponga que 21, 3.3 y 1.9 son estimaciones para los valores propios de la matriz A que se muestra a continuación. Encuentre el valor propio más pequeño, exacto a seis decimales.

$$A = \begin{bmatrix} 10 & -8 & -4 \\ -8 & 13 & 4 \\ -4 & 5 & 4 \end{bmatrix}$$

SOLUCIÓN Los dos valores propios menores parecen estar cercanos entre sí, de modo que se emplea el método de potencias inverso para $A - 1.9I$. La tabla 3 muestra los resultados de los cálculos obtenidos con MATLAB. Aquí $\mathbf{x}_0$ se eligió arbitrariamente, $\mathbf{y}_k = (A - 1.9I)^{-1}\mathbf{x}_k$, μ_k es la entrada más grande en $\mathbf{y}_k$, $\mathbf{v}_k = 1.9 + 1/\mu_k$ y $\mathbf{x}_{k+1} = (1/\mu_k)\mathbf{y}_k$. Resulta que la estimación inicial del valor propio fue muy buena y la secuencia de potencias inversa convergió de manera rápida. El valor propio más pequeño es exactamente 2. ■

TABLA 3 Método de potencias inverso

k	0	1	2	3	4
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5736 \\ .0646 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5054 \\ .0045 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .5004 \\ .0003 \\ 1 \end{bmatrix}$	$\begin{bmatrix} .50003 \\ .00002 \\ 1 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 4.45 \\ .50 \\ 7.76 \end{bmatrix}$	$\begin{bmatrix} 5.0131 \\ .0442 \\ 9.9197 \end{bmatrix}$	$\begin{bmatrix} 5.0012 \\ .0031 \\ 9.9949 \end{bmatrix}$	$\begin{bmatrix} 5.0001 \\ .0002 \\ 9.9996 \end{bmatrix}$	$\begin{bmatrix} 5.000006 \\ .000015 \\ 9.999975 \end{bmatrix}$
μ_k	7.76	9.9197	9.9949	9.9996	9.999975
ν_k	2.03	2.0008	2.00005	2.000004	2.0000002

Si no se dispone de una estimación para el valor propio más pequeño de una matriz, se puede simplemente tomar $\alpha = 0$ en el método de potencias inverso. Esta elección de α funciona razonablemente bien, cuando el valor propio más pequeño está mucho más cerca de cero que los demás valores propios.

Los dos algoritmos presentados en esta sección son herramientas prácticas para muchas situaciones simples, y brindan una introducción al problema de la estimación de valores propios. Un método iterativo más robusto y ampliamente utilizado es el algoritmo QR. Por ejemplo, este es el corazón del comando `eig(A)` de MATLAB, que rápidamente calcula valores propios y vectores propios de A. En los ejercicios de la sección 5.2 se presenta una breve descripción del algoritmo QR. En casi todos los libros modernos de análisis numérico se presentan más detalles.

PROBLEMA DE PRÁCTICA

¿Cómo se podría decidir si un vector $\mathbf{x}$ dado es una buena aproximación a un vector propio de una matriz A? Si lo es, ¿cómo se estimaría el valor propio correspondiente? Experimente con

$$A = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} 1.0 \\ -4.3 \\ 8.1 \end{bmatrix}$$

5.8 EJERCICIOS

En los ejercicios 1 a 4, la matriz A va seguida por una secuencia $\{\mathbf{x}_k\}$ generada por el método de potencias. Utilice esos datos para estimar el mayor valor propio de A, y dé un vector propio correspondiente.

1. $A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .25 \end{bmatrix}, \begin{bmatrix} 1 \\ .3158 \end{bmatrix}, \begin{bmatrix} 1 \\ .3298 \end{bmatrix}, \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$

2. $A = \begin{bmatrix} 1.8 & -.8 \\ -3.2 & 4.2 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -.5625 \\ 1 \end{bmatrix}, \begin{bmatrix} -.3021 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2601 \\ 1 \end{bmatrix}, \begin{bmatrix} -.2520 \\ 1 \end{bmatrix}$

3. $A = \begin{bmatrix} .5 & .2 \\ .4 & .7 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ .8 \end{bmatrix}, \begin{bmatrix} .6875 \\ 1 \end{bmatrix}, \begin{bmatrix} .5577 \\ 1 \end{bmatrix}, \begin{bmatrix} .5188 \\ 1 \end{bmatrix}$

4. $A = \begin{bmatrix} 4.1 & -6 \\ 3 & -4.4 \end{bmatrix}$;

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .7368 \end{bmatrix}, \begin{bmatrix} 1 \\ .7541 \end{bmatrix}, \begin{bmatrix} 1 \\ .7490 \end{bmatrix}, \begin{bmatrix} 1 \\ .7502 \end{bmatrix}$

5. Sea $A = \begin{bmatrix} 15 & 16 \\ -20 & -21 \end{bmatrix}$. Los vectores $\mathbf{x}, \ldots, A^5\mathbf{x}$ son $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\begin{bmatrix} 31 \\ -41 \end{bmatrix}, \begin{bmatrix} -191 \\ 241 \end{bmatrix}, \begin{bmatrix} 991 \\ -1241 \end{bmatrix}, \begin{bmatrix} -4991 \\ 6241 \end{bmatrix}, \begin{bmatrix} 24991 \\ -31241 \end{bmatrix}$.

Encuentre un vector con un 1 en la segunda entrada que sea cercano a un vector propio de A. Use cuatro lugares decimales. Compruebe su respuesta y dé una estimación para el valor propio dominante de A.

6. Sea $A = \begin{bmatrix} -2 & -3 \\ 6 & 7 \end{bmatrix}$. Repita el ejercicio 5, utilizando la siguiente secuencia $\mathbf{x}, A\mathbf{x}, \ldots, A^5\mathbf{x}$.

$\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 13 \end{bmatrix}, \begin{bmatrix} -29 \\ 61 \end{bmatrix}, \begin{bmatrix} -125 \\ 253 \end{bmatrix}, \begin{bmatrix} -509 \\ 1021 \end{bmatrix}, \begin{bmatrix} -2045 \\ 4093 \end{bmatrix}$

[**M**] Los ejercicios 7 a 12 requieren MATLAB u otra herramienta computacional. En los ejercicios 7 y 8, utilice el método de potencias con $\mathbf{x}_0$ dado. Liste $\{\mathbf{x}_k\}$ y $\{\mu_k\}$ para $k = 1, \ldots, 5$. En los ejercicios 9 y 10, indique μ_5 y μ_6.

7. $A = \begin{bmatrix} 6 & 7 \\ 8 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

8. $A = \begin{bmatrix} 2 & 1 \\ 4 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

9. $A = \begin{bmatrix} 8 & 0 & 12 \\ 1 & -2 & 1 \\ 0 & 3 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 9 \\ 0 & 1 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Otra estimación se puede realizar para un valor propio cuando está disponible un vector propio aproximado. Observe que si $A\mathbf{x} = \lambda\mathbf{x}$, entonces $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(\lambda\mathbf{x}) = \lambda(\mathbf{x}^T\mathbf{x})$, y el **cociente de Rayleigh**

$$R(\mathbf{x}) = \frac{\mathbf{x}^T A\mathbf{x}}{\mathbf{x}^T\mathbf{x}}$$

es igual a λ. Si $\mathbf{x}$ es cercano a un vector propio para λ, entonces este cociente es cercano a λ. Cuando A es una matriz simétrica ($A^T = A$), el cociente de Rayleigh $R(\mathbf{x}_k) = (\mathbf{x}_k^T A\mathbf{x}_k)/(\mathbf{x}_k^T\mathbf{x}_k)$ tendrá aproximadamente el doble de dígitos de exactitud que el factor de escalamiento μ_k en el método de potencias. En los ejercicios 11 y 12 compruebe el aumento de precisión mediante el cálculo de μ_k y $R(\mathbf{x}_k)$ para $k = 1, \ldots, 4$.

11. $A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

12. $A = \begin{bmatrix} -3 & 2 \\ 2 & 0 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

Los ejercicios 13 y 14 se aplican a una matriz A de 3×3 cuyos valores propios se estiman como 4, -4 y 3.

13. Si se conoce que los valores propios cercanos a 4 y -4 tienen diferentes valores absolutos, ¿funcionará el método de potencias? ¿El método continúa siendo útil?

14. Suponga que se conoce que los valores propios cercanos a 4 y -4 tienen exactamente el mismo valor absoluto. Describa cómo se podría obtener una secuencia que estime el valor propio cercano a 4.

15. Suponga que $A\mathbf{x} = \lambda\mathbf{x}$ con $\mathbf{x} \neq \mathbf{0}$. Sean α un escalar distinto de los valores propios de A, y $B = (A - \alpha I)^{-1}$. Reste $\alpha\mathbf{x}$ en ambos miembros de la ecuación $A\mathbf{x} = \lambda\mathbf{x}$, y utilice álgebra para demostrar que $1/(\lambda - \alpha)$ es un valor propio de B, con $\mathbf{x}$ como un vector propio correspondiente.

16. Suponga que μ es un valor propio de B en el ejercicio 15, y que $\mathbf{x}$ es un vector propio correspondiente, de manera que $(A - \alpha I)^{-1}\mathbf{x} = \mu\mathbf{x}$. Utilice esta ecuación para encontrar un valor propio de A en términos de μ y α. [*Nota:* $\mu \neq 0$ porque B es invertible].

17. [**M**] Utilice el método de potencias inverso para estimar el valor propio intermedio de la A del ejemplo 3, con una exactitud a cuatro lugares decimales. Haga $\mathbf{x}_0 = (1, 0, 0)$.

18. [**M**] Sea A como en el ejercicio 9. Aplique el método de potencias inverso con $\mathbf{x}_0 = (1, 0, 0)$ para estimar el valor propio de A cercano a $\alpha = -1.4$, con una exactitud a cuatro decimales.

[**M**] En los ejercicios 19 y 20, determine: *a*) el valor propio más grande y *b*) el valor propio más cercano a cero. En cada caso, $\mathbf{x}_0 = (1, 0, 0, 0)$ y realice las aproximaciones hasta que la secuencia de acercamiento parezca exacta con cuatro decimales. Incluya el vector propio aproximado.

19. $A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}$

20. $A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 2 & 12 & 13 & 11 \\ -2 & 3 & 0 & 2 \\ 4 & 5 & 7 & 2 \end{bmatrix}$

21. Un error común es que si A tiene un valor propio estrictamente dominante, entonces, para cualquier k suficientemente grande, el vector $A^k\mathbf{x}$ es aproximadamente igual a un vector propio de A. Para las tres matrices que se muestran a continuación, estudie qué sucede a $A^k\mathbf{x}$ cuando $\mathbf{x} = (.5, .5)$, e intente obtener conclusiones generales (para una matriz de 2×2).

a) $A = \begin{bmatrix} .8 & 0 \\ 0 & .2 \end{bmatrix}$ *b*) $A = \begin{bmatrix} 1 & 0 \\ 0 & .8 \end{bmatrix}$

c) $A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Para A y $\mathbf{x}$ dados,

$$A\mathbf{x} = \begin{bmatrix} 5 & 8 & 4 \\ 8 & 3 & -1 \\ 4 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1.00 \\ -4.30 \\ 8.10 \end{bmatrix} = \begin{bmatrix} 3.00 \\ -13.00 \\ 24.50 \end{bmatrix}$$

Si $A\mathbf{x}$ es cercanamente un múltiplo de $\mathbf{x}$, entonces las razones de las entradas correspondientes en los dos vectores deberían ser casi constantes. Entonces, calcule:

{entrada en $A\mathbf{x}$}	$\div$	{entrada en $\mathbf{x}$}	=	{razón}
3.00		1.00		3.000
−13.00		−4.30		3.023
24.50		8.10		3.025

WEB

Cada entrada en $A\mathbf{x}$ es alrededor de tres veces la entrada correspondiente en $\mathbf{x}$, de manera que $\mathbf{x}$ es cercano a un vector propio de A. Cualesquiera de las razones anteriores es una estimación del valor propio. (Con cinco decimales, el valor propio es 3.02409).

CAPÍTULO 5 EJERCICIOS COMPLEMENTARIOS

En estos ejercicios suplementarios, A y B representan matrices cuadradas de tamaño adecuado.

1. Marque cada enunciado como verdadero o falso. Justifique cada respuesta.

a) Si A es invertible y 1 es uno de sus valores propios, entonces 1 también es valor propio de A^{-1}.

b) Si A es equivalente por filas a la matriz identidad I, entonces A es diagonalizable.

c) Si A contiene una fila o una columna de ceros, entonces 0 es un valor propio de A.

d) Cada valor propio de A también lo es de A^2.

e) Cada vector propio de A también lo es de A^2.

f) Cada vector propio de una matriz invertible A también lo es de A^{-1}.

g) Los valores propios deben ser escalares diferentes de cero.

h) Los vectores propios deben ser vectores distintos de cero.

i) Dos vectores propios correspondientes al mismo valor propio siempre son linealmente dependientes.

j) Las matrices similares siempre tienen exactamente los mismos valores propios.

k) Las matrices similares siempre tienen exactamente los mismos vectores propios.

l) La suma de dos vectores propios de una matriz A también es vector propio de esta.

m) Los valores propios de una matriz triangular A superior son exactamente las entradas distintas de cero sobre la diagonal de A.

n) Las matrices A y A^T tienen los mismos valores propios, contando sus multiplicidades.

o) Si una matriz A de 5×5 tiene menos de 5 valores propios distintos, entonces A no es diagonalizable.

p) Existe una matriz de 2×2 que no tiene vectores propios en $\mathbb{R}^2$.

q) Si A es diagonalizable, entonces las columnas de A son linealmente independientes.

r) Un vector distinto de cero no puede corresponder a dos valores propios diferentes de A.

s) Una matriz (cuadrada) A es invertible si y solo si existe un sistema de coordenadas donde la transformación $\mathbf{x} \mapsto A\mathbf{x}$ esté representada por una matriz diagonal.

t) Si cada vector $\mathbf{e}_j$ en la base estándar para $\mathbb{R}^n$ es un vector propio de A, entonces A es una matriz diagonal.

u) Si A es similar a una matriz diagonalizable B, entonces A también es diagonalizable.

v) Si A y B son matrices de $n \times n$ invertibles, entonces AB es similar a BA.

w) Una matriz de $n \times n$ con n vectores propios linealmente independientes es invertible.

x) Si A es una matriz de $n \times n$ diagonalizable, entonces cada vector en $\mathbb{R}^n$ se escribe como una combinación lineal de vectores propios de A.

2. Demuestre que si $\mathbf{x}$ es un vector propio de la matriz producto AB y $B\mathbf{x} \neq \mathbf{0}$, entonces $B\mathbf{x}$ es un vector propio de BA.

3. Suponga que $\mathbf{x}$ es un vector propio de A correspondiente a un valor propio λ.

a) Pruebe que $\mathbf{x}$ es un vector propio de $5I - A$. ¿Cuál es el valor propio asociado?

b) Demuestre que $\mathbf{x}$ es un vector propio de $5I - 3A + A^2$. ¿Cuál es el valor propio asociado?

4. Utilice inducción matemática para demostrar que si λ es un valor propio de una matriz A de $n \times n$, con $\mathbf{x}$ como un vector propio correspondiente, entonces, para cada entero positivo m, λ^m es un valor propio de A^m con $\mathbf{x}$ como un vector propio correspondiente.

5. Si $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$, defina $p(A)$ como la matriz formada al remplazar cada potencia de t en $p(t)$ por la potencia correspondiente de A (con $A^0 = I$). Es decir,

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n$$

Demuestre que si λ es un valor propio de A, entonces un valor propio de $p(A)$ es $p(\lambda)$.

6. Suponga que $A = PDP^{-1}$, donde P es 2×2 y $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.

a) Sea $B = 5I - 3A + A^2$. Demuestre que B es diagonalizable encontrando una factorización conveniente de B.

b) Dados $p(t)$ y $p(A)$ como en el ejercicio 5, demuestre que $p(A)$ es diagonalizable.

7. Suponga que A es diagonalizable y que $p(t)$ es el polinomio característico de A. Defina $p(A)$ como en el ejercicio 5, y pruebe que $p(A)$ es la matriz cero. Este hecho, el cual es válido para *cualquier* matriz cuadrada, se conoce como *teorema de Cayley-Hamilton*.

8. *a*) Sea A una matriz de $n \times n$ diagonalizable. Demuestre que si la multiplicidad de un valor propio λ es n, entonces $A = \lambda I$.

b) Utilice el inciso *a*) para demostrar que la matriz $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ no es diagonalizable.

9. Demuestre que $I - A$ es invertible cuando todos los valores propios de A son menores que 1 en magnitud. [*Sugerencia:* ¿Qué sería cierto si $I - A$ no fuera invertible?]

10. Demuestre que si A es diagonalizable, con todos los valores propios menores que 1 en magnitud, entonces A^k tiende a la matriz cero conforme $k \to \infty$. [*Sugerencia:* Considere $A^k\mathbf{x}$ donde $\mathbf{x}$ representa cualquiera de las columnas de I].

11. Sean $\mathbf{u}$ un vector propio de A correspondiente a un valor propio λ, y H la recta en $\mathbb{R}^n$ que pasa por $\mathbf{u}$ y por el origen.

a) Explique por qué H es invariante bajo A en el sentido de que $A\mathbf{x}$ está en H siempre que $\mathbf{x}$ esté en H.

b) Sea K el subespacio unidimensional de $\mathbb{R}^n$ que es invariante bajo A. Explique por qué K contiene un vector propio de A.

12. Sea $G = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$. Utilice la fórmula (1) para el determinante dado en la sección 5.2 para explicar por qué det $G =$ (det A)(det B). De esto, deduzca que el polinomio característico de G es el producto de los polinomios característicos de A y B.

Utilice el ejercicio 12 para encontrar los valores propios de las matrices en los ejercicios 13 y 14.

13. $A = \begin{bmatrix} 3 & -2 & 8 \\ 0 & 5 & -2 \\ 0 & -4 & 3 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 5 & -6 & -7 \\ 2 & 4 & 5 & 2 \\ 0 & 0 & -7 & -4 \\ 0 & 0 & 3 & 1 \end{bmatrix}$

15. Sea J la matriz de $n \times n$ con todas las entradas igual a 1, y considere $A = (a - b)I + bJ$; es decir,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Con los resultados del ejercicio 16 de los ejercicios complementarios del capítulo 3 demuestre que los valores propios de A son $a - b$ y $a + (n - 1)b$. ¿Cuáles son las multiplicidades de esos valores propios?

16. Aplique el resultado del ejercicio 15 para obtener los valores propios de las matrices $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$

y $\begin{bmatrix} 7 & 3 & 3 & 3 & 3 \\ 3 & 7 & 3 & 3 & 3 \\ 3 & 3 & 7 & 3 & 3 \\ 3 & 3 & 3 & 7 & 3 \\ 3 & 3 & 3 & 3 & 7 \end{bmatrix}$.

17. Sea $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Recuerde del ejercicio 25 de la sección 5.4 que tr A (la traza de A) es la suma de las entradas diagonales en A. Demuestre que el polinomio característico de A es

$\lambda^2 - (\text{tr } A)\lambda + \det A$

Entonces, demuestre que los valores propios de una matriz A de 2×2 son reales si y solo si $\det A \leq \left(\frac{\text{tr } A}{2}\right)^2$.

18. Sea $A = \begin{bmatrix} .4 & -.3 \\ .4 & 1.2 \end{bmatrix}$. Explique por qué A^k se aproxima a $\begin{bmatrix} -.5 & -.75 \\ 1.0 & 1.50 \end{bmatrix}$ cuando $k \to \infty$.

Los ejercicios 19 a 23 conciernen al polinomio

$p(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + t^n$

y a una matriz C_p de $n \times n$ llamada **matriz compañera** de p:

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

19. Escriba la matriz compañera C_p para $p(t) = 6 - 5t + t^2$ y, luego, determine el polinomio característico de C_p.

20. Sea $p(t) = (t - 2)(t - 3)(t - 4) = -24 + 26t - 9t^2 + t^3$. Escriba la matriz compañera para $p(t)$, y utilice las técnicas del capítulo 3 para obtener su polinomio característico.

21. Utilice inducción matemática y pruebe que para $n \geq 2$,

$$\begin{aligned} \det(C_p - \lambda I) &= (-1)^n(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) \\ &= (-1)^n p(\lambda) \end{aligned}$$

[*Sugerencia:* Desarrolle por cofactores sobre la primera columna, para demostrar que det $(C_p - \lambda I)$ tiene la forma $(-\lambda)B + (-1)^n a_0$, donde B es cierto polinomio (por la suposición de inducción)].

22. Sean $p(t) = a_0 + a_1 t + a_2 t^2 + t^3$, y λ un cero de p.

a) Escriba la matriz compañera para p.

b) Explique por qué $\lambda^3 = -a_0 - a_1\lambda - a_2\lambda^2$, y demuestre que $(1, \lambda, \lambda^2)$ es un vector propio de la matriz compañera de p.

23. Sea p el polinomio del ejercicio 22, y suponga que la ecuación $p(t) = 0$ tiene las raíces distintas $\lambda_1, \lambda_2, \lambda_3$. Sea V la matriz de Vandermonde

$$V = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

(La transpuesta de V se consideró en el ejercicio complementario 11 del capítulo 2). Utilice el ejercicio 22 y un teorema de este capítulo para deducir que V es invertible (pero no calcule V^{-1}). Después, explique por qué $V^{-1}C_pV$ es una matriz diagonal.

24. [**M**] El comando `roots(p)` de MATLAB calcula las raíces de la ecuación polinomial $p(t) = 0$. Lea un manual de MATLAB y, luego, describa la idea básica que sustenta el algoritmo para el comando `roots`.

25. [**M**] Aplique un programa matricial para diagonalizar

$$A = \begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$$

si es posible. Utilice el comando de valores propios para construir la matriz diagonal D. Si el programa tiene un comando que genere vectores propios, úselo para crear una matriz invertible P. Después, determine $AP - PD$ y PDP^{-1}. Analice sus resultados.

26. [**M**] Repita el ejercicio 25 para $A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$.

6

Ortogonalidad y mínimos cuadrados

EJEMPLO INTRODUCTORIO

Base de datos geográficos de Norteamérica y sistema de navegación GPS

Imagine que inicia un enorme proyecto que, según las estimaciones, durará 10 años y requiere los esfuerzos de muchas personas para construir y resolver un sistema de 1,800,000 por 900,000 ecuaciones lineales. Esto es exactamente lo que se hizo en 1974 con la encuesta geodésica nacional de Estados Unidos, cuando se actualizó el North American Datum (NAD), una red de 268,000 puntos de referencia localizados de manera precisa que generan el territorio completo de América del Norte, junto con Groenlandia, Hawai, las Islas Vírgenes, Puerto Rico y otras islas del Caribe.

Las latitudes y longitudes registradas en el NAD se deben determinar con una precisión dentro del rango de unos cuantos centímetros, ya que constituyen la base de encuestas, mapas, límites legales de propiedad y planos de proyectos de ingeniería civil, como los de carreteras e instalaciones de servicios públicos. Sin embargo, se tuvieron que agregar más de 200,000 nuevos puntos a los ya existentes desde la última actualización en 1927, y se acumularon errores con el paso de los años debido a mediciones imprecisas y a corrimientos de la corteza terrestre. En 1983 concluyó la recolección de datos para el reajuste del NAD.

El sistema de ecuaciones del NAD no tenía solución de la manera habitual, pero sí una *solución de mínimos cuadrados*, que asignó latitudes y longitudes a los puntos de referencia en la forma que mejor correspondía a las 1.8 millones de observaciones. En 1986 se encontró la solución de mínimos cuadrados al resolver un sistema relacionado de las llamadas *ecuaciones normales*, lo que implicó 928,735 ecuaciones con 928,735 variables.[1]

Más recientemente, el conocimiento de puntos de referencia en el suelo se ha vuelto crucial para determinar la ubicación exacta de satélites con el *Sistema de posicionamiento global* (*Global Positioning System*, GPS). Un satélite GPS calcula su posición en relación con la Tierra midiendo el tiempo que tardan las señales en llegar desde tres transmisores terrestres. Para hacer esto, los satélites utilizan relojes atómicos precisos que se han sincronizado con estaciones terrestres (cuyas ubicaciones se conocen con exactitud gracias al NAD).

El *Sistema de posicionamiento global* se utiliza para determinar las ubicaciones de los nuevos puntos de referencia terrestres y también para encontrar la ubicación del usuario sobre el suelo tomando como base los mapas ya existentes. Cuando el conductor de un automóvil (o un alpinista) enciende un receptor GPS, este último mide los tiempos de llegada de señales provenientes de al menos tres satélites. Esta información, junto con los datos transmitidos sobre las ubicaciones de los satélites y los tiempos del mensaje, se utiliza para ajustar el tiempo del receptor GPS y así determinar su ubicación aproximada sobre la Tierra. Con la información proveniente de un cuarto satélite, el receptor GPS puede incluso establecer su altura aproximada.

[1] Se presenta un análisis matemático de la estrategia de solución (junto con detalles de todo el proyecto NAD) en *North American Datum of 1983,* Charles R. Schwarz (ed.), National Geodetic Survey, National Oceanic and Atmospheric Administration (NOAA) Professional Paper NOS 2, 1989.

Los problemas del NAD y el GPS se resuelven encontrando un vector que "satisface aproximadamente" un sistema de ecuaciones inconsistente. Una cuidadosa explicación de esta aparente contradicción requerirá de las ideas desarrolladas en las primeras cinco secciones de este capítulo.

WEB

Para obtener una solución aproximada de un sistema de ecuaciones inconsistente que carece de solución, se necesita una idea bien definida de cercanía. La sección 6.1 introduce los conceptos de distancia y ortogonalidad en un espacio vectorial. Las secciones 6.2 y 6.3 muestran cómo se puede emplear la ortogonalidad para identificar aquel punto dentro de un subespacio W que es el más cercano al punto $\mathbf{y}$ externo a W. Considerando a W como el espacio columna de una matriz, la sección 6.5 desarrolla un método para obtener soluciones aproximadas ("mínimos cuadrados") de sistemas lineales inconsistentes, tal como el sistema resuelto para el informe NAD.

La sección 6.4 brinda otra oportunidad para ver proyecciones ortogonales en acción, creando una factorización matricial ampliamente utilizada en álgebra lineal numérica. Las secciones restantes examinan algunos de los muchos problemas de mínimos cuadrados que surgen en las aplicaciones, incluyendo aquellas en espacios más generales que $\mathbb{R}^n$.

6.1 PRODUCTO INTERIOR, LONGITUD Y ORTOGONALIDAD

Aquí se definen para $\mathbb{R}^n$ los conceptos geométricos de longitud, distancia y perpendicularidad, que son bien conocidos para $\mathbb{R}^2$ y $\mathbb{R}^3$. Esos conceptos ofrecen poderosas herramientas geométricas para resolver muchos problemas aplicados, incluyendo los problemas ya mencionados de mínimos cuadrados. Los tres conceptos se definen en términos del producto interior de dos vectores.

Producto interior

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$, entonces $\mathbf{u}$ y $\mathbf{v}$ se consideran matrices de $n \times 1$. La transpuesta $\mathbf{u}^T$ es una matriz de $1 \times n$, y el producto matricial $\mathbf{u}^T\mathbf{v}$ es una matriz de 1×1, que se representa como un solo número real (un escalar) sin corchetes. El número $\mathbf{u}^T\mathbf{v}$ se llama el **producto interior** de $\mathbf{u}$ y $\mathbf{v}$, y con frecuencia se representa como $\mathbf{u} \cdot \mathbf{v}$. Este producto interior, que se mencionó en los ejercicios de la sección 2.1, también se llama **producto punto**. Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces, el producto interior de $\mathbf{u}$ y $\mathbf{v}$ es

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

EJEMPLO 1 Calcule $\mathbf{u}\cdot\mathbf{v}$ y $\mathbf{v}\cdot\mathbf{u}$ para $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix}$.

SOLUCIÓN

$$\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = [\,2 \quad -5 \quad -1\,]\begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} = (2)(3) + (-5)(2) + (-1)(-3) = -1$$

$$\mathbf{v}\cdot\mathbf{u} = \mathbf{v}^T\mathbf{u} = [\,3 \quad 2 \quad -3\,]\begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix} = (3)(2) + (2)(-5) + (-3)(-1) = -1$$ ■

A partir de los cálculos en el ejemplo 1 resulta claro por qué $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$. La conmutatividad del producto interior es válida en general. Las siguientes propiedades del producto interior se deducen fácilmente de las características de la operación de transponer estudiada en la sección 2.1. (Véase los ejercicios 21 y 22 al final de esta sección).

TEOREMA 1

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$, y c un escalar. Entonces,

a) $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$

b) $(\mathbf{u} + \mathbf{v})\cdot\mathbf{w} = \mathbf{u}\cdot\mathbf{w} + \mathbf{v}\cdot\mathbf{w}$

c) $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\cdot(c\mathbf{v})$

d) $\mathbf{u}\cdot\mathbf{u} \geq 0$, y $\mathbf{u}\cdot\mathbf{u} = 0$ si y solo si $\mathbf{u} = \mathbf{0}$

Las propiedades *b*) y *c*) se pueden combinar varias veces para obtener la siguiente regla útil:

$$(c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{w} = c_1(\mathbf{u}_1\cdot\mathbf{w}) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{w})$$

Longitud de un vector

Si $\mathbf{v}$ está en $\mathbb{R}^n$, con entradas $v_1,\ldots, v_n$, entonces la raíz cuadrada de $\mathbf{v}\cdot\mathbf{v}$ está definida, ya que $\mathbf{v}\cdot\mathbf{v}$ es no negativo.

DEFINICIÓN

La **longitud** (o **norma**) de $\mathbf{v}$ es el escalar no negativo $\|\mathbf{v}\|$ definido por

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \quad \text{y} \quad \|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$$

Suponga que $\mathbf{v}$ está en $\mathbb{R}^2$, es decir, $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Como es usual, si se identifica a $\mathbf{v}$ con un punto geométrico en el plano, entonces $\|\mathbf{v}\|$ coincide con el concepto estándar de la longitud del segmento de recta que va del origen a $\mathbf{v}$. Esto se deduce a partir del teorema de Pitágoras aplicado a un triángulo como el de la figura 1.

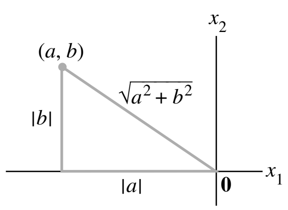


FIGURA 1
Interpretación de $\|\mathbf{v}\|$ como longitud.

Un cálculo similar con la diagonal de una caja rectangular demuestra que la definición de longitud de un vector $\mathbf{v}$ en $\mathbb{R}^3$ coincide con el concepto habitual de longitud.

Para cualquier escalar c, la longitud de $c\mathbf{v}$ es $|c|$ veces la longitud de $\mathbf{v}$. Es decir,

$$\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$$

(Para ver esto, calcule $\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2\mathbf{v}\cdot\mathbf{v} = c^2\|\mathbf{v}\|^2$ y obtenga raíces cuadradas).

Un **vector unitario** es un vector cuya longitud es 1. Si un vector **v** distinto de cero se *divide* entre su longitud —es decir, se multiplica por $1/\|\mathbf{v}\|$—, se obtiene un vector unitario **u** ya que la longitud de **u** es $(1/\|\mathbf{v}\|)\|\mathbf{v}\|$. El proceso de crear **u** a partir de **v** en ocasiones se llama **normalización de v**, y se dice que **u** está *en la misma dirección* que **v**.

Los ejemplos que se presentan a continuación emplean notación de vectores (columnas), para ahorrar espacio.

EJEMPLO 2 Sea $\mathbf{v} = (1, -2, 2, 0)$. Encuentre un vector unitario **u** en la misma dirección que **v**.

SOLUCIÓN Primero, calcule la longitud de **v**:

$$\begin{aligned}\|\mathbf{v}\|^2 &= \mathbf{v}\cdot\mathbf{v} = (1)^2 + (-2)^2 + (2)^2 + (0)^2 = 9\\ \|\mathbf{v}\| &= \sqrt{9} = 3\end{aligned}$$

Después, multiplique **v** por $1/\|\mathbf{v}\|$ para obtener

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v} = \frac{1}{3}\mathbf{v} = \frac{1}{3}\begin{bmatrix}1\\-2\\2\\0\end{bmatrix} = \begin{bmatrix}1/3\\-2/3\\2/3\\0\end{bmatrix}$$

Para comprobar que $\|\mathbf{u}\| = 1$ es suficiente probar que $\|\mathbf{u}\|^2 = 1$.

$$\begin{aligned}\|\mathbf{u}\|^2 &= \mathbf{u}\cdot\mathbf{u} = \left(\tfrac{1}{3}\right)^2 + \left(-\tfrac{2}{3}\right)^2 + \left(\tfrac{2}{3}\right)^2 + (0)^2\\ &= \tfrac{1}{9} + \tfrac{4}{9} + \tfrac{4}{9} + 0 = 1\end{aligned}$$

■

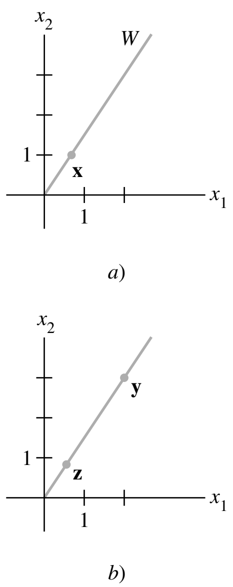


FIGURA 2
Normalización de un vector para obtener un vector unitario.

EJEMPLO 3 Sea W el subespacio de $\mathbb{R}^2$ generado por $\mathbf{x} = (\frac{2}{3}, 1)$. Obtenga un vector unitario **z** que sea una base para W.

SOLUCIÓN W consiste en todos los múltiplos de **x**, como el de la figura 2*a*). Cualquier vector distinto de cero en W es una base para W. Para simplificar los cálculos, "escale" **x** para eliminar fracciones. Es decir, multiplique **x** por 3 para obtener

$$\mathbf{y} = \begin{bmatrix}2\\3\end{bmatrix}$$

Ahora calcule $\|\mathbf{y}\|^2 = 2^2 + 3^2 = 13$, $\|\mathbf{y}\| = \sqrt{13}$, y normalice **y** para obtener

$$\mathbf{z} = \frac{1}{\sqrt{13}}\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}2/\sqrt{13}\\3/\sqrt{13}\end{bmatrix}$$

Véase la figura 2*b*). Otro vector unitario es $(-2/\sqrt{13}, -3/\sqrt{13})$. ■

Distancia en $\mathbb{R}^n$

Ahora ya estamos listos para describir qué tan cerca está un vector de otro. Recuerde que si a y b son números reales, la distancia sobre la recta numérica entre a y b es el número $|a - b|$. En la figura 3 se ilustran dos ejemplos. Esta definición de distancia en $\mathbb{R}$ tiene una analogía directa en $\mathbb{R}^n$.

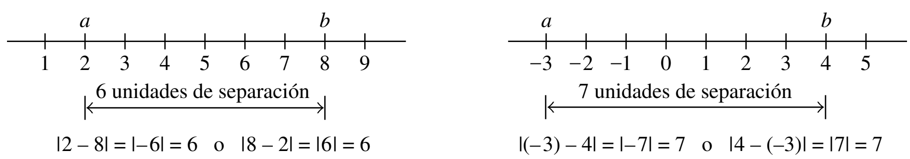


FIGURA 3 Distancias en $\mathbb{R}$.

DEFINICIÓN

> Para **u** y **v** en $\mathbb{R}^n$, la **distancia entre u y v**, que se escribe como dist(**u**, **v**), es la longitud del vector **u** − **v**. Es decir,
>
> $$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

En $\mathbb{R}^2$ y $\mathbb{R}^3$, esta definición de distancia coincide con las fórmulas usuales de la distancia euclidiana entre dos puntos, como muestran los siguientes dos ejemplos.

EJEMPLO 4 Determine la distancia entre los vectores $\mathbf{u} = (7, 1)$ y $\mathbf{v} = (3, 2)$.

SOLUCIÓN Calcule

$$\mathbf{u} - \mathbf{v} = \begin{bmatrix} 7 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{4^2 + (-1)^2} = \sqrt{17}$$

La figura 4 muestra los vectores **u**, **v** y **u** − **v**. Cuando el vector **u** − **v** se suma a **v**, el resultado es **u**. Observe que en la figura 4 el paralelogramo revela que la distancia de **u** a **v** es igual a la distancia de **u** − **v** a **0**. ■

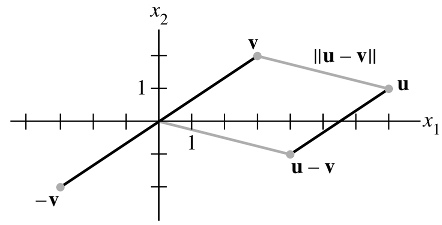


FIGURA 4 La distancia entre **u** y **v** es la longitud de **u** − **v**.

EJEMPLO 5 Si $\mathbf{u} = (u_1, u_2, u_3)$ y $\mathbf{v} = (v_1, v_2, v_3)$, entonces

$$\begin{aligned} \text{dist}(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} \\ &= \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + (u_3 - v_3)^2} \end{aligned}$$

■

Vectores ortogonales

El resto de este capítulo depende del hecho de que el concepto de rectas perpendiculares de geometría euclidiana tiene un análogo en $\mathbb{R}^n$.

Considere $\mathbb{R}^2$ o $\mathbb{R}^3$ y dos rectas que pasan por el origen determinado por los vectores **u** y **v**. Las dos rectas que se muestran en la figura 5 son geométricamente perpendiculares si y solo si la distancia de **u** a **v** es la misma que la distancia de **u** a −**v**. Esto es lo mismo que requerir que los cuadrados de las distancias sean iguales. Ahora

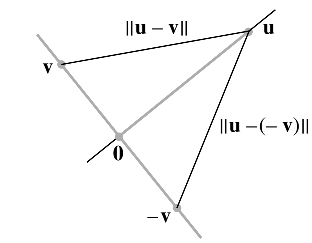


FIGURA 5

$$\begin{aligned} [\,\text{dist}(\mathbf{u}, -\mathbf{v})\,]^2 &= \|\mathbf{u} - (-\mathbf{v})\|^2 = \|\mathbf{u} + \mathbf{v}\|^2 \\ &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot (\mathbf{u} + \mathbf{v}) + \mathbf{v} \cdot (\mathbf{u} + \mathbf{v}) && \text{Teorema 1(b)} \\ &= \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v} && \text{Teorema 1(a), (b)} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathbf{u} \cdot \mathbf{v} && \text{Teorema 1(a)} \qquad (1) \end{aligned}$$

Los mismos cálculos con **v** y –**v** intercambiados indican que

$$\begin{aligned}[\text{dist}(\mathbf{u},\mathbf{v})]^2 &= \|\mathbf{u}\|^2 + \|-\mathbf{v}\|^2 + 2\mathbf{u}\cdot(-\mathbf{v}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\mathbf{u}\cdot\mathbf{v}\end{aligned}$$

Las dos distancias al cuadrado son iguales si y solo si $2\mathbf{u}\cdot\mathbf{v} = -2\mathbf{u}\cdot\mathbf{v}$, lo que ocurre si y solo si $\mathbf{u}\cdot\mathbf{v} = 0$.

Este cálculo demuestra que cuando los vectores **u** y **v** se identifican con puntos geométricos, las rectas correspondientes que pasan por los puntos y por el origen son perpendiculares si y solo si $\mathbf{u}\cdot\mathbf{v} = 0$. La siguiente definición generaliza a $\mathbb{R}^n$ este concepto de perpendicularidad (u *ortogonalidad*, como se le llama con frecuencia en álgebra lineal).

DEFINICIÓN

Dos vectores **u** y **v** en $\mathbb{R}^n$ son **ortogonales** (entre sí) si $\mathbf{u}\cdot\mathbf{v} = 0$.

Observe que el vector cero es ortogonal a todo vector en $\mathbb{R}^n$ porque $\mathbf{0}^T\mathbf{v} = 0$ para toda **v**.

El siguiente teorema expone un dato útil en relación con los vectores ortogonales. La demostración se deduce inmediatamente de los cálculos en la ecuación (1) y de la definición de ortogonalidad. El triángulo rectángulo que se ilustra en la figura 6 permite visualizar las longitudes que aparecen en el teorema.

TEOREMA 2

Teorema de Pitágoras

Dos vectores **u** y **v** son ortogonales si y solo si $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

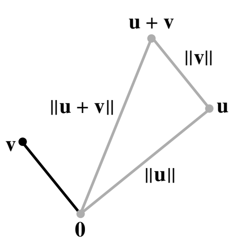


FIGURA 6

Complementos ortogonales

Para practicar el uso de productos interiores, aquí se introduce un concepto que será útil en la sección 6.3 y en todo el capítulo. Si un vector **z** es ortogonal a todo vector en un subespacio W de $\mathbb{R}^n$, entonces se dice que **z** es **ortogonal a** W. El conjunto de todos los vectores **z** que son ortogonales a W se llama **complemento ortogonal** de W y se denota con $W^\perp$ (que se lee "perpendicular a W").

EJEMPLO 6 Sean W un plano a través del origen en $\mathbb{R}^3$, y L la recta que pasa por el origen y es perpendicular a W. Si **z** y **w** son distintos de cero, **z** está sobre L y **w** está en W, entonces el segmento de recta de **0** a **z** es perpendicular al segmento de recta de **0** a **w**; es decir, $\mathbf{z}\cdot\mathbf{w} = 0$. Véase la figura 7. De manera que todo vector sobre L es ortogonal a cada **w** en W. En efecto, L consiste en *todos* los vectores que son ortogonales a todos los vectores **w** en W, y W consiste en todos los vectores ortogonales a todos los vectores **z** en L. Es decir,

$$L = W^\perp \quad \text{y} \quad W = L^\perp \qquad \blacksquare$$

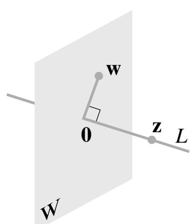


FIGURA 7
Un plano y una recta que pasan por **0** como complementos ortogonales.

Los siguientes dos hechos sobre $W^\perp$, con W como un subespacio de $\mathbb{R}^n$, se necesitarán más adelante en este capítulo. Las demostraciones se sugieren en los ejercicios 29 y 30. Los ejercicios 27 a 31 ofrecen una excelente oportunidad de practicar con las propiedades del producto interior.

1. Un vector **x** está en $W^\perp$ si y solo si **x** es ortogonal a cada vector de cualquier conjunto que genera a W.
2. $W^\perp$ es un subespacio de $\mathbb{R}^n$.

El siguiente teorema y el ejercicio 31 comprueban las afirmaciones hechas en la sección 4.6 en relación con los subespacios que se muestran en la figura 8. (Véase también el ejercicio 28 de la sección 4.6).

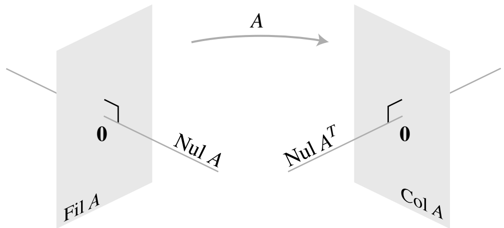


FIGURA 8 Los subespacios fundamentales determinados por una matriz A de $m \times n$.

TEOREMA 3

> Sea A una matriz de $m \times n$. El complemento ortogonal del espacio fila de A es el espacio nulo de A, y el complemento ortogonal del espacio columna de A es el espacio nulo de A^T:
>
> $$(\text{Fil } A)^{\perp} = \text{Nul } A \quad \text{y} \quad (\text{Col } A)^{\perp} = \text{Nul } A^T$$

DEMOSTRACIÓN La regla fila-columna para calcular $A\mathbf{x}$ indica que si $\mathbf{x}$ está en Nul A, entonces $\mathbf{x}$ es ortogonal a cada fila de A (con las filas consideradas como vectores en $\mathbb{R}^n$). Como las filas de A generan el espacio fila, entonces $\mathbf{x}$ es ortogonal a Fil A. Y a la inversa, si $\mathbf{x}$ es ortogonal a Fil A, entonces $\mathbf{x}$, evidentemente, es ortogonal a cada fila de A, y por lo tanto, $A\mathbf{x} = \mathbf{0}$. Esto demuestra el primer enunciado del teorema. Ya que este enunciado es verdadero para cualquier matriz, es válido para A^T. Es decir, el complemento ortogonal del espacio fila de A^T es el espacio nulo de A^T. Esto demuestra la segunda afirmación porque Fil A^T = Col A. ■

Ángulos en $\mathbb{R}^2$ y $\mathbb{R}^3$ (opcional)

Si $\mathbf{u}$ y $\mathbf{v}$ son vectores distintos de cero en $\mathbb{R}^2$ o $\mathbb{R}^3$, entonces existe una agradable conexión entre su producto interior y el ángulo ϑ entre los dos segmentos de recta que van del origen a los puntos identificados con $\mathbf{u}$ y $\mathbf{v}$. La fórmula es

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta \tag{2}$$

Si se desea comprobar esta fórmula para vectores en $\mathbb{R}^2$, considere el triángulo que se ilustra en la figura 9, con lados de longitudes $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ y $\|\mathbf{u} - \mathbf{v}\|$. De acuerdo con la ley de los cosenos,

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta$$

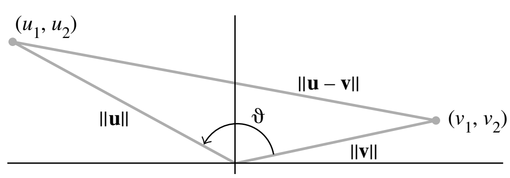


FIGURA 9 El ángulo entre dos vectores.

que se puede reordenar para obtener

$$\begin{aligned}\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta &= \frac{1}{2}\left[\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right] \\ &= \frac{1}{2}\left[u_1^2 + u_2^2 + v_1^2 + v_2^2 - (u_1 - v_1)^2 - (u_2 - v_2)^2\right] \\ &= u_1v_1 + u_2v_2 \\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned}$$

La comprobación es similar para $\mathbb{R}^3$. Cuando $n > 3$, la fórmula (2) se puede utilizar para *definir* el ángulo entre dos vectores en $\mathbb{R}^n$. En estadística, por ejemplo, el valor de cos ϑ definido por la ecuación (2) para vectores convenientes **u** y **v** es lo que los especialistas en estadística llaman *coeficiente de correlación*.

PROBLEMAS DE PRÁCTICA

Sean $\mathbf{a} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix}$ y $\mathbf{d} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}$.

1. Calcule $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$ y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a}$.
2. Encuentre un vector unitario **u** en la dirección de **c**.
3. Demuestre que **d** es ortogonal a **c**.
4. Utilice los resultados de los problemas de práctica 2 y 3 para explicar por qué **d** debe ser ortogonal al vector unitario **u**.

6.1 EJERCICIOS

En los ejercicios 1 a 8, determine las cantidades indicadas utilizando los vectores

$$\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix},\quad \mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix},\quad \mathbf{w} = \begin{bmatrix} 3 \\ -1 \\ -5 \end{bmatrix},\quad \mathbf{x} = \begin{bmatrix} 6 \\ -2 \\ 3 \end{bmatrix}$$

1. $\mathbf{u}\cdot\mathbf{u}$, $\mathbf{v}\cdot\mathbf{u}$ y $\dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$
2. $\mathbf{w}\cdot\mathbf{w}$, $\mathbf{x}\cdot\mathbf{w}$ y $\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}$
3. $\dfrac{1}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}$
4. $\dfrac{1}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$
5. $\left(\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$
6. $\left(\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{x}\cdot\mathbf{x}}\right)\mathbf{x}$
7. $\|\mathbf{w}\|$
8. $\|\mathbf{x}\|$

En los ejercicios 9 a 12, obtenga un vector unitario en la dirección del vector indicado.

9. $\begin{bmatrix} -30 \\ 40 \end{bmatrix}$
10. $\begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}$
11. $\begin{bmatrix} 7/4 \\ 1/2 \\ 1 \end{bmatrix}$
12. $\begin{bmatrix} 8/3 \\ 2 \end{bmatrix}$

13. Encuentre la distancia entre $\mathbf{x} = \begin{bmatrix} 10 \\ -3 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} -1 \\ -5 \end{bmatrix}$.

14. Calcule la distancia entre $\mathbf{u} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$ y $\mathbf{z} = \begin{bmatrix} -4 \\ -1 \\ 8 \end{bmatrix}$.

En los ejercicios 15 a 18, determine qué pares de vectores son ortogonales.

15. $\mathbf{a} = \begin{bmatrix} 8 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$
16. $\mathbf{u} = \begin{bmatrix} 12 \\ 3 \\ -5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}$
17. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -4 \\ 1 \\ -2 \\ 6 \end{bmatrix}$
18. $\mathbf{y} = \begin{bmatrix} -3 \\ 7 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} 1 \\ -8 \\ 15 \\ -7 \end{bmatrix}$

En los ejercicios 19 y 20, todos los vectores están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

19. *a*) $\mathbf{v}\cdot\mathbf{v} = \|\mathbf{v}\|^2$.
    b) Para cualquier escalar c, $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$.
    c) Si la distancia de **u** a **v** es igual a la distancia de **u** a $-\mathbf{v}$, entonces **u** y **v** son ortogonales.
    d) Para una matriz cuadrada A, los vectores en Col A son ortogonales a los vectores en Nul A.
    e) Si los vectores $\mathbf{v}_1,\ldots, \mathbf{v}_p$ generan un subespacio W y si **x** es ortogonal a cada $\mathbf{v}_j$ para $j = 1,\ldots, p$, entonces **x** está en $W^{\perp}$.

20. *a*) $\mathbf{u}\cdot\mathbf{v}$ -$\mathbf{v}\cdot\mathbf{u} = 0$.

b) Para cualquier escalar c, $\|c\mathbf{v}\| = c\|\mathbf{v}\|$.

c) Si $\mathbf{x}$ es ortogonal a cada vector en un subespacio W, entonces $\mathbf{x}$ está en $W^\perp$.

d) Si $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

e) Si A es una matriz de $m \times n$, los vectores en el espacio nulo de A son ortogonales a los vectores en el espacio fila de A.

21. Utilice la definición de transpuesta del producto interior para comprobar los incisos *b*) y *c*) del teorema 1. Mencione los hechos pertinentes del capítulo 2.

22. Sea $\mathbf{u} = (u_1, u_2, u_3)$. Explique por qué $\mathbf{u}\cdot\mathbf{u} \geq 0$. ¿Cuándo $\mathbf{u}\cdot\mathbf{u} = 0$?

23. Sean $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$. Calcule y compare $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$, $\|\mathbf{v}\|^2$ y $\|\mathbf{u} + \mathbf{v}\|^2$. No utilice el teorema de Pitágoras.

24. Compruebe la *ley del paralelogramo* para vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$:

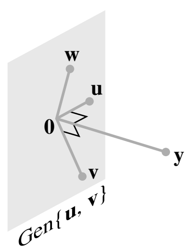
$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

25. Sea $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Describa el conjunto H de vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ que son ortogonales a $\mathbf{v}$. [*Sugerencia:* Considere $\mathbf{v} = \mathbf{0}$ y $\mathbf{v} \neq \mathbf{0}$].

26. Sean $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, y W el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^3$ tales que $\mathbf{u}\cdot\mathbf{x} = 0$. ¿Qué teorema del capítulo 4 se puede utilizar para demostrar que W es un subespacio de $\mathbb{R}^3$? Describa a W en lenguaje geométrico.

27. Suponga que un vector $\mathbf{y}$ que es ortogonal a los vectores $\mathbf{u}$ y $\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal al vector $\mathbf{u} + \mathbf{v}$.

28. Suponga que $\mathbf{y}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal a cada $\mathbf{w}$ en Gen $\{\mathbf{u}, \mathbf{v}\}$. [*Sugerencia:* Una $\mathbf{w}$ arbitraria en Gen $\{\mathbf{u}, \mathbf{v}\}$ tiene la forma $\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$. Demuestre que $\mathbf{y}$ es ortogonal a dicho vector $\mathbf{w}$].



29. Sea W = Gen $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Demuestre que si $\mathbf{x}$ es ortogonal a cada $\mathbf{v}_j$, para $1 \leq j \leq p$, entonces $\mathbf{x}$ es ortogonal a todo vector en W.

30. Sean W un subespacio de $\mathbb{R}^n$, y $W^\perp$ el conjunto de todos los vectores ortogonales a W. Demuestre que $W^\perp$ es un subespacio de $\mathbb{R}^n$, considerando los siguientes pasos.

a) Tome $\mathbf{z}$ en $W^\perp$, y sea $\mathbf{u}$ cualquier elemento de W. Entonces $\mathbf{z}\cdot\mathbf{u} = 0$. Tome cualquier escalar c y demuestre que $c\mathbf{z}$ es ortogonal a $\mathbf{u}$. (Puesto que $\mathbf{u}$ era un elemento arbitrario de W, esto demostrará que $c\mathbf{z}$ está en $W^\perp$).

b) Tome $\mathbf{z}_1$ y $\mathbf{z}_2$ en $W^\perp$, y sea $\mathbf{u}$ cualquier elemento de W. Demuestre que $\mathbf{z}_1 + \mathbf{z}_2$ es ortogonal a $\mathbf{u}$. ¿Qué se puede concluir acerca de $\mathbf{z}_1 + \mathbf{z}_2$? ¿Por qué?

c) Termine la demostración de que $W^\perp$ es un subespacio de $\mathbb{R}^n$.

31. Demuestre que si $\mathbf{x}$ está en W y $W^\perp$, entonces $\mathbf{x} = \mathbf{0}$.

32. [**M**] Construya un par $\mathbf{u}$, $\mathbf{v}$ de vectores aleatorios en $\mathbb{R}^4$, y sea

$$A = \begin{bmatrix} .5 & .5 & .5 & .5 \\ .5 & .5 & -.5 & -.5 \\ .5 & -.5 & .5 & -.5 \\ .5 & -.5 & -.5 & .5 \end{bmatrix}$$

a) Denote las columnas de A como $\mathbf{a}_1, \ldots, \mathbf{a}_4$. Obtenga la longitud de cada columna, y calcule $\mathbf{a}_1\cdot\mathbf{a}_2$, $\mathbf{a}_1\cdot\mathbf{a}_3$, $\mathbf{a}_1\cdot\mathbf{a}_4$, $\mathbf{a}_2\cdot\mathbf{a}_3$, $\mathbf{a}_2\cdot\mathbf{a}_4$ y $\mathbf{a}_3\cdot\mathbf{a}_4$.

b) Calcule y compare las longitudes de $\mathbf{u}$, $A\mathbf{u}$, $\mathbf{v}$ y $A\mathbf{v}$.

c) Utilice la ecuación (2) de esta sección para calcular el coseno del ángulo entre $\mathbf{u}$ y $\mathbf{v}$. Compare esto con el coseno del ángulo entre $A\mathbf{u}$ y $A\mathbf{v}$.

d) Repita los incisos *b*) y *c*) para otros dos pares de vectores aleatorios. ¿Qué se puede suponer del efecto de A sobre los vectores?

33. [**M**] Genere vectores aleatorios $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{v}$ en $\mathbb{R}^4$ con entradas enteras (y $\mathbf{v} \neq \mathbf{0}$), y calcule las cantidades

$$\left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \left(\frac{\mathbf{y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v},\ \frac{(10\mathbf{x})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

Repita los cálculos con nuevos vectores aleatorios $\mathbf{x}$ y $\mathbf{y}$. ¿Qué se puede suponer acerca del mapeo $\mathbf{x} \mapsto T(\mathbf{x}) = \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$ (para $\mathbf{v} \neq \mathbf{0}$)? Compruebe su suposición algebraicamente.

34. [**M**] Sea $A = \begin{bmatrix} -6 & 3 & -27 & -33 & -13 \\ 6 & -5 & 25 & 28 & 14 \\ 8 & -6 & 34 & 38 & 18 \\ 12 & -10 & 50 & 41 & 23 \\ 14 & -21 & 49 & 29 & 33 \end{bmatrix}$. Construya una matriz N cuyas columnas formen una base para Nul A, y elabore una matriz R cuyas *filas* formen una base para Fil A (para detalles, véase la sección 4.6). Efectúe un cálculo matricial con N y R que muestre un hecho del teorema 3.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $\mathbf{a}\cdot\mathbf{b} = 7$, $\mathbf{a}\cdot\mathbf{a} = 5$. Así, $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}} = \dfrac{7}{5}$ y $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a} = \dfrac{7}{5}\mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Escale c, multiplicando por 3 para obtener $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Calcule $\|\mathbf{y}\|^2 = 29$ y $\|\mathbf{y}\| = \sqrt{29}$.

   El vector unitario en la dirección de $\mathbf{c}$ y $\mathbf{y}$ es $\mathbf{u} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. $\mathbf{d}$ es ortogonal a $\mathbf{c}$, porque

$$\mathbf{d}\cdot\mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}\cdot\begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. $\mathbf{d}$ es ortogonal a $\mathbf{u}$ porque $\mathbf{u}$ tiene la forma $k\mathbf{c}$ para alguna k, y

$$\mathbf{d}\cdot\mathbf{u} = \mathbf{d}\cdot(k\mathbf{c}) = k(\mathbf{d}\cdot\mathbf{c}) = k(0) = 0$$

6.2 CONJUNTOS ORTOGONALES

Se dice que un conjunto de vectores $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ en $\mathbb{R}^n$ es un **conjunto ortogonal** si cada par de distintos vectores del conjunto es ortogonal, es decir, si $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ siempre que $i \neq j$.

EJEMPLO 1 Demuestre que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal, donde

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

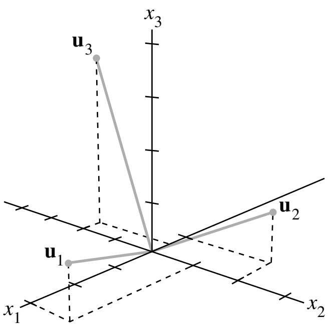


FIGURA 1

SOLUCIÓN Considere los tres posibles pares de vectores distintos, a saber, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$ y $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{u}_2 &= 3(-1) + 1(2) + 1(1) = 0 \\ \mathbf{u}_1\cdot\mathbf{u}_3 &= 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0 \\ \mathbf{u}_2\cdot\mathbf{u}_3 &= -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0 \end{aligned}$$

Cada par de vectores diferentes es ortogonal, y así $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortogonal. Véase la figura 1; ahí los tres segmentos de recta son mutuamente perpendiculares. ■

TEOREMA 4

> Si $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ es un conjunto ortogonal de vectores distintos de cero en $\mathbb{R}^n$, entonces S es linealmente independiente y, por lo tanto, es una base para el subespacio generado por S.

DEMOSTRACIÓN Si $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ para algunos escalares $c_1,\ldots,c_p$, entonces

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

ya que $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2,\ldots,\mathbf{u}_p$. Como $\mathbf{u}_1$ es diferente de cero, entonces $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero y así $c_1 = 0$. De manera similar, $c_2,\ldots,c_p$ deben ser cero. Por lo tanto, S es linealmente independiente. ■

DEFINICIÓN

Una **base ortogonal** para un subespacio W de $\mathbb{R}^n$ es una base para W que también es un conjunto ortogonal.

El siguiente teorema sugiere por qué una base ortogonal es mucho más agradable que otras bases. Los pesos en una combinación lineal se pueden calcular fácilmente.

TEOREMA 5

Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ una base ortogonal para un subespacio W de $\mathbb{R}^n$. Para cada $\mathbf{y}$ en W, los pesos en la combinación lineal

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

están dados por

$$c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \ldots, p)$$

DEMOSTRACIÓN Como en la demostración anterior, la ortogonalidad de $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ demuestra que

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Puesto que $\mathbf{u}_1\cdot\mathbf{u}_1$ no es cero, en la ecuación anterior se puede despejar c_1. Para encontrar c_j, considerando $j = 2, \ldots, p$, calcule $\mathbf{y}\cdot\mathbf{u}_j$ y despeje c_j. ■

EJEMPLO 2 El conjunto $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ del ejemplo 1 es una base ortogonal para $\mathbb{R}^3$. Exprese el vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ como una combinación lineal de los vectores en S.

SOLUCIÓN Calcule

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= 11, & \mathbf{y}\cdot\mathbf{u}_2 &= -12, & \mathbf{y}\cdot\mathbf{u}_3 &= -33 \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 11, & \mathbf{u}_2\cdot\mathbf{u}_2 &= 6, & \mathbf{u}_3\cdot\mathbf{u}_3 &= 33/2 \end{aligned}$$

Por el teorema 5,

$$\begin{aligned} \mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\ &= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\ &= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3 \end{aligned}$$

■

Observe qué fácil es obtener los pesos necesarios para construir $\mathbf{y}$ a partir de una base ortogonal. Si la base no fuera ortogonal, se necesitaría resolver un sistema de ecuaciones lineales para encontrar los pesos, como en el capítulo 1.

A continuación se presenta una construcción que será un paso clave en muchos cálculos que implican ortogonalidad, y que conducirá a una interpretación geométrica del teorema 5.

Una proyección ortogonal

Dado un vector $\mathbf{u}$ distinto de cero en $\mathbb{R}^n$, considere el problema de descomponer un vector $\mathbf{y}$ en $\mathbb{R}^n$ en la suma de dos vectores, siendo uno un múltiplo de $\mathbf{u}$ y el otro ortogonal a $\mathbf{u}$. Se desea escribir

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

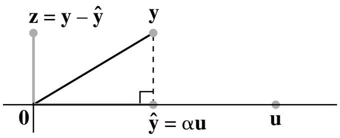


FIGURA 2
Determinación de α para hacer que $\mathbf{y} - \hat{\mathbf{y}}$ sea ortogonal a $\mathbf{u}$.

donde $\hat{\mathbf{y}} = \alpha\mathbf{u}$ para algún escalar α, y $\mathbf{z}$ es algún vector ortogonal a $\mathbf{u}$. Véase la figura 2. Dado cualquier escalar α, sea $\mathbf{z} = \mathbf{y} - \alpha\mathbf{u}$, tal que la ecuación (1) se satisface. Entonces $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\mathbf{u}$ si y solo si

$$0 = (\mathbf{y} - \alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - (\alpha\mathbf{u})\cdot\mathbf{u} = \mathbf{y}\cdot\mathbf{u} - \alpha(\mathbf{u}\cdot\mathbf{u})$$

Es decir, la ecuación (1) se satisface con $\mathbf{z}$ ortogonal a $\mathbf{u}$ si y solo si $\alpha = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ y $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. El vector $\hat{\mathbf{y}}$ es la **proyección ortogonal de y sobre u**, y el vector $\mathbf{z}$ es la **componente de y ortogonal a u**.

Si c es cualquier escalar distinto de cero y si $\mathbf{u}$ se remplaza por $c\mathbf{u}$ en la definición de $\hat{\mathbf{y}}$, entonces la proyección ortogonal de $\mathbf{y}$ sobre $c\mathbf{u}$ coincide exactamente con la proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ (ejercicio 31). De ahí que esta proyección esté determinada por el *subespacio* L generado por $\mathbf{u}$ (la recta que pasa por $\mathbf{u}$ y $\mathbf{0}$). Algunas veces $\hat{\mathbf{y}}$ se denota como $\text{proy}_L\,\mathbf{y}$, y se llama la **proyección ortogonal de y sobre** L. Es decir,

$$\boxed{\hat{\mathbf{y}} = \text{proj}_L\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}} \qquad (2)$$

EJEMPLO 3 Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. Encuentre la proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$. Después escriba $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en Gen $\{\mathbf{u}\}$ y el otro ortogonal a $\mathbf{u}$.

SOLUCIÓN Calcule

$$\mathbf{y}\cdot\mathbf{u} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 40$$

$$\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}\cdot\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 20$$

La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ es

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

y la componente de $\mathbf{y}$ ortogonal a $\mathbf{u}$ es

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

La suma de esos dos vectores es $\mathbf{y}$. Es decir,

$$\underset{\substack{\uparrow \\ \mathbf{y}}}{\begin{bmatrix} 7 \\ 6 \end{bmatrix}} = \underset{\substack{\uparrow \\ \hat{\mathbf{y}}}}{\begin{bmatrix} 8 \\ 4 \end{bmatrix}} + \underset{\substack{\uparrow \\ (\mathbf{y} - \hat{\mathbf{y}})}}{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}$$

En la figura 3 se muestra esta descomposición de $\mathbf{y}$. *Nota:* Si son correctos los cálculos anteriores, entonces $\{\hat{\mathbf{y}}, \mathbf{y} - \hat{\mathbf{y}}\}$ será un conjunto ortogonal. Para comprobarlo, determine

$$\hat{\mathbf{y}}\cdot(\mathbf{y} - \hat{\mathbf{y}}) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -8 + 8 = 0 \qquad \blacksquare$$

Como el segmento de recta de la figura 3 entre $\mathbf{y}$ y $\hat{\mathbf{y}}$ es perpendicular a L, por construcción de $\hat{\mathbf{y}}$, entonces el punto identificado con $\hat{\mathbf{y}}$ es el punto más cercano de L a $\mathbf{y}$. (Esto se puede demostrar con geometría. Ahora se supondrá esto para $\mathbb{R}^2$ y en la sección 6.3 se demostrará para $\mathbb{R}^n$).

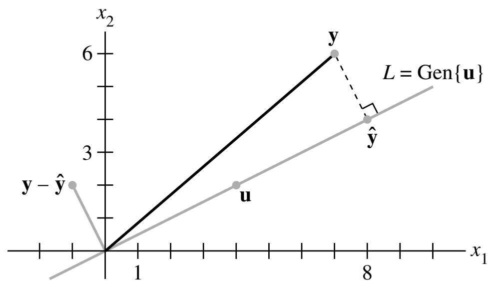


FIGURA 3 La proyección ortogonal de $\mathbf{y}$ sobre una recta L que pasa por el origen.

EJEMPLO 4 En la figura 3 encuentre la distancia de $\mathbf{y}$ a L.

SOLUCIÓN La distancia de $\mathbf{y}$ a L es la longitud del segmento de recta perpendicular de $\mathbf{y}$ a la proyección ortogonal $\hat{\mathbf{y}}$. Esta longitud es igual a la longitud de $\mathbf{y} - \hat{\mathbf{y}}$. Así, la distancia es

$$\|\mathbf{y} - \hat{\mathbf{y}}\| = \sqrt{(-1)^2 + 2^2} = \sqrt{5}$$ ■

Una interpretación geométrica del teorema 5

En la ecuación (2), la fórmula para la proyección ortogonal $\hat{\mathbf{y}}$ tiene la misma apariencia que cada uno de los términos en el teorema 5. Así, el teorema 5 descompone un vector $\mathbf{y}$ en una suma de proyecciones ortogonales sobre subespacios unidimensionales.

Es fácil visualizar el caso en el cual $W = \mathbb{R}^2 = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, con $\mathbf{u}_1$ y $\mathbf{u}_2$ ortogonales. Cualquier $\mathbf{y}$ en $\mathbb{R}^2$ se puede escribir en la forma

$$\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \tag{3}$$

El primer término en (3) es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_1$ (la recta que pasa por $\mathbf{u}_1$ y por el origen), y el segundo término es la proyección de $\mathbf{y}$ sobre el subespacio generado por $\mathbf{u}_2$. Así, la ecuación (3) expresa a $\mathbf{y}$ como la suma de sus proyecciones sobre los ejes (ortogonales) determinados por $\mathbf{u}_1$ y $\mathbf{u}_2$. Véase la figura 4.

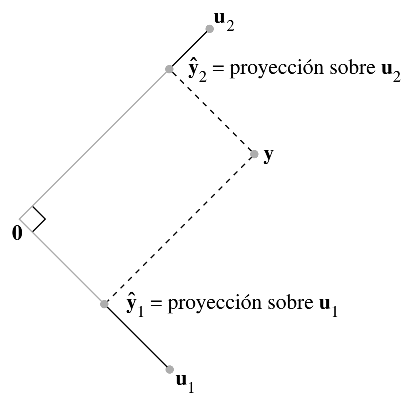


FIGURA 4 Un vector descompuesto en la suma de dos proyecciones.

El teorema 5 descompone a cada $\mathbf{y}$ en $\text{Gen}\,\{\mathbf{u}_1,\ldots, \mathbf{u}_p\}$ en la suma de p proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales.

Descomposición de una fuerza en sus componentes

La descomposición que se ilustra en la figura 4 se presenta en física cuando se aplica una fuerza a un objeto. La elección de un sistema de coordenadas adecuado permite que la fuerza se represente con un vector **y** en $\mathbb{R}^2$ o $\mathbb{R}^3$. Con frecuencia el problema implica alguna dirección particular de interés, que queda representada por otro vector **u**. Por ejemplo, si el objeto se está moviendo en línea recta cuando se aplica la fuerza, entonces el vector **u** podría apuntar en la dirección de movimiento, como se muestra en la figura 5. Un paso clave en el problema es descomponer la fuerza en una componente en la dirección de **u** y en una componente ortogonal a **u**. Los cálculos serían análogos a los realizados en el ejemplo 3 anterior.

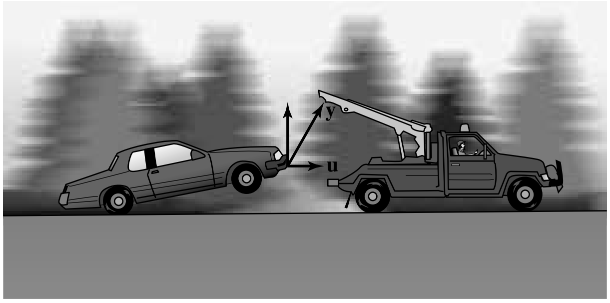


FIGURA 5

Conjuntos ortonormales

Un conjunto $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ es un **conjunto ortonormal** si es un conjunto ortogonal de vectores unitarios. Si W es el subespacio generado por tal conjunto, entonces $\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}$ es una **base ortonormal** para W, porque el conjunto es linealmente independiente de manera automática, de acuerdo con al teorema 4.

El ejemplo más sencillo de un conjunto ortonormal es la base estándar $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ para $\mathbb{R}^n$. Cualquier subconjunto no vacío de $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ también es ortonormal. A continuación se presenta un ejemplo más complicado.

EJEMPLO 5 Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortonormal de $\mathbb{R}^3$, donde

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUCIÓN Calcule

$$\begin{aligned}
\mathbf{v}_1\cdot\mathbf{v}_2 &= -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0 \\
\mathbf{v}_1\cdot\mathbf{v}_3 &= -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0 \\
\mathbf{v}_2\cdot\mathbf{v}_3 &= 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0
\end{aligned}$$

Así $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal. Además,

$$\begin{aligned}
\mathbf{v}_1\cdot\mathbf{v}_1 &= 9/11 + 1/11 + 1/11 = 1 \\
\mathbf{v}_2\cdot\mathbf{v}_2 &= 1/6 + 4/6 + 1/6 = 1 \\
\mathbf{v}_3\cdot\mathbf{v}_3 &= 1/66 + 16/66 + 49/66 = 1
\end{aligned}$$

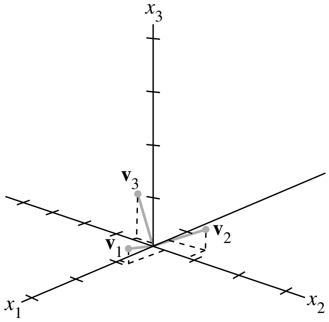


FIGURA 6

que muestra que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ son vectores unitarios. Así, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortonormal. Como el conjunto es linealmente independiente, entonces sus tres vectores forman una base para $\mathbb{R}^3$. Véase la figura 6. ■

Cuando los vectores en un conjunto ortogonal de vectores distintos de cero *se normalizan* para que tengan longitud unitaria, entonces los nuevos vectores seguirán siendo ortogonales, por lo que el nuevo conjunto será un conjunto ortonormal. Véase el ejercicio 32. En la figura 6 (ejemplo 5) es sencillo comprobar que los vectores son simplemente los vectores unitarios en las direcciones de los vectores de la figura 1 (ejemplo 1).

Las matrices cuyas columnas forman un conjunto ortonormal son importantes en aplicaciones y en algoritmos computacionales para cálculos matriciales. Sus propiedades fundamentales se exponen en los teoremas 6 y 7.

TEOREMA 6 Una matriz U de $m \times n$ tiene columnas ortonormales si y solo si $U^TU = I$.

DEMOSTRACIÓN Para simplificar la notación, suponga que U solo tiene tres columnas, cada una un vector en $\mathbb{R}^m$. En esencia, la demostración del caso general es la misma. Sea $U = [\mathbf{u}_1\ \ \mathbf{u}_2\ \ \mathbf{u}_3]$ y calcule

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

Las entradas en la matriz de la derecha son productos interiores, empleando la notación transpuesta. Las columnas de U son ortogonales si y solo si

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

Todas las columnas de U tienen longitud unitaria si y solo si

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

El teorema se deduce inmediatamente de las ecuaciones (4) a (6). ■

TEOREMA 7 Sea U una matriz de $m \times n$ con columnas ortonormales, y $\mathbf{x}$ y $\mathbf{y}$ están en $\mathbb{R}^n$. Entonces,

a) $\|U\mathbf{x}\| = \|\mathbf{x}\|$

b) $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$

c) $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ si y solo si $\mathbf{x}\cdot\mathbf{y} = 0$

Las propiedades *a*) y *c*) dicen que el mapeo lineal $\mathbf{x} \mapsto U\mathbf{x}$ preserva longitudes y ortogonalidad. Esas propiedades son importantes para muchos algoritmos computacionales. Véase el ejercicio 25 para la demostración del teorema 7.

EJEMPLO 6 Sean $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Observe que U tiene columnas ortonormales y

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Compruebe que $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUCIÓN

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

■

Los teoremas 6 y 7 son particularmente útiles cuando se aplican a matrices *cuadradas*. Una **matriz ortogonal** es una matriz cuadrada invertible U tal que $U^{-1} = U^T$. De acuerdo con el teorema 6, tal matriz tiene columnas ortonormales.[1] Es fácil ver que cualquier matriz *cuadrada* con columnas ortonormales es una matriz ortogonal. De manera sorprendente, dicha matriz también debe tener *filas* ortonormales. Véase los ejercicios 27 y 28. En el capítulo 7 se presentarán con frecuencia matrices ortogonales.

EJEMPLO 7 La matriz

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

es una matriz ortogonal porque es cuadrada y porque sus columnas son ortonormales, de acuerdo con el ejemplo 5. Compruebe que las filas ¡también sean ortonormales! ■

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Demuestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal para $\mathbb{R}^2$.
2. Sean $\mathbf{y}$ y L como en el ejemplo 3 y en la figura 3. Calcule la proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre L utilizando $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ en lugar de $\mathbf{u}$ del ejemplo 3.
3. Sean U y $\mathbf{x}$ como en el ejemplo 6, y $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Compruebe que $U\mathbf{x}\cdot U\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$.

6.2 EJERCICIOS

En los ejercicios 1 a 6, determine qué conjuntos de vectores son ortogonales.

1. $\begin{bmatrix} -1 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ -7 \end{bmatrix}$

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

4. $\begin{bmatrix} 2 \\ -5 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$

5. $\begin{bmatrix} 3 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 8 \\ 7 \\ 0 \end{bmatrix}$

6. $\begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix}$

En los ejercicios 7 a 10, demuestre que $\{\mathbf{u}_1, \mathbf{u}_2\}$ o $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ son una base ortogonal para $\mathbb{R}^2$ o $\mathbb{R}^3$, respectivamente. Después exprese $\mathbf{x}$ como una combinación lineal de los vectores $\mathbf{u}$.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

[1] Un mejor nombre sería *matriz ortonormal*, y este término se emplea en algunos textos de estadística. Sin embargo, *matriz ortogonal* es el término estándar en álgebra lineal.

8. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

9. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}$

10. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

11. Calcule la proyección ortogonal de $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ sobre la recta que pasa por $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ y por el origen.

12. Calcule la proyección ortogonal de $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ sobre la recta que pasa por $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ y por el origen.

13. Sean $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de dos vectores ortogonales, uno en Gen $\{\mathbf{u}\}$ y el otro ortogonal a $\mathbf{u}$.

14. Sean $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Escriba $\mathbf{y}$ como la suma de un vector en Gen $\{\mathbf{u}\}$ y un vector ortogonal a $\mathbf{u}$.

15. Sean $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Calcule la distancia de $\mathbf{y}$ a la recta que pasa por $\mathbf{u}$ y el origen.

16. Sean $\mathbf{y} = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Calcule la distancia de $\mathbf{y}$ a la recta que pasa por $\mathbf{u}$ y el origen.

En los ejercicios 17 a 22, determine cuáles conjuntos de vectores son ortonormales. Si un conjunto solamente es ortogonal, normalice los vectores para obtener un conjunto ortonormal.

17. $\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 0 \\ 1/2 \end{bmatrix}$

18. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$

19. $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}, \begin{bmatrix} .8 \\ .6 \end{bmatrix}$

20. $\begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \begin{bmatrix} 1/3 \\ 2/3 \\ 0 \end{bmatrix}$

21. $\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{20} \\ 3/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{20} \\ -1/\sqrt{20} \end{bmatrix}, \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

22. $\begin{bmatrix} 1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

En los ejercicios 23 y 24, todos los vectores están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

23. *a*) No todo conjunto linealmente independiente en $\mathbb{R}^n$ es un conjunto ortogonal.

b) Si $\mathbf{y}$ es una combinación lineal de vectores distintos de cero de un conjunto ortogonal, entonces los pesos en la combinación lineal se pueden calcular sin operaciones de fila sobre una matriz.

c) Si se normalizan los vectores en un conjunto ortogonal de vectores distintos de cero, entonces tal vez algunos de los nuevos vectores no sean ortogonales.

d) Una matriz con columnas ortonormales es una matriz ortogonal.

e) Si L es una recta que pasa por $\mathbf{0}$ y si $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre L, entonces $\|\hat{\mathbf{y}}\|$ da la distancia de $\mathbf{y}$ a L.

24. *a*) No todo conjunto ortogonal en $\mathbb{R}^n$ es linealmente independiente.

b) Si un conjunto $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ tiene la propiedad de que $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ siempre que $i \neq j$, entonces S es un conjunto ortonormal.

c) Si las columnas de una matriz A de $m \times n$ son ortonormales, entonces el mapeo lineal $\mathbf{x} \mapsto A\mathbf{x}$ preserva longitudes.

d) La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{v}$ es igual a la proyección ortogonal de $\mathbf{y}$ sobre $c\mathbf{v}$ siempre que $c \neq 0$.

e) Una matriz ortogonal es invertible.

25. Demuestre el teorema 7. [*Sugerencia:* Para *a*), calcule $\|U\mathbf{x}\|^2$, o primero demuestre *b*)].

26. Suponga que W es un subespacio de $\mathbb{R}^n$ generado por n vectores ortogonales distintos de cero. Explique por qué $W = \mathbb{R}^n$.

27. Sea U una matriz cuadrada con columnas ortonormales. Explique por qué U es invertible. (Mencione los teoremas que utilice).

28. Sea U una matriz ortogonal de $n \times n$. Demuestre que las filas de U forman una base ortonormal de $\mathbb{R}^n$.

29. Sean U y V matrices ortogonales de $n \times n$. Explique por qué UV es una matriz ortogonal. [Es decir, explique por qué UV es invertible y su inversa es $(UV)^T$].

30. Sea U una matriz ortogonal; construya V intercambiando algunas de las columnas de U. Explique por qué V es una matriz ortogonal.

31. Demuestre que la proyección ortogonal de un vector $\mathbf{y}$ sobre una recta L que pasa por el origen en $\mathbb{R}^2$ no depende de la elección del vector $\mathbf{u}$ diferente de cero en L que se emplea en la fórmula para $\hat{\mathbf{y}}$. Para hacer esto, suponga que $\mathbf{y}$ y $\mathbf{u}$ están dados y que $\hat{\mathbf{y}}$ se calculó mediante la fórmula (2) de esta sección. Sustituya $\mathbf{u}$ en tal fórmula por $c\mathbf{u}$, donde c es un escalar distinto de cero no especificado. Demuestre que la nueva fórmula da el mismo valor para $\hat{\mathbf{y}}$.

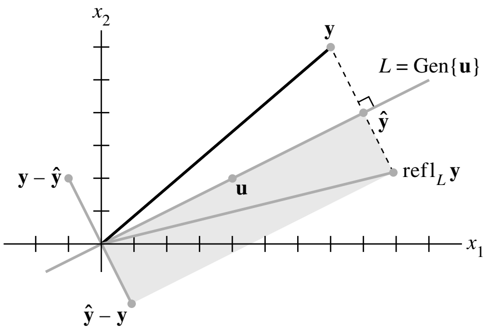
32. Sean $\{\mathbf{v}_1, \mathbf{v}_2\}$ un conjunto ortogonal de vectores distintos de cero, y c_1, c_2 escalares diferentes de cero. Demuestre que $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ también es un conjunto ortogonal. Como la ortogonalidad de un conjunto está definida en términos de pares de vectores, esto demuestra que si se normalizan los vectores en un conjunto ortogonal, el nuevo conjunto seguirá siendo ortogonal.

33. Dado $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\,\{\mathbf{u}\}$. Demuestre que el mapeo $\mathbf{x} \mapsto \text{proy}_L\,\mathbf{x}$ es una transformación lineal.

34. Dado $\mathbf{u} \neq \mathbf{0}$ en $\mathbb{R}^n$, sea $L = \text{Gen}\,\{\mathbf{u}\}$. Para $\mathbf{y}$ en $\mathbb{R}^n$, la **reflexión de y en** L es el punto $\text{refl}_L\,\mathbf{y}$ definido por

$\text{refl}_L\,\mathbf{y} = 2\cdot\text{proy}_L\,\mathbf{y} - \mathbf{y}$.

Véase la figura, que indica que $\text{refl}_L\,\mathbf{y}$ es la suma de $\hat{\mathbf{y}} = \text{proy}_L\,\mathbf{y}$ y $\hat{\mathbf{y}} - \mathbf{y}$. Demuestre que el mapeo $\mathbf{y} \mapsto \text{refl}_L$ es una transformación lineal.



La reflexión de $\mathbf{y}$ en una recta que pasa por el origen.

35. [**M**] Demuestre que las columnas de la matriz A son ortogonales haciendo el cálculo matricial adecuado. Especifique el cálculo que realizó.

$$A = \begin{bmatrix} -6 & -3 & 6 & 1 \\ -1 & 2 & 1 & -6 \\ 3 & 6 & 3 & -2 \\ 6 & -3 & 6 & -1 \\ 2 & -1 & 2 & 3 \\ -3 & 6 & 3 & 2 \\ -2 & -1 & 2 & -3 \\ 1 & 2 & 1 & 6 \end{bmatrix}$$

36. [**M**] En los incisos *a*) a *d*), sea U la matriz formada por la normalización de cada columna de la matriz A del ejercicio 35.

a) Calcule U^TU y UU^T. ¿En qué difieren?

b) Genere un vector aleatorio $\mathbf{y}$ en $\mathbb{R}^8$, y determine $\mathbf{p} = UU^T\mathbf{y}$ y $\mathbf{z} = \mathbf{y} - \mathbf{p}$. Explique por qué $\mathbf{p}$ está en Col A. Compruebe que $\mathbf{z}$ sea ortogonal a $\mathbf{p}$.

c) Compruebe que $\mathbf{z}$ es ortogonal a cada columna de U.

d) Observe que $\mathbf{y} = \mathbf{p} + \mathbf{z}$, con $\mathbf{p}$ en Col A. Explique por qué $\mathbf{z}$ está en $(\text{Col } A)^{\perp}$. (En la siguiente sección se explicará el significado de esta descomposición de $\mathbf{y}$).

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Los vectores son ortogonales porque

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = -2/5 + 2/5 = 0$$

Son unitarios porque

$$\|\mathbf{u}_1\|^2 = (-1/\sqrt{5})^2 + (2/\sqrt{5})^2 = 1/5 + 4/5 = 1$$
$$\|\mathbf{u}_2\|^2 = (2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 4/5 + 1/5 = 1$$

En particular, el conjunto $\{\mathbf{u}_1, \mathbf{u}_2\}$ es linealmente independiente y, por lo tanto, es una base para $\mathbb{R}^2$ porque hay dos vectores en el conjunto.

2. Cuando $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{20}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Esta es la misma $\hat{\mathbf{y}}$ que se encontró en el ejemplo 3. La proyección ortogonal no parece depender de la $\mathbf{u}$ seleccionada sobre la recta. Véase el ejercicio 31.

3. $$U\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$$

También, del ejemplo 6, $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$ y $U\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$. Por lo tanto,

$$U\mathbf{x}\cdot U\mathbf{y} = 3 + 7 + 2 = 12 \quad \text{y} \quad \mathbf{x}\cdot\mathbf{y} = -6 + 18 = 12$$

6.3 PROYECCIONES ORTOGONALES

La proyección ortogonal de un punto en $\mathbb{R}^2$ sobre una recta que pasa por el origen tiene una importante analogía en $\mathbb{R}^n$. Dado un vector $\mathbf{y}$ y un subespacio W en $\mathbb{R}^n$, existe un vector $\hat{\mathbf{y}}$ en W tal que **1.** $\hat{\mathbf{y}}$ es el único vector en W para el cual $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W, y **2.** $\hat{\mathbf{y}}$ es el único vector en W más cercano a $\mathbf{y}$. Véase la figura 1. Estas dos propiedades de $\hat{\mathbf{y}}$ dan la clave para encontrar las soluciones de mínimos cuadrados de sistemas lineales, que se mencionaron en el ejemplo introductorio de este capítulo. En la sección 6.5 se contará la historia completa.

Como preparación para el primer teorema, observe que siempre que un vector $\mathbf{y}$ se representa como una combinación lineal de vectores $\mathbf{u}_1,\dots,\mathbf{u}_n$ en $\mathbb{R}^n$, los términos en la suma para $\mathbf{y}$ se pueden agrupar en dos partes de manera que $\mathbf{y}$ se representa como

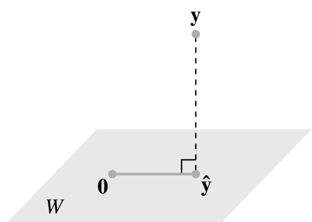


FIGURA 1

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$

donde $\mathbf{z}_1$ es una combinación lineal de algunas $\mathbf{u}_i$, y $\mathbf{z}_2$ es una combinación lineal de las $\mathbf{u}_i$ restantes. Esta idea es útil en particular cuando $\{\mathbf{u}_1,\dots,\mathbf{u}_n\}$ es una base ortogonal. De la sección 6.1 recuerde que $W^\perp$ denota el conjunto de todos los vectores ortogonales a un subespacio W.

EJEMPLO 1 Sea $\{\mathbf{u}_1,\dots,\mathbf{u}_5\}$ una base ortogonal para $\mathbb{R}^5$ y

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Considere el subespacio $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, y escriba $\mathbf{y}$ como la suma de un vector $\mathbf{z}_1$ en W y un vector $\mathbf{z}_2$ en $W^\perp$.

SOLUCIÓN Escriba

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

donde $\quad \mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \quad$ está en Gen $\{\mathbf{u}_1, \mathbf{u}_2\}$

y $\quad \mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5 \quad$ está en Gen $\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

Para demostrar que $\mathbf{z}_2$ está en $W^\perp$, es suficiente probar que $\mathbf{z}_2$ es ortogonal a los vectores en la base $\{\mathbf{u}_1, \mathbf{u}_2\}$ para W. (Véase la sección 6.1). Utilizando propiedades del producto interior, calcule

$$\begin{aligned}\mathbf{z}_2\cdot\mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5)\cdot\mathbf{u}_1 \\ &= c_3\mathbf{u}_3\cdot\mathbf{u}_1 + c_4\mathbf{u}_4\cdot\mathbf{u}_1 + c_5\mathbf{u}_5\cdot\mathbf{u}_1 \\ &= 0\end{aligned}$$

ya que $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_3$, $\mathbf{u}_4$ y $\mathbf{u}_5$. Un cálculo similar indica que $\mathbf{z}_2\cdot\mathbf{u}_2 = 0$. Así, $\mathbf{z}_2$ está en $W^\perp$. ■

El siguiente teorema demuestra que la descomposición $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$ del ejemplo 1 se puede calcular sin tener una base ortogonal para $\mathbb{R}^n$. Es suficiente con tener una base ortogonal solo para W.

TEOREMA 8 **Teorema de descomposición ortogonal**

Sea W un subespacio de $\mathbb{R}^n$. Entonces toda $\mathbf{y}$ en $\mathbb{R}^n$ se puede escribir de forma única como

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

donde $\hat{\mathbf{y}}$ está en W y $\mathbf{z}$ está en $W^\perp$. De hecho, si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es cualquier base ortogonal de W, entonces

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \tag{2}$$

y $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ en la ecuación (1) es la **proyección ortogonal de y sobre** W, y con frecuencia se escribe como $\text{proy}_W\,\mathbf{y}$. Véase la figura 2. Cuando W es un subespacio unidimensional, la fórmula para $\hat{\mathbf{y}}$ coincide con la fórmula que se presentó en la sección 6.2.

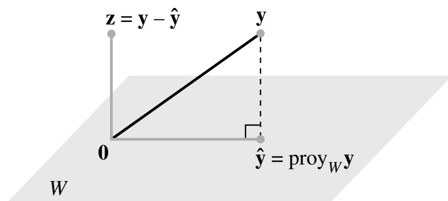


FIGURA 2 Proyección ortogonal de **y** sobre W.

DEMOSTRACIÓN Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ una base ortogonal para W, y defina $\hat{\mathbf{y}}$ con la ecuación (2).[1] Entonces $\hat{\mathbf{y}}$ está en W porque $\hat{\mathbf{y}}$ es una combinación lineal de la base $\mathbf{u}_1, \ldots, \mathbf{u}_p$. Sea $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Como $\mathbf{u}_1$ es ortogonal a $\mathbf{u}_2, \ldots, \mathbf{u}_p$, entonces a partir de la ecuación (2) se deduce que

$$\begin{aligned}\mathbf{z}\cdot\mathbf{u}_1 = (\mathbf{y} - \hat{\mathbf{y}})\cdot\mathbf{u}_1 &= \mathbf{y}\cdot\mathbf{u}_1 - \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1\cdot\mathbf{u}_1 - 0 - \cdots - 0 \\ &= \mathbf{y}\cdot\mathbf{u}_1 - \mathbf{y}\cdot\mathbf{u}_1 = 0\end{aligned}$$

Por lo tanto, $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$. De manera similar, $\mathbf{z}$ es ortogonal a cada $\mathbf{u}_j$ en la base para W. Por consiguiente, $\mathbf{z}$ es ortogonal para todo vector en W. Es decir, $\mathbf{z}$ está en $W^\perp$.

Para demostrar que la descomposición en la ecuación (1) es única, suponga que $\mathbf{y}$ se puede escribir como $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ con $\hat{\mathbf{y}}_1$ en W, y $\mathbf{z}_1$ en $W^\perp$. Entonces $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$ (ya que ambos lados son iguales a $\mathbf{y}$), y así

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

Esta igualdad indica que el vector $\mathbf{v} = \mathbf{y} - \hat{\mathbf{y}}_1$ está en W y en $W^\perp$ (porque $\mathbf{z}_1$ y $\mathbf{z}$ están ambos en $W^\perp$, y $W^\perp$ es un subespacio). Por lo tanto, $\mathbf{v}\cdot\mathbf{v} = 0$, lo que demuestra que $\mathbf{v} = \mathbf{0}$. Esto prueba que $\mathbf{y} = \hat{\mathbf{y}}_1$ y también que $\mathbf{z}_1 = \mathbf{z}$. ■

La unicidad de la descomposición (1) demuestra que la proyección ortogonal $\hat{\mathbf{y}}$ solo depende de W y no de la base particular empleada en la ecuación (2).

[1] Se puede suponer que W no es el subespacio cero porque, de otra forma, $W^\perp = \mathbb{R}^n$ y la ecuación (1) simplemente sería $\mathbf{y} = \mathbf{0} + \mathbf{y}$. En la siguiente sección se demostrará que cualquier subespacio distinto de cero de $\mathbb{R}^n$ tiene una base ortogonal.

EJEMPLO 2 Sean $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Escriba $\mathbf{y}$ como la suma de un vector en W y de un vector ortogonal a W.

SOLUCIÓN La proyección ortogonal de $\mathbf{y}$ sobre W es

$$\begin{aligned}
\hat{\mathbf{y}} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 \\
&= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}
\end{aligned}$$

Además,

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

El teorema 8 asegura que $\mathbf{y} - \hat{\mathbf{y}}$ está en $W^{\perp}$. Sin embargo, para comprobar los cálculos, es una buena idea comprobar que $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$, y por lo tanto a toda W. La descomposición deseada de $\mathbf{y}$ es

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

■

Interpretación geométrica de la proyección ortogonal

Cuando W es un subespacio unidimensional, la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ solo contiene un término. Así, cuando $\dim W > 1$, cada término de la ecuación (2) es en sí mismo una proyección ortogonal de $\mathbf{y}$ sobre el subespacio unidimensional generado por uno de los vectores $\mathbf{u}$ en la base para W. La figura 3 muestra esto cuando W es un subespacio de $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Aquí, $\hat{\mathbf{y}}_1$ y $\hat{\mathbf{y}}_2$ denotan las proyecciones de $\mathbf{y}$ sobre las rectas generadas por $\mathbf{u}_1$ y $\mathbf{u}_2$, respectivamente. La proyección ortogonal $\hat{\mathbf{y}}$ de $\mathbf{y}$ sobre W es la suma de las proyecciones de $\mathbf{y}$ sobre subespacios unidimensionales que son ortogonales entre sí. El vector $\hat{\mathbf{y}}$ en la figura 3 corresponde al vector $\mathbf{y}$ de la figura 4 de la sección 6.2, porque ahora es $\hat{\mathbf{y}}$ el que está en W.

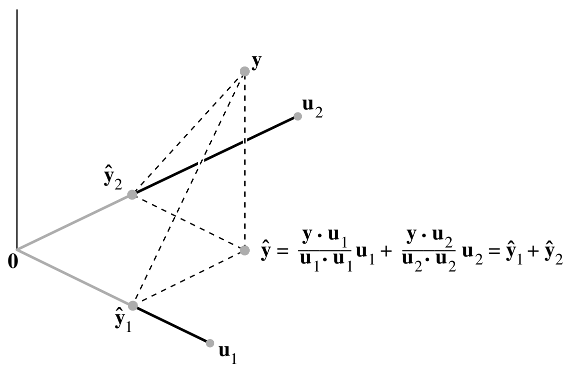


FIGURA 3 La proyección ortogonal de $\mathbf{y}$ es la suma de sus proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales.

Propiedades de proyecciones ortogonales

Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortogonal para W y resulta que $\mathbf{y}$ está en W, entonces la fórmula para $\text{proy}_W\, \mathbf{y}$ es exactamente la misma que la representación de $\mathbf{y}$ en el teorema 5 de la sección 6.2. En este caso, $\text{proy}_W\, \mathbf{y} = \mathbf{y}$.

Si $\mathbf{y}$ está en $W = \text{Gen}\,\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, entonces $\text{proy}_W\, \mathbf{y} = \mathbf{y}$.

Este resultado también se deduce del siguiente teorema.

TEOREMA 9

Teorema de la mejor aproximación

W es un subespacio de $\mathbb{R}^n$, $\mathbf{y}$ es cualquier vector en $\mathbb{R}^n$, y $\hat{\mathbf{y}}$ es la proyección ortogonal de $\mathbf{y}$ sobre W. Entonces $\hat{\mathbf{y}}$ en W es el punto más cercano a $\mathbf{y}$, en el sentido de que

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{3}$$

para toda $\mathbf{v}$ en W diferente de $\hat{\mathbf{y}}$.

El vector $\hat{\mathbf{y}}$ del teorema 9 se llama **la mejor aproximación a y por elementos de *W***. En secciones posteriores del libro se examinarán problemas donde una $\mathbf{y}$ dada se debe remplazar, o *aproximar*, mediante un vector $\mathbf{v}$ en algún subespacio W fijo. La distancia de $\mathbf{y}$ a $\mathbf{v}$, dada por $\|\mathbf{y} - \mathbf{v}\|$, se puede considerar como el "error" de usar $\mathbf{v}$ en lugar de $\mathbf{y}$. El teorema 9 afirma que este error se minimiza cuando $\mathbf{v} = \hat{\mathbf{y}}$.

La desigualdad (3) conduce a una nueva demostración de que $\hat{\mathbf{y}}$ no depende de la base ortogonal particular empleada para calcularlo. Si se utilizara una base ortogonal diferente para W con la finalidad de construir una proyección ortogonal de $\mathbf{y}$, entonces esta proyección también sería el punto en W más cercano a $\mathbf{y}$, a saber, $\hat{\mathbf{y}}$.

DEMOSTRACIÓN Tome $\mathbf{v}$ en W diferente de $\hat{\mathbf{y}}$. Véase la figura 4. Entonces $\hat{\mathbf{y}} - \mathbf{v}$ está en W. De acuerdo con el teorema de descomposición ortogonal, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W. En particular, $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a $\hat{\mathbf{y}} - \mathbf{v}$ (que está en W). Puesto que

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v})$$

entonces, al utilizar el teorema de Pitágoras, se obtiene

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$$

(Véase el triángulo rectángulo a la derecha de la figura 4. Se indica la longitud de cada lado). Ahora $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ porque $\hat{\mathbf{y}} - \mathbf{v} \neq 0$, y así se deduce inmediatamente la desigualdad (3). ■

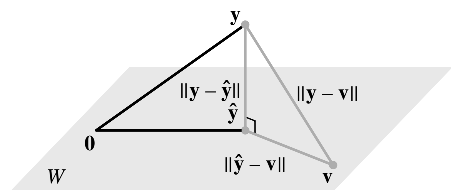


FIGURA 4 La proyección ortogonal de $\mathbf{y}$ sobre W es el punto en W más cercano a $\mathbf{y}$.

EJEMPLO 3 Si $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, como en el ejemplo 2, entonces el punto en W más cercano a $\mathbf{y}$ es

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$ ■

EJEMPLO 4 La distancia de un punto $\mathbf{y}$ en $\mathbb{R}^n$ a un subespacio W se define como la distancia de $\mathbf{y}$ al punto más cercano en W. Encuentre la distancia de $\mathbf{y}$ a $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, donde

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

SOLUCIÓN De acuerdo con el teorema de la mejor aproximación, la distancia de $\mathbf{y}$ a W es $\|\mathbf{y} - \hat{\mathbf{y}}\|$, donde $\hat{\mathbf{y}} = \text{proy}_W\,\mathbf{y}$. Ya que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortogonal para W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$

La distancia de $\mathbf{y}$ a W es $\sqrt{45} = 3\sqrt{5}$. ■

El teorema final en esta sección muestra cómo la fórmula (2) para $\text{proy}_W\,\mathbf{y}$ se simplifica cuando la base para W es un conjunto ortonormal.

TEOREMA 10

Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortonormal para un subespacio W de $\mathbb{R}^n$, entonces

$$\text{proy}_W\,\mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p \tag{4}$$

Si $U = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_p]$, entonces

$$\text{proy}_W\,\mathbf{y} = UUT\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \tag{5}$$

DEMOSTRACIÓN La fórmula (4) se obtiene inmediatamente de la ecuación (2) del teorema 8. Además, la ecuación (4) indica que $\text{proy}_W\,\mathbf{y}$ es una combinación lineal de las columnas de U empleando los pesos $\mathbf{y}\cdot\mathbf{u}_1$, $\mathbf{y}\cdot\mathbf{u}_2, \ldots, \mathbf{y}\cdot\mathbf{u}_p$. Los pesos se pueden representar como $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, probando que son las entradas en $U^T\mathbf{y}$, y justificando la ecuación (5). ■

WEB

Suponga que U es una matriz de $n \times p$ con columnas ortonormales, y sea W el espacio columna de U. Entonces,

$$U^TU\mathbf{x} = I_p\,\mathbf{x} = \mathbf{x} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{R}^p \quad \text{Teorema 6}$$
$$UU^T\mathbf{y} = \text{proy}_W\,\mathbf{y} \quad \text{para toda } \mathbf{y} \text{ en } \mathbb{R}^n \quad \text{Teorema 10}$$

Si U es una matriz (cuadrada) de $n \times n$ con columnas ortonormales, entonces U es una matriz *ortogonal*, el espacio columna W es todo de $\mathbb{R}^n$, y $UU^T\mathbf{y} = I\mathbf{y} = \mathbf{y}$ para toda $\mathbf{y}$ en $\mathbb{R}^n$.

Aunque la fórmula (4) es importante para fines teóricos, en la práctica generalmente implica algunos cálculos con raíces cuadradas de números (en las entradas de $\mathbf{u}_i$). La fórmula (2) se recomienda para cálculos a mano.

PROBLEMA DE PRÁCTICA

Sean $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$ y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$. Con base en el hecho de que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales, calcule $\text{proy}_W\, \mathbf{y}$.

6.3 EJERCICIOS

En los ejercicios 1 y 2, se puede suponer que $\{\mathbf{u}_1, \ldots, \mathbf{u}_4\}$ es una base ortogonal para $\mathbb{R}^4$.

1. $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -4 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 5 \\ -3 \\ -1 \\ 1 \end{bmatrix}$

   $\mathbf{x} = \begin{bmatrix} 10 \\ -8 \\ 2 \\ 0 \end{bmatrix}$. Escriba $\mathbf{x}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ y el otro en $\text{Gen}\{\mathbf{u}_4\}$.

2. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$.

   $\mathbf{v} = \begin{bmatrix} 4 \\ 5 \\ -3 \\ 3 \end{bmatrix}$. Escriba $\mathbf{v}$ como la suma de dos vectores, uno en $\text{Gen}\{\mathbf{u}_1\}$ y el otro en $\text{Gen}\{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$.

En los ejercicios 3 a 6, compruebe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ es un conjunto ortogonal, y luego encuentre la proyección ortogonal de $\mathbf{y}$ sobre $\text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

3. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

4. $\mathbf{y} = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$

5. $\mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$

6. $\mathbf{y} = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -4 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 7 a 10, sea W el subespacio generado por los vectores $\mathbf{u}$, y escriba $\mathbf{y}$ como la suma de un vector en W y un vector ortogonal a W.

7. $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$

8. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$

9. $\mathbf{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{y} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

En los ejercicios 11 y 12, determine el punto más cercano a $\mathbf{y}$ en el subespacio W generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

11. $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

12. $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

En los ejercicios 13 y 14, calcule la mejor aproximación a $\mathbf{z}$ con vectores de la forma $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

13. $\mathbf{z} = \begin{bmatrix} 3 \\ -7 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$

14. $\mathbf{z} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -2 \\ 4 \\ 2 \end{bmatrix}$

15. Sean $\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Encuentre la distancia de $\mathbf{y}$ al plano en $\mathbb{R}^3$ generado por $\mathbf{u}_1$ y $\mathbf{u}_2$.

16. Sean $\mathbf{y}$, $\mathbf{v}_1$ y $\mathbf{v}_2$ como en el ejercicio 12. Encuentre la distancia de $\mathbf{y}$ al subespacio de $\mathbb{R}^4$ generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

17. Sean $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ y $W = \text{Gen}\{\mathbf{u}_1, \mathbf{u}_2\}$.

a) Sea $U = [\mathbf{u}_1 \quad \mathbf{u}_2]$. Calcule U^TU y UU^T.

b) Calcule $\text{proy}_W\,\mathbf{y}$ y $(UU^T)\mathbf{y}$.

18. Sean $\mathbf{y} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$ y $W = \text{Gen}\,\{\mathbf{u}_1\}$.

a) Sea U la matriz de 2×1 cuya única columna es $\mathbf{u}_1$. Calcule U^TU y UU^T.

b) Determine $\text{proy}_W\,\mathbf{y}$ y $(UU^T)\mathbf{y}$.

19. Sean $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$ y $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Observe que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales, pero que $\mathbf{u}_3$ no es ortogonal a $\mathbf{u}_1$ o $\mathbf{u}_2$. Es posible demostrar que $\mathbf{u}_3$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Con base en este hecho, construya un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

20. Sean $\mathbf{u}_1$ y $\mathbf{u}_2$ como en el ejercicio 19, y $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Es posible demostrar que $\mathbf{u}_4$ no está en el subespacio W generado por $\mathbf{u}_1$ y $\mathbf{u}_2$. Con base en este hecho, obtenga un vector $\mathbf{v}$ diferente de cero en $\mathbb{R}^3$ que sea ortogonal a $\mathbf{u}_1$ y $\mathbf{u}_2$.

En los ejercicios 21 y 22, todos los vectores y subespacios están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Si $\mathbf{z}$ es ortogonal a $\mathbf{u}_1$ y a $\mathbf{u}_2$, y si $W = \text{Gen}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, entonces $\mathbf{z}$ debe estar en $W^\perp$.

b) Para cada $\mathbf{y}$ y cada subespacio W, el vector $\mathbf{y} - \text{proy}_W\,\mathbf{y}$ es ortogonal a W.

c) Algunas veces $\hat{\mathbf{y}}$, la proyección ortogonal de $\mathbf{y}$ sobre un subespacio W, puede depender de la base ortogonal de W empleada para calcular $\hat{\mathbf{y}}$.

d) Si $\mathbf{y}$ está en un subespacio W, entonces la proyección ortogonal de $\mathbf{y}$ sobre W es $\mathbf{y}$ misma.

e) Si las columnas de una matriz U de $n \times p$, son ortonormales, entonces $UU^T\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre el espacio columna de U.

22. *a*) Si W es un subespacio de $\mathbb{R}^n$ y si $\mathbf{v}$ está en W y en $W^\perp$, entonces $\mathbf{v}$ debe ser el vector cero.

b) En el teorema de descomposición ortogonal, cada término en la fórmula (2) para $\hat{\mathbf{y}}$ es en sí mismo una proyección de $\mathbf{y}$ sobre un subespacio de W.

c) Si $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$, donde $\mathbf{z}_1$ está en un subespacio W y $\mathbf{z}_2$ está en $W^\perp$, entonces $\mathbf{z}_1$ debe ser la proyección ortogonal de $\mathbf{y}$ sobre W.

d) La mejor aproximación a $\mathbf{y}$ por elementos de un subespacio W está dada por el vector $\mathbf{y} - \text{proy}_W\,\mathbf{y}$.

e) Si una matriz U de $n \times p$, tiene columnas ortonormales, entonces $UU^T\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

23. Sea A una matriz de $m \times n$. Demuestre que cada vector $\mathbf{x}$ en $\mathbb{R}^n$ se puede escribir en la forma $\mathbf{x} = \mathbf{p} + \mathbf{u}$, donde $\mathbf{p}$ está en Fil A y $\mathbf{u}$ pertenece a Nul A. Además, demuestre que si la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces existe una única $\mathbf{p}$ en Fil A tal que $A\mathbf{p} = \mathbf{b}$.

24. Sea W un subespacio de $\mathbb{R}^n$ con una base ortogonal $\{\mathbf{w}_1, \ldots, \mathbf{w}_p\}$, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_q\}$ una base ortogonal para $W^\perp$.

a) Explique por qué $\{\mathbf{w}_1, \ldots, \mathbf{w}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q\}$ es un conjunto ortogonal.

b) Explique por qué el conjunto del inciso *a*) genera a $\mathbb{R}^n$.

c) Demuestre que $\dim W + \dim W^\perp = n$.

25. [**M**] Sea U la matriz de 8×4 del ejercicio 36 de la sección 6.2. Encuentre el punto más cercano a $\mathbf{y} = (1, 1, 1, 1, 1, 1, 1, 1)$ en Col U. Escriba las instrucciones que utilizó para resolver este problema.

26. [**M**] Sea U la matriz del ejercicio 25. Obtenga la distancia de $\mathbf{b} = (1, 1, 1, 1, -1, -1, -1, -1)$ a Col U.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Calcule

$$\text{proy}_W\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{88}{66}\mathbf{u}_1 + \frac{-2}{6}\mathbf{u}_2$$

$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = \mathbf{y}$$

En este caso, $\mathbf{y}$ resulta ser una combinación lineal de $\mathbf{u}_1$ y $\mathbf{u}_2$, de manera que $\mathbf{y}$ está en W. El punto en W más cerca de $\mathbf{y}$ es $\mathbf{y}$ misma.

6.4 PROCESO DE GRAM-SCHMIDT

El proceso de Gram-Schmidt es un sencillo algoritmo para obtener una base ortogonal u ortonormal para cualquier subespacio diferente de cero de $\mathbb{R}^n$. Los dos primeros ejemplos de este proceso son para realizarse a mano.

EJEMPLO 1 Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$. Construya una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$ para W.

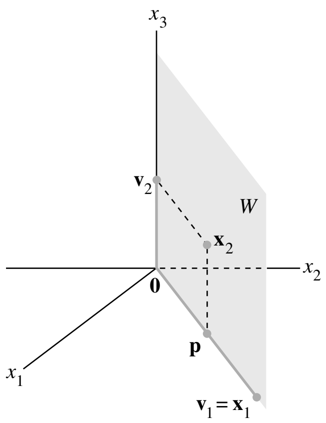


FIGURA 1 Construcción de una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2\}$.

SOLUCIÓN En la figura 1 se ilustra el subespacio W, junto con $\mathbf{x}_1$, $\mathbf{x}_2$ y la proyección $\mathbf{p}$ de $\mathbf{x}_2$ sobre $\mathbf{x}_1$. La componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$ es $\mathbf{x}_2 - \mathbf{p}$, que está en W porque se forma a partir de $\mathbf{x}_2$ y de un múltiplo de $\mathbf{x}_1$. Sea $\mathbf{v}_1 = \mathbf{x}_1$ y

$$\mathbf{v}_2 = \mathbf{x}_2 - \mathbf{p} = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{x}_1}{\mathbf{x}_1 \cdot \mathbf{x}_1}\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} - \frac{15}{45}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortogonal de vectores distintos de cero en W. Como $\dim W = 2$, entonces el conjunto $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base para W. ■

El siguiente ejemplo ilustra plenamente el proceso de Gram-Schmidt. Estúdielo con cuidado.

EJEMPLO 2 Sean $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Entonces, $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ es, a todas luces, linealmente independiente y, por consiguiente, es una base para un subespacio W de $\mathbb{R}^4$. Construya una base ortogonal para W.

SOLUCIÓN

Paso 1. Sean $\mathbf{v}_1 = \mathbf{x}_1$ y $W_1 = \text{Gen}\,\{\mathbf{x}_1\} = \text{Gen}\,\{\mathbf{v}_1\}$.

Paso 2. Sea $\mathbf{v}_2$ el vector producido al restar de $\mathbf{x}_2$ su proyección sobre el subespacio W_1. Es decir, sea

$$\begin{aligned} \mathbf{v}_2 &= \mathbf{x}_2 - \text{proy}_{W_1}\,\mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Ya que } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} \end{aligned}$$

Como en el ejemplo 1, $\mathbf{v}_2$ es la componente de $\mathbf{x}_2$ ortogonal a $\mathbf{x}_1$, y $\{\mathbf{v}_1, \mathbf{v}_2\}$ es una base ortogonal para el subespacio W_2 generado por $\mathbf{x}_1$ y $\mathbf{x}_2$.

Paso 2′ (opcional). Si es pertinente, escale $\mathbf{v}_2$ para simplificar cálculos posteriores. Como $\mathbf{v}_2$ tiene entradas fraccionales, entonces es conveniente escalarlo por un factor de 4 y sustituir $\{\mathbf{v}_1, \mathbf{v}_2\}$ por la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Paso 3. Sea $\mathbf{v}_3$ el vector obtenido restando de $\mathbf{x}_3$ su proyección sobre el subespacio W_2. Utilice la base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$ para calcular esta proyección sobre W_2:

$$\operatorname{proy}_{W_2} \mathbf{x}_3 = \overset{\substack{\text{Proyección} \\ \text{de } \mathbf{x}_3 \text{ sobre } \mathbf{v}_1 \\ \downarrow}}{\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1} + \overset{\substack{\text{Proyección de} \\ \mathbf{x}_3 \text{ sobre } \mathbf{v}_2' \\ \downarrow}}{\frac{\mathbf{x}_3 \cdot \mathbf{v}_2'}{\mathbf{v}_2' \cdot \mathbf{v}_2'} \mathbf{v}_2'} = \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Entonces, $\mathbf{v}_3$ es la componente de $\mathbf{x}_3$ ortogonal a W_2, a saber,

$$\mathbf{v}_3 = \mathbf{x}_3 - \operatorname{proy}_{W_2} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Véase en la figura 2 un diagrama de esta construcción. Observe que $\mathbf{v}_3$ está en W porque $\mathbf{x}_3$ y $\operatorname{proy}_{W_2}\mathbf{x}_3$ están en W. Así, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es un conjunto ortogonal de vectores diferentes de cero y, por lo tanto, un conjunto linealmente independiente en W. Observe que W es tridimensional porque está definido con una base de tres vectores. Por lo tanto, de acuerdo con el teorema de la base de la sección 4.5, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ es una base ortogonal para W. ■

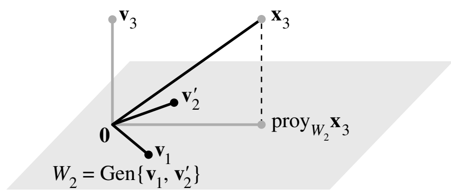


FIGURA 2 Construcción de $\mathbf{v}_3$ a partir de $\mathbf{x}_3$ y W_2.

La demostración del siguiente teorema revela que esta estrategia realmente funciona. No se menciona el escalamiento de vectores porque solo se usa para simplificar los cálculos a mano.

TEOREMA 11

Proceso de Gram-Schmidt

A partir de una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ para un subespacio W de $\mathbb{R}^n$, se define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Entonces, $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es una base ortogonal para W. Además,

$$\operatorname{Gen}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{Gen}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \quad \text{para } 1 \leq k \leq p \tag{1}$$

DEMOSTRACIÓN Para $1 \leq k \leq p$, sea $W_k = \text{Gen}\,\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$. Establezca que $\mathbf{v}_1 = \mathbf{x}_1$, por lo que Gen $\{\mathbf{v}_1\}$ = Gen $\{\mathbf{x}_1\}$. Suponga que, para alguna $k < p$, se construyeron $\mathbf{v}_1,\ldots,\mathbf{v}_k$, de manera que $\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ es una base ortogonal para W_k. Se define

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - \text{proy}_{W_k}\,\mathbf{x}_{k+1} \tag{2}$$

De acuerdo con el teorema de descomposición ortogonal, $\mathbf{v}_{k+1}$ es ortogonal a W_k. Observe que $\text{proy}_{W_k}\,\mathbf{x}_{k+1}$ está en W_k y, por lo tanto, también en W_{k+1}. Como $\mathbf{x}_{k+1}$ está en W_{k+1}, entonces también $\mathbf{v}_{k+1}$ lo está (ya que W_{k+1} es un subespacio y es cerrado bajo la resta). Además, $\mathbf{v}_{k+1} \neq \mathbf{0}$ porque $\mathbf{x}_{k+1}$ no está en $W_k = \text{Gen}\,\{\mathbf{x}_1,\ldots,\mathbf{x}_k\}$. Por lo tanto, $\{\mathbf{v}_1,\ldots,\mathbf{v}_{k+1}\}$ es un conjunto ortogonal de vectores distintos de cero en el espacio $(k+1)$-dimensional W_{k+1}. De acuerdo con el teorema de la base de la sección 4.5, este conjunto es una base ortogonal para W_{k+1}. Por lo tanto, $W_{k+1} = \text{Gen}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_{k+1}\}$. El proceso se detiene cuando $k + 1 = p$. ■

El teorema 11 indica que cualquier subespacio W distinto de cero de $\mathbb{R}^n$ tiene una base ortogonal, porque una base ordinaria $\{\mathbf{x}_1,\ldots,\mathbf{x}_p\}$ siempre está disponible (de acuerdo con el teorema 11 de la sección 4.5), y el proceso de Gram-Schmidt solo depende de la existencia de proyecciones ortogonales sobre subespacios de W que ya tengan bases ortogonales.

Bases ortonormales

Una base ortonormal se construye con facilidad a partir de una base ortogonal $\{\mathbf{v}_1,\ldots,\mathbf{v}_p\}$: simplemente se normalizan (es decir, se "escalan") todas las $\mathbf{v}_k$. Cuando se trabajan problemas a mano, esto es más fácil que normalizar cada $\mathbf{v}_k$ conforme se vayan encontrando (porque evita la escritura innecesaria de raíces cuadradas).

EJEMPLO 3 En el ejemplo 1 se construyó la base ortogonal

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Una base ortonormal es

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{45}}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

■

Factorización QR de matrices

WEB

Si una matriz A de $m \times n$ tiene columnas linealmente independientes $\mathbf{x}_1,\ldots,\mathbf{x}_n$, la aplicación del proceso de Gram-Schmidt (con normalizaciones) a $\mathbf{x}_1,\ldots,\mathbf{x}_n$ equivale a *factorizar* A, como se describe en el siguiente teorema. Esta factorización se utiliza ampliamente en algoritmos computacionales para diversos cálculos, como la resolución de ecuaciones (que se analizó en la sección 6.5) y la determinación de valores propios (que se mencionó en los ejercicios de la sección 5.2).

TEOREMA 12 **La factorización QR**

Si A es una matriz de $m \times n$ con columnas linealmente independientes, entonces A se puede factorizar como $A = QR$, donde Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para Col A, y R es una matriz triangular superior invertible de $n \times n$ con entradas positivas en su diagonal.

DEMOSTRACIÓN Las columnas de A forman una base $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ para Col A. Construya una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ para $W = \text{Col } A$ con la propiedad (1) del teorema 11. Esta base se puede construir mediante el proceso de Gram-Schmidt o con alguna otra técnica. Sea

$$Q = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n\,]$$

Para $k = 1, \ldots, n$, $\mathbf{x}_k$ está en Gen $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ = Gen $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. Entonces existen constantes, $r_{1k}, \ldots, r_{kk}$, tales que

$$\mathbf{x}_k = r_{1k}\mathbf{u}_1 + \cdots + r_{kk}\mathbf{u}_k + 0 \cdot \mathbf{u}_{k+1} + \cdots + 0 \cdot \mathbf{u}_n$$

Podemos suponer que $r_{kk} \geq 0$. (Si $r_{kk} < 0$, multiplique r_{kk} y $\mathbf{u}_k$ por -1). Esto muestra que $\mathbf{x}_k$ es una combinación lineal de las columnas de Q empleando como pesos las entradas del vector

$$\mathbf{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Es decir, $\mathbf{x}_k = Q\mathbf{r}_k$ para $k = 1, \ldots, n$. Sea $R = [\mathbf{r}_1 \cdots \mathbf{r}_n]$. Entonces,

$$A = [\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n\,] = [\,Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_n\,] = QR$$

El hecho de que R es invertible se deduce fácilmente del hecho de que las columnas de A son linealmente independientes (ejercicio 19). Como resulta evidente que R es triangular superior, sus entradas diagonales no negativas deben ser positivas. ■

EJEMPLO 4 Encuentre una factorización QR de $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

SOLUCIÓN Las columnas de A son los vectores $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$ del ejemplo 2. En ese ejemplo, se encontró una base ortogonal para Col A = Gen $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Para simplificar la aritmética que sigue, escale $\mathbf{v}_3$ dejando que $\mathbf{v}_3' = 3\mathbf{v}_3$. Después, normalice los tres vectores para obtener $\mathbf{u}_1$, $\mathbf{u}_2$ y $\mathbf{u}_3$, y utilice esos vectores como las columnas de Q:

$$Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}$$

Por construcción, las primeras k columnas de Q son una base ortonormal de Gen $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. De la demostración del teorema 12, $A = QR$ para alguna R. Para encontrar R, observe que $Q^TQ = I$, porque las columnas de Q son ortonormales. De ahí que

$$Q^TA = Q^T(QR) = IR = R$$

y

$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix} \qquad \blacksquare$$

NOTAS NUMÉRICAS

1. Cuando el proceso de Gram-Schmidt se ejecuta en una computadora, el error por redondeo puede crecer conforme se van calculando los vectores $\mathbf{u}_k$, uno por uno. Cuando j y k son grandes, pero diferentes, los productos interiores $\mathbf{u}_j^T\mathbf{u}_k$ quizá no sean suficientemente cercanos a cero. Esta pérdida de ortogonalidad se puede reducir de manera sustancial reordenando los cálculos.[1] Sin embargo, en vez de este método de Gram-Schmidt modificado se prefiere el método de la factorización QR basado en computadora porque conduce a bases ortonormales más exactas, aun cuando la factorización requiere casi el doble de aritmética.
2. Para obtener una factorización QR de una matriz A, un programa de cómputo generalmente multiplica A por la izquierda por una secuencia de matrices ortogonales hasta que A se transforme en una matriz triangular superior. Esta construcción es análoga a la multiplicación por la izquierda con matrices elementales que produce una factorización LU de A.

PROBLEMA DE PRÁCTICA

Sea $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, donde $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Construya una base ortonormal para W.

6.4 EJERCICIOS

En los ejercicios 1 a 6, el conjunto indicado es una base para un subespacio W. Utilice el proceso de Gram-Schmidt con la finalidad de obtener una base ortogonal para W.

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 8 \\ 5 \\ -6 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$

4. $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -7 \\ -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ -9 \\ 3 \end{bmatrix}$

[1] Véase *Fundamentals of Matrix Computations*, de David S. Watkins (Nueva York: John Wiley & Sons, 1991), pp. 167-180.

7. Encuentre una base ortonormal del subespacio generado por los vectores del ejercicio 3.

8. Obtenga una base ortonormal del subespacio generado por los vectores del ejercicio 4.

En los ejercicios 9 a 12, determine una base ortogonal para el espacio columna de cada matriz.

9. $\begin{bmatrix} 3 & -5 & 1 \\ 1 & 1 & 1 \\ -1 & 5 & -2 \\ 3 & -7 & 8 \end{bmatrix}$ **10.** $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{bmatrix}$ **12.** $\begin{bmatrix} 1 & 3 & 5 \\ -1 & -3 & 1 \\ 0 & 2 & 3 \\ 1 & 5 & 2 \\ 1 & 5 & 8 \end{bmatrix}$

En los ejercicios 13 y 14, las columnas de Q se obtuvieron aplicando el proceso de Gram-Schmidt a las columnas de A. Encuentre una matriz R triangular superior tal que $A = QR$. Compruebe su resultado.

13. $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$

14. $A = \begin{bmatrix} -2 & 3 \\ 5 & 7 \\ 2 & -2 \\ 4 & 6 \end{bmatrix}, Q = \begin{bmatrix} -2/7 & 5/7 \\ 5/7 & 2/7 \\ 2/7 & -4/7 \\ 4/7 & 2/7 \end{bmatrix}$

15. Encuentre una factorización QR de la matriz del ejercicio 11.

16. Obtenga una factorización QR de la matriz del ejercicio 12.

En los ejercicios 17 y 18, todos los vectores y subespacios están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal para W, entonces la multiplicación de $\mathbf{v}_3$ por un escalar c da una nueva base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b) El proceso de Gram-Schmidt aplicado a un conjunto linealmente independiente $\{\mathbf{x}_1,\dots, \mathbf{x}_p\}$ produce un conjunto ortogonal $\{\mathbf{v}_1,\dots, \mathbf{v}_p\}$ con la propiedad de que para cada k, los vectores $\mathbf{v}_1,\dots, \mathbf{v}_k$ generan el mismo subespacio que se originó por $\mathbf{x}_1,\dots, \mathbf{x}_k$.

c) Si $A = QR$, donde Q tiene columnas ortonormales, entonces $R = Q^TA$.

18. *a*) Si $W = \text{Gen}\,\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ con $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linealmente independiente, y si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal en W, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para W.

b) Si $\mathbf{x}$ no está en un subespacio W, entonces $\mathbf{x} - \text{proy}_W\, \mathbf{x}$ no es cero.

c) En una factorización QR, por ejemplo, $A = QR$ (cuando A tiene columnas linealmente independientes), las columnas de Q forman una base ortonormal para el espacio columna de A.

19. Suponga que $A = QR$, donde Q es de $m \times n$ y R es de $n \times n$. Demuestre que si las columnas de A son linealmente independientes, entonces R debe ser invertible. [*Sugerencia:* Estudie la ecuación $R\mathbf{x} = \mathbf{0}$ y considere el hecho de que $A = QR$].

20. Suponga que $A = QR$, donde R es una matriz invertible. Demuestre que A y Q tienen el mismo espacio columna. [*Sugerencia:* Dada $\mathbf{y}$ en Col A, demuestre que $\mathbf{y} = Q\mathbf{x}$ para alguna $\mathbf{x}$. Además, dada $\mathbf{y}$ en Col Q, demuestre que $\mathbf{y} = A\mathbf{x}$ para alguna $\mathbf{x}$].

21. Dada $A = QR$ como en el teorema 12, describa cómo encontrar una matriz ortogonal Q_1 de $m \times m$ y una matriz triangular superior invertible R de $n \times n$ tales que

$$A = Q_1 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

La instrucción qr de **MATLAB** proporciona esta factorización QR “completa” cuando el rango de $A = n$.

22. Sean $\mathbf{u}_1,\dots, \mathbf{u}_p$ una base ortogonal para un subespacio W de $\mathbb{R}^n$, y $T : \mathbb{R}^n \to \mathbb{R}^n$ definida por $T(\mathbf{x}) = \text{proy}_W\, \mathbf{x}$. Demuestre que T es una transformación lineal.

23. Suponga que $A = QR$ es una factorización QR de una matriz A de $m \times n$ (con columnas linealmente independientes). Particione A como $[A_1\ A_2]$, donde A_1 tiene p columnas. Muestre cómo obtener una factorización QR de A_1, y explique por qué su factorización tiene las características adecuadas.

24. [**M**] Utilice el proceso de Gram-Schmidt, como en el ejemplo 2, y obtenga una base ortogonal para el espacio columna de

$$A = \begin{bmatrix} -10 & 13 & 7 & -11 \\ 2 & 1 & -5 & 3 \\ -6 & 3 & 13 & -3 \\ 16 & -16 & -2 & 5 \\ 2 & 1 & -5 & -7 \end{bmatrix}$$

25. [**M**] Aplique el método expuesto en esta sección para obtener una factorización de la matriz del ejercicio 24.

26. [**M**] En un programa de matrices, el proceso de Gram-Schmidt funciona mucho mejor con vectores ortonormales. Comenzando con $\mathbf{x}_1,\dots, \mathbf{x}_p$ como en el teorema 11, sea $A = [\mathbf{x}_1 \cdots x_p]$. Suponga que Q es una matriz de $n \times k$ cuyas columnas forman una base ortonormal para el subespacio W_k generado por las primeras k columnas de A. Entonces para $\mathbf{x}$ en $\mathbb{R}^n$, $QQ^T\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre W_k (teorema 10 de la sección 6.3). Si $\mathbf{x}_{k+1}$ es la siguiente columna de A, entonces la ecuación (2) en la demostración del teorema 11 se convierte en

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - Q(Q^T\mathbf{x}_{k+1})$$

(Los paréntesis en la ecuación anterior reducen el número de operaciones aritméticas). Sea $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}/\|\mathbf{v}_{k+1}\|$. La nueva Q para el siguiente paso es $[Q\quad \mathbf{u}_{k+1}]$. Utilice este procedimiento para calcular la factorización QR de la matriz del ejercicio 24. Escriba las instrucciones que utilice.

WEB

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Sean $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 0\mathbf{v}_1 = \mathbf{x}_2$. Así, $\{\mathbf{x}_1, \mathbf{x}_2\}$ ya es ortogonal. Ahora todo lo que se necesita es normalizar los vectores. Sea

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

En vez de normalizar $\mathbf{v}_2$ directamente, se normaliza $\mathbf{v}_2' = 3\mathbf{v}_2$:

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2'\|}\mathbf{v}_2' = \frac{1}{\sqrt{1^2 + 1^2 + (-2)^2}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

Entonces, $\{\mathbf{u}_1, \mathbf{u}_2\}$ es una base ortonormal para W.

6.5 PROBLEMAS DE MÍNIMOS CUADRADOS

El ejemplo introductorio a este capítulo describió un enorme problema del tipo $A\mathbf{x} = \mathbf{b}$ que no tuvo solución. En las aplicaciones son frecuentes los sistemas inconsistentes, aunque por lo general no aparecen matrices de coeficientes tan grandes. Cuando se pide una solución y esta no existe, lo mejor que se puede hacer es encontrar una $\mathbf{x}$ tal que $A\mathbf{x}$ esté lo más cerca posible de $\mathbf{b}$.

Piense que $A\mathbf{x}$ es una *aproximación* a $\mathbf{b}$. Cuanto menor sea la distancia entre $\mathbf{b}$ y $A\mathbf{x}$, dada por $\|\mathbf{b} - A\mathbf{x}\|$, mucho mejor será la aproximación. El **problema general de mínimos cuadrados** consiste en encontrar una $\mathbf{x}$ que haga a $\|\mathbf{b} - A\mathbf{x}\|$ tan pequeña como sea posible. El adjetivo “mínimos cuadrados” se origina en el hecho de que $\|\mathbf{b} - A\mathbf{x}\|$ es la raíz cuadrada de una suma de cuadrados.

DEFINICIÓN

Si A es de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$, una **solución de mínimos cuadrados** de $A\mathbf{x} = \mathbf{b}$ es un $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

para toda $\mathbf{x}$ en $\mathbb{R}^n$.

El aspecto más importante del problema de mínimos cuadrados es que sin importar cuál $\mathbf{x}$ se seleccione, el vector $A\mathbf{x}$ necesariamente estará en el espacio columna, Col A. Así que se busca un vector $\mathbf{x}$ que haga que $A\mathbf{x}$ sea el punto en Col A más cercano a $\mathbf{b}$. Véase la figura 1. (Desde luego, si $\mathbf{b}$ resulta estar en Col A, entonces $\mathbf{b}$ *es* $A\mathbf{x}$ para alguna $\mathbf{x}$, y tal $\mathbf{x}$ es una “solución de mínimos cuadrados”).

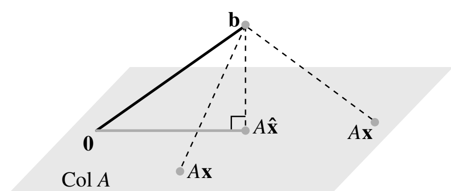


FIGURA 1 El vector $\mathbf{b}$ está más cerca de $A\hat{\mathbf{x}}$ que de $A\mathbf{x}$ para otra $\mathbf{x}$.

Solución del problema general de mínimos cuadrados

Con base en los A y $\mathbf{b}$ anteriores, aplique el teorema de la mejor aproximación que se expuso en la sección 6.3 al subespacio Col A. Sea

$$\hat{\mathbf{b}} = \text{proy}_{\text{Col } A}\, \mathbf{b}$$

Como $\hat{\mathbf{b}}$ está en el espacio columna de A, la ecuación $A\mathbf{x} = \hat{\mathbf{b}}$ *es* consistente, y existe una $\hat{\mathbf{b}}$ en $\mathbb{R}^n$ tal que

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}} \tag{1}$$

Como $\hat{\mathbf{b}}$ es el punto en Col A más cercano a $\mathbf{b}$, un vector $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ si y solo si $\hat{\mathbf{x}}$ satisface la ecuación (1). Tal $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ es un lista de pesos para construir $\hat{\mathbf{b}}$ a partir de las columnas de A. Véase la figura 2. [Existen muchas soluciones de la ecuación (1) si tiene variables libres].

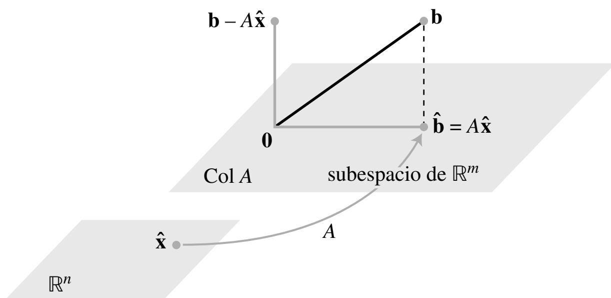


FIGURA 2 La solución de mínimos cuadrados $\hat{\mathbf{x}}$ está en $\mathbb{R}^n$.

Suponga que $\hat{\mathbf{x}}$ satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Según el teorema de descomposición ortogonal de la sección 6.3, la proyección $\hat{\mathbf{b}}$ tiene la propiedad de que $\mathbf{b} - \hat{\mathbf{b}}$ es ortogonal a Col A, de manera que $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a cada columna de A. Si $\mathbf{a}_j$ es cualquier columna de A, entonces $\mathbf{a}_j \cdot (\mathbf{b} - A\hat{\mathbf{x}}) = 0$ y $\mathbf{a}_j^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$. Como cada $\mathbf{a}_j^T$ es una fila de A^T,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0} \tag{2}$$

(Esta ecuación también se deduce del teorema 3 de la sección 6.1). Así,

$$\begin{aligned} A^T\mathbf{b} - A^TA\hat{\mathbf{x}} &= \mathbf{0} \\ A^TA\hat{\mathbf{x}} &= A^T\mathbf{b} \end{aligned}$$

Estos cálculos indican que cada solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ satisface la ecuación

$$\boxed{A^TA\mathbf{x} = A^T\mathbf{b}} \tag{3}$$

La ecuación matricial (3) representa un sistema de ecuaciones llamado las **ecuaciones normales** para $A\mathbf{x} = \mathbf{b}$. Es frecuente que cada solución del sistema (3) se denote con $\hat{\mathbf{x}}$.

TEOREMA 13

El conjunto de soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ coincide con el conjunto no vacío de soluciones de las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$.

DEMOSTRACIÓN Como ya se mostró antes, el conjunto de soluciones de mínimos cuadrados es no vacío y cada solución de mínimos cuadrados $\hat{\mathbf{x}}$ satisface las ecuaciones normales. A la inversa, suponga que $\hat{\mathbf{x}}$ satisface $A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$. Entonces, $\hat{\mathbf{x}}$ satisface la ecuación (2) anterior, lo que demuestra que $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a las filas de A^T y, por lo tanto, ortogonal a las columnas de A. Como las columnas de A generan Col A, entonces el vector $\mathbf{b} - A\hat{\mathbf{x}}$ es ortogonal a todo elemento de Col A. Por consiguiente, la ecuación

$$\mathbf{b} = A\hat{\mathbf{x}} + (\mathbf{b} - A\hat{\mathbf{x}})$$

es una descomposición de $\mathbf{b}$ en la suma de un vector en Col A y de un vector ortogonal a Col A. Por la singularidad de la descomposición ortogonal, $A\hat{\mathbf{x}}$ debe ser la proyección ortogonal de $\mathbf{b}$ sobre Col A. Es decir, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, y $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados. ■

EJEMPLO 1 Encuentre una solución de mínimos cuadrados del sistema inconsistente $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

SOLUCIÓN Para emplear las ecuaciones normales (3), calcule

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Entonces, la ecuación $A^TA\mathbf{x} = A^T\mathbf{b}$ se convierte en

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Pueden utilizarse operaciones de fila para resolver este sistema, pero como A^TA es de 2×2 e invertible, tal vez sea más rápido calcular

$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

y luego resolver $A^TA\mathbf{x} = A^T\mathbf{b}$ como

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$ ■

En muchos cálculos, A^TA es invertible, pero este no siempre es el caso. El siguiente ejemplo implica a una matriz como las que suelen presentarse en estadística, en problemas de *análisis de varianza*.

EJEMPLO 2 Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

SOLUCIÓN Calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

La matriz aumentada para $A^TA\mathbf{x} = A^T\mathbf{b}$ es

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución general es $x_1 = 3 - x_4$, $x_2 = -5 + x_4$, $x_3 = -2 + x_4$, y x_4 es libre. Así, la solución general de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ tiene la forma

$$\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$ ■

El siguiente teorema brinda útiles criterios para determinar cuándo existe solamente una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (Desde luego, la proyección ortogonal $\hat{\mathbf{b}}$ siempre es única).

TEOREMA 14

Sea A una matriz de $m \times n$. Los siguientes enunciados son lógicamente equivalentes:

a) La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución de mínimos cuadrados única para cada $\mathbf{b}$ en $\mathbb{R}^m$.

b) Las columnas de A son linealmente independientes.

c) La matriz A^TA es invertible.

Cuando estos enunciados son verdaderos, la solución $\hat{\mathbf{x}}$ de mínimos cuadrados está dada por

$$\hat{\mathbf{x}} = (A^TA)^{-1} A^T\mathbf{b} \tag{4}$$

En los ejercicios 19 a 21 se indican los elementos principales para demostrar el teorema 14; en tales ejercicios también se revisan conceptos del capítulo 4. La fórmula (4) para $\hat{\mathbf{x}}$ es útil sobre todo para fines teóricos y cálculos a mano cuando A^TA es una matriz invertible de 2×2.

Cuando se usa una solución $\hat{\mathbf{x}}$ de mínimos cuadrados para producir $A\hat{\mathbf{x}}$ como una aproximación a $\mathbf{b}$, entonces la distancia de $\mathbf{b}$ a $A\hat{\mathbf{x}}$ se llama el **error de mínimos cuadrados** de esta aproximación.

EJEMPLO 3 Dados A y $\mathbf{b}$ como en el ejemplo 1, determine el error de mínimos cuadrados en la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

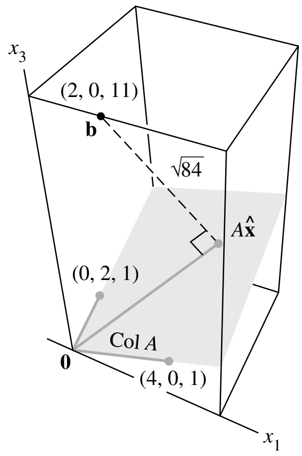


FIGURA 3

SOLUCIÓN A partir del ejemplo 1,

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{y} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

Por lo tanto,

$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

y

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

El error de mínimos cuadrados es $\sqrt{84}$. Para cualquier $\mathbf{x}$ en $\mathbb{R}^2$, la distancia entre $\mathbf{b}$ y el vector $A\mathbf{x}$ es al menos $\sqrt{84}$. Véase la figura 3. Observe que la solución $\hat{\mathbf{x}}$ de mínimos cuadrados no se presenta en la figura. ■

Cálculos alternativos de soluciones de mínimos cuadrados

El siguiente ejemplo muestra cómo encontrar una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortogonales. Con frecuencia tales matrices se presentan en problemas de regresión lineal, como los que se analizarán en la siguiente sección.

EJEMPLO 4 Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

SOLUCIÓN Como las columnas $\mathbf{a}_1$ y $\mathbf{a}_2$ de A son ortogonales, la proyección ortogonal de $\mathbf{b}$ sobre Col A está dada por

$$\begin{aligned} \hat{\mathbf{b}} &= \frac{\mathbf{b}\cdot\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b}\cdot\mathbf{a}_2}{\mathbf{a}_2\cdot\mathbf{a}_2}\mathbf{a}_2 = \frac{8}{4}\mathbf{a}_1 + \frac{45}{90}\mathbf{a}_2 \\ &= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix} \end{aligned} \tag{5}$$

Ahora que se conoce $\hat{\mathbf{b}}$, se puede resolver $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Pero esto es trivial, porque ya se sabe que los pesos a colocar sobre las columnas de A producen $\hat{\mathbf{b}}$. A partir de la ecuación (5) es claro que

$$\hat{\mathbf{x}} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$ ■

En algunos casos, es posible que las ecuaciones normales para un problema de mínimos cuadrados estén *mal condicionadas*; es decir, en ocasiones, pequeños errores en los cálculos de las entradas de A^TA causan grandes errores en la solución $\hat{\mathbf{x}}$. Si las columnas de A son linealmente independientes, la solución de mínimos cuadrados con frecuencia se puede calcular de manera más confiable con una factorización QR de A (descrita en la sección 6.4).[1]

[1] El método QR se compara con el método de la ecuación normal estándar en G. Golub y C. Van Loan, *Matrix Computations*, 3a. ed. (Baltimore: Johns Hopkins Press, 1996), pp. 230-231.

TEOREMA 15

Dada una matriz A de $m \times n$, con columnas linealmente independientes, sea $A = QR$ una factorización QR de A como en el teorema 12. Entonces, para cada $\mathbf{b}$ en $\mathbb{R}^m$, la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución de mínimos cuadrados única, dada por

$$\hat{\mathbf{x}} = R^{-1} Q^T \mathbf{b} \tag{6}$$

DEMOSTRACIÓN Sea $\hat{\mathbf{x}} = R^{-1} Q^T \mathbf{b}$. Entonces,

$$A\hat{\mathbf{x}} = QR\hat{\mathbf{x}} = QRR^{-1}Q^T\mathbf{b} = QQ^T\mathbf{b}$$

De acuerdo con el teorema 12, las columnas de Q forman una base ortonormal para Col A. Por lo tanto, según el teorema 10, $QQ^T\mathbf{b}$ es la proyección ortogonal $\hat{\mathbf{b}}$ de $\mathbf{b}$ sobre Col A. Así, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, lo que muestra que $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. La unicidad de $\hat{\mathbf{x}}$ se deduce del teorema 14. ■

NOTA NUMÉRICA

Como en el teorema 15, R es triangular superior, entonces $\hat{\mathbf{x}}$ se debería calcular como la solución exacta de la ecuación

$$R\mathbf{x} = Q^T\mathbf{b} \tag{7}$$

Es mucho más rápido resolver (7), por sustitución hacia atrás o mediante operaciones de fila, que calcular R^{-1} y utilizar la ecuación (6).

EJEMPLO 5 Encuentre la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix}$$

SOLUCIÓN La factorización QR de A se obtiene como en la sección 6.4

$$A = QR = \begin{bmatrix} 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

Luego,

$$Q^T\mathbf{b} = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 7 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

La solución de mínimos cuadrados $\hat{\mathbf{x}}$ satisface $R\mathbf{x} = Q^T\mathbf{b}$; es decir,

$$\begin{bmatrix} 2 & 4 & 5 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \\ 4 \end{bmatrix}$$

Esta ecuación se resuelve fácilmente y conduce a $\hat{\mathbf{x}} = \begin{bmatrix} 10 \\ -6 \\ 2 \end{bmatrix}$. ■

PROBLEMAS DE PRÁCTICA

1. Sean $A = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$. Encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, y calcule el error de mínimos cuadrados asociado.
2. ¿Qué se puede decir acerca de la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ es ortogonal a las columnas de A?

6.5 EJERCICIOS

En los ejercicios 1 a 4, encuentre una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ mediante *a*) la construcción de las ecuaciones normales para $\hat{\mathbf{x}}$ y *b*) el despeje de $\hat{\mathbf{x}}$.

1. $A = \begin{bmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$

2. $A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \\ 2 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & -2 \\ -1 & 2 \\ 0 & 3 \\ 2 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ -4 \\ 2 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}$

En los ejercicios 5 y 6, describa todas las soluciones de mínimos cuadrados de la ecuación $A\mathbf{x} = \mathbf{b}$.

5. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 8 \\ 2 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{bmatrix}$

7. Calcule el error de mínimos cuadrados asociado con la solución de mínimos cuadrados que encontró en el ejercicio 3.

8. Determine el error de mínimos cuadrados asociado con la solución de mínimos cuadrados que encontró en el ejercicio 4.

En los ejercicios 9 a 12, encuentre *a*) la proyección ortogonal de $\mathbf{b}$ sobre Col A y *b*) una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

9. $A = \begin{bmatrix} 1 & 5 \\ 3 & 1 \\ -2 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ -2 \\ -3 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$

11. $A = \begin{bmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 6 \end{bmatrix}$

13. Sean $A = \begin{bmatrix} 3 & 4 \\ -2 & 1 \\ 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 11 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$. Calcule $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Podría ser $\mathbf{u}$ una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin calcular una solución de mínimos cuadrados).

14. Sean $A = \begin{bmatrix} 2 & 1 \\ -3 & -4 \\ 3 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Determine $A\mathbf{u}$ y $A\mathbf{v}$, y compárelos con $\mathbf{b}$. ¿Es posible que al menos uno de los dos entre $\mathbf{u}$ o $\mathbf{v}$ sea una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$? (Responda sin obtener una solución de mínimos cuadrados).

En los ejercicios 15 y 16, utilice la factorización $A = QR$ para encontrar la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

15. $A = \begin{bmatrix} 2 & 3 \\ 2 & 4 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 \\ 2/3 & 2/3 \\ 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$

16. $A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$

En los ejercicios 17 y 18, A es una matriz de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

17. *a*) El problema de mínimos cuadrados general consiste en obtener una $\mathbf{x}$ que haga que $A\mathbf{x}$ esté tan cerca de $\mathbf{b}$ como sea posible.

b) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ que satisface $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, donde $\hat{\mathbf{b}}$ es la proyección ortogonal de $\mathbf{b}$ sobre Col A.

c) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es un vector $\hat{\mathbf{x}}$ tal que $\|\mathbf{b} - A\mathbf{x}\| \leq \|\mathbf{b} - A\hat{\mathbf{x}}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

d) Cualquier solución de $A^TA\mathbf{x} = A^T\mathbf{b}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

e) Si las columnas de A son linealmente independientes, entonces la ecuación $A\mathbf{x} = \mathbf{b}$ tiene exactamente una solución de mínimos cuadrados.

18. *a*) Si $\mathbf{b}$ está en el espacio columna de A, entonces cada solución de $A\mathbf{x} = \mathbf{b}$ es una solución de mínimos cuadrados.

b) La solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el punto en el espacio columna de A más cercano a $\mathbf{b}$.

c) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es una lista de pesos que, cuando se aplica a las columnas de A, produce la proyección ortogonal de $\mathbf{b}$ sobre Col A.

d) Si $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, entonces $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

e) Las ecuaciones normales siempre ofrecen un método confiable para calcular soluciones de mínimos cuadrados.

f) Si A tiene una factorización QR, por ejemplo, $A = QR$, entonces la mejor manera de obtener una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es calcular $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.

19. Sea A una matriz de $m \times n$. Realice los siguientes pasos para demostrar que un vector $\mathbf{x}$ en $\mathbb{R}^n$ satisface $A\mathbf{x} = \mathbf{0}$ si y solo si $A^TA\mathbf{x} = \mathbf{0}$. Esto demostrará que Nul A = Nul A^TA.

a) Demuestre que si $A\mathbf{x} = \mathbf{0}$, entonces $A^TA\mathbf{x} = \mathbf{0}$.

b) Suponga que $A^TA\mathbf{x} = \mathbf{0}$. Explique por qué $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$, y use esto para probar que $A\mathbf{x} = \mathbf{0}$.

20. Sea A una matriz de $m \times n$ tal que A^TA es invertible. Demuestre que las columnas de A son linealmente independientes. [*Precaución:* No suponga que A sea invertible; es más, tal vez ni siquiera sea cuadrada].

21. Sea A una matriz de $m \times n$ cuyas columnas son linealmente independientes. [*Precaución: A* no necesariamente es cuadrada].

a) Con base en el ejercicio 19, demuestre que A^TA es una matriz invertible.

b) Explique por qué A debe tener, al menos, tantas filas como columnas.

c) Determine el rango de A.

22. Con base en el ejercicio 19, demuestre que rango A^TA = rango A. [*Sugerencia:* ¿Cuántas columnas tiene A^TA? ¿Cómo se relaciona esto con el rango de A^TA?].

23. Suponga que A es de $m \times n$ con columnas linealmente independientes y que $\mathbf{b}$ está en $\mathbb{R}^m$. Utilice las ecuaciones normales con la finalidad de obtener una fórmula para $\hat{\mathbf{b}}$, la proyección de $\mathbf{b}$ sobre Col A. [*Sugerencia:* Primero encuentre $\hat{\mathbf{x}}$. La fórmula no requiere una base ortogonal para Col A].

24. Obtenga una fórmula para la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ cuando las columnas de A son ortonormales.

25. Describa todas las soluciones de mínimos cuadrados del sistema

$$x + y = 2$$
$$x + y = 4$$

26. [**M**] El ejemplo 3 de la sección 4.8 mostró un filtro lineal pasa bajos que cambió la señal $\{y_k\}$ en $\{y_{k+1}\}$ y transformó la señal de alta frecuencia $\{w_k\}$ a una señal cero, donde $y_k = \cos(\pi k/4)$ y $w_k = \cos(3\pi k/4)$. Los siguientes cálculos diseñarán un filtro con aproximadamente esas propiedades. La ecuación del filtro es

$$a_0 y_{k+2} + a_1 y_{k+1} + a_2 y_k = z_k \quad \text{para toda } k \tag{8}$$

Como las señales son periódicas, con periodo 8, basta con estudiar la ecuación (8) para $k = 0, \ldots, 7$. La acción sobre las dos señales descritas se traduce en dos conjuntos de ocho ecuaciones, que se muestran a continuación:

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array} \\ \begin{bmatrix} 0 & .7 & 1 \\ -.7 & 0 & .7 \\ -1 & -.7 & 0 \\ -.7 & -1 & -.7 \\ 0 & -.7 & -1 \\ .7 & 0 & -.7 \\ 1 & .7 & 0 \\ .7 & 1 & .7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{array}{c} y_{k+1} \\ \begin{bmatrix} .7 \\ 0 \\ -.7 \\ -1 \\ -.7 \\ 0 \\ .7 \\ 1 \end{bmatrix} \end{array}$$

$$\begin{array}{c} \\ k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\begin{array}{c} \begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array} \\ \begin{bmatrix} 0 & -.7 & 1 \\ .7 & 0 & -.7 \\ -1 & .7 & 0 \\ .7 & -1 & .7 \\ 0 & .7 & -1 \\ -.7 & 0 & .7 \\ 1 & -.7 & 0 \\ -.7 & 1 & -.7 \end{bmatrix} \end{array}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Escriba una ecuación $A\mathbf{x} = \mathbf{b}$, donde A es una matriz de 16×3 formada por las dos matrices de coeficientes anteriores y donde $\mathbf{b}$ en $\mathbb{R}^{16}$ se forma a partir de los dos lados derechos de las ecuaciones. Encuentre a_0, a_1 y a_2 dadas por la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. (El .7 en los datos anteriores se empleó como una aproximación a $\sqrt{2}/2$, para ilustrar cómo se procede en un cálculo típico en un problema aplicado. Si, en vez de ello, se utilizara .707, los coeficientes del filtro resultante concordarían, al menos, en siete decimales con $\sqrt{2}/4$, $1/2$ y $\sqrt{2}/4$, los valores obtenidos mediante cálculos aritméticos exactos).

WEB

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Primero calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

Después, reduzca por filas la matriz aumentada para las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución de mínimos cuadrados general es $x_1 = 2 + \frac{3}{2}x_3$, $x_2 = -1 - \frac{1}{2}x_3$, con x_3 libre. Para una solución específica, tome $x_3 = 0$ (por ejemplo), y obtenga

$$\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

Para encontrar el error de mínimos cuadrados, calcule

$$\hat{\mathbf{b}} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$$

Resulta que $\hat{\mathbf{b}} = \mathbf{b}$, de manera que $\|\mathbf{b} - \hat{\mathbf{b}}\| = 0$. El error de mínimos cuadrados es cero porque $\mathbf{b}$ está en Col A.

2. Si $\mathbf{b}$ es ortogonal a las columnas de A, entonces la proyección de $\mathbf{b}$ sobre el espacio columna de A es $\mathbf{0}$. En este caso, una solución de mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ satisface $A\hat{\mathbf{x}} = \mathbf{0}$.

6.6 APLICACIONES A MODELOS LINEALES

Una tarea común en ciencia e ingeniería es analizar y entender relaciones entre diferentes cantidades que varían. Esta sección describe una variedad de situaciones en las que los datos se emplean para construir o comprobar una fórmula que predice el valor de una variable como función de otras variables. En cada caso, el reto será equivalente a resolver un problema de mínimos cuadrados.

Para facilitar la aplicación del análisis a problemas reales que encontrará en sus actividades profesionales, se elige la notación que comúnmente se utiliza en el análisis estadístico de datos científicos y de ingeniería. En vez de escribir $A\mathbf{x} = \mathbf{b}$, se escribe $X\boldsymbol{\beta} = \mathbf{y}$, donde X es la **matriz de diseño**, $\boldsymbol{\beta}$ es el **vector de parámetros** y $\mathbf{y}$ es el **vector de observaciones**.

Rectas de mínimos cuadrados

La relación más sencilla entre dos variables x y y es la ecuación lineal $y = \beta_0 + \beta_1 x$.[1] Con frecuencia los datos experimentales generan los puntos $(x_1, y_1), \ldots, (x_n, y_n)$ que, cuando

[1] Es común utilizar esta notación para rectas de mínimos cuadrados en vez de $y = mx + b$.

se grafican, parecen estar cerca de una recta. Se desea determinar los parámetros β_0 y β_1 para hacer que la recta esté lo más "cerca" posible de dichos puntos.

Suponga que β_0 y β_1 están fijos, y considere la recta $y = \beta_0 + \beta_1 x$ de la figura 1. Para cada dato (x_j, y_j) existe un punto $(x_j, \beta_0 + \beta_1 x_j)$ sobre la recta con la misma coordenada x. Por otro lado, y_j es el valor *observado* de y, y $\beta_0 + \beta_1 x_j$ es el valor *predicho* para y (determinado por la recta). La diferencia entre los valores observado y predicho para y se llama *residuo*.

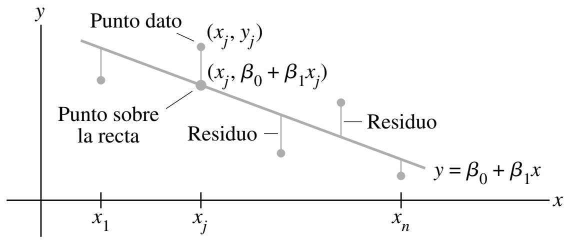


FIGURA 1 Ajuste de datos experimentales a una recta.

Existen varias maneras de medir qué tan "cerca" está la recta respecto de los datos. La elección habitual (sobre todo porque los cálculos matemáticos son sencillos) es sumar los cuadrados de los residuos. La **recta de mínimos cuadrados** es la recta $y = \beta_0 + \beta_1 x$ que minimiza la suma de los cuadrados de los residuos**.** A esta recta también se le conoce como **recta de regresión de y sobre x**, porque se supone que cualquier error en los datos solo ocurre en las coordenadas y. Los coeficientes β_0, β_1 de la recta son los **coeficientes de regresión**.[2]

Si los puntos de los datos estuvieran sobre una recta, los parámetros β_0 y β_1 satisfarían las ecuaciones

Valor predicho de *y*		**Valor observado de *y***
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

Este sistema se puede representar como

$$X\beta = \mathbf{y} \quad \text{donde} \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \tag{1}$$

Desde luego, si los puntos de datos no están sobre una recta, entonces no existen parámetros β_0, β_1 para los cuales los valores predichos de y en $X\beta$ sean iguales a los valores observados de y en $\mathbf{y}$, y $X\beta = \mathbf{y}$ no tiene solución. Este es un problema de mínimos cuadrados, $A\mathbf{x} = \mathbf{b}$, ¡con diferente notación!

El cuadrado de la distancia entre los vectores $X\beta$ y $\mathbf{y}$ es precisamente la suma de los cuadrados de los residuos. El β que minimiza esta suma también minimiza la distancia entre $X\beta$ y $\mathbf{y}$. *Calcular la solución de mínimos cuadrados de $X\beta = \mathbf{y}$ equivale a encontrar el vector β que determina la recta de mínimos cuadrados de la figura 1.*

[2] Si los errores de medición estuvieran en x y no en y, simplemente se intercambiarían las coordenadas de los datos (x_j, y_j) antes de trazar la gráfica de los puntos y calcular la recta de regresión. Si ambas coordenadas están sujetas a posibles errores, entonces se podría elegir la recta que minimiza la suma de los cuadrados de las distancias *ortogonales* (perpendiculares) de los puntos a la recta. Véase los problemas de práctica de la sección 7.5.

EJEMPLO 1 Encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la recta de mínimos cuadrados que mejor se ajuste a los puntos de datos (2, 1), (5, 2), (7, 3) y (8, 3).

SOLUCIÓN Utilice las coordenadas x de los datos para construir la matriz de diseño X en la ecuación (1) y las coordenadas y para construir el vector de observaciones $\mathbf{y}$:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

Para la solución de mínimos cuadrados de $X\beta = \mathbf{y}$, obtenga las ecuaciones normales (con la nueva notación):

$$X^T X \beta = X^T \mathbf{y}$$

Es decir, calcule

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Las ecuaciones normales son

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Por lo tanto,

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

De manera que la recta de mínimos cuadrados tiene la ecuación

$$y = \frac{2}{7} + \frac{5}{14} x$$

Véase la figura 2. ■

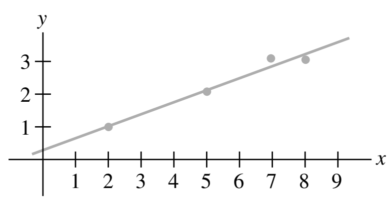


FIGURA 2 La recta de mínimos cuadrados $y = \frac{2}{7} + \frac{5}{14}x$.

Una práctica común, antes de calcular la recta de mínimos cuadrados, consiste en calcular el promedio $\bar{x}$ de los valores x originales y formar una nueva variable $x^* = x - \bar{x}$. Se dice que los nuevos datos x quedan en su **forma de desviación media**. En este caso, las dos columnas de la matriz de diseño serán ortogonales. Se simplifica la solución de las ecuaciones normales, justo como en el ejemplo 4 de la sección 6.5. Véase los ejercicios 17 y 18.

Modelo lineal general

En algunas aplicaciones, es necesario ajustar puntos de datos a algo diferente de una línea recta. En los ejemplos que siguen, la ecuación matricial continúa siendo $X\beta = \mathbf{y}$, pero la forma específica de X cambia de un problema a otro. Por lo general, los especialistas en estadística introducen un **vector residual** $\boldsymbol{\epsilon}$, definido como $\boldsymbol{\epsilon} = \mathbf{y} - X\beta$, y que se escribe

$$\mathbf{y} = X\beta + \boldsymbol{\epsilon}$$

Cualquier ecuación de esta forma es un **modelo lineal**. Una vez que X y $\mathbf{y}$ están determinadas, el objetivo es minimizar la longitud de $\boldsymbol{\epsilon}$, lo que equivale a encontrar una solución de mínimos cuadrados de $X\beta = \mathbf{y}$. En cada caso, la solución de mínimos cuadrados $\hat{\beta}$ es una solución de las ecuaciones normales

$$X^T X\beta = X^T\mathbf{y}$$

Ajuste de otras curvas con mínimos cuadrados

Cuando los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ en una gráfica de dispersión no se encuentran cerca de una recta, tal vez resulte pertinente postular alguna otra relación funcional entre x y y.

Los siguientes dos ejemplos muestran cómo ajustar datos mediante curvas que tienen la forma general

$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x) \tag{2}$$

donde $f_0, \ldots, f_k$ son funciones conocidas y $\beta_0, \ldots, \beta_k$ son parámetros que se deben determinar. Como se verá, la ecuación (2) describe un modelo lineal porque es lineal en los parámetros desconocidos.

Para un valor particular de x, la ecuación (2) da un valor predicho, o "ajustado", de y. La diferencia entre el valor observado y el valor predicho es el residuo. Los parámetros $\beta_0, \ldots, \beta_k$ se deben determinar para minimizar la suma de los cuadrados de los residuos.

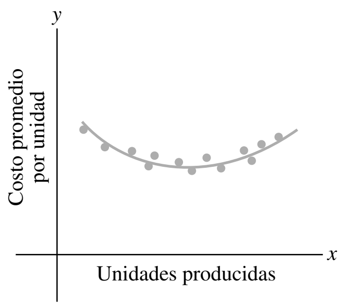


FIGURA 3
Curva de costo promedio.

EJEMPLO 2 Suponga que los puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ parecen estar sobre una parábola y no sobre una recta. Por ejemplo, si la coordenada x denota el nivel de producción de una compañía, y y representa el costo promedio por unidad de operación en un nivel de x unidades por día, entonces una curva típica de costo promedio parece una parábola que se abre hacia arriba (figura 3). En ecología, se usa una curva parabólica que se abre hacia abajo para modelar la producción primaria neta de nutrientes en una planta, como una función del área superficial del follaje (figura 4). Suponga que deseamos aproximar los datos mediante una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 \tag{3}$$

Describa el modelo lineal que produce un "ajuste de mínimos cuadrados" de los datos, empleando la ecuación (3).

SOLUCIÓN La ecuación (3) describe la relación ideal. Suponga que los valores reales de los parámetros son $\beta_0, \beta_1, \beta_2$. Entonces las coordenadas del primer punto de datos (x_1, y_1) satisfacen una ecuación de la forma

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1$$

donde ϵ_1 es el error residual entre el valor observado y_1 y el valor predicho $\beta_0 + \beta_1 x_1 + \beta_2 x_1^2$ para y. Cada punto de datos determina una ecuación similar:

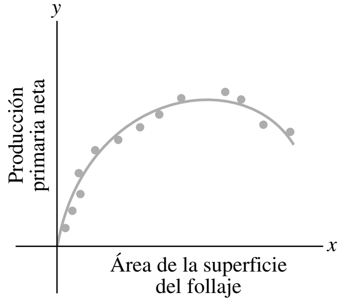


FIGURA 4
Producción de nutrientes.

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \epsilon_2 \\ &\vdots \qquad\qquad \vdots \\ y_n &= \beta_0 + \beta_1 x_n + \beta_2 x_n^2 + \epsilon_n \end{aligned}$$

Es sencillo escribir este sistema de ecuaciones en la forma $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$. Para encontrar X, inspeccione las primeras filas del sistema y observe el patrón.

$$\begin{array}{ccccccc} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} & = & \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix} & & \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} & + & \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \\ \mathbf{y} & = & X & & \beta & + & \boldsymbol{\epsilon} \end{array}$$

■

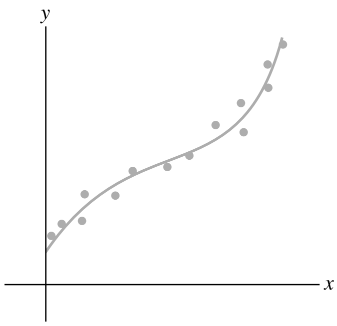


FIGURA 5
Puntos de datos a lo largo de una curva cúbica.

EJEMPLO 3 Si los puntos de datos tienden a seguir un patrón como en la figura 5, entonces un modelo adecuado podría ser una ecuación de la forma

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$$

Dichos datos podrían ser, por ejemplo, los costos totales de una compañía, como una función del nivel de producción. Describa el modelo lineal que da un ajuste de mínimos cuadrados de este tipo para los datos $(x_1, y_1), \ldots, (x_n, y_n)$.

SOLUCIÓN Mediante un análisis similar al efectuado en el ejemplo 2, se obtiene

$$\begin{array}{cccc} \text{Vector de observaciones} & \text{Matriz de diseño} & \text{Vector de parámetros} & \text{Vector residual} \\ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, & X = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 \end{bmatrix}, & \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, & \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \end{array}$$

■

Regresión múltiple

Suponga que en un experimento hay dos variables independientes (por ejemplo, u y v) y una variable dependiente, y. Una ecuación sencilla para predecir y a partir de u y v tiene la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v \tag{4}$$

Una ecuación predictiva más general podría tener la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v + \beta_3 u^2 + \beta_4 uv^2 + \beta_5 v^2 \tag{5}$$

Esta ecuación se utiliza en geología, por ejemplo, para modelar superficies de erosión, glaciares, el pH del suelo y otras cantidades. En tales casos, el ajuste por mínimos cuadrados se llama *superficie de tendencia*.

Tanto la ecuación (4) como la (5) conducen a un modelo lineal porque son lineales en los parámetros desconocidos (aun cuando u y v están multiplicados). En general, un modelo lineal surgirá siempre que y se prediga mediante una ecuación de la forma

$$y = \beta_0 f_0(u, v) + \beta_1 f_1(u, v) + \cdots + \beta_k f_k(u, v)$$

con las funciones conocidas $f_0, \ldots, f_k$ y los pesos desconocidos $\beta_0, \ldots, \beta_k$.

EJEMPLO 4 En geografía, los modelos locales del terreno se construyen mediante los datos $(u_1, v_1, y_1), \ldots, (u_n, v_n, y_n)$, donde u_j, v_j y y_j son latitud, longitud y altitud, respectivamente. La solución es el *plano de mínimos cuadrados*. Véase la figura 6.

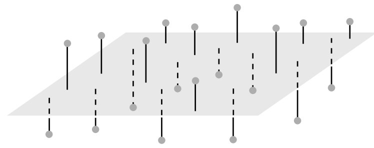


FIGURA 6 Un plano de mínimos cuadrados.

SOLUCIÓN Se espera que los datos satisfagan las siguientes ecuaciones

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \epsilon_2 \\ &\ \vdots \qquad\qquad\qquad \vdots \\ y_n &= \beta_0 + \beta_1 u_n + \beta_2 v_n + \epsilon_n \end{aligned}$$

Este sistema tiene la forma matricial $\mathbf{y} = X\beta + \epsilon$, donde

$$\begin{array}{cccc} \text{Vector de observaciones} & \text{Matriz de diseño} & \text{Vector de parámetros} & \text{Vector residual} \\ \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, & X = \begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}, & \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, & \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \end{array}$$

■

El ejemplo 4 revela que el modelo lineal para regresión múltiple tiene la misma forma abstracta que el modelo para regresión simple de los primeros ejemplos. El álgebra lineal nos permite entender el principio general subyacente en todos los modelos lineales. Una vez que X se define adecuadamente, las ecuaciones normales para β tienen la misma forma matricial, sin importar cuántas variables estén implicadas. Así, para cualquier modelo lineal donde X^TX sea invertible, el $\hat{\beta}$ de mínimos cuadrados está dado por $(X^TX)^{-1}X^T\mathbf{y}$.

Lecturas adicionales

Ferguson, J., *Introduction to Linear Algebra in Geology* (Nueva York: Chapman & Hall, 1994).

Krumbein, W. C. y F. A. Graybill, *An Introduction to Statistical Models in Geology* (Nueva York: McGraw-Hill, 1965).

Legendre, P. y L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in Modern Geography, No. 5 (Norwich, Inglaterra: Geo Books, 1975).

PROBLEMA DE PRÁCTICA

Cuando las ventas mensuales de un producto están sujetas a las fluctuaciones estacionales, entonces una curva que aproxime los datos de venta podría tener la forma

$$y = \beta_0 + \beta_1 x + \beta_2 \operatorname{sen}(2\pi x/12)$$

donde x es el tiempo en meses. El término $\beta_0 + \beta_1 x$ da la tendencia básica de ventas, y el término seno refleja los cambios estacionales en las ventas. Determine la matriz de diseño y el vector de parámetros para el modelo lineal que conduce a un ajuste de mínimos cuadrados de la ecuación anterior. Suponga que los datos son $(x_1, y_1), \ldots, (x_n, y_n)$.

6.6 EJERCICIOS

En los ejercicios 1 a 4, encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la recta de mínimos cuadrados que mejor se ajuste a los puntos de datos indicados.

1. (0, 1), (1, 1), (2, 2), (3, 2)

2. (1, 0), (2, 1), (4, 2), (5, 3)

3. (-1), (0, 1), (1, 2), (2, 4)

4. (2, 3), (3, 2), (5, 1), (6, 0)

5. Sea X la matriz de diseño empleada para determinar la recta de mínimos cuadrados que se ajusta a los datos $(x_1, y_1), \ldots, (x_n, y_n)$. Utilice un teorema de la sección 6.5 para demostrar que las ecuaciones normales tienen solución única si y solo si los datos incluyen al menos dos puntos de datos con diferente coordenada x.

6. Sea X la matriz de diseño del ejemplo 2 correspondiente a un ajuste de mínimos cuadrados de los datos $(x_1, y_1), \ldots, (x_n, y_n)$ a una curva parabólica. Suponga que x_1, x_2 y x_3 son distintas. Explique por qué solo existe una parábola que mejor se ajusta a los datos, en el sentido de mínimos cuadrados. (Véase el ejercicio 5).

7. Un cierto experimento genera los datos (1, 1.8), (2, 2.7), (3, 3.4), (4, 3.8), (5, 3.9). Describa el modelo que produce un ajuste de mínimos cuadrados de esos puntos mediante una función de la forma

$$y = \beta_1 x + \beta_2 x^2$$

Se podría presentar dicha función, por ejemplo, como el ingreso derivado de la venta de x unidades de un producto, cuando la cantidad ofrecida para la venta afecta el precio del producto.

a) Determine la matriz de diseño, el vector de observaciones y el vector de parámetros desconocidos.

b) [**M**] Encuentre la curva de mínimos cuadrados asociada con los datos.

8. Una curva sencilla que con frecuencia es un buen modelo para los costos variables de una compañía, como una función del nivel x de ventas, tiene la forma $y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3$. No existe término constante porque no se incluyen los costos fijos.

a) Determine la matriz de diseño y el vector de parámetros para el modelo lineal que conduce a un ajuste de mínimos cuadrados de la ecuación anterior, con los datos $(x_1, y_1), \ldots, (x_n, y_n)$.

b) [**M**] Encuentre la curva de mínimos cuadrados de la forma anterior para ajustar los datos (4, 1.58), (6, 2.08), (8, 2.5), (10, 2.8), (12, 3.1), (14, 3.4), (16, 3.8) y (18, 4.32), con valores en miles. Si es posible, realice una gráfica que muestre los puntos de datos y la curva de aproximación cúbica.

9. Un cierto experimento genera los datos (1, 7.9), (2, 5.4) y (3, $-.9$). Describa el modelo que da un ajuste de mínimos cuadrados de esos puntos mediante una función de la forma

$$y = A \cos x + B \operatorname{sen} x$$

10. Suponga que las sustancias radiactivas A y B tienen constantes de decaimiento de .02 y .07, respectivamente. Si en el momento $t = 0$ una mezcla de esas dos sustancias contiene M_A gramos de A y M_B gramos de B, entonces un modelo para la cantidad total y de la mezcla presente en el momento t es

$$y = M_A e^{-.02t} + M_B e^{-.07t}$$

Suponga que no se conocen las cantidades iniciales M_A y M_B, pero un científico logra medir las cantidades totales presentes en diferentes momentos y registra los siguientes puntos (t_i, y_i): (10, 21.34), (11, 20.68), (12, 20.05), (14, 18.87) y (15, 18.30).

a) Describa un modelo lineal que se pueda utilizar para estimar M_A y M_B.

b) [**M**] Encuentre la curva de mínimos cuadrados basada en la ecuación (6).

En 1986 fue la última aparición del cometa Halley, el cual reaparecerá en el año 2061.

11. [**M**] De acuerdo con la segunda ley de Kepler, un cometa debería tener una órbita elíptica, parabólica o hiperbólica (ignorando las atracciones gravitacionales de los planetas). En convenientes coordenadas polares, la posición (r, ϑ) de un cometa satisface una ecuación de la forma

$$r = \beta + e(r \cdot \cos \vartheta)$$

donde β es una constante y e es la *excentricidad* de la órbita, con $0 \le e < 1$ para una elipse, $e = 1$ para una parábola, y $e > 1$ para una hipérbola. Suponga que los siguientes datos corresponden a las observaciones de un cometa recién descubierto. Determine el tipo de la órbita e indique dónde estará el cometa cuando $\vartheta = 4.6$ radianes.[3]

ϑ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

12. [**M**] La presión sanguínea sistólica p (en milímetros de mercurio) de un niño saludable y su peso w (en libras) están relacionados aproximadamente mediante la ecuación

$$\beta_0 + \beta_1 \ln w = p$$

Utilice los siguientes datos experimentales para estimar la presión sanguínea sistólica de un niño saludable que pesa 100 libras.

[3] La idea básica del ajuste de mínimos cuadrados de los datos se debe a K. F. Gauss (e, independientemente, a A. Legendre), cuya fama despuntó en 1801 cuando utilizó el método para obtener la trayectoria del asteroide *Ceres*. Cuarenta días después de que el asteroide fue descubierto, desapareció detrás del Sol. Gauss predijo que Ceres aparecería 10 meses después y precisó su ubicación. La exactitud de la predicción asombró a la comunidad científica europea.

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.39	4.73	4.88
p	91	98	103	110	112

13. [**M**] Para medir el desempeño de un avión durante el despegue, cada segundo se midió su posición horizontal, de $t = 0$ a $t = 12$. Las posiciones (en pies) fueron: 0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, y 809.2.

a) Encuentre la curva cúbica de mínimos cuadrados $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ para esos datos.

b) Con base en el resultado del inciso *a*), estime la velocidad del avión cuando $t = 4.5$ segundos.

14. Sean $\overline{x} = \frac{1}{n}(x_1 + \cdots + x_n)$ y $\overline{y} = \frac{1}{n}(y_1 + \cdots + y_n)$. Demuestre que la recta de mínimos cuadrados para los datos $(x_1, y_1), \ldots, (x_n, y_n)$ debe pasar a través de $(\bar{x}, \bar{y})$. Es decir, demuestre que $\bar{x}$ y $\bar{y}$ satisfacen la ecuación lineal $\bar{y} = \beta_0 + \beta_1\bar{x}$. [*Sugerencia:* Deduzca esta ecuación mediante la ecuación vectorial $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$. Denote la primera columna de X con $\mathbf{1}$. Con base en el hecho de que el vector residual $\boldsymbol{\epsilon}$ es ortogonal al espacio columna de X y, por lo tanto, es ortogonal a $\mathbf{1}$].

Considerando los datos para un problema de mínimos cuadrados, $(x_1, y_1), \ldots, (x_n, y_n)$, son útiles las siguientes abreviaciones:

$$\sum x = \sum_{i=1}^n x_i, \quad \sum x^2 = \sum_{i=1}^n x_i^2,$$
$$\sum y = \sum_{i=1}^n y_i, \quad \sum xy = \sum_{i=1}^n x_i y_i$$

Las ecuaciones normales para una recta de mínimos cuadrados $y = \hat{\beta}_0 + \hat{\beta}_1 x$ se pueden escribir en la forma

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \sum x &= \sum y \\ \hat{\beta}_0 \sum x + \hat{\beta}_1 \sum x^2 &= \sum xy \end{aligned} \qquad (7)$$

15. Deduzca las ecuaciones normales (7) a partir de la forma matricial presentada en esta sección.

16. Utilice una matriz inversa para resolver el sistema de ecuaciones (7) y así obtener las fórmulas para $\hat{\beta}_0$ y $\hat{\beta}_1$ que aparecen en muchos libros de estadística.

17. *a*) Rescriba los datos del ejemplo 1 con nuevas coordenadas x en la forma de desviación media. Sea X la matriz de diseño asociada. ¿Por qué son ortogonales las columnas de X?

b) Escriba las ecuaciones normales para los datos del inciso *a*), y resuélvalas para encontrar la recta de mínimos cuadrados, $y = \beta_0 + \beta_1 x^*$, donde $x^* = x - 5.5$.

18. Suponga que las coordenadas x de los datos $(x_1, y_1), \ldots, (x_n, y_n)$ están en la forma de desviación media, así que $\sum x_i = 0$. Demuestre que si X es la matriz de diseño para la recta de mínimos cuadrados de este caso, entonces X^TX es una matriz diagonal.

Los ejercicios 19 y 20 implican a una matriz de diseño X con dos o más columnas y una solución de mínimos cuadrados $\hat{\beta}$ de $\mathbf{y} = X\beta$. Considere los siguientes números.

i. $\|X\hat{\beta}\|^2$, la suma de los cuadrados del "término de regresión". Denote este número con SS(R).

ii. $\|\mathbf{y} - X\hat{\beta}\|^2$, la suma de los cuadrados para el término de error. Denote este número con SS(E).

iii. $\|\mathbf{y}\|^2$, la suma "total" de los cuadrados de los valores y. Denote este número con SS(T).

Todos los libros de estadística analizan el tema de la regresión, y el modelo lineal $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$ introduce esos números, aunque la notación y la terminología tal vez varíen de un texto a otro. Para simplificar el asunto, suponga que la media de los valores y es cero. En este caso, SS(T) es proporcional a la *varianza* del conjunto de valores y.

19. Justifique la ecuación SS(T) = SS(R) + SS(E). [*Sugerencia:* Utilice un teorema, y explique por qué se satisfacen las hipótesis del teorema]. Esta ecuación es extremadamente importante en estadística, en la teoría de regresión y en el análisis de varianza.

20. Demuestre que $\|X\hat{\beta}\|^2 = \hat{\beta}^T X^T\mathbf{y}$. [*Sugerencia:* Rescriba el lado izquierdo y considere el hecho de que $\hat{\beta}$ satisface las ecuaciones normales]. Esta fórmula para SS(R) se emplea en estadística. Con base en esto y en el ejercicio 19, obtenga la fórmula estándar para SS(E):

$$\text{SS(E)} = \mathbf{y}^T\mathbf{y} - \hat{\beta}^T X^T\mathbf{y}$$

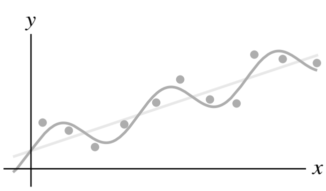
SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Construya X y β tal que la k-ésima fila de $X\beta$ sea el valor de y predicho correspondiente al punto de datos (x_k, y_k), a saber,

$$\beta_0 + \beta_1 x_k + \beta_2 \text{ sen } (2\pi\, x_k/12)$$

Debería ser claro que

$$X = \begin{bmatrix} 1 & x_1 & \text{sen}(2\pi x_1/12) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \text{sen}(2\pi x_n/12) \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$



Tendencia de ventas con fluctuaciones estacionales.

6.7 ESPACIOS CON PRODUCTO INTERIOR

Los conceptos de longitud, distancia y ortogonalidad son esenciales en aplicaciones que implican un espacio vectorial. Para $\mathbb{R}^n$, esos conceptos se basaron en las propiedades del producto interior indicadas en el teorema 1 de la sección 6.1. Para otros espacios, se necesitan analogías del producto interior con las mismas propiedades. Ahora las conclusiones del teorema 1 se convierten en *axiomas* en la siguiente definición.

DEFINICIÓN

Un **producto interior** sobre un espacio vectorial V es una función que asocia un número real $\langle \mathbf{u}, \mathbf{v} \rangle$ para cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ en V, y satisface los siguientes axiomas, para toda $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ en V, y todos los escalares c:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \quad \text{y} \quad \langle \mathbf{u}, \mathbf{u} \rangle = 0$ si y solo si $\mathbf{u} = \mathbf{0}$

Un espacio vectorial con un producto interior se llama **espacio con producto interior**.

El espacio vectorial $\mathbb{R}^n$ con el producto interior estándar es un espacio con producto interior, y casi todo lo analizado en este capítulo para $\mathbb{R}^n$ también es válido para los espacios con producto interior. Los ejemplos en esta sección y la siguiente establecen el fundamento para una variedad de aplicaciones que se estudian en cursos de ingeniería, física, matemáticas y estadística.

EJEMPLO 1 Elija dos números positivos cualesquiera, por ejemplo, 4 y 5; para los vectores $\mathbf{u} = (u_1, u_2)$ y $\mathbf{v} = (v_1, v_2)$ en $\mathbb{R}^2$, establezca

$$\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 \tag{1}$$

Demuestre que la ecuación (1) define un producto interior.

SOLUCIÓN Sin duda, el axioma 1 se satisface, porque $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 + 5u_2v_2 = 4v_1u_1 + 5v_2u_2 = \langle v, u \rangle$. Si $\mathbf{w} = (w_1, w_2)$, entonces,

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle &= 4(u_1 + v_1)w_1 + 5(u_2 + v_2)w_2 \\ &= 4u_1w_1 + 5u_2w_2 + 4v_1w_1 + 5v_2w_2 \\ &= \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

Esto comprueba el axioma 2. Para el axioma 3, calcule

$$\langle c\mathbf{u}, \mathbf{v} \rangle = 4(cu_1)v_1 + 5(cu_2)v_2 = c(4u_1v_1 + 5u_2v_2) = c\langle \mathbf{u}, \mathbf{v} \rangle$$

Para el axioma 4, observe que $\langle \mathbf{u}, \mathbf{u} \rangle = 4u_1^2 + 5u_2^2 \geq 0$ y $4u_1^2 + 5u_2^2 = 0$ solo si $u_1 = u_2 = 0$, es decir, si $\mathbf{u} = \mathbf{0}$. Además, $\langle \mathbf{0}, \mathbf{0} \rangle = 0$. Así, (1) define un producto interior sobre $\mathbb{R}^2$. ■

Sobre $\mathbb{R}^n$ se pueden definir productos interiores similares a (1), que surgen naturalmente en problemas de "mínimos cuadrados ponderados", en los cuales los pesos se asignan a las diversas entradas en la suma para el producto interior de manera que se dé mayor importancia a las mediciones más confiables.

De ahora en adelante, cuando un espacio con producto interior implique polinomios u otras funciones, se escribirán las funciones en la forma usual, en vez de emplear negritas para los vectores. Sin duda, es importante recordar que cada función *es* un vector cuando se trata como un elemento de un espacio vectorial.

EJEMPLO 2 Sean $t_0,\ldots, t_n$ números reales distintos. Para p y q en $\mathbb{P}_n$, defina

$$\langle p, q\rangle = p(t_0)q(t_0) + p(t_1)q(t_1) + \cdots + p(t_n)q(t_n) \tag{2}$$

Es sencillo comprobar los axiomas 1 a 3 del producto interior. Para el axioma 4, observe que

$$\langle p, p\rangle = [p(t_0)]^2 + [p(t_1)]^2 + \cdots + [p(t_n)]^2 \geq 0$$

Además, $\langle \mathbf{0}, \mathbf{0}\rangle = 0$. (Aquí el cero en "negritas" denota el polinomio cero, el vector cero en $\mathbb{P}_n$). Si $\langle p, p\rangle = 0$, entonces p se debe anular en $n + 1$ puntos: $t_0,\ldots, t_n$. Esto solo es posible si p es el polinomio cero, porque el grado de p es menor que $n + 1$. Así, (2) define un producto interior en $\mathbb{P}_n$. ■

EJEMPLO 3 Considere que V está en $\mathbb{P}_2$, con el producto interior del ejemplo 2, donde $t_0 = 0$, $t_1 = \frac{1}{2}$, y $t_2 = 1$. Sean $p(t) = 12t^2$ y $q(t) = 2t - 1$. Calcule $\langle p, q\rangle$ y $\langle q, q\rangle$.

SOLUCIÓN

$$\begin{aligned}\langle p, q\rangle &= p(0)q(0) + p\left(\tfrac{1}{2}\right)q\left(\tfrac{1}{2}\right) + p(1)q(1)\\ &= (0)(-1) + (3)(0) + (12)(1) = 12\\ \langle q, q\rangle &= [q(0)]^2 + [q\left(\tfrac{1}{2}\right)]^2 + [q(1)]^2\\ &= (-1)^2 + (0)^2 + (1)^2 = 2\end{aligned}$$

■

Longitudes, distancias y ortogonalidad

Sea V un espacio con producto interior, con el producto interior denotado con $\langle \mathbf{u}, \mathbf{v}\rangle$. Igual que en $\mathbb{R}^n$, defina la **longitud**, o **norma**, de un vector $\mathbf{v}$ como el escalar

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$$

De manera equivalente, $\|\mathbf{v}\|^2 = \langle \mathbf{v}, \mathbf{v}\rangle$. (Esta definición tiene sentido porque $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$, pero la definición *no* dice que $\langle \mathbf{v}, \mathbf{v}\rangle$ sea una "suma de cuadrados" porque $\mathbf{v}$ no necesita ser un elemento de $\mathbb{R}^n$).

Un **vector unitario** es aquel cuya longitud es 1. La **distancia entre u y v** es $\|\mathbf{u} - \mathbf{v}\|$. Los vectores $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

EJEMPLO 4 Suponga que $\mathbb{P}_2$ tiene el producto interior (2) del ejemplo 3. Calcule las longitudes de los vectores $p(t) = 12t^2$ y $q(t) = 2t - 1$.

SOLUCIÓN

$$\begin{aligned}\|p\|^2 &= \langle p, p\rangle = [p(0)]^2 + \left[p\left(\tfrac{1}{2}\right)\right]^2 + [p(1)]^2\\ &= 0 + [3]^2 + [12]^2 = 153\\ \|p\| &= \sqrt{153}\end{aligned}$$

Del ejemplo 3, $\langle q, q\rangle = 2$. Por lo tanto, $\|q\| = \sqrt{2}$. ■

Proceso de Gram-Schmidt

La existencia de bases ortogonales para subespacios de dimensión finita de un espacio vectorial con producto interior se puede establecer con el proceso de Gram-Schmidt, al igual que en $\mathbb{R}^n$. Mediante este proceso, es posible construir ciertas bases ortogonales que surgen a menudo en las aplicaciones.

La proyección ortogonal de un vector sobre un subespacio W con una base ortogonal se puede construir en la forma habitual. La proyección no depende de la base ortogonal seleccionada, y tiene las propiedades descritas en los teoremas de descomposición ortogonal y de la mejor aproximación.

EJEMPLO 5 Considere que V está en P$_4$ con el producto interior del ejemplo 2, lo que implica evaluación de polinomios en $-2, -1, 0, 1$ y 2, y que $\mathbb{P}_2$ es un subespacio de V. Obtenga una base ortogonal para $\mathbb{P}_2$ aplicando el proceso de Gram-Schmidt a los polinomios 1, t y t^2.

SOLUCIÓN El producto interior sólo depende de los valores de un polinomio en $-2,\ldots, 2$, así que se listan los valores de cada polinomio como un vector en $\mathbb{R}^5$, debajo del nombre de cada polinomio:[1]

$$\begin{array}{cccc} \text{Polinomio:} & 1 & t & t^2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 4 \\ 1 \\ 0 \\ 1 \\ 4 \end{bmatrix} \end{array}$$

El producto interior de dos polinomios en V es igual al producto interior (estándar) de sus vectores correspondientes en $\mathbb{R}^5$. Observe que t es ortogonal a la función constante 1. Así, se toman $p_0(t) = 1$ y $p_1(t) = t$. Para p_2, utilice los vectores en $\mathbb{R}^5$ para calcular la proyección de t^2 sobre Gen $\{p_0, p_1\}$:

$$\begin{aligned} \langle t^2, p_0\rangle &= \langle t^2, 1\rangle = 4 + 1 + 0 + 1 + 4 = 10 \\ \langle p_0, p_0\rangle &= 5 \\ \langle t^2, p_1\rangle &= \langle t^2, t\rangle = -8 + (-1) + 0 + 1 + 8 = 0 \end{aligned}$$

La proyección ortogonal de t^2 sobre Gen $\{1, t\}$ es $\frac{10}{5}p_0 + 0p_1$. Por consiguiente,

$$p_2(t) = t^2 - 2p_0(t) = t^2 - 2$$

Una base ortogonal para el subespacio $\mathbb{P}_2$ de V es:

$$\begin{array}{cccc} \text{Polinomio:} & p_0 & p_1 & p_2 \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix} \end{array} \qquad (3)$$

■

Mejor aproximación en espacios con producto interior

Un problema común en matemáticas aplicadas implica un espacio vectorial V cuyos elementos son funciones. El problema es aproximar una función f en V mediante una función g de un subespacio especificado W de V. La "cercanía" de la aproximación de f depende de la manera en que se defina $\|f - g\|$. Solo se considerará el caso en que la distancia entre f y g esté determinada por un producto interior. En tal caso, la *mejor aproximación a f mediante funciones en W* es la proyección ortogonal de f sobre el subespacio W.

EJEMPLO 6 Considere que V está en $\mathbb{P}_4$ con el producto interior del ejemplo 5; considere también que p_0, p_1 y p_2 constituyen la base ortogonal que se encontró en el ejemplo 5 para el subespacio $\mathbb{P}_2$. Determine la mejor aproximación a $p(t) = 5 - \frac{1}{2}t^4$ mediante polinomios en $\mathbb{P}_2$.

---

[1] Cada polinomio en $\mathbb{P}_4$ está determinado de manera unívoca por su valor en los cinco números $-2,\ldots, 2$. De hecho, la correspondencia entre p y su vector de valores es un isomorfismo, es decir, un mapeo uno a uno sobre $\mathbb{R}^5$ que preserva combinaciones lineales.

SOLUCIÓN Los valores de p_0, p_1 y p_2 en los números $-2, -1, 0, 1$ y 2 están listados en los vectores de $\mathbb{R}^5$ en (3). Los valores correspondientes para p son $-3, 9/2, 5, 9/2$ y -3. Calcule

$$\begin{aligned} &\langle p, p_0\rangle = 8, \qquad \langle p, p_1\rangle = 0, \qquad \langle p, p_2\rangle = -31 \\ &\langle p_0, p_0\rangle = 5, \qquad\qquad\qquad\qquad\;\; \langle p_2, p_2\rangle = 14 \end{aligned}$$

Así, la mejor aproximación en V a p mediante polinomios en $\mathbb{P}_2$ es

$$\begin{aligned} \hat{p} = \text{proy}_{\mathbb{P}_2}\, p &= \frac{\langle p, p_0\rangle}{\langle p_0, p_0\rangle} p_0 + \frac{\langle p, p_1\rangle}{\langle p_1, p_1\rangle} p_1 + \frac{\langle p, p_2\rangle}{\langle p_2, p_2\rangle} p_2 \\ &= \tfrac{8}{5} p_0 + \tfrac{-31}{14} p_2 = \tfrac{8}{5} - \tfrac{31}{14}(t^2 - 2). \end{aligned}$$

De todos los polinomios en $\mathbb{P}_2$, este es el más cercano a p cuando la distancia entre polinomios se mide solo en $-2, -1, 0, 1$ y 2. Véase la figura 1. ■

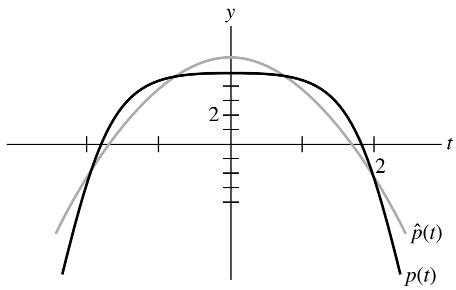


FIGURA 1

Los polinomios p_0, p_1 y p_2 de los ejemplos 5 y 6 pertenecen a una clase de polinomios que en estadística se denominan *polinomios ortogonales*.[2] La ortogonalidad se refiere al tipo de producto interior descrito en el ejemplo 2.

Dos desigualdades

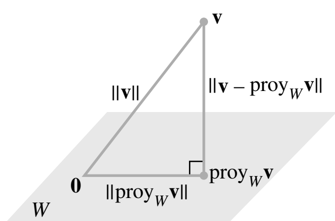


FIGURA 2
La hipotenusa es el lado más largo.

Considerando un vector $\mathbf{v}$ en un espacio V con producto interior y dado un subespacio W de dimensión finita, se aplica el teorema de Pitágoras a la descomposición ortogonal de $\mathbf{v}$ con respecto a W y se obtiene

$$\|\mathbf{v}\|^2 = \|\text{proy}_W\, \mathbf{v}\|^2 + \|\mathbf{v} - \text{proy}_W\, \mathbf{v}\|^2$$

Véase la figura 2. En particular, esto muestra que la norma de la proyección de $\mathbf{v}$ sobre W no excede la propia norma de $\mathbf{v}$. Esta sencilla observación conduce a la siguiente importante desigualdad.

TEOREMA 16

Desigualdad de Cauchy-Schwarz

Para toda $\mathbf{u}$ y $\mathbf{v}$ en V,

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\| \tag{4}$$

[2] Véase *Statistics and Experimental Design in Engineering and the Physical Sciences*, 2a. ed., de Norman L. Johnson y Fred C. Leone (Nueva York: John Wiley & Sons, 1977). En este libro las tablas listan los “polinomios ortogonales”, que son simplemente los valores de los polinomios en números como $-2, -1, 0, 1$ y 2.

DEMOSTRACIÓN Si $\mathbf{u} = \mathbf{0}$, entonces ambos lados de (4) son iguales a cero, por lo tanto, la desigualdad es verdadera en este caso. (Véase el problema de práctica 1). Si $\mathbf{u} \neq \mathbf{0}$, sea W el subespacio generado por $\mathbf{u}$. Recuerde que $\|c\mathbf{u}\| = |c|\,\|\mathbf{u}\|$ para cualquier escalar c. Así,

$$\|\text{proy}_W\,\mathbf{v}\| = \left\| \frac{\langle \mathbf{v}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} \right\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{|\langle \mathbf{u}, \mathbf{u}\rangle|}\|\mathbf{u}\| = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{\|\mathbf{u}\|^2}\|\mathbf{u}\| = \frac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|}$$

Puesto que $\|\text{proy}_W\,\mathbf{v}\| \leq \|\mathbf{v}\|$, se tiene $\dfrac{|\langle \mathbf{u}, \mathbf{v}\rangle|}{\|\mathbf{u}\|} \leq \|\mathbf{v}\|$, lo que da la ecuación (4). ■

La desigualdad de Cauchy-Schwarz es útil en muchas ramas de las matemáticas. En los ejercicios se presentan unas cuantas aplicaciones. Esta desigualdad es necesaria para probar otra desigualdad fundamental que implica normas de vectores. Véase la figura 3.

TEOREMA 17

Desigualdad del triángulo

Para toda $\mathbf{u}$ y $\mathbf{v}$ en V,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

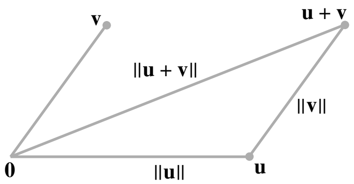


FIGURA 3
Las longitudes de los lados de un triángulo.

DEMOSTRACIÓN

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle \\
&\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \qquad \text{Cauchy–Schwarz} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

La desigualdad del triángulo se deduce inmediatamente al sacar la raíz cuadrada en ambos lados. ■

Un producto interior para $C[a, b]$ (se requiere cálculo)

Quizás el espacio con producto interior más ampliamente utilizado en aplicaciones sea el espacio vectorial $C[a, b]$ de todas las funciones continuas en un intervalo $a \leq t \leq b$, con un producto interior que enseguida se describirá.

Se inicia considerando un polinomio p y cualquier entero n mayor o igual que el grado de p. Entonces p está en $\mathbb{P}_n$, y se puede calcular una "longitud" para p utilizando el producto interior del ejemplo 2 que implica evaluación en $n + 1$ puntos en $[a, b]$. Sin embargo, esta longitud de p solo capta el comportamiento en esos $n + 1$ puntos. Como p está en $\mathbb{P}_n$ para todas las n grandes, se podría utilizar un valor de n bastante grande, con mucho más puntos para la "evaluación" asociada con el producto interior. Véase la figura 4.

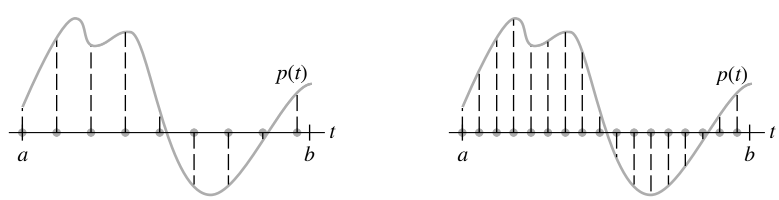
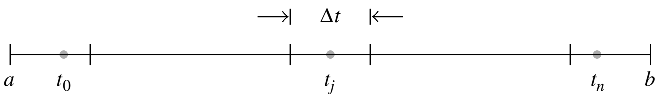


FIGURA 4 Uso de diferentes números de puntos de evaluación en $[a, b]$ para calcular $\|p\|^2$.

Se particiona $[a, b]$ en $n + 1$ subintervalos de longitud $\Delta t = (b - a)/(n + 1)$, y sean $t_0,\ldots, t_n$ puntos arbitrarios en esos subintervalos.



Si n es grande, el producto interior sobre $\mathbb{P}_n$ determinado por $t_0,\ldots, t_n$ tenderá a dar un gran valor para $\langle p, p\rangle$, así que se reduce a escala y se divide entre $n + 1$. Observe que $1/(n + 1) = \Delta t/(b - a)$, y se define

$$\langle p, q\rangle = \frac{1}{n+1}\sum_{j=0}^{n} p(t_j)q(t_j) = \frac{1}{b-a}\left[\sum_{j=0}^{n} p(t_j)q(t_j)\Delta t\right]$$

Ahora, n crece sin límite. Puesto que los polinomios p y q son funciones continuas, entonces la expresión entre corchetes es una suma de Riemann que se aproxima a una integral definida, lo que conduce a considerar el *valor promedio de* $p(t)q(t)$ sobre el intervalo $[a, b]$:

$$\frac{1}{b-a}\int_a^b p(t)q(t)\,dt$$

Esta cantidad está definida para polinomios de cualquier grado (de hecho, para todas las funciones continuas), y tiene todas las propiedades de un producto interior, como lo muestra el siguiente ejemplo. El factor de escala $1/(b - a)$ no es esencial y con frecuencia se omite con la finalidad de simplificar.

EJEMPLO 7 Para f, g en $C[a, b]$, sea

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt \tag{5}$$

Demuestre que la ecuación (5) define un producto interior sobre $C[a, b]$.

SOLUCIÓN Los axiomas 1 a 3 del producto interior se deducen de las propiedades elementales de las integrales definidas. Para el axioma 4, observe que

$$\langle f, f\rangle = \int_a^b [f(t)]^2\,dt \ge 0$$

La función $[f(t)]^2$ es continua y no negativa en $[a, b]$. Si la integral definida de $[f(t)]^2$ es cero, entonces $[f(t)]^2$ debe ser idénticamente cero sobre $[a, b]$, de acuerdo con un teorema en cálculo avanzado; en tal caso, f es la función cero. Así $\langle f, f\rangle = 0$ implica que f es la función cero en $[a, b]$. Por lo tanto, (5) define un producto interior sobre $C[a, b]$. ■

EJEMPLO 8 Sean V el espacio $C[0, 1]$ con el producto interior del ejemplo 7, y W el subespacio generado por los polinomios $p_1(t) = 1$, $p_2(t) = 2t - 1$ y $p_3(t) = 12t^2$. Utilice el proceso de Gram-Schmidt y encuentre una base ortogonal para W.

SOLUCIÓN Sea $q_1 = p_1$, y calcule

$$\langle p_2, q_1\rangle = \int_0^1 (2t-1)(1)\,dt = (t^2 - t)\Big|_0^1 = 0$$

De manera que p_2 ya es ortogonal a q_1, y se puede tomar $q_2 = p_2$. Para la proyección de p_3 sobre $W_2 = \text{Gen}\,\{q_1, q_2\}$, se calcula

$$\begin{aligned}
\langle p_3, q_1\rangle &= \int_0^1 12t^2 \cdot 1\, dt = 4t^3\Big|_0^1 = 4 \\
\langle q_1, q_1\rangle &= \int_0^1 1 \cdot 1\, dt = t\Big|_0^1 = 1 \\
\langle p_3, q_2\rangle &= \int_0^1 12t^2(2t-1)\, dt = \int_0^1 (24t^3 - 12t^2)\, dt = 2 \\
\langle q_2, q_2\rangle &= \int_0^1 (2t-1)^2\, dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}
\end{aligned}$$

Entonces

$$\text{proy}_{W_2}\, p_3 = \frac{\langle p_3, q_1\rangle}{\langle q_1, q_1\rangle} q_1 + \frac{\langle p_3, q_2\rangle}{\langle q_2, q_2\rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{1/3} q_2 = 4q_1 + 6q_2$$

y

$$q_3 = p_3 - \text{proy}_{W_2}\, p_3 = p_3 - 4q_1 - 6q_2$$

Como una función, $q_3(t) = 12t^2 - 4 - 6(2t-1) = 12t^2 - 12t + 2$. La base ortogonal para el subespacio W es $\{q_1, q_2, q_3\}$. ■

PROBLEMAS DE PRÁCTICA

Aplique los axiomas del producto interior para comprobar los siguientes enunciados.

1. $\langle \mathbf{v}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle = 0$.
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$.

6.7 EJERCICIOS

1. Suponga que $\mathbb{R}^2$ tiene el producto interior del ejemplo 1, y sean $\mathbf{x} = (1, 1)$ y $\mathbf{y} = (5, -1)$.
   a) Encuentre $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ y $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$.
   b) Describa todos los vectores (z_1, z_2) que son ortogonales a $\mathbf{y}$.

2. Considere que $\mathbb{R}^2$ tiene el producto interior del ejemplo 1. Demuestre que la desigualdad de Cauchy-Schwarz es válida para $\mathbf{x} = (3, -2)$ y $\mathbf{y} = (-2, 1)$. (*Sugerencia:* Estudie $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$).

Los ejercicios 3 a 8 se refieren a $\mathbb{P}_2$ con el producto interior dado por evaluación en -1, 0 y 1. (Véase el ejemplo 2).

3. Calcule $\langle p, q\rangle$, donde $p(t) = 4 + t$, $q(t) = 5 - 4t^2$.
4. Calcule $\langle p, q\rangle$, donde $p(t) = 3t - t^2$, $q(t) = 3 + 2t^2$.
5. Obtenga $\|p\|$ y $\|q\|$, para p y q del ejercicio 3.
6. Obtenga $\|p\|$ y $\|q\|$, para p y q del ejercicio 4.
7. Determine la proyección ortogonal de q sobre el subespacio generado por p, para p y q en el ejercicio 3.
8. Obtenga la proyección ortogonal de q sobre el subespacio generado por p, para p y q en el ejercicio 4.
9. Considere que $\mathbb{P}_3$ tiene el producto interior dado por evaluación en $-3, -1, 1$ y 3. Sean $p_0(t) = 1$, $p_1(t) = t$ y $p_2(t) = t^2$.
   a) Calcule la proyección ortogonal de p_2 sobre el subespacio generado por p_0 y p_1.
   b) Encuentre un polinomio q que sea ortogonal a p_0 y p_1, tal que $\{p_0, p_1, q\}$ sea una base ortogonal para Gen $\{p_0, p_1, p_2\}$. Escale el polinomio q de modo que su vector de valores en $(-3, -1, 1, 3)$ sea $(1, -1, -1, 1)$.
10. Considere que $\mathbb{P}_3$ tiene el producto interior del ejercicio 9, con los polinomios p_0, p_1 y q ahí descritos. Determine la mejor aproximación a $p(t) = t^3$ mediante polinomios en Gen $\{p_0, p_1, q\}$.
11. Suponga que p_0, p_1 y p_2 son los polinomios ortogonales descritos en el ejemplo 5, donde el producto interior sobre $\mathbb{P}_4$ está dado por evaluación en $-2, -1, 0, 1$ y 2. Encuentre la proyección ortogonal de t^3 sobre Gen $\{p_0, p_1, p_2\}$.
12. Obtenga un polinomio p_3 tal que $\{p_0, p_1, p_2, p_3\}$ (véase el ejercicio 11) sea una base ortogonal para el subespacio $\mathbb{P}_3$ de $\mathbb{P}_4$. Escale el polinomio p_3 de manera que su vector de valores sea $(-1, 2, 0, -2, 1)$.

13. Sea A una matriz invertible de $n \times n$. Demuestre que para $\mathbf{u}$, $\mathbf{v}$ en $\mathbb{R}^n$, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{u})^T(A\mathbf{v})$ define un producto interior sobre $\mathbb{R}^n$.

14. Sea T una transformación lineal uno a uno de un espacio vectorial V en $\mathbb{R}^n$. Demuestre que para $\mathbf{u}$, $\mathbf{v}$ en V, la fórmula $\langle \mathbf{u}, \mathbf{v} \rangle = T(\mathbf{u}) \cdot T(\mathbf{v})$ define un producto interior sobre V.

Utilice los axiomas del producto interior y otros resultados de esta sección para comprobar los enunciados de los ejercicios 15 a 18.

15. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ para todos los escalares c.

16. Si $\{\mathbf{u}, \mathbf{v}\}$ es un conjunto ortonormal en V, entonces $\|\mathbf{u} - \mathbf{v}\| = \sqrt{2}$.

17. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$

18. $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$

19. Dados $a \geq 0$ y $b \geq 0$, sean $\mathbf{u} = \begin{bmatrix} \sqrt{a} \\ \sqrt{b} \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix}$.

    Aplique la desigualdad de Cauchy-Schwarz para comparar la media geométrica $\sqrt{ab}$ con la media aritmética $(a + b)/2$.

20. Sean $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Utilice la desigualdad de Cauchy-Schwarz para demostrar que

    $$\left(\frac{a+b}{2}\right)^2 \leq \frac{a^2 + b^2}{2}$$

Los ejercicios 21 a 24 se refieren a $V = C[0, 1]$, con el producto interior dado por una integral, como en el ejemplo 7.

21. Calcule $\langle f, g \rangle$, donde $f(t) = 1 - 3t^2$ y $g(t) = t - t^3$.

22. Determine $\langle f, g \rangle$, donde $f(t) = 5t - 3$ y $g(t) = t^3 - t^2$.

23. Obtenga $\|f\|$ para la f del ejercicio 21.

24. Calcule $\|g\|$ para la g del ejercicio 22.

25. Sea V el espacio $C[-1, 1]$ con el producto interior del ejemplo 7. Encuentre una base ortogonal para el subespacio generado por los polinomios 1, t y t^2. Los polinomios en esta base se llaman *polinomios de Legendre*.

26. Sea V el espacio $C[-2, 2]$ con el producto interior del ejemplo 7. Obtenga una base ortogonal para el subespacio generado por los polinomios 1, t y t^2.

27. **[M]** Considere que $\mathbb{P}_4$ tiene el producto interior como en el ejemplo 5, y sean p_0, p_1, p_2 los polinomios ortogonales de ese ejemplo. Con su programa de matrices, aplique el proceso de Gram-Schmidt al conjunto $\{p_0, p_1, p_2, t^3, t^4\}$ y cree una base ortogonal para $\mathbb{P}_4$.

28. **[M]** Sea V el espacio $C[0, 2\pi]$ con el producto interior del ejemplo 7. Aplique el proceso de Gram-Schmidt con la finalidad de crear una base ortogonal para el subespacio generado por $\{1, \cos t, \cos^2 t, \cos^3 t\}$. Utilice un programa de matrices o computacional para calcular las integrales definidas adecuadas.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Por el axioma 1, $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle$. Entonces $\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0\langle \mathbf{v}, \mathbf{v} \rangle$, por el axioma 3, de manera que $\langle \mathbf{0}, \mathbf{v} \rangle = 0$.
2. Por los axiomas 1 y 2, y después 1 otra vez, $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.

6.8 APLICACIONES DE ESPACIOS CON PRODUCTO INTERIOR

Los ejemplos de esta sección ilustran cómo surgen en problemas prácticos los espacios con producto interior definidos en la sección 6.7. El primer ejemplo se relaciona con el inmenso problema de mínimos cuadrados de actualización del North American Datum, descrito en el ejemplo introductorio de este capítulo.

Mínimos cuadrados ponderados

Sea $\mathbf{y}$ un vector de n observaciones, $y_1, \ldots, y_n$, y suponga que se desea aproximar $\mathbf{y}$ mediante un vector $\hat{\mathbf{y}}$ que pertenece a determinado subespacio de $\mathbb{R}^n$. (En la sección 6.5, $\hat{\mathbf{y}}$ se escribió como $A\mathbf{x}$, de manera que $\hat{\mathbf{y}}$ estuviera en el espacio columna de A). Denote las entradas en $\hat{\mathbf{y}}$ como $\hat{y}_1, \ldots, \hat{y}_n$. Entonces la *suma de los cuadrados de los errores*, o SS(E), al aproximar $\mathbf{y}$ por $\hat{\mathbf{y}}$ es

$$\text{SS(E)} = (y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2 \tag{1}$$

Esto es simplemente $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$, utilizando la longitud estándar en $\mathbb{R}^n$.

Ahora suponga que las mediciones que generaron las entradas en $\mathbf{y}$ no son igualmente confiables. (Este fue el caso para el North American Datum, porque las mediciones se realizaron durante un periodo de 140 años). Como otro ejemplo, las entradas en $\mathbf{y}$ se podrían calcular a partir de varias muestras de medidas, con tamaños de muestra desiguales). Entonces, resulta adecuado ponderar los errores cuadráticos de la ecuación (1) de tal manera que se dé mayor importancia a las mediciones más confiables.[1] Si los pesos se denotan como $w_1^2, \dots, w_n^2$, entonces la suma ponderada de los errores cuadráticos es

$$\text{SS(E) ponderada} = w_1^2(y_1 - \hat{y}_1)^2 + \cdots + w_n^2(y_n - \hat{y}_n)^2 \tag{2}$$

Este es el cuadrado de la longitud de $\mathbf{y} - \hat{\mathbf{y}}$, donde la longitud se deduce a partir de un producto interior análogo al del ejemplo 1 de la sección 6.7, a saber,

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n$$

Algunas veces es conveniente transformar un problema de mínimos cuadrados ponderados en un problema equivalente de mínimos cuadrados ordinario. Sea W la matriz diagonal con $w_1, \dots, w_n$ (positivos) en su diagonal, de manera que

$$W\mathbf{y} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & w_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 y_1 \\ w_2 y_2 \\ \vdots \\ w_n y_n \end{bmatrix}$$

con una expresión similar para $W\hat{\mathbf{y}}$. Observe que el j-ésimo término en la ecuación (2) se puede escribir como

$$w_j^2(y_j - \hat{y}_j)^2 = (w_j y_j - w_j \hat{y}_j)^2$$

Se deduce que la SS(E) ponderada en la ecuación (2) es el cuadrado de la longitud ordinaria en $\mathbb{R}^n$ de $W\mathbf{y} - W\hat{\mathbf{y}}$, que se puede representar en la forma $\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2$.

Ahora suponga que el vector de aproximación $\hat{\mathbf{y}}$ se construirá mediante las columnas de A. Entonces se busca una $\hat{\mathbf{x}}$ que haga que $A\hat{\mathbf{x}} = \hat{\mathbf{y}}$ esté tan cerca de $\mathbf{y}$ como sea posible. Sin embargo, la medida de cercanía es el error ponderado,

$$\|W\mathbf{y} - W\hat{\mathbf{y}}\|^2 = \|W\mathbf{y} - WA\hat{\mathbf{x}}\|^2$$

Así, $\hat{\mathbf{x}}$ es la solución (ordinaria) de mínimos cuadrados de la ecuación

$$WA\mathbf{x} = W\mathbf{y}$$

La ecuación normal para la solución de mínimos cuadrados es

$$(WA)^T WA\mathbf{x} = (WA)^T W\mathbf{y}$$

EJEMPLO 1 Encuentre la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que mejor se ajuste a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$ y $(2, 3)$. Suponga que los errores al medir los valores y de los dos últimos puntos de datos son más grandes que para los demás puntos. Pondere estos datos a la mitad en relación con los datos restantes.

[1] Nota para lectores con conocimientos de estadística: Suponga que los errores al medir y_i son variables aleatorias independientes con media igual a cero y varianzas $\sigma_1^2, \dots, \sigma_n^2$. Entonces, los pesos adecuados en la ecuación (2) son $w_i^2 = 1/\sigma_i^2$. Cuanto mayor sea la varianza del error, menor será el peso.

SOLUCIÓN Como en la sección 6.6, escriba X para la matriz A y β para el vector $\mathbf{x}$, y obtenga

$$X = \begin{bmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix}$$

Para una matriz de ponderación, seleccione W con entradas diagonales 2, 2, 2, 1 y 1. Al multiplicar por la izquierda por W se escalan las filas de X y $\mathbf{y}$:

$$WX = \begin{bmatrix} 2 & -4 \\ 2 & -2 \\ 2 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \quad W\mathbf{y} = \begin{bmatrix} 6 \\ 10 \\ 10 \\ 4 \\ 3 \end{bmatrix}$$

Para la ecuación normal, calcule

$$(WX)^T WX = \begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \quad \text{y} \quad (WX)^T W\mathbf{y} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

y resuelva

$$\begin{bmatrix} 14 & -9 \\ -9 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 59 \\ -34 \end{bmatrix}$$

La solución de la ecuación normal es (a dos dígitos significativos) $\beta_0 = 4.3$ y $\beta_1 = .20$. La recta deseada es

$$y = 4.3 + .20x$$

En contraste, la recta de mínimos cuadrados ordinaria para estos datos es

$$y = 4.0 - .10x$$

En la figura 1 se representan ambas rectas. ■

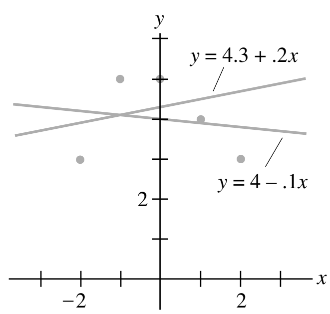


FIGURA 1
Rectas de mínimos cuadrados ordinaria y ponderada.

Análisis de tendencia de datos

Sea f una función desconocida cuyos valores se conocen (quizá solo aproximadamente) en $t_0, \ldots, t_n$. Si existe una "tendencia lineal" en los datos $f(t_0), \ldots, f(t_n)$, entonces se espera poder aproximar los valores de f mediante una función de la forma $\beta_0 + \beta_1 t$. Si hay una "tendencia cuadrática" de los datos, podría intentarse una función con la estructura $\beta_0 + \beta_1 t + \beta_2 t^2$. Esto se analizó en la sección 6.6, pero desde otro punto de vista.

En algunos problemas estadísticos, es importante poder separar las tendencias lineal y cuadrática (y posiblemente cúbica o de mayor orden). Por ejemplo, suponga que unos ingenieros están analizando el desempeño de un nuevo automóvil, y $f(t)$ representa la distancia entre el vehículo (en el momento t) y algún punto de referencia. Si el auto viaja a velocidad constante, entonces la gráfica de $f(t)$ debería ser una recta cuya pendiente es la velocidad del auto. Si se presiona el acelerador repentinamente, entonces la gráfica de $f(t)$ cambiará para incluir un término cuadrático y posiblemente un término cúbico (debido a la aceleración). Para analizar la capacidad del auto para rebasar a otro, por ejemplo, los ingenieros tal vez quieran separar las componentes cuadrática y cúbica del término lineal.

Si la función es aproximada por una curva de la forma $y = \beta_0 + \beta_1 t + \beta_2 t^2$, es posible que el coeficiente β^2 no brinde la información deseada sobre la tendencia cuadrática en los datos, porque quizá no sea "independiente" (en un sentido estadístico) de los otros β_i.

Para efectuar el **análisis de tendencia** de los datos, se introduce un producto interior sobre el espacio $\mathbb{P}_n$ análogo al del ejemplo 2 de la sección 6.7. Para p, q en $\mathbb{P}_n$, se define

$$\langle p, q \rangle = p(t_0)q(t_0) + \cdots + p(t_n)q(t_n)$$

En la práctica, los especialistas en estadística rara vez necesitan considerar tendencias en los datos de grado mayor a tres o cuatro. Así que sean p_0, p_1, p_2, p_3 una base ortogonal del subespacio $\mathbb{P}_3$ de $\mathbb{P}_n$, obtenida al aplicar el proceso de Gram-Schmidt a los polinomios 1, t, t^2 y t^3. De acuerdo con el ejercicio complementario 11 del capítulo 2, existe un polinomio g en $\mathbb{P}_n$ cuyos valores en $t_0, \ldots, t_n$ coinciden con aquellos de la función desconocida f. Sea $\hat{g}$ la proyección ortogonal (con respecto al producto interior dado) de g sobre P_3, digamos,

$$\hat{g} = c_0 p_0 + c_1 p_1 + c_2 p_2 + c_3 p_3$$

Entonces $\hat{g}$ se llama una **función de tendencia** cúbica, y $c_0, \ldots, c_3$ son los **coeficientes de tendencia** de los datos. El coeficiente c_1 mide la tendencia lineal, c_2 la tendencia cuadrática, y c_3 la tendencia cúbica. Resulta que si los datos tienen ciertas propiedades, entonces esos coeficientes son estadísticamente independientes.

Como $p_0, \ldots, p_3$ son ortogonales, los coeficientes de tendencia se pueden calcular uno por uno, independientemente de los otros. (Recuerde que $c_i = \langle g, p_i \rangle / \langle p_i, p_i \rangle$). Es posible ignorar p_3 y c_3 si solo se desea la tendencia cuadrática. Y si, por ejemplo, se necesita determinar la tendencia cuártica, solo se tendría que encontrar (por medio de Gram-Schmidt) un polinomio p_4 en $\mathbb{P}_4$ ortogonal a $\mathbb{P}_3$ y calcular $\langle g, p_4 \rangle / \langle p_4, p_4 \rangle$.

EJEMPLO 2 El uso más sencillo y común del análisis de tendencia ocurre cuando los puntos $t_0, \ldots, t_n$ se pueden ajustar para quedar equidistantes entre sí y su suma sea cero. Ajuste una función de tendencia cuadrática a los datos $(-2, 3)$, $(-1, 5)$, $(0, 5)$, $(1, 4)$ y $(2, 3)$.

SOLUCIÓN Las coordenadas t se escalan convenientemente para emplear los polinomios ortogonales encontrados en el ejemplo 5 de la sección 6.7:

$$\begin{array}{lcccc} \text{Polinomio:} & p_0 & p_1 & p_2 & \text{Datos: } g \\ \text{Vector de valores:} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, & \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 2 \\ -1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, & \begin{bmatrix} 3 \\ 5 \\ 5 \\ 4 \\ 3 \end{bmatrix} \end{array}$$

Los cálculos solo necesitan esos vectores, y no las fórmulas específicas para los polinomios ortogonales. La mejor aproximación a los datos mediante polinomios en $\mathbb{P}_2$ es la proyección ortogonal dada por

$$\begin{aligned} \hat{p} &= \frac{\langle g, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0 + \frac{\langle g, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1 + \frac{\langle g, p_2 \rangle}{\langle p_2, p_2 \rangle} p_2 \\ &= \tfrac{20}{5} p_0 - \tfrac{1}{10} p_1 - \tfrac{7}{14} p_2 \end{aligned}$$

y

$$\hat{p}(t) = 4 - .1t - .5(t^2 - 2) \tag{3}$$

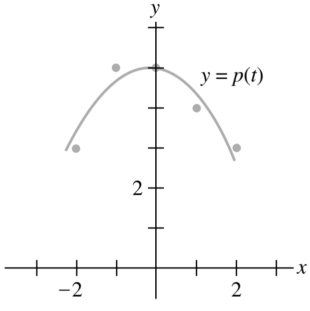


FIGURA 2
Aproximación por una función de tendencia cuadrática.

Como el coeficiente de p_2 no es extremadamente pequeño, sería razonable concluir que la tendencia es al menos cuadrática. Esto se confirma con la gráfica de la figura 2. ■

Serie de Fourier (se requiere cálculo)

Con frecuencia, las funciones continuas se aproximan mediante combinaciones lineales de funciones seno y coseno. Por ejemplo, una función continua podría representar una onda sonora, una señal eléctrica del algún tipo, o el movimiento de un sistema mecánico vibratorio.

Para simplificar, considere funciones en $0 \leq t \leq 2\pi$. Resulta que cualquier función en $C[0, 2\pi]$ se puede aproximar tan cerca como se requiera con una función de la forma

$$\frac{a_0}{2} + a_1 \cos t + \cdots + a_n \cos nt + b_1 \operatorname{sen} t + \cdots + b_n \operatorname{sen} nt \tag{4}$$

para un valor de n suficientemente grande. La función (4) es un **polinomio trigonométrico**. Si a_n y b_n no son ambas cero, entonces el polinomio es de **orden n**. La conexión entre polinomios trigonométricos y otras funciones en $C[0, 2\pi]$ depende del hecho de que para cualquier $n \geq 1$, el conjunto

$$\{1, \cos t, \cos 2t, \ldots, \cos nt, \operatorname{sen} t, \operatorname{sen} 2t, \ldots, \operatorname{sen} nt\} \tag{5}$$

sea ortogonal con respecto al producto interior

$$\langle f, g \rangle = \int_0^{2\pi} f(t) g(t)\, dt \tag{6}$$

Esta ortogonalidad se comprueba como en el siguiente ejemplo y en los ejercicios 5 y 6.

EJEMPLO 3 Considere que $C[0, 2\pi]$ tiene el producto interior (6), y sean m y n enteros positivos diferentes. Demuestre que $\cos mt$ y $\cos nt$ son ortogonales.

SOLUCIÓN Utilice una identidad trigonométrica. Cuando $m \neq n$,

$$\begin{aligned}
\langle \cos mt, \cos nt \rangle &= \int_0^{2\pi} \cos mt \cos nt\, dt \\
&= \frac{1}{2} \int_0^{2\pi} [\cos(mt + nt) + \cos(mt - nt)]\, dt \\
&= \frac{1}{2} \left[\frac{\operatorname{sen}(mt + nt)}{m + n} + \frac{\operatorname{sen}(mt - nt)}{m - n} \right] \Bigg|_0^{2\pi} = 0 \qquad \blacksquare
\end{aligned}$$

Sea W el subespacio de $C[0, 2\pi]$ generado por las funciones de la ecuación (5). Dada f en $C[0, 2\pi]$, la mejor aproximación a f mediante funciones en W es la **aproximación de Fourier de n-ésimo orden** sobre $[0, 2\pi]$. Puesto que las funciones de la ecuación (5) son ortogonales, la mejor aproximación está dada por la proyección ortogonal sobre W. En este caso, los coeficientes a_k y b_k de la ecuación (4) son los **coeficientes de Fourier** de f. La fórmula estándar para una proyección ortogonal indica que

$$a_k = \frac{\langle f, \cos kt \rangle}{\langle \cos kt, \cos kt \rangle}, \quad b_k = \frac{\langle f, \operatorname{sen} kt \rangle}{\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle}, \quad k \geq 1$$

En el ejercicio 7 se le pide demostrar que $\langle \cos kt, \cos kt \rangle = \pi$ y $\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle = \pi$. Así,

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt\, dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \operatorname{sen} kt\, dt \tag{7}$$

El coeficiente de la función (constante) 1 en la proyección ortogonal es

$$\frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_0^{2\pi} f(t) \cdot 1\, dt = \frac{1}{2} \left[\frac{1}{\pi} \int_0^{2\pi} f(t) \cos(0 \cdot t)\, dt \right] = \frac{a_0}{2}$$

donde a_0 está definida por (7) para $k = 0$. Esto explica por qué el término constante en (4) se escribe como $a_0/2$.

EJEMPLO 4 Encuentre la aproximación de Fourier de n-ésimo orden a la función $f(t) = t$ sobre el intervalo $[0, 2\pi]$.

SOLUCIÓN Calcule

$$\frac{a_0}{2} = \frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi} t\,dt = \frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right] = \pi$$

y para $k > 0$, empleando integración por partes,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} t\cos kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\cos kt + \frac{t}{k}\operatorname{sen} kt\right]_0^{2\pi} = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} t\operatorname{sen} kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\operatorname{sen} kt - \frac{t}{k}\cos kt\right]_0^{2\pi} = -\frac{2}{k}$$

Así, la aproximación de Fourier de n-ésimo orden a $f(t) = t$ es

$$\pi - 2\operatorname{sen} t - \operatorname{sen} 2t - \frac{2}{3}\operatorname{sen} 3t - \cdots - \frac{2}{n}\operatorname{sen} nt$$

La figura 3 muestra las aproximaciones de Fourier de tercer y cuarto órdenes a f. ■

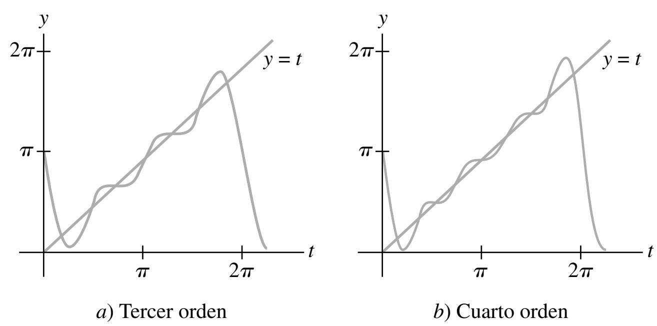


FIGURA 3 Aproximaciones de Fourier de la función $f(t) = t$.

La norma de la diferencia entre f y una aproximación de Fourier es el **error cuadrático medio** de la aproximación. (El término *medio* se refiere al hecho de que la norma está determinada por una integral). Es posible demostrar que el error cuadrático medio se aproxima a cero conforme se incrementa el orden de la aproximación de Fourier. Por esa razón, es común escribir

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty}(a_m\cos mt + b_m\operatorname{sen} mt)$$

Esta expresión para $f(t)$ es la **serie de Fourier** para f sobre $[0, 2\pi]$. El término $a_m\cos mt$, por ejemplo, es la proyección de f sobre un subespacio unidimensional generado por $\cos mt$.

PROBLEMAS DE PRÁCTICA

1. Sean $q_1(t) = 1$, $q_2(t) = t$ y $q_3(t) = 3t^2 - 4$. Compruebe que $\{q_1, q_2, q_3\}$ es un conjunto ortogonal en $C[-2, 2]$ con el producto interior del ejemplo 7 de la sección 6.7 (integración de -2 a 2).
2. Encuentre las aproximaciones de Fourier de primer y tercer órdenes a

$$f(t) = 3 - 2\operatorname{sen} t + 5\operatorname{sen} 2t - 6\cos 2t$$

6.8 EJERCICIOS

1. Encuentre la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$ que mejor se ajuste a los datos $(-2, 0)$, $(-1, 0)$, $(0, 2)$, $(1, 4)$ y $(2, 4)$, suponiendo que el primero y el último de los puntos de datos son menos confiables. Pondérelos a la mitad en relación con los tres restantes puntos interiores.

2. En un problema de mínimos cuadrados ponderados, suponga que 5 de 25 puntos de datos tienen una medición de y que es menos confiable que las demás, y en consecuencia se deben ponderar a la mitad en relación con los otros 20 puntos. Un método consiste en ponderar los 20 puntos por un factor de 1, y los otros 5 por un factor de $\frac{1}{2}$. Un segundo método consiste en ponderar los 20 puntos por un factor de 2, y los otros 5 por un factor de 1. ¿Estos dos métodos producen resultados diferentes? Explique su respuesta.

3. Ajuste una función de tendencia cúbica a los datos del ejemplo 2. El polinomio cúbico ortogonal es $p_3(t) = \frac{5}{6}t^3 - \frac{17}{6}t$.

4. Para realizar un análisis de tendencia de seis puntos de datos igualmente espaciados, se pueden emplear polinomios ortogonales con respecto a la evaluación en los puntos $t = -5, -3, -1, 1, 3$ y 5.

   a) Demuestre que los primeros tres polinomios ortogonales son

   $$p_0(t) = 1, \quad p_1(t) = t \quad \text{y} \quad p_2(t) = \tfrac{3}{8}t^2 - \tfrac{35}{8}$$

   (El polinomio p_2 se escaló para que sus valores en los puntos de evaluación fueran enteros pequeños).

   b) Ajuste una función de tendencia cuadrática a los datos

   $$(-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)$$

En los ejercicios 5 a 14, el espacio es $C[0, 2\pi]$ con el producto interior (6).

5. Demuestre que sen mt y sen nt son ortogonales cuando $m \neq n$.

6. Demuestre que sen mt y cos nt son ortogonales para todos los enteros positivos m y n.

7. Demuestre que $\|\cos kt\|^2 = \pi$ y $\|\text{sen } kt\|^2$ para $k > 0$.

8. Encuentre la aproximación de Fourier de tercer orden a $f(t) = t - 1$.

9. Obtenga la aproximación de Fourier de tercer orden a $f(t) = 2\pi - t$.

10. Determine la aproximación de Fourier de tercer orden a la *función de onda cuadrada*, $f(t) = 1$ para $0 \leq t < \pi$ y $f(t) = -1$ para $\pi \leq t < 2\pi$.

11. Encuentre la aproximación de Fourier de tercer orden a $\text{sen}^2 t$, sin efectuar cálculos con integrales.

12. Obtenga la aproximación de Fourier de tercer orden a $\cos^3 t$, sin efectuar cálculos con integrales.

13. Explique por qué un coeficiente de Fourier de la suma de dos funciones es la suma de los coeficientes de Fourier correspondientes de las dos funciones.

14. Suponga que los primeros pocos coeficientes de Fourier de alguna función f en $C[0, 2\pi]$ son a_0, a_1, a_2, y b_1, b_2, b_3. ¿Cuál de los siguientes polinomios trigonométricos es más cercano a f? Argumente su respuesta.

    $$g(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \text{ sen } t$$

    $$h(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \text{ sen } t + b_2 \text{ sen } 2t$$

15. [**M**] Consulte los datos del ejercicio 13 de la sección 6.6, concernientes al despegue de un avión. Suponga que los errores de las posibles mediciones aumentan conforme se incrementa la rapidez del avión, y sea W la matriz ponderada diagonal cuyas entradas diagonales son 1, 1, 1, .9, .9, .8, .7, .6, .5, .4, .3, .2 y .1. Encuentre la curva cúbica que se ajuste a los datos con error de mínimos cuadrados ponderados, y úsela para estimar la velocidad del avión cuando $t = 4.5$ segundos.

16. [**M**] Sean f_4 y f_5 las aproximaciones de Fourier de cuarto y quinto órdenes en $C[0, 2\pi]$ a la función de onda cuadrada del ejercicio 10. Elabore gráficas separadas de f_4 y f_5 en el intervalo $[0, 2\pi]$, y trace una gráfica de f_5 en $[-2\pi, 2\pi]$.

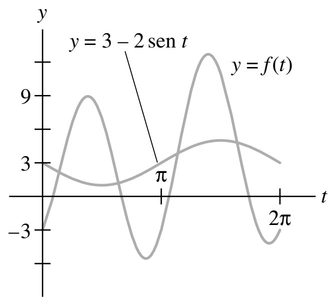
SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Calcule

$$\langle q_1, q_2 \rangle = \int_{-2}^{2} 1 \cdot t \, dt = \frac{1}{2}t^2 \Big|_{-2}^{2} = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^{2} 1 \cdot (3t^2 - 4) \, dt = (t^3 - 4t) \Big|_{-2}^{2} = 0$$

$$\langle q_2, q_3 \rangle = \int_{-2}^{2} t \cdot (3t^2 - 4) \, dt = \left(\frac{3}{4}t^4 - 2t^2\right) \Big|_{-2}^{2} = 0$$



Aproximaciones de primer y tercer órdenes a $f(t)$.

2. La aproximación de Fourier de tercer orden a f es la mejor aproximación en $C[0, 2\varpi]$ a f mediante funciones (vectores) en el subespacio generado por 1, cos t, cos $2t$, cos $3t$, sen t, sen $2t$ y sen $3t$. Pero f está evidentemente *en* este subespacio, así que f es su propia mejor aproximación:

$$f(t) = 3 - 2 \text{ sen } t + 5 \text{ sen } 2t - 6 \cos 2t$$

Para la aproximación de primer orden, la función más cercana a f en el subespacio $W = \text{Gen }\{1, \cos t, \text{sen } t\}$ es $3 - 2$ sen t. Los otros dos términos en la fórmula para $f(t)$ son ortogonales a las funciones en W, así que no contribuyen a las integrales que dan los coeficientes de Fourier para la aproximación de primer orden.

CAPÍTULO 6 EJERCICIOS COMPLEMENTARIOS

1. Los siguientes enunciados se refieren a vectores en $\mathbb{R}^n$ (o $\mathbb{R}^m$) con el producto interior estándar. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

a) La longitud de cada vector es un número positivo.

b) Un vector $\mathbf{v}$ y su negativo $-\mathbf{v}$ tienen iguales longitudes.

c) La distancia entre $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} - \mathbf{v}\|$.

d) Si r es cualquier escalar, entonces $\|r\mathbf{v}\| = r\|\mathbf{v}\|$.

e) Si dos vectores son ortogonales, entonces son linealmente independientes.

f) Si $\mathbf{x}$ es ortogonal a $\mathbf{u}$ y $\mathbf{v}$, entonces $\mathbf{x}$ debe ser ortogonal a $\mathbf{u} - \mathbf{v}$.

g) Si $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

h) Si $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

i) La proyección ortogonal de $\mathbf{y}$ sobre $\mathbf{u}$ es un múltiplo escalar de $\mathbf{y}$.

j) Si un vector $\mathbf{y}$ coincide con su proyección ortogonal sobre un subespacio W, entonces $\mathbf{y}$ está en W.

k) El conjunto de todos los vectores en $\mathbb{R}^n$ ortogonales a un vector fijo es un subespacio de $\mathbb{R}^n$.

l) Si W es un subespacio de $\mathbb{R}^n$, entonces W y $W^\perp$ no tienen vectores en común.

m) Si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal y si c_1, c_2 y c_3 son escalares, entonces $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, c_3\mathbf{v}_3\}$ es un conjunto ortogonal.

n) Si una matriz U tiene columnas ortonormales, entonces $UU^T = I$.

o) Una matriz cuadrada con columnas ortogonales es una matriz ortogonal.

p) Si una matriz cuadrada tiene columnas ortonormales, entonces también tiene filas ortonormales.

q) Si W es un subespacio, entonces $\|\text{proy}_W \mathbf{v}\|^2 + \|\mathbf{v} - \text{proy}_W \mathbf{v}\|^2 = \|\mathbf{v}\|^2$.

r) Una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es el vector $A\hat{\mathbf{x}}$ en Col A más cercano a $\mathbf{b}$, de manera que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\hat{\mathbf{x}}\|$ para toda $\mathbf{x}$.

s) Las ecuaciones normales para una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ están dadas por $\hat{\mathbf{x}} = (A^TA)^{-1} A^T\mathbf{b}$.

2. Sea $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ un conjunto ortonormal. Compruebe la siguiente igualdad mediante inducción, iniciando con $p = 2$. Si $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, entonces

$$\|\mathbf{x}\|^2 = |c_1|^2 + \cdots + |c_p|^2$$

3. Sea $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ un conjunto ortonormal en $\mathbb{R}^n$. Compruebe la siguiente *desigualdad de Bessel*, que es válida para toda $\mathbf{x}$ en $\mathbb{R}^n$:

$$\|\mathbf{x}\|^2 \geq |\mathbf{x}\cdot\mathbf{v}_1|^2 + |\mathbf{x}\cdot\mathbf{v}_2|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_p|^2$$

4. Sea U una matriz ortogonal de $n \times n$. Demuestre que si $\{\mathbf{v}_1,\ldots, \mathbf{v}_n\}$ es una base ortonormal para $\mathbb{R}^n$, entonces también lo es $\{U\mathbf{v}_1,\ldots, U\mathbf{v}_n\}$.

5. Demuestre que si una matriz U de $n \times n$ satisface $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ para todas las $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, entonces U es una matriz ortogonal.

6. Demuestre que si U es una matriz ortogonal, entonces cualquier valor propio real de U debe ser ± 1.

7. Una *matriz de Householder*, o un *reflector elemental*, tiene la forma $Q = I - 2\mathbf{u}\mathbf{u}^T$, donde $\mathbf{u}$ es un vector unitario. (Véase el ejercicio 13 en los ejercicios complementarios del capítulo 2). Demuestre que Q es una matriz ortogonal. (Los reflectores elementales se utilizan a menudo en programas computacionales para construir una factorización QR de una matriz A. Si A tiene columnas linealmente independientes, entonces la multiplicación por la izquierda mediante una secuencia de reflectores elementales puede generar una matriz triangular superior).

8. Sea $T : \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal que preserva longitudes; es decir, $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

a) Demuestre que T también preserva ortogonalidad; es decir, $T(\mathbf{x})\cdot T(\mathbf{y}) = 0$ siempre que $\mathbf{x}\cdot\mathbf{y} = 0$.

b) Demuestre que la matriz estándar de T es una matriz ortogonal.

9. Considere que $\mathbf{u}$ y $\mathbf{v}$ representan vectores linealmente independientes en $\mathbb{R}^n$ que *no* son ortogonales. Describa cómo encontrar la mejor aproximación a $\mathbf{z}$ en $\mathbb{R}^n$ mediante vectores de la forma $x_1\mathbf{u} + x_2\mathbf{v}$ sin construir primero una base ortogonal para Gen $\{\mathbf{u}, \mathbf{v}\}$.

10. Suponga que las columnas de A son linealmente independientes. Determine qué pasa con la solución de mínimos cuadrados $\hat{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ cuando $\mathbf{b}$ se remplaza por $c\mathbf{b}$ para algún escalar c diferente de cero.

11. Si a, b y c son números distintos, entonces el siguiente sistema es inconsistente porque las gráficas de las ecuaciones son planos paralelos. Demuestre que el conjunto de todas las soluciones de mínimos cuadrados del sistema es precisamente el plano cuya ecuación es $x - 2y + 5z = (a + b + c)/3$.

$$\begin{aligned} x - 2y + 5z &= a \\ x - 2y + 5z &= b \\ x - 2y + 5z &= c \end{aligned}$$

12. Considere el problema de encontrar un valor propio de una matriz A de $n \times n$ cuando se conoce un vector propio $\mathbf{v}$ aproximado. Puesto que $\mathbf{v}$ no es exactamente correcto, la ecuación

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

probablemente no tendrá una solución. Sin embargo, λ se puede estimar mediante una solución de mínimos cuadrados cuando la ecuación (1) se analiza adecuadamente. Piense en $\mathbf{v}$ como en una matriz V de $n \times 1$, y considere λ un vector en $\mathbb{R}^1$, y denote el vector $A\mathbf{v}$ con el símbolo $\mathbf{b}$. Luego, (1) se reduce a $\mathbf{b} = \lambda V$, que también se puede escribir como $V\lambda = \mathbf{b}$. Obtenga la solución de mínimos cuadrados de este sistema de n ecuaciones con la única incógnita λ, y escriba esta solución empleando los símbolos originales. La estimación resultante para λ se denomina *cociente de Rayleigh*. Véase los ejercicios 11 y 12 de la sección 5.8.

13. Siga los pasos descritos a continuación para probar las siguientes relaciones entre los cuatro subespacios fundamentales determinados por una matriz A de $m \times n$.

$$\text{Fil}\,A = (\text{Nul}\,A)^\perp, \quad \text{Col}\,A = (\text{Nul}\,A^T)^\perp$$

a) Demuestre que Fil A está contenida en $(\text{Nul}\,A)^\perp$. (Demuestre que si $\mathbf{x}$ está en Fil A, entonces $\mathbf{x}$ es ortogonal a cada $\mathbf{u}$ en Nul A).

b) Suponga que rango $A = r$. Encuentre dim Nul A y dim $(\text{Nul}\,A)^\perp$, y entonces deduzca del inciso *a*) que Fil $A = (\text{Nul}\,A)^\perp$. [*Sugerencia:* Estudie los ejercicios de la sección 6.3].

c) Explique por qué Col $A = (\text{Nul}\,A)^\perp$.

14. Explique por qué una ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución si y solo si $\mathbf{b}$ es ortogonal a todas las soluciones de la ecuación $A^T\mathbf{x} = \mathbf{0}$.

Los ejercicios 15 y 16 conciernen a la *factorización de Schur* (real) de una matriz A de $n \times n$ en la forma $A = URU^T$, donde U es una matriz ortogonal y R es una matriz triangular superior de $n \times n$.[1]

15. Demuestre que si A admite una factorización de Schur (real), $A = URU^T$, entonces A tiene n valores propios reales, contando multiplicidades.

16. Sea A una matriz de $n \times n$ con n valores propios reales, contando multiplicidades, denotados con $\lambda_1,\ldots, \lambda_n$. Es posible demostrar que A admite una factorización de Schur (real). Los incisos *a*) y *b*) dan las ideas clave de la demostración. El resto de la demostración equivale a repetir *a*) y *b*) para matrices sucesivamente más pequeñas, y luego unir todos los resultados.

a) Sean $\mathbf{u}_1$ un vector propio unitario correspondiente a λ_1, y $\mathbf{u}_2,\ldots, \mathbf{u}_n$ cualesquiera otros vectores tales que $\{\mathbf{u}_1,\ldots, \mathbf{u}_n\}$ sea una base ortonormal para $\mathbb{R}^n$, y entonces sea $U = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_n]$. Demuestre que la primera columna de $U^T AU$ es $\lambda_1\mathbf{e}_1$, donde $\mathbf{e}_1$ es la primera columna de la matriz identidad de $n \times n$.

b) El inciso *a*) implica que U^TAU tiene la forma que se muestra abajo. Explique por qué los valores propios de A_1 son $\lambda_2,\ldots, \lambda_n$. [*Sugerencia:* Véase los ejercicios complementarios del capítulo 5].

$$U^TAU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[**M**] Cuando el lado derecho de una ecuación $A\mathbf{x} = \mathbf{b}$ se modifica ligeramente, por ejemplo, a $A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$ para algún vector $\Delta\mathbf{b}$, la solución cambia de $\mathbf{x}$ a $\mathbf{x} + \Delta\mathbf{x}$, donde $\Delta\mathbf{x}$ satisface $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$. El cociente $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ es el **cambio relativo** en $\mathbf{b}$ (o el **error relativo** en $\mathbf{b}$ cuando $\Delta\mathbf{b}$ representa el error posible en las entradas de $\mathbf{b}$). El cambio relativo en la solución es $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$. Cuando A es invertible, el **número de condición** de A, representado como cond(A), aporta un límite de la magnitud del cambio relativo en $\mathbf{x}$:

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \leq \text{cond}(A)\cdot\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \tag{2}$$

En los ejercicios 17 a 20, resuelva $A\mathbf{x} = \mathbf{b}$ y $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$, y demuestre que la desigualdad (2) es válida en cada caso. (Véase el análisis de matrices *mal condicionadas* en los ejercicios 41 a 43 de la sección 2.3).

17. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}, \Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

18. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} .500 \\ -1.407 \end{bmatrix}, \Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

[1] Si se permiten números complejos, *cada matriz A de n × n* admite una factorización de Schur (compleja), $A = URU^{-1}$, donde R es triangular superior y U^{-1} es la transpuesta *conjugada* de U. Este hecho de gran utilidad se analiza en la obra *Matrix Analysis*, de Roger A. Horn y Charles R. Johnson (Cambridge: Cambridge University Press, 1985), pp.79-100.

19. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .100 \\ 2.888 \\ -1.404 \\ 1.462 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .49 \\ -1.28 \\ 5.78 \\ 8.04 \end{bmatrix}$

20. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4.230 \\ -11.043 \\ 49.991 \\ 69.536 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .27 \\ 7.76 \\ -3.77 \\ 3.93 \end{bmatrix}$

7 Matrices simétricas y formas cuadráticas

EJEMPLO INTRODUCTORIO

Procesamiento de imágenes multicanal

Alrededor del mundo, en poco más de 80 *minutos*, los dos satélites Landsat cruzan silenciosamente el firmamento en órbitas casi polares, grabando imágenes del terreno y las líneas costeras, en franjas de 185 kilómetros de ancho. Luego de 16 días, cada satélite habrá transitado por encima de casi todos los kilómetros cuadrados de la superficie de la Tierra, de manera que cualquier zona se puede monitorizar cada ocho días.

Las imágenes Landsat son útiles para muchos objetivos. Los diseñadores y planificadores urbanos las utilizan para estudiar la rapidez y dirección del crecimiento urbano, el desarrollo industrial y otros cambios en el uso del suelo. Esas imágenes permiten analizar la humedad del suelo en las zonas rurales, clasificar la vegetación de regiones remotas, y localizar arroyos y lagos tierra adentro. Los gobiernos pueden detectar y evaluar los daños de desastres naturales, como incendios forestales, flujos de lava, inundaciones y huracanes. Las agencias de protección ambiental pueden identificar contaminación por emisiones de humo y medir la temperatura del agua en lagos y ríos cerca de plantas de energía.

Los sensores satelitales toman siete imágenes simultáneas de cada región del suelo que se va a estudiar. Los sensores registran energía de bandas de longitud de onda separadas: tres en el espectro de luz visible y cuatro en las bandas infrarroja y térmica. Cada imagen se digitaliza y se almacena como un arreglo rectangular de números, donde cada número indica la intensidad de la señal en un punto pequeño (o *pixel*) correspondiente en la imagen. Cada una de las siete imágenes es un canal de una *imagen multicanal* o *multiespectral*.

Las siete imágenes Landsat de una región fija, por lo general, contienen mucha información redundante, ya que algunos aspectos aparecen en varias imágenes. Sin embargo, otros aspectos, por su color o temperatura, reflejan luz que solo se registra en uno o dos de los sensores. Un objetivo del procesamiento de imágenes multiespectrales es el de visualizar los datos en una forma que permita extraer información de mejor calidad que cuando se estudia cada imagen por separado.

El *análisis de componentes principales* es una manera eficaz de suprimir información redundante y concentrar, en tan solo una o dos imágenes compuestas, la mayor parte de la información de los datos iniciales. A grandes rasgos, el objetivo es encontrar una combinación lineal especial de las imágenes, es decir, una lista de pesos que en cada pixel combinen los siete valores correspondientes de las imágenes en un nuevo valor. Los pesos se eligen de manera que permitan ampliar, en comparación con cualquiera de las imágenes originales, el rango de las intensidades de luz (la *varianza de la escena*) en la imagen compuesta (llamada la *primera componente principal*). También es posible obtener imágenes adicionales de *componentes*, mediante un criterio que se explicará en la sección 7.5.

En las fotografías tomadas sobre el Valle Railroad, Nevada, que se muestran a continuación se ilustra el análisis de componentes principales. En las fotografías *a*), *b*) y *c*) se observan imágenes de tres bandas espectrales Landsat. La información total en las tres bandas se reacomoda en las tres imágenes de componentes principales en las fotografías *d*), *e*) y *f*). La primera componente *d*) presenta (o "explica") el 93.5% de la varianza de la escena presente en los datos iniciales. De esta forma, los datos iniciales de tres canales se reducen a datos de un canal, con un pérdida (en algún sentido) de solo el 6.5% de la varianza de la escena.

La empresa Earth Satellite Corporation, en Rockville, Maryland, que amablemente proporcionó las fotos que aquí se incluyen, está experimentando con imágenes de 224 bandas espectrales separadas. El análisis de componentes principales, esencial para esos conjuntos enormes de datos, por lo general reduce los datos a unas 15 componentes principales utilizables.

WEB

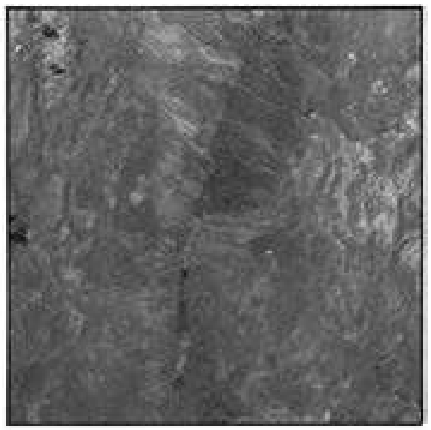
a) Banda espectral 1: Azul visible.

b) Banda espectral 4: Casi infrarrojo.

c) Banda espectral 7: Infrarrojo medio.

d) Componente principal 1: 93.5%.

e) Componente principal 2: 5.3%.



f) Componente principal 3: 1.2%.

En las aplicaciones, las matrices simétricas surgen, de una u otra manera, con mayor frecuencia que cualquier otra clase de matrices. La teoría es bella y rica, dependiendo en esencia de la diagonalización y la ortogonalidad, temas que se estudiaron en los capítulos 5 y 6, respectivamente. La diagonalización de una matriz simétrica, descrita en la sección 7.1, es el fundamento del análisis en las secciones 7.2 y 7.3 concernientes a formas cuadráticas. A la vez, la sección 7.3 es necesaria para las dos últimas secciones referentes a la descomposición en valores singulares y al procesamiento de imágenes descrito en el ejemplo introductorio. En este capítulo todos los vectores y las matrices tienen entradas reales.

7.1 DIAGONALIZACIÓN DE MATRICES SIMÉTRICAS

Una matriz **simétrica** es una matriz A tal que $A^T = A$. Esta matriz es necesariamente cuadrada. Sus entradas en la diagonal principal son arbitrarias, pero sus otras entradas se presentan por pares, en lados opuestos de la diagonal principal.

EJEMPLO 1 De las siguientes matrices, solo las primeras tres son simétricas:

$$\text{Simétricas:}\quad \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}, \quad \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$$

$$\text{No simétricas:}\quad \begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$ ■

Para comenzar el estudio de matrices simétricas, es útil revisar el proceso de diagonalización de la sección 5.3.

EJEMPLO 2 Si es posible, diagonalice la matriz $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$.

SOLUCIÓN La ecuación característica de A es

$$0 = -\lambda^3 + 17\lambda^2 - 90\lambda + 144 = -(\lambda - 8)(\lambda - 6)(\lambda - 3)$$

Los cálculos estándar producen una base para cada espacio propio:

$$\lambda = 8:\ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = 6:\ \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda = 3:\ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Estos tres vectores conforman una base para $\mathbb{R}^3$. De hecho, es fácil comprobar que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base *ortogonal* de $\mathbb{R}^3$. La experiencia del capítulo 6 sugiere que una base *ortonormal* sería útil en los cálculos, así que a continuación se presentan los vectores propios normalizados a uno:

$$\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Sean

$$P = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad D = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Entonces $A = PDP^{-1}$, como es usual. Pero esta vez, como P es cuadrada y tiene columnas ortonormales, resulta que P es una matriz *ortogonal*, y P^{-1} es simplemente P^T. (Véase la sección 6.2). ■

El teorema 1 explica por qué los vectores propios del ejemplo 2 son ortogonales, ya que corresponden a distintos valores propios.

TEOREMA 1

Si A es simétrica, entonces dos vectores propios cualesquiera de diferentes espacios propios son ortogonales.

DEMOSTRACIÓN Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ los vectores propios asociados a los valores propios diferentes λ_1 y λ_2. Para demostrar que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, calcule

$$\begin{aligned}
\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 = (A\mathbf{v}_1)^T \mathbf{v}_2 && \text{Puesto que } \mathbf{v}_1 \text{ es un vector propio} \\
&= (\mathbf{v}_1^T A^T)\mathbf{v}_2 = \mathbf{v}_1^T (A\mathbf{v}_2) && \text{Debido a que } A^T = A \\
&= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2) && \text{Porque } \mathbf{v}_2 \text{ es un vector propio} \\
&= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2
\end{aligned}$$

Por lo tanto, $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Pero $(\lambda_1 - \lambda_2) \neq 0$, así que $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. ■

El tipo especial de diagonalización del ejemplo 2 es crucial para la teoría de matrices simétricas. Se dice que una matriz A de $n \times n$ es **diagonalizable ortogonalmente** si existen una matriz ortogonal P (con $P^{-1} = P^T$) y una matriz diagonal D tales que

$$A = PDP^T = PDP^{-1} \tag{1}$$

Esta diagonalización requiere n vectores propios ortonormales linealmente independientes. ¿Cuándo es posible esto? Si A es diagonalizable ortogonalmente como en la ecuación (1), entonces

$$A^T = (PDP^T)^T = P^{TT} D^T P^T = PDP^T = A$$

Por lo tanto, ¡A es simétrica! El teorema 2 que se presenta a continuación, muestra que, a la inversa, cada matriz simétrica es diagonalizable ortogonalmente. La demostración es mucho más difícil y se omite; la idea principal para la demostración se presentará después del teorema 3.

TEOREMA 2

Una matriz A de $n \times n$ es diagonalizable ortogonalmente si y solo si A es una matriz simétrica.

Este teorema es sorprendente, ya que el trabajo realizado en el capítulo 5 sugeriría que es prácticamente imposible decir cuándo una matriz es diagonalizable. Pero este no es el caso para las matrices simétricas.

El siguiente ejemplo se refiere a una matriz cuyos valores propios no son todos distintos.

EJEMPLO 3 Diagonalice ortogonalmente la matriz $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$, cuya ecuación característica es

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2)$$

SOLUCIÓN Los cálculos usuales producen bases para los espacios propios:

$$\lambda = 7: \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda = -2: \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

Aunque $\mathbf{v}_1$ y $\mathbf{v}_2$ son linealmente independientes, no son ortogonales. Recuerde de la sección 6.2 que la proyección de $\mathbf{v}_2$ sobre $\mathbf{v}_1$ es $\dfrac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$, y que la componente de $\mathbf{v}_2$ ortogonal a $\mathbf{v}_1$ es

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \frac{-1/2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{z}_2\}$ es un conjunto ortogonal en el espacio propio para $\lambda = 7$. (Observe que $\mathbf{z}_2$ es una combinación lineal de los vectores propios $\mathbf{v}_1$ y $\mathbf{v}_2$, así que $\mathbf{z}_2$ está en el espacio propio. Esta construcción de $\mathbf{z}_2$ es justamente el proceso de Gram-Schmidt de la sección 6.4). Como el espacio propio tiene dos dimensiones (con la base $\mathbf{v}_1$, $\mathbf{v}_2$), el conjunto ortogonal $\{\mathbf{v}_1, \mathbf{z}_2\}$ es una *base ortogonal* para el espacio propio, de acuerdo con el teorema de la base. (Véase la sección 2.9 o la 4.5).

Al normalizar $\mathbf{v}_1$ y $\mathbf{z}_2$ se obtiene la siguiente base ortonormal del espacio propio para $\lambda = 7$:

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$$

Una base ortonormal del espacio propio para $\lambda = -2$ es

$$\mathbf{u}_3 = \frac{1}{\|2\mathbf{v}_3\|} 2\mathbf{v}_3 = \frac{1}{3}\begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$$

De acuerdo con el teorema 1, $\mathbf{u}_3$ es ortogonal a los vectores propios $\mathbf{u}_1$ y $\mathbf{u}_2$. Por lo tanto, $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ es un conjunto ortonormal. Sean

$$P = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3\,] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Entonces P diagonaliza ortogonalmente a A, y a $A = PDP^{-1}$. ■

En el ejemplo 3, el valor propio 7 tiene multiplicidad dos y el espacio propio es bidimensional. Este hecho no es fortuito, como lo muestra el siguiente teorema.

Teorema espectral

Algunas veces el conjunto de valores propios de una matriz A se denomina el *espectro* de A, y la siguiente descripción de los valores propios se conoce como *teorema espectral*.

TEOREMA 3

Teorema espectral para matrices simétricas

Una matriz simétrica A de $n \times n$ tiene las siguientes propiedades:

a) A tiene n valores propios reales, contando las multiplicidades.

b) La dimensión del espacio propio para cada valor propio λ es igual a la multiplicidad de λ como una raíz de la ecuación característica.

c) Los espacios propios son mutuamente ortogonales, en el sentido de que los vectores propios correspondientes a diferentes valores propios son ortogonales.

d) A es diagonalizable ortogonalmente.

El inciso *a*) se deduce del ejercicio 24 de la sección 5.5. El inciso *b*) se deduce fácilmente del inciso *d*). (Véase el ejercicio 31). El inciso *c*) es el teorema 1. A partir del inciso *a*), puede hacerse una demostración para *d*) utilizando el ejercicio 32 y la factorización de Schur analizada en el ejercicio complementario 16 del capítulo 6. Se omiten los detalles.

Descomposición espectral

Suponga que $A = PDP^{-1}$, donde las columnas de P son los vectores propios ortonormales $\mathbf{u}_1,\ldots,\mathbf{u}_n$ de A, y los valores propios correspondientes $\lambda_1,\ldots,\lambda_n$ están en la matriz diagonal D. Entonces, como $P^{-1} = P^T$,

$$A = PDP^T = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{bmatrix}$$

Utilizando la expansión columna-fila de un producto (teorema 10 de la sección 2.4), se puede escribir

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \lambda_2\mathbf{u}_2\mathbf{u}_2^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T \tag{2}$$

A esta representación de A se le llama **descomposición espectral** de A porque la divide en partes determinadas por su espectro (valores propios). Cada término en (2) es una matriz de $n \times n$ de rango 1. Por ejemplo, cada columna de $\lambda_1\mathbf{u}_1\mathbf{u}_1^T$ es un múltiplo de $\mathbf{u}_1$. Además, cada matriz $\mathbf{u}_j\mathbf{u}_j^T$ es una **matriz de proyección** en el sentido de que para cada $\mathbf{x}$ en $\mathbb{R}^n$, el vector $(\mathbf{u}_j\mathbf{u}_j^T)\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre el subespacio generado por $\mathbf{u}_j$. (Véase el ejercicio 35).

EJEMPLO 4 Construya una descomposición espectral de la matriz A que tiene la diagonalización ortogonal

$$A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

SOLUCIÓN Denote las columnas de P como $\mathbf{u}_1$ y $\mathbf{u}_2$. Entonces,

$$A = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T$$

Para comprobar esta descomposición de A, calcule

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix}$$

y

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{bmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{bmatrix} + \begin{bmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A \qquad \blacksquare$$

NOTA NUMÉRICA

Cuando A es simétrica y no demasiado grande, los modernos algoritmos computacionales de gran desempeño calculan valores propios y vectores propios con gran precisión. Aplican una secuencia de transformaciones de semejanza sobre A que implican matrices ortogonales. Las entradas diagonales de las matrices transformadas convergen rápidamente a los valores propios de A. (Véase las notas numéricas de la sección 5.2). Por lo general, el empleo de matrices ortogonales evita el acumulamiento de errores numéricos durante el proceso. Cuando A es simétrica, la secuencia de matrices ortogonales se combina para formar una matriz ortogonal cuyas columnas son vectores propios de A.

Una matriz no simétrica no puede tener un conjunto completo de vectores propios ortogonales, pero el algoritmo continuará produciendo valores propios más o menos exactos. Después de eso, se requieren técnicas no ortogonales para calcular los vectores propios.

PROBLEMAS DE PRÁCTICA

1. Demuestre que si A es una matriz simétrica, entonces A^2 también es simétrica.
2. Demuestre que si A es diagonalizable ortogonalmente, entonces A^2 también lo es.

7.1 EJERCICIOS

En los ejercicios 1 a 6, determine cuáles matrices son simétricas.

1. $\begin{bmatrix} 3 & 5 \\ 5 & -7 \end{bmatrix}$ **2.** $\begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 2 \\ 4 & 4 \end{bmatrix}$ **4.** $\begin{bmatrix} 0 & 8 & 3 \\ 8 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix}$

5. $\begin{bmatrix} -6 & 2 & 0 \\ 0 & -6 & 2 \\ 0 & 0 & -6 \end{bmatrix}$ **6.** $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \end{bmatrix}$

En los ejercicios 7 a 12, determine qué matrices son ortogonales. Si alguna es ortogonal, encuentre su inversa.

7. $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$ **8.** $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

9. $\begin{bmatrix} -5 & 2 \\ 2 & 5 \end{bmatrix}$ **10.** $\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{bmatrix}$

11. $\begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 0 & 1/\sqrt{5} & -2/\sqrt{5} \\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45} \end{bmatrix}$

12. $\begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}$

En los ejercicios 13 a 22, diagonalice ortogonalmente las matrices, dando una matriz ortogonal P y una matriz diagonal D. Para ahorrarle tiempo, los valores propios en los ejercicios 17 a 22 son: (17) 5, 2, -2; (18) 25, 3, -50; (19) 7, -2; (20) 13, 7, 1; (21) 9, 5, 1; (22) 2, 0.

13. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$ **16.** $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ **18.** $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

19. $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ **20.** $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$

21. $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$ **22.** $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

23. Sean $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Compruebe que 2 es un valor propio de A y $\mathbf{v}$ un vector propio. Después, diagonalice A ortogonalmente.

24. Sean $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

Compruebe que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A. Después diagonalice A ortogonalmente.

En los ejercicios 25 y 26, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

25. *a*) Una matriz de $n \times n$ que es diagonalizable ortogonalmente debe ser simétrica.

b) Si $A^T = A$ y si los vectores $\mathbf{u}$ y $\mathbf{v}$ satisfacen $A\mathbf{u} = 3\mathbf{u}$ y $A\mathbf{v} = 4\mathbf{v}$, entonces $\mathbf{u}\cdot\mathbf{v} = 0$.

c) Una matriz simétrica de $n \times n$ tiene n valores propios reales distintos.

d) Para $\mathbf{v}$ diferente de cero en $\mathbb{R}^n$, la matriz $\mathbf{v}\mathbf{v}^T$ es una matriz de proyección.

26. *a*) Cada matriz simétrica es diagonalizable ortogonalmente.

b) Si $B = PDP^T$, donde $P^T = P^{-1}$ y D es una matriz diagonal, entonces B es una matriz simétrica.

c) Una matriz ortogonal es diagonalizable ortogonalmente.

d) La dimensión de un espacio propio de una matriz simétrica es igual a la multiplicidad del valor propio correspondiente.

27. Suponga que A es una matriz simétrica de $n \times n$, y B es cualquier matriz de $n \times m$. Demuestre que B^TAB, B^TB y BB^T son matrices simétricas.

28. Demuestre que si A es una matriz simétrica de $n \times n$, entonces $(A\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot(A\mathbf{y})$ para toda $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{R}^n$.

29. Suponga que A es invertible y diagonalizable ortogonalmente. Explique por qué A^{-1} también es diagonalizable ortogonalmente.

30. Suponga que tanto A como B son diagonalizables ortogonalmente y que $AB = BA$. Explique por qué AB también es diagonalizable ortogonalmente.

31. Sean $A = PDP^{-1}$, donde P es ortogonal y D es diagonal, y λ es un valor propio de A de multiplicidad k. Entonces λ se presenta k veces sobre la diagonal de D. Explique por qué la dimensión del espacio propio para λ es k.

32. Suponga que $A = PRP^{-1}$, donde P es ortogonal y R es triangular superior. Demuestre que si A es simétrica, entonces R es simétrica y, por lo tanto, realmente es una matriz diagonal.

33. Construya una descomposición espectral de A del ejemplo 2.

34. Construya una descomposición espectral de A del ejemplo 3.

35. Sea $\mathbf{u}$ el vector unitario en $\mathbb{R}^n$, y sea $B = \mathbf{u}\mathbf{u}^T$.

a) Dada cualquier $\mathbf{x}$ en $\mathbb{R}^n$, calcule $B\mathbf{x}$ y demuestre que $B\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre $\mathbf{u}$, como se describió en la sección 6.2.

b) Demuestre que B es una matriz simétrica y que $B^2 = B$.

c) Demuestre que $\mathbf{u}$ es un vector propio de B. ¿Cuál es el valor propio correspondiente?

36. Sea B una matriz simétrica de $n \times n$ tal que $B^2 = B$. Cualquier matriz de este tipo se conoce como **matriz de proyección** (o **matriz de proyección ortogonal**). Para cualquier $\mathbf{y}$ en $\mathbb{R}^n$, sean $\hat{\mathbf{y}} = B\mathbf{y}$ y $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

a) Demuestre que $\mathbf{z}$ es ortogonal a $\hat{\mathbf{y}}$.

b) Sea W el espacio columna de B. Demuestre que $\mathbf{y}$ es la suma de un vector en W y de un vector en $W^\perp$. ¿Por qué esto demuestra que $B\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre el espacio columna de B?

[**M**] En los ejercicios 37 a 40, diagonalice ortogonalmente las matrices indicadas. Para practicar los métodos de esta sección, no utilice la rutina de vectores propios de su programa de matrices. En vez de ello, aplique el programa para encontrar los valores propios, y para cada valor propio λ, obtenga una base ortonormal para $\mathrm{Nul}(A - \lambda I)$, como en los ejemplos 2 y 3.

37. $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

38. $\begin{bmatrix} .38 & -.18 & -.06 & -.04 \\ -.18 & .59 & -.04 & .12 \\ -.06 & -.04 & .47 & -.12 \\ -.04 & .12 & -.12 & .41 \end{bmatrix}$

39. $\begin{bmatrix} .31 & .58 & .08 & .44 \\ .58 & -.56 & .44 & -.58 \\ .08 & .44 & .19 & -.08 \\ .44 & -.58 & -.08 & .31 \end{bmatrix}$

40. $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. $(A^2)^T = (AA)^T = A^TA^T$, por una propiedad de la transpuesta. Por hipótesis, $A^T = A$. Así que $(A^2)^T = AA = A^2$, lo que demuestra que A^2 es simétrica.

2. Si A es diagonalizable ortogonalmente, entonces A es simétrica de acuerdo con el teorema 2. Según el problema de práctica 1, A^2 es simétrica y, por lo tanto, es diagonalizable ortogonalmente (teorema 2).

7.2 FORMAS CUADRÁTICAS

Hasta ahora en este libro nos hemos concentrado en ecuaciones lineales, excepto por la sumas de cuadrados que se encontraron en el capítulo 6 al calcular $\mathbf{x}^T\mathbf{x}$. Estas sumas y expresiones más generales, llamadas *formas cuadráticas*, se presentan con frecuencia en aplicaciones del álgebra lineal a la ingeniería (en criterios de diseño y optimización) y el procesamiento de señales (como potencia de ruido de salida). Las aplicaciones también surgen, por ejemplo, en física (como energías potencial y cinética), geometría diferencial (como la curvatura normal de superficies), economía (como funciones de utilidad) y estadística (en elipsoides de confianza). Parte de las bases matemáticas para dichas aplicaciones fluye fácilmente de nuestro trabajo de matrices simétricas.

Una **forma cuadrática** en $\mathbb{R}^n$ es una función Q definida sobre $\mathbb{R}^n$ cuyo valor en un vector $\mathbf{x}$ de $\mathbb{R}^n$ se puede calcular mediante una expresión de la forma $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde A es una matriz simétrica de $n \times n$. La matriz A se denomina **matriz de la forma cuadrática**.

El ejemplo más sencillo de una forma cuadrática diferente de cero es $Q(\mathbf{x}) = \mathbf{x}^TI\mathbf{x} = \|\mathbf{x}\|^2$. Los ejemplos 1 y 2 revelan la conexión entre cualquier matriz simétrica A y la forma cuadrática $\mathbf{x}^TA\mathbf{x}$.

EJEMPLO 1 Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Calcule $\mathbf{x}^TA\mathbf{x}$ para las siguientes matrices:

a) $A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$ *b*) $A = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$

SOLUCIÓN

a) $\mathbf{x}^TA\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 4x_1 \\ 3x_2 \end{bmatrix} = 4x_1^2 + 3x_2^2$

b) En A existen dos entradas -2. Observe cómo participan en los cálculos. La entrada (1, 2) en A está en negritas.

$$\begin{aligned}\mathbf{x}^TA\mathbf{x} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 3 & \mathbf{-2} \\ -2 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 3x_1 - \mathbf{2}x_2 \\ -2x_1 + 7x_2 \end{bmatrix} \\ &= x_1(3x_1 - \mathbf{2}x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - \mathbf{2}x_1x_2 - 2x_2x_1 + 7x_2^2 \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2\end{aligned}$$ ■

La presencia de $-4x_1x_2$ en la forma cuadrática del ejemplo 1*b*) se debe a las entradas -2 fuera de la diagonal en la matriz A. En contraste, la forma cuadrática asociada con la matriz diagonal A del ejemplo 1*a*) no tiene el *producto cruzado* x_1x_2.

EJEMPLO 2 Para $\mathbf{x}$ en $\mathbb{R}^3$, sea $Q(\mathbf{x}) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3$. Escriba esta forma cuadrática como $\mathbf{x}^TA\mathbf{x}$.

SOLUCIÓN Los coeficientes de x_1^2, x_2^2, x_3^2 están en la diagonal de A. Para que A sea simétrica, el coeficiente de x_ix_j con $i \neq j$ debe dividirse en partes iguales entre las entradas (i, j) y (j, i) en A. El coeficiente de x_1x_3 es 0. Es sencillo comprobar que:

$$Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}\begin{bmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$ ■

EJEMPLO 3 Sea $Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$. Calcule el valor de $Q(\mathbf{x})$ para $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

SOLUCIÓN

$$\begin{aligned} Q(-3,1) &= (-3)^2 - 8(-3)(1) - 5(1)^2 = 28 \\ Q(2,-2) &= (2)^2 - 8(2)(-2) - 5(-2)^2 = 16 \\ Q(1,-3) &= (1)^2 - 8(1)(-3) - 5(-3)^2 = -20 \end{aligned}$$ ■

En algunos casos, las formas cuadráticas son más fáciles de usar cuando no tienen productos cruzados, es decir, cuando la matriz de la forma cuadrática es una matriz diagonal. Por fortuna, el producto cruzado se puede eliminar realizando un adecuado cambio de variable.

Cambio de variable en una forma cuadrática

Si $\mathbf{x}$ representa un vector variable en $\mathbb{R}^n$, entonces un **cambio de variable** es una ecuación de la forma

$$\mathbf{x} = P\mathbf{y}, \qquad \text{o, de manera equivalente,} \quad \mathbf{y} = P^{-1}\mathbf{x} \tag{1}$$

donde P es una matriz invertible y $\mathbf{y}$ es un nuevo vector variable en $\mathbb{R}^n$. Aquí $\mathbf{y}$ es el vector de coordenadas de $\mathbf{x}$ con respecto a la base de $\mathbb{R}^n$ determinada por las columnas de P. (Véase la sección 4.4).

Si el cambio de variable (1) se efectúa en una forma cuadrática $\mathbf{x}^TA\mathbf{x}$, entonces

$$\mathbf{x}^TA\mathbf{x} = (P\mathbf{y})^TA(P\mathbf{y}) = \mathbf{y}^TP^TAP\mathbf{y} = \mathbf{y}^T(P^TAP)\mathbf{y} \tag{2}$$

y la nueva matriz de la forma cuadrática es P^TAP. Como A es simétrica, entonces el teorema 2 garantiza que existe una matriz *ortogonal* P tal que P^TAP es una matriz diagonal D, y así la forma cuadrática en (2) se convierte en $\mathbf{y}^TD\mathbf{y}$. Esta es la estrategia del siguiente ejemplo.

EJEMPLO 4 Realice un cambio de variable que transforme la forma cuadrática del ejemplo 3 en una forma cuadrática sin productos cruzados.

SOLUCIÓN La matriz de la forma cuadrática del ejemplo 3 es

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

El primer paso es que A se diagonalice ortogonalmente. Sus valores propios resultan ser $\lambda = 3$ y $\lambda = -7$. Los vectores propios unitarios asociados son:

$$\lambda = 3: \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}; \qquad \lambda = -7: \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

Esos vectores son automáticamente ortogonales (porque corresponden a diferentes valores propios), de manera que brindan una base ortonormal para $\mathbb{R}^2$. Sean

$$P = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \qquad D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

Entonces, $A = PDP^{-1}$ y $D = P^{-1}AP = P^TAP$, como se indicó antes. Un cambio de variable conveniente es

$$\mathbf{x} = P\mathbf{y}, \qquad \text{donde } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

De esta forma,

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) \\ &= \mathbf{y}^T P^T AP\mathbf{y} = \mathbf{y}^T D\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned} \qquad \blacksquare$$

Para ilustrar el significado de la igualdad de las formas cuadráticas del ejemplo 4, se puede calcular $Q(\mathbf{x})$ para $\mathbf{x} = (2, -2)$ empleando la nueva forma cuadrática. Primero, como $\mathbf{x} = P\mathbf{y}$,

$$\mathbf{y} = P^{-1}\mathbf{x} = P^T\mathbf{x}$$

de manera que

$$\mathbf{y} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 6/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

Por lo tanto,

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5) \\ &= 80/5 = 16 \end{aligned}$$

Este es el valor de $Q(\mathbf{x})$ en el ejemplo 3 cuando $\mathbf{x} = (2, -2)$. Véase la figura 1.

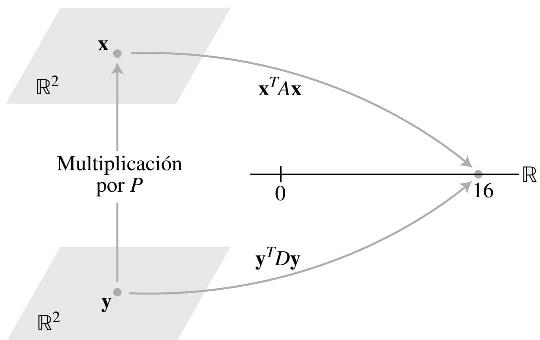


FIGURA 1 Cambio de variable en $\mathbf{x}^T A\mathbf{x}$.

El ejemplo 4 ilustra el siguiente teorema. En esencia, la demostración del teorema se presentó antes del ejemplo 4.

TEOREMA 4 **Teorema de los ejes principales**

Sea A una matriz simétrica de $n \times n$. Luego, existe un cambio de variable ortogonal, $\mathbf{x} = P\mathbf{y}$, que convierte la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ en una forma cuadrática $\mathbf{y}^T D\mathbf{y}$ sin productos cruzados.

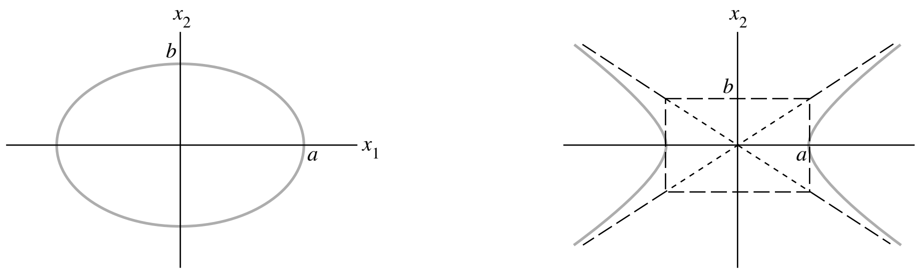
En el teorema, las columnas de P se llaman **ejes principales** de la forma cuadrática $\mathbf{x}^T A\mathbf{x}$. El vector $\mathbf{y}$ es el vector de coordenadas de $\mathbf{x}$ respecto de la base ortonormal de $\mathbb{R}^n$ que dan esos ejes principales.

Ejes principales desde un punto de vista geométrico

Suponga que $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, donde A es una matriz simétrica invertible de 2×2, y sea c una constante. Es posible demostrar que el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^2$ que satisfacen

$$\mathbf{x}^T A\mathbf{x} = c \tag{3}$$

corresponde a una elipse (o circunferencia), a una hipérbola, a dos rectas que se intersecan, a un solo punto, o quizá no contenga puntos. Si A es una matriz diagonal, la gráfica está en *posición estándar*, como se muestra en la figura 2. Si A no es una matriz diagonal, la gráfica de la ecuación (3) ha dado un giro respecto de la posición estándar, como se muestra en la



$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1, \; a > b > 0$

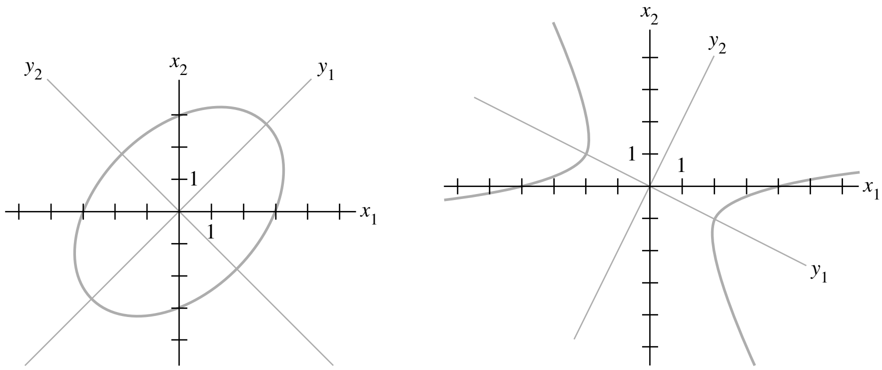
elipse

$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1, \; a > b > 0$

hipérbola

FIGURA 2 Una elipse y una hipérbola en posición estándar.

figura 3. Encontrar los *ejes principales* (determinados por los vectores propios de A) equivale a encontrar un nuevo sistema de coordenadas con respecto al cual la gráfica está en posición estándar.



a) $5x_1^2 - 4x_1x_2 + 5x_2^2 = 48$

b) $x_1^2 - 8x_1x_2 - 5x_2^2 = 16$

FIGURA 3 Una elipse y una hipérbola que *no* están en posición estándar.

La hipérbola en la figura 3*b*) es la gráfica de la ecuación $\mathbf{x}^T A\mathbf{x} = 16$, donde A es la matriz del ejemplo 4. El eje y_1 positivo en la figura 3*b*) está en la dirección de la primera columna de la matriz P del ejemplo 4, y el eje y_2 positivo se encuentra en la dirección de la segunda columna de P.

EJEMPLO 5 La elipse de la figura 3*a*) es la gráfica de la ecuación $5x_1^2 - 4x_1x_2 + 5x_1^2 = 48$. Encuentre un cambio de variable que elimine el producto cruzado de la ecuación.

SOLUCIÓN La matriz de la forma cuadrática es $A = \begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$. Los valores propios de A resultan ser 3 y 7, con los vectores propios unitarios correspondientes

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Sea $P = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. Entonces, P diagonaliza ortogonalmente a A, de manera que el cambio de variable $\mathbf{x} = P\mathbf{y}$ produce la forma cuadrática $\mathbf{y}^T D\mathbf{y} = 3y_1^2 + 7y_2^2$. La figura 3*a*) muestra los nuevos ejes para este cambio de variable. ■

Clasificación de formas cuadráticas

Cuando A es una matriz de $n \times n$, la forma cuadrática $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ es una función de valores reales con dominio $\mathbb{R}^n$. La figura 4 presenta las gráficas de cuatro formas cuadráticas con dominio $\mathbb{R}^2$. Para cada punto $\mathbf{x} = (x_1, x_2)$ en el dominio de una forma cuadrática Q, la gráfica muestra el punto (x_1, x_2, z), donde $z = Q(\mathbf{x})$. Observe que, excepto en $\mathbf{x} = \mathbf{0}$, todos los valores de $Q(\mathbf{x})$ son positivos en la figura 4*a*) y todos son negativos en la figura 4*d*). Las secciones transversales horizontales de las gráficas son elipses en la figura 4*a*) y 4*d*), y son hipérbolas en la figura 4*c*).

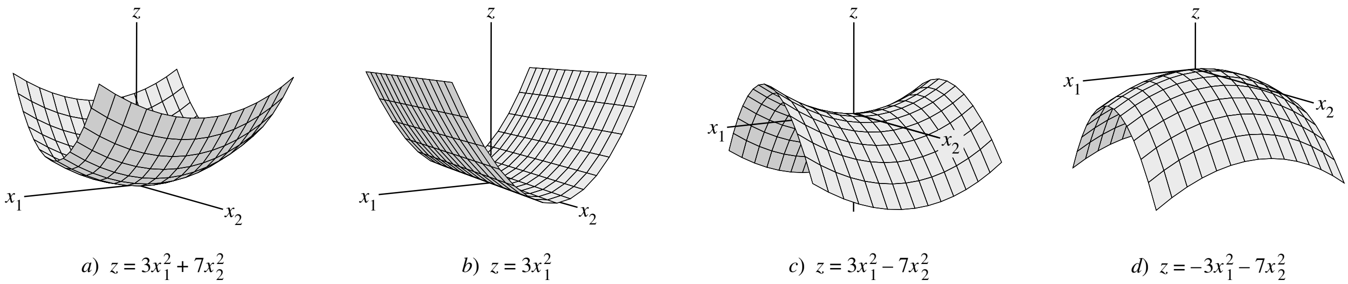


FIGURA 4 Gráficas de formas cuadráticas.

Los ejemplos sencillos de 2×2 de la figura 4 ponen de manifiesto las siguientes definiciones.

DEFINICIÓN

Una forma cuadrática Q es:

a) **positiva definida** si $Q(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$,

b) **negativa definida** si $Q(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$,

c) **indefinida** si $Q(\mathbf{x})$ toma valores positivos y negativos.

También, se dice que Q es **positiva semidefinida** si $Q(\mathbf{x}) \geq 0$ para toda $\mathbf{x}$, y **negativa semidefinida** si $Q(\mathbf{x}) \leq 0$ para toda $\mathbf{x}$. Las formas cuadráticas en los incisos *a*) y *b*) de la figura 4 son semidefinidas positivas, pero la forma en *a*) se describe mejor como positiva definida.

El teorema 5 caracteriza algunas formas cuadráticas en términos de valores propios.

TEOREMA 5

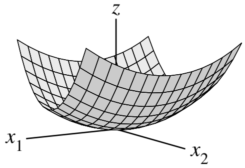
Formas cuadráticas y valores propios

Sea A una matriz simétrica de $n \times n$. Así, una forma cuadrática $\mathbf{x}^T A\mathbf{x}$ es:

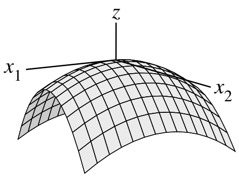
a) positiva definida si y solo si todos los valores propios de A son positivos,

b) negativa definida si y solo si todos los valores propios de A son negativos, o

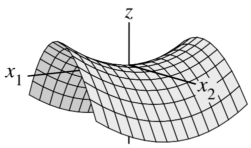
c) indefinida si y solo si A tiene valores propios positivos y negativos.



Positiva definida



Negativa definida



Indefinida

DEMOSTRACIÓN De acuerdo con el teorema de los ejes principales, existe un cambio ortogonal de variable $\mathbf{x} = P\mathbf{y}$ tal que

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{4}$$

donde $\lambda_1,\ldots, \lambda_n$ son los valores propios de A. Como P es invertible, existe una correspondencia uno a uno entre todas las $\mathbf{x}$ y todas las $\mathbf{y}$ diferentes de cero. Por lo tanto, los valores de $Q(\mathbf{x})$ para $\mathbf{x} \neq \mathbf{0}$ coinciden con los valores de la expresión en el miembro derecho de la ecuación (4), la cual, como es evidente, está controlada por los signos de los valores propios $\lambda_1,\ldots, \lambda_n$, en las tres maneras descritas en el teorema. ■

EJEMPLO 6 ¿$Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ es positiva definida?

SOLUCIÓN A causa de todos los signos positivos, esta forma “parece” positiva definida. Pero la matriz de la forma es

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{bmatrix}$$

y los valores propios de A resultan ser 5, 2 y -1. Así, Q es una forma cuadrática indefinida, y no positiva definida. ■

WEB

Con frecuencia, la clasificación de una forma cuadrática se realiza sobre la matriz de la forma. Por consiguiente, una **matriz** A **positiva definida** es una matriz *simétrica* para la cual la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ es positiva definida. Otros términos, tales como **matriz positiva semidefinida**, se definen de manera análoga.

NOTA NUMÉRICA

Una manera rápida de determinar si una matriz simétrica es positiva definida consiste en intentar factorizar A en la forma $A = \mathbb{R}^T R$, donde $\mathbb{R}$ es triangular superior con entradas diagonales positivas. (Un enfoque consiste en utilizar un algoritmo ligeramente modificado para la factorización LU). Tal *factorización de Cholesky* es posible si y solo si A es positiva definida. Véase el ejercicio complementario 7 al final del capítulo 7.

PROBLEMA DE PRÁCTICA

WEB

Describa una matriz A positiva semidefinida en términos de sus valores propios.

7.2 EJERCICIOS

1. Calcule la forma cuadrática $\mathbf{x}^T A\mathbf{x}$, cuando $A = \begin{bmatrix} 5 & 1/3 \\ 1/3 & 1 \end{bmatrix}$ y

a) $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ *b)* $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$ *c)* $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

2. Determine la forma cuadrática $\mathbf{x}^T A\mathbf{x}$, para $A = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ y

a) $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *b)* $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$ *c)* $\mathbf{x} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

3. Encuentre la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^2$.

a) $10x_1^2 - 6x_1x_2 - 3x_2^2$ *b)* $5x_1^2 + 3x_1x_2$

4. Obtenga la matriz de la forma cuadrática. Suponga que $\mathbf{x}$ está en $\mathbb{R}^2$.

a) $20x_1^2 + 15x_1x_2 - 10x_2^2$ *b)* x_1x_2

5. Determine la matriz de la forma cuadrática. Suponga que **x** está en $\mathbb{R}^3$.

a) $8x_1^2 + 7x_2^2 - 3x_3^2 - 6x_1x_2 + 4x_1x_3 - 2x_2x_3$

b) $4x_1x_2 + 6x_1x_3 - 8x_2x_3$

6. Encuentre la matriz de la forma cuadrática. Suponga que **x** está en $\mathbb{R}^3$.

a) $5x_1^2 - x_2^2 + 7x_3^2 + 5x_1x_2 - 3x_1x_3$

b) $x_3^2 - 4x_1x_2 + 4x_2x_3$

7. Realice un cambio de variable, $\mathbf{x} = P\mathbf{y}$, que transforme la forma cuadrática $x_1^2 + 10x_1x_2 + x_2^2$ en una forma cuadrática sin producto cruzado. Determine P y la nueva forma cuadrática.

8. Sea A la matriz de la forma cuadrática

$9x_1^2 + 7x_2^2 + 11x_3^2 - 8x_1x_2 + 8x_1x_3$

Es posible demostrar que los valores propios de A son 3, 9 y 15. Encuentre una matriz ortogonal P tal que el cambio de variable $\mathbf{x} = P\mathbf{y}$ transforme $\mathbf{x}^TA\mathbf{x}$ en una forma cuadrática sin productos cruzados. Determine P y la nueva forma cuadrática.

En los ejercicios 9 a 18, clasifique las formas cuadráticas. Después realice un cambio de variable, $\mathbf{x} = P\mathbf{y}$, que convierta la forma cuadrática en una que no incluya productos cruzados. Escriba la nueva forma cuadrática. Construya P utilizando los métodos de la sección 7.1.

9. $3x_1^2 - 4x_1x_2 + 6x_2^2$

10. $9x_1^2 - 8x_1x_2 + 3x_2^2$

11. $2x_1^2 + 10x_1x_2 + 2x_2^2$

12. $-5x_1^2 + 4x_1x_2 - 2x_2^2$

13. $x_1^2 - 6x_1x_2 + 9x_2^2$

14. $8x_1^2 + 6x_1x_2$

15. [**M**] $-2x_1^2 - 6x_2^2 - 9x_3^2 - 9x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 6x_3x_4$

16. [**M**] $4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 3x_1x_2 + 3x_3x_4 - 4x_1x_4 + 4x_2x_3$

17. [**M**] $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 9x_1x_2 - 12x_1x_4 + 12x_2x_3 + 9x_3x_4$

18. [**M**] $11x_1^2 - x_2^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - 2x_3x_4$

19. ¿Cuál es el valor más grande posible de la forma cuadrática $5x_1^2 + 8x_2^2$ si $\mathbf{x} = (x_1, x_2)$ y $\mathbf{x}^T\mathbf{x}=1$, es decir, si $x_1^2 + x_2^2 = 1$? (Intente algunos ejemplos de **x**).

20. ¿Cuál es el valor más grande de la forma cuadrática $5x_1^2 - 3x_2^2$ si $\mathbf{x}^T\mathbf{x} = 1$?

En los ejercicios 21 y 22, las matrices son de $n \times n$ y los vectores están en $\mathbb{R}^n$. Marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) La matriz de una forma cuadrática es una matriz simétrica.

b) Una forma cuadrática no tiene productos cruzados si y solo si la matriz de la forma cuadrática es una matriz diagonal.

c) Los ejes principales de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$ son vectores propios de A.

d) Una forma cuadrática positiva definida Q satisface $Q(\mathbf{x}) > 0$ para toda **x** en $\mathbb{R}^n$.

e) Si todos los valores propios de una matriz simétrica A son positivos, entonces la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es positiva definida.

f) Una factorización de Cholesky de una matriz simétrica A tiene la forma $A = R^TR$, para una matriz triangular superior R con entradas diagonales positivas.

22. *a*) La expresión $\|\mathbf{x}\|^2$ es una forma cuadrática.

b) Si A es simétrica y P es una matriz ortogonal, entonces el cambio de variable $\mathbf{x} = P\mathbf{y}$ convierte $\mathbf{x}^TA\mathbf{x}$ en una forma cuadrática sin productos cruzados.

c) Si A es una matriz simétrica de 2×2, entonces el conjunto de **x** tal que $\mathbf{x}^TA\mathbf{x} = c$ (para una constante c) corresponde a un círculo, una elipse o una hipérbola.

d) Una forma cuadrática indefinida es positiva semidefinida o negativa semidefinida.

e) Si A es simétrica y la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ solo tiene valores negativos para $\mathbf{x} \neq \mathbf{0}$, entonces todos los valores propios de A son negativos.

Los ejercicios 23 y 24 muestran cómo clasificar una forma cuadrática $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, cuando $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ y $\det A \neq 0$, sin obtener los valores propios de A.

23. Si λ_1 y λ_2 son los valores propios de A, entonces el polinomio característico de A se puede escribir de dos maneras: $\det(A - \lambda I)$ y $(\lambda - \lambda_1)(\lambda - \lambda_2)$. Con base en este hecho, demuestre que $\lambda_1 + \lambda_2 = a + d$ (las entradas diagonales de A) y $\lambda_1\lambda_2 = \det A$.

24. Compruebe los siguientes enunciados.

a) Q es positiva definida si $\det A > 0$ y $a > 0$.

b) Q es negativa definida si $\det A > 0$ y $a < 0$.

c) Q es indefinida si $\det A < 0$.

25. Demuestre que si B es de $m \times n$, entonces B^TB es positiva semidefinida; y si B es de $n \times n$ e invertible, entonces B^TB es positiva definida.

26. Demuestre que si una matriz A de $n \times n$ es positiva definida, entonces existe una matriz B positiva definida tal que $A = B^TB$. [*Sugerencia:* Escriba $A = PDP^T$, con $P^T = P^{-1}$. Construya una matriz C diagonal tal que $D = C^TC$, y sea $B = PCP^T$. Demuestre que B funciona].

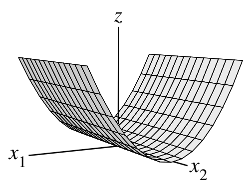
27. Sean A y B matrices simétricas de $n \times n$ cuyos valores propios son todos positivos. Demuestre que todos los valores propios de $A + B$ son positivos. [*Sugerencia:* Considere formas cuadráticas].

28. Sea A una matriz simétrica e invertible de $n \times n$. Demuestre que si la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es positiva definida, entonces también lo es la forma cuadrática $\mathbf{x}^TA^{-1}\mathbf{x}$. [*Sugerencia:* Considere valores propios].

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Efectúe un cambio de variable ortogonal $\mathbf{x} = P\mathbf{y}$, y escriba

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$



Positiva semidefinida

como en la ecuación (4). Si un valor propio —por ejemplo, λ_i— fuera negativo, entonces $\mathbf{x}^T A\mathbf{x}$ sería negativo para la $\mathbf{x}$ correspondiente a $\mathbf{y} = \mathbf{e}_i$ (la i-ésima columna de I_n). Así, todos los valores propios de una forma cuadrática positiva semidefinida deben ser no negativos. A la inversa, si los valores propios son positivos, la expansión anterior indica que $\mathbf{x}^T A\mathbf{x}$ debe ser positiva semidefinida.

7.3 OPTIMIZACIÓN RESTRINGIDA

Es frecuente que ingenieros, economistas, científicos y matemáticos necesiten encontrar el valor máximo o mínimo de una forma cuadrática $Q(\mathbf{x})$ para $\mathbf{x}$ en algún conjunto específico. Por lo común, el problema se puede adaptar de tal manera que $\mathbf{x}$ varíe sobre un conjunto de vectores unitarios. Este *problema de optimización restringida* tiene una solución interesante y elegante. El ejemplo 6 y el análisis de la sección 7.5 ilustrarán cómo se presentan en la práctica este tipo de problemas.

El requisito de que un vector $\mathbf{x}$ en $\mathbb{R}^n$ sea unitario se puede establecer en varias formas equivalentes:

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

y

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1 \tag{1}$$

La versión ampliada de la ecuación (1) de $\mathbf{x}^T\mathbf{x} = 1$ se utiliza comúnmente en las aplicaciones.

Cuando una forma cuadrática Q no tiene productos cruzados, es fácil encontrar el máximo y el mínimo de $Q(\mathbf{x})$ para $\mathbf{x}^T\mathbf{x} = 1$.

EJEMPLO 1 Obtenga los valores máximo y mínimo de $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ con la restricción $\mathbf{x}^T\mathbf{x} = 1$.

SOLUCIÓN Como x_2^2 y x_3^2 son no negativos, observe que

$$4x_2^2 \leq 9x_2^2 \quad \text{y} \quad 3x_3^2 \leq 9x_3^2$$

y por lo tanto,

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\leq 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

siempre que $x_1^2 + x_2^2 + x_3^2 = 1$. Así, el valor máximo de $Q(\mathbf{x})$ no puede exceder 9 cuando $\mathbf{x}$ es un vector unitario. Además, $Q(\mathbf{x}) = 9$ cuando $\mathbf{x} = (1, 0, 0)$. Por consiguiente, 9 es el valor máximo de $Q(\mathbf{x})$ para $\mathbf{x}^T\mathbf{x} = 1$.

Para encontrar el valor mínimo de $Q(\mathbf{x})$, observe que

$$9x_1^2 \geq 3x_1^2, \qquad 4x_2^2 \geq 3x_2^2$$

y en consecuencia,

$$Q(\mathbf{x}) \geq 3x_1^2 + 3x_2^2 + 3x_3^2 = 3(x_1^2 + x_2^2 + x_3^2) = 3$$

siempre que $x_1^2 + x_2^2 + x_3^2 = 1$. Además, $Q(\mathbf{x}) = 3$ cuando $x_1 = 0$, $x_2 = 0$, y $x_3 = 1$. Así, 3 es el valor mínimo de $Q(\mathbf{x})$ cuando $\mathbf{x}^T\mathbf{x} = 1$. ■

En el ejemplo 1 es fácil ver que la matriz de la forma cuadrática Q tiene valores propios 9, 4 y 3, y que los valores propios mayor y menor son iguales al máximo y mínimo (restringidos) de $Q(\mathbf{x})$, respectivamente. Se verá que lo mismo es válido para cualquier forma cuadrática.

EJEMPLO 2 Sean $A = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$, y $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ para $\mathbf{x}$ en $\mathbb{R}^2$. La figura 1 muestra la gráfica de Q. La figura 2 solo muestra la parte de la gráfica dentro de un cilindro; la intersección del cilindro con la superficie es el conjunto de puntos (x_1, x_2, z) tales que $z = Q(x_1, x_2)$ y $x_1^2 + x_2^2 = 1$. Las "alturas" de esos puntos son los valores restringidos de $Q(\mathbf{x})$. Geométricamente, el problema de optimización restringido es localizar los puntos más alto y más bajo en la curva de intersección.

Los dos puntos más altos sobre la curva están 7 unidades arriba del plano x_1x_2, en $x_1 = 0$ y $x_2 = \pm 1$. Esos puntos corresponden al valor propio 7 de A y a los vectores propios $\mathbf{x} = (0, 1)$ y $-\mathbf{x} = (0, -1)$. De manera similar, los dos puntos más bajos sobre la curva están 3 unidades arriba del plano x_1x_2, y corresponden al valor propio 3 y a los vectores propios $(1, 0)$ y $(-1, 0)$. ■

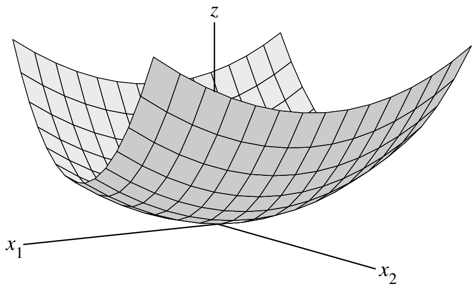


FIGURA 1 $z = 3x_1^2 + 7x_2^2$.

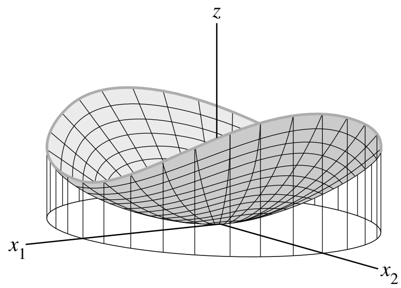


FIGURA 2 La intersección de $z = 3x_1^2 + 7x_2^2$ y el cilindro $x_1^2 + x_2^2 = 1$.

En la figura 2 cada punto de la curva de intersección tiene una coordenada z entre 3 y 7, y para cualquier número t entre 3 y 7, existe un vector unitario $\mathbf{x}$ tal que $Q(\mathbf{x}) = t$. En otras palabras, el conjunto de todos los posibles valores de $\mathbf{x}^T A\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es el intervalo cerrado $3 \le t \le 7$.

Es posible demostrar que para cualquier matriz simétrica A, el conjunto de todos los posibles valores de $\mathbf{x}^T A\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es un intervalo cerrado sobre el eje real. (Véase el ejercicio 13). Denote los puntos extremos izquierdo y derecho de este intervalo con m y M, respectivamente. Es decir, sean

$$m = \text{mín}\,\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}, \quad M = \text{máx}\,\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\} \tag{2}$$

En el ejercicio 12 se le pide demostrar que si λ es un valor propio de A, entonces $m \le \lambda \le M$. El siguiente teorema dice que m y M son en sí mismos valores propios de A, justo como en el ejemplo 2.[1]

TEOREMA 6

Sea A una matriz simétrica, y m y M se definen como en la ecuación (2). Entonces M es el valor propio más grande λ_1 de A y m es el valor propio más pequeño de A. El valor de $\mathbf{x}^T A\mathbf{x}$ es M cuando $\mathbf{x}$ es un vector propio unitario $\mathbf{u}_1$ correspondiente a M. El valor de $\mathbf{x}^T A\mathbf{x}$ es m cuando $\mathbf{x}$ es un vector propio unitario correspondiente a m.

[1] Los términos *mínimo* y *máximo* en la ecuación (2), y *menor* y *mayor* en el teorema, se refieren al ordenamiento natural de los números reales, no a magnitudes.

DEMOSTRACIÓN A se diagonaliza ortogonalmente como PDP^{-1}. Se sabe que

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} \quad \text{cuando } \mathbf{x} = P\mathbf{y} \tag{3}$$

También,

$$\|\mathbf{x}\| = \|P\mathbf{y}\| = \|\mathbf{y}\| \quad \text{para toda } \mathbf{y}$$

porque $P^TP = I$ y $\|P\mathbf{y}\|^2 = (P\mathbf{y})^T(P\mathbf{y}) = \mathbf{y}^TP^TP\mathbf{y} = \mathbf{y}^T\mathbf{y} = \|\mathbf{y}\|^2$. En particular, $\|\mathbf{y}\| = 1$ si y solo si $\|\mathbf{x}\| = 1$. Así, $\mathbf{x}^TA\mathbf{x}$ y $\mathbf{y}^TD\mathbf{y}$ asumen el mismo conjunto de valores conforme $\mathbf{x}$ y $\mathbf{y}$ varían sobre el conjunto de todos los vectores unitarios.

Para simplificar la notación, suponga que A es una matriz de 3×3 con valores propios $a \geq b \geq c$. Arregle las columnas (los vectores propios) de P de manera que $P = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3]$ y

$$D = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

Dado cualquier vector unitario $\mathbf{y}$ en $\mathbb{R}^3$ con coordenadas y_1, y_2, y_3, observe que

$$\begin{aligned} ay_1^2 &= ay_1^2 \\ by_2^2 &\leq ay_2^2 \\ cy_3^2 &\leq ay_3^2 \end{aligned}$$

y se obtienen estas desigualdades:

$$\begin{aligned} \mathbf{y}^TD\mathbf{y} &= ay_1^2 + by_2^2 + cy_3^2 \\ &\leq ay_1^2 + ay_2^2 + ay_3^2 \\ &= a(y_1^2 + y_2^2 + y_3^2) \\ &= a\|\mathbf{y}\|^2 = a \end{aligned}$$

De esta manera, $M \leq a$, por definición de M. Sin embargo, $\mathbf{y}^TD\mathbf{y} = a$ cuando $\mathbf{y} = \mathbf{e}_1 = (1, 0, 0)$, así que en efecto $M = a$. De acuerdo con la ecuación (3), la $\mathbf{x}$ que corresponde a $\mathbf{y} = \mathbf{e}_1$ es el vector propio $\mathbf{u}_1$ de A, porque

$$\mathbf{x} = P\mathbf{e}_1 = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{u}_1$$

Por lo tanto, $M = a = \mathbf{e}_1^T D\mathbf{e}_1 = \mathbf{u}_1^T A\mathbf{u}_1$, lo que demuestra el enunciado sobre M. Un argumento similar demuestra que m es el valor propio más pequeño, c, y ese valor de $\mathbf{x}^TA\mathbf{x}$ se obtiene cuando $\mathbf{x} = P\mathbf{e}_3 = \mathbf{u}_3$. ■

EJEMPLO 3 Sea $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}$. Encuentre el valor máximo de la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ sujeta a la restricción $\mathbf{x}^T\mathbf{x} = 1$, y obtenga un vector unitario en el que se alcance dicho valor máximo.

SOLUCIÓN Según el teorema 6, el valor máximo deseado es el valor propio más grande de A. La ecuación característica resulta ser

$$0 = -\lambda^3 + 10\lambda^2 - 27\lambda + 18 = -(\lambda - 6)(\lambda - 3)(\lambda - 1)$$

El valor propio más grande es 6.

El máximo restringido de $\mathbf{x}^TA\mathbf{x}$ se obtiene cuando $\mathbf{x}$ es un vector propio unitario para $\lambda = 6$. Resuelva $(A - 6I)\mathbf{x} = \mathbf{0}$ y encuentre un vector propio $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Establezca $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$. ■

En el teorema 7 y en aplicaciones posteriores, los valores de $\mathbf{x}^T A\mathbf{x}$ se calculan con restricciones adicionales sobre el vector unitario $\mathbf{x}$.

TEOREMA 7

Sean A, λ_1 y $\mathbf{u}_1$ como en el teorema 6. Así, el valor máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las restricciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0$$

es el segundo valor propio más grande, λ_2, y este máximo se logra cuando $\mathbf{x}$ es un vector propio $\mathbf{u}_2$ correspondiente a λ_2.

El teorema 7 se demuestra mediante un argumento similar al anterior donde el teorema se reduce al caso en que la matriz de la forma cuadrática es diagonal. El siguiente ejemplo da una idea de la demostración para el caso de una matriz diagonal.

EJEMPLO 4 Encuentre el valor máximo de $9x_1^2 + 4x_2^2 + 3x_3^2$ con las restricciones $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^T\mathbf{u}_1 = 0$, donde $\mathbf{u}_1 = (1, 0, 0)$. Observe que $\mathbf{u}_1$ es un vector propio unitario correspondiente al valor propio más grande $\lambda = 9$ de la matriz de la forma cuadrática.

SOLUCIÓN Si las coordenadas de $\mathbf{x}$ son x_1, x_2, x_3, entonces la restricción $\mathbf{x}^T\mathbf{u}_1 = 0$ significa simplemente que $x_1 = 0$. Para tal vector unitario, $x_2^2 + x_3^2 = 1$, y

$$\begin{aligned} 9x_1^2 + 4x_2^2 + 3x_3^2 &= 4x_2^2 + 3x_3^2 \\ &\leq 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4 \end{aligned}$$

Así, el máximo restringido de la forma cuadrática no excede 4. Y este valor se logra para $\mathbf{x} = (0, 1, 0)$, que es un vector propio para el segundo valor propio más grande de la matriz de la forma cuadrática. ■

EJEMPLO 5 Sean A la matriz del ejemplo 3 y $\mathbf{u}_1$ un vector propio unitario correspondiente al valor propio más grande de A. Encuentre el valor máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las condiciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0 \tag{4}$$

SOLUCIÓN A partir del ejemplo 3, el segundo valor propio más grande de A es $\lambda = 3$. Resuelva $(A - 3I)\mathbf{x} = \mathbf{0}$ para encontrar un vector propio, y normalícelo para obtener

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

El vector $\mathbf{u}_2$ es automáticamente ortogonal a $\mathbf{u}_1$ porque los vectores corresponden a diferentes valores propios. Así, el máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las restricciones en la ecuación (4) es 3, que se obtiene cuando $\mathbf{x} = \mathbf{u}_2$. ■

El siguiente teorema generaliza el teorema 7 y, junto con el teorema 6, ofrece una útil caracterización de *todos* los valores propios de A. Se omite la demostración.

TEOREMA 8 Sea A una matriz simétrica de $n \times n$ con una diagonalización ortogonal $A = PDP^{-1}$, donde las entradas sobre la diagonal en D están ordenadas de tal manera que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ y las columnas de P son los vectores propios unitarios correspondientes $\mathbf{u}_1, \dots, \mathbf{u}_n$. Entonces, para $k = 2, \dots, n$, el valor máximo de $\mathbf{x}^T A\mathbf{x}$ sujeto a las restricciones

$$\mathbf{x}^T\mathbf{x} = 1, \quad \mathbf{x}^T\mathbf{u}_1 = 0, \quad \dots, \quad \mathbf{x}^T\mathbf{u}_{k-1} = 0$$

es el valor propio λ_k, y este máximo se logra en $\mathbf{x} = \mathbf{u}_k$.

El teorema 8 será útil en las secciones 7.4 y 7.5. La siguiente aplicación solo requiere del teorema 6.

EJEMPLO 6 El gobierno de un condado planea reparar x cientos de millas de caminos públicos y puentes, y mejorar y cientos de acres de parques y áreas recreativas durante el próximo año. El condado debe decidir cómo distribuir sus recursos (fondos, equipo, mano de obra, etc.) entre esos dos proyectos. Si es más eficaz en términos de costos trabajar simultáneamente en ambos proyectos en lugar de trabajar en uno solo, entonces x y y podrían satisfacer una *restricción* del tipo

$$4x^2 + 9y^2 \leq 36$$

Véase la figura 3. Cada punto (x, y) en el *conjunto factible* sombreado representa una posible obra pública programada para el año. Los puntos sobre la curva de restricción, $4x^2 + 9y^2 = 36$, utilizan las cantidades máximas de recursos disponibles.

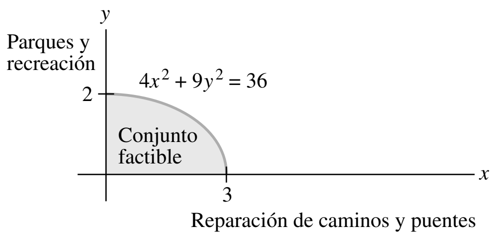


FIGURA 3 Programas de trabajos públicos.

Al elegir su programa de obras públicas, el condado quiere considerar las opiniones de sus residentes. Para medir el valor, o *utilidad*, que los residentes de un lugar asignarían a los diversos programas de trabajo (x, y), algunas veces los economistas emplean una función de la forma

$$q(x, y) = xy$$

El conjunto de puntos (x, y) en el cual $q(x, y)$ es una constante se denomina *curva de indiferencia*. En la figura 4 se muestran tres de estas curvas. Los puntos sobre una curva de indiferencia corresponden a las alternativas que los residentes del condado, como grupo, encontrarían igualmente valiosas.[2] Encuentre el programa de obras públicas que maximice la función utilidad q.

SOLUCIÓN La ecuación de restricción $4x^2 + 9y^2 = 36$ no describe un conjunto de vectores unitarios, pero un cambio de variable puede resolver el problema. Rescriba la restricción en la forma

$$\left(\frac{x}{3}\right)^2 + \left(\frac{y}{2}\right)^2 = 1$$

[2] Las curvas de indiferencia se analizan en Michael D. Intriligator, Ronald G. Bodkin y Cheng Hsiao, *Econometric Models, Techniques, and Applications* (Upper Saddle River, NJ: Prentice-Hall, 1996).

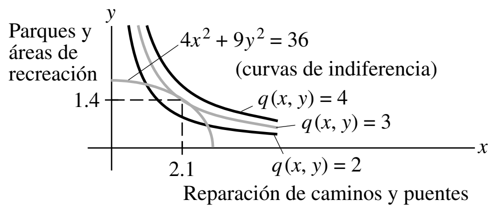


FIGURA 4 El programa de obras públicas óptimo es (2.1, 1.4).

y defina

$$x_1 = \frac{x}{3}, \quad x_2 = \frac{y}{2}, \quad \text{es decir,} \quad x = 3x_1 \quad \text{y} \quad y = 2x_2$$

Entonces, la ecuación de restricción se convierte en

$$x_1^2 + x_2^2 = 1$$

y la función utilidad se convierte en $q(3x_1, 2x_2) = (3x_1)(2x_2) = 6x_1x_2$. Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Entonces, el problema es maximizar $Q(\mathbf{x}) = 6x_1x_2$ sujeto a $\mathbf{x}^T\mathbf{x} = 1$. Observe que $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde

$$A = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}$$

Los valores propios de A son ± 3, con vectores propios $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ para $\lambda = 3$ y $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ para $\lambda = -3$. Así, el valor máximo de $Q(\mathbf{x}) = q(x_1, x_2)$ es 3, que se alcanza cuando $x_1 = 1/\sqrt{2}$ y $x_2 = 1/\sqrt{2}$.

En términos de las variables originales, el programa de trabajos públicos óptimo es $x = 3x_1 = 3/\sqrt{2} \approx 2.1$ cientos de millas de caminos y puentes, y $y = 2x_2 = \sqrt{2} \approx 1.4$ cientos de acres de parques y áreas de recreación. El programa de obras públicas óptimo es justo el punto donde se encuentran la curva de restricción y la curva de indiferencia $q(x, y) = 3$. Los puntos (x, y) con la utilidad más alta están sobre las curvas de indiferencia que no tocan la curva de restricción. Véase la figura 4. ■

PROBLEMAS DE PRÁCTICA

1. Sea $Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Encuentre un cambio de variable que transforme Q en una forma cuadrática sin el producto cruzado, y determine la nueva forma cuadrática.
2. Con Q como en el problema 1, encuentre el valor máximo de $Q(\mathbf{x})$ con la restricción $\mathbf{x}^T\mathbf{x} = 1$, y determine un vector unitario en el que se logre el máximo.

7.3 EJERCICIOS

En los ejercicios 1 y 2, encuentre el cambio de variable $\mathbf{x} = P\mathbf{y}$ que transforme la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ en $\mathbf{y}^TD\mathbf{y}$ como se muestra.

1. $5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3 = 9y_1^2 + 6y_2^2 + 3y_3^2$
2. $3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 = 5y_1^2 + 2y_2^2$

   [*Sugerencia:* $\mathbf{x}$ y $\mathbf{y}$ deben tener el mismo número de coordenadas, de manera que la forma cuadrática que se muestra aquí debe tener un coeficiente cero para y_3^2].

En los ejercicios 3 a 6, encuentre: *a*) el valor máximo de $Q(\mathbf{x})$ con la restricción $\mathbf{x}^T\mathbf{x} = 1$, *b*) un vector unitario $\mathbf{u}$ donde se logre este máximo, y *c*) el máximo de $Q(\mathbf{x})$ con las condiciones $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^T\mathbf{u} = 0$.

3. $Q(\mathbf{x}) = 5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3$

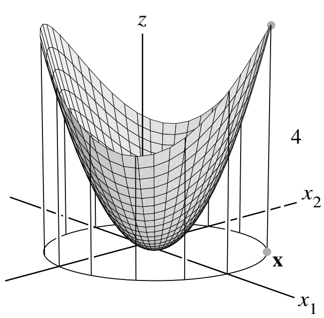
   (Véase el ejercicio 1).

4. $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$ (Véase el ejercicio 2).
5. $Q(\mathbf{x}) = 5x_1^2 + 5x_2^2 - 4x_1x_2$
6. $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 + 3x_1x_2$
7. Sea $Q(\mathbf{x}) = -2x_1^2 - x_2^2 + 4x_1x_2 + 4x_2x_3$. Determine un vector unitario $\mathbf{x}$ en $\mathbb{R}^3$ en el cual $Q(\mathbf{x})$ sea máximo, con la restricción $\mathbf{x}^T\mathbf{x} = 1$. [*Sugerencia:* Los valores propios de la matriz de la forma cuadrática Q son 2, -1 y -4].
8. Sea $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$. Obtenga un vector unitario $\mathbf{x}$ en $\mathbb{R}^3$ en el que $Q(\mathbf{x})$ sea máximo, con la restricción $\mathbf{x}^T\mathbf{x} = 1$. [*Sugerencia:* Los valores propios de la matriz de la forma cuadrática Q son 9 y -3].
9. Encuentre el valor máximo de $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 - 2x_1x_2$, con la restricción $x_1^2 + x_2^2 = 1$. (No continúe para encontrar un vector donde se alcance el máximo).
10. Obtenga el valor máximo de $Q(\mathbf{x}) = -3x_1^2 + 5x_2^2 - 2x_1x_2$, con la restricción $x_1^2 + x_2^2 = 1$. (No continúe para encontrar un vector donde se alcance el máximo).
11. Suponga que $\mathbf{x}$ es un vector propio unitario de una matriz A correspondiente a un valor propio 3. ¿Cuál es el valor de $\mathbf{x}^TA\mathbf{x}$?
12. Sea λ cualquier valor propio de una matriz simétrica A. Justifique el enunciado hecho en esta sección de que $m \le \lambda \le M$, donde m y M se definen como en la ecuación (2). [*Sugerencia:* Encuentre una $\mathbf{x}$ tal que $\lambda = \mathbf{x}^TA\mathbf{x}$].
13. Considere que A es una matriz simétrica de $n \times n$, y que M y m representan los valores máximo y mínimo de la forma cuadrática $\mathbf{x}^TA\mathbf{x}$; denote los vectores propios unitarios correspondientes como $\mathbf{u}_1$ y $\mathbf{u}_n$. Los siguientes cálculos demuestran que dado cualquier número t entre M y m, existe un vector unitario $\mathbf{x}$ tal que $t = \mathbf{x}^TA\mathbf{x}$. Compruebe que $t = (1 - \alpha)m + \alpha M$ para algún número α entre 0 y 1. Después deje que $\mathbf{x} = \sqrt{1-\alpha}\mathbf{u}_n + \sqrt{\alpha}\mathbf{u}_1$, y demuestre que $\mathbf{x}^T\mathbf{x} = 1$ y $\mathbf{x}^TA\mathbf{x} = t$.

[M] En los ejercicios 14 a 17, siga las instrucciones para los ejercicios 3 a 6.

14. $x_1x_2 + 3x_1x_3 + 30x_1x_4 + 30x_2x_3 + 3x_2x_4 + x_3x_4$
15. $3x_1x_2 + 5x_1x_3 + 7x_1x_4 + 7x_2x_3 + 5x_2x_4 + 3x_3x_4$
16. $4x_1^2 - 6x_1x_2 - 10x_1x_3 - 10x_1x_4 - 6x_2x_3 - 6x_2x_4 - 2x_3x_4$
17. $-6x_1^2 - 10x_2^2 - 13x_3^2 - 13x_4^2 - 4x_1x_2 - 4x_1x_3 - 4x_1x_4 + 6x_3x_4$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA



El valor máximo de $Q(\mathbf{x})$ sujeto a $\mathbf{x}^T\mathbf{x} = 1$ es 4.

1. La matriz de la forma cuadrática es $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Es fácil encontrar los valores propios, 4 y 2, y los vectores propios unitarios correspondientes, $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ y $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. Así que el cambio deseado de variable es $\mathbf{x} = P\mathbf{y}$, donde $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (Aquí un error común es olvidar normalizar los vectores propios). La nueva forma cuadrática es $\mathbf{y}^TD\mathbf{y} = 4y_1^2 + 2y_2^2$.
2. El máximo de $Q(\mathbf{x})$ para un vector unitario $\mathbf{x}$ es 4, y el máximo se logra en el vector propio unitario $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. [Una respuesta incorrecta común es $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Este vector maximiza la forma cuadrática $\mathbf{y}^TD\mathbf{y}$ en vez de $Q(\mathbf{x})$].

7.4 DESCOMPOSICIÓN EN VALORES SINGULARES

Los teoremas de diagonalización de las secciones 5.3 y 7.1 intervienen en muchas aplicaciones interesantes. Por desgracia, como se sabe, no todas las matrices se pueden factorizar como $A = PDP^{-1}$ con D diagonal. Sin embargo, ¡una factorización $A = QDP^{-1}$ *es* posible para cualquier matriz A de $m \times n$! Una factorización especial de este tipo, llamada *descomposición en valores singulares*, es una de las más útiles factorizaciones matriciales en álgebra lineal aplicada.

La descomposición en valores singulares se basa en la siguiente propiedad de la diagonalización ordinaria que se puede imitar para matrices rectangulares: los valores absolutos de los valores propios de una matriz simétrica A miden las cantidades que A estira o comprime ciertos vectores (los vectores propios). Si $A\mathbf{x} = \lambda\mathbf{x}$ y $\|\mathbf{x}\| = 1$, entonces

$$\|A\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\| = |\lambda| \tag{1}$$

Si λ_1 es el valor propio con la mayor magnitud, entonces un vector propio unitario $\mathbf{v}_1$ correspondiente identifica una dirección en la cual el efecto de estiramiento de A es máximo. Es decir, la longitud de $A\mathbf{x}$ se maximiza cuando $\mathbf{x} = \mathbf{v}_1$, y $\|A\mathbf{v}_1\| = |\lambda_1|$, por la ecuación (1). Esta descripción de $\mathbf{v}_1$ y $|\lambda_1|$ tiene un análogo para matrices rectangulares que conducirán a la descomposición en valores singulares.

EJEMPLO 1 Si $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$, entonces la transformación lineal $\mathbf{x} \mapsto A\mathbf{x}$ mapea la esfera unitaria $\{\mathbf{x} : \|\mathbf{x}\| = 1\}$ en $\mathbb{R}^3$ sobre una elipse en $\mathbb{R}^2$, como se muestra en la figura 1. Encuentre un vector unitario $\mathbf{x}$ en el que la longitud $\|A\mathbf{x}\|$ se maximiza, y calcule esa longitud máxima.

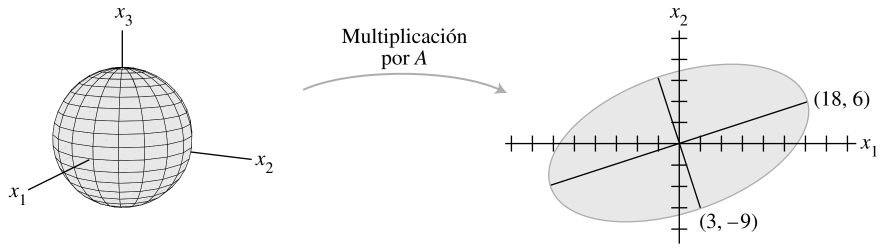


FIGURA 1 Una transformación de $\mathbb{R}^3$ a $\mathbb{R}^2$.

SOLUCIÓN La cantidad $\|A\mathbf{x}\|^2$ se maximiza con la misma $\mathbf{x}$ que maximiza $\|A\mathbf{x}\|$, y $\|A\mathbf{x}\|^2$ es más fácil de estudiar. Observe que

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^T A^T A\mathbf{x} = \mathbf{x}^T(A^T A)\mathbf{x}$$

Además, A^TA es una matriz simétrica, ya que $(A^TA)^T = A^TA^{TT} = A^TA$. Así que ahora el problema es maximizar la forma cuadrática $\mathbf{x}^T(A^TA)\mathbf{x}$ sujeta a la restricción $\|\mathbf{x}\| = 1$. Por el teorema 6 de la sección 7.3, el valor máximo es el valor propio más grande λ_1 de A^TA. Además, el valor máximo se alcanza en un vector propio unitario de A^TA correspondiente a λ_1.

Para la matriz A en este ejemplo,

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix}\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

Los valores propios de A^TA son $\lambda_1 = 360$, $\lambda_2 = 90$ y $\lambda_3 = 0$. Los vectores propios unitarios correspondientes son, respectivamente,

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

El valor máximo de $\|A\mathbf{x}\|^2$ es 360, que se alcanza cuando $\mathbf{x}$ es el vector unitario $\mathbf{v}_1$. El vector $A\mathbf{v}_1$ es el punto sobre la elipse de la figura 1 más alejado del origen, a saber,

$$A\mathbf{v}_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$$

Para $\|\mathbf{x}\| = 1$, el valor máximo de $\|A\mathbf{x}\|$ es $\|A\mathbf{v}_1\| = \sqrt{360} = 6\sqrt{10}$. ■

El ejemplo 1 sugiere que el efecto de A sobre la esfera unitaria en $\mathbb{R}^3$ está relacionado con la forma cuadrática $\mathbf{x}^T(A^TA)\mathbf{x}$. De hecho, el comportamiento geométrico general de la transformación $\mathbf{x} \mapsto A\mathbf{x}$ se refleja en esta forma cuadrática, como se verá más adelante.

Los valores singulares de una matriz de $m \times n$

Sea A una matriz de $m \times n$. A^TA es simétrica y se puede diagonalizar ortogonalmente. Sea $\{\mathbf{v}_1,\dots, \mathbf{v}_n\}$ una base ortonormal para $\mathbb{R}^n$ que consiste en vectores propios de A^TA, y sean $\lambda_1,\dots, \lambda_n$ los valores propios asociados de A^TA. Por lo tanto, para $1 \le i \le n$,

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i \\ &= \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) && \text{Puesto que } \mathbf{v}_i \text{ es un vector propio de } A^TA \\ &= \lambda_i && \text{Porque } \mathbf{v}_i \text{ es un vector unitario} \end{aligned} \tag{2}$$

De esta forma, todos los valores propios de A^TA son no negativos. Al volver a numerar, si es necesario, se puede suponer que los valores propios están arreglados de manera que

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n \ge 0$$

Los **valores singulares** de A son las raíces cuadradas de los valores propios de A^TA, que se denotan con $\sigma_1,\dots, \sigma_n$, y se arreglan en orden decreciente. Es decir, $\sigma_i = \sqrt{\lambda_i}$ para $1 \le i \le n$. De acuerdo con la ecuación (2), *los valores singulares de A son las longitudes de los vectores* $A\mathbf{v}_1,\dots, A\mathbf{v}_n$.

EJEMPLO 2 Sea A la matriz del ejemplo 1. Como los valores propios de A^TA son 360, 90 y 0, entonces los valores singulares de A son

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = 0$$

A partir del ejemplo 1, el primer valor singular de A es el máximo de $\|A\mathbf{x}\|$ sobre todos los vectores unitarios, y se alcanza el máximo en el vector propio unitario $\mathbf{v}_1$. El teorema 7 de la sección 7.3 indica que el segundo valor singular de A es el máximo de $\|A\mathbf{x}\|$ sobre todos los vectores unitarios que sean *ortogonales* a $\mathbf{v}_1$, y este máximo se logra en el segundo vector propio unitario, $\mathbf{v}_2$ (ejercicio 22). Para $\mathbf{v}_2$ del ejemplo 1,

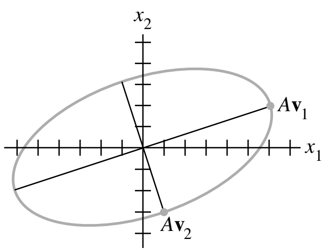


FIGURA 2

$$A\mathbf{v}_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$$

Este punto está sobre el eje menor de la elipse de la figura 1, así como $A\mathbf{v}_1$ está sobre el eje mayor. (Véase la figura 2). Los primeros dos valores singulares de A son las longitudes de los semiejes mayor y menor de la elipse. ■

En la figura 2, no es casual que $A\mathbf{v}_1$ y $A\mathbf{v}_2$ sean ortogonales, como muestra el siguiente teorema.

TEOREMA 9

Suponga que $\{\mathbf{v}_1,\dots, \mathbf{v}_n\}$ es una base ortonormal de $\mathbb{R}^n$ que consiste en vectores propios de A^TA, arreglados de tal forma que los valores propios correspondientes de A^TA satisfacen $\lambda_1 \ge \cdots \ge \lambda_n$, y suponga que A tiene r valores singulares diferentes de cero. Entonces, $\{A\mathbf{v}_1,\dots, A\mathbf{v}_r\}$ es una base ortogonal para Col A, y rango $A = r$.

DEMOSTRACIÓN Como $\mathbf{v}_i$ y $\lambda_j\mathbf{v}_j$ son ortogonales para $i \ne j$,

$$(A\mathbf{v}_i)^T(A\mathbf{v}_j) = \mathbf{v}_i^T A^T A\mathbf{v}_j = \mathbf{v}_i^T(\lambda_j \mathbf{v}_j) = 0$$

Por lo tanto, $\{A\mathbf{v}_1,\dots, A\mathbf{v}_n\}$ es un conjunto ortogonal. Además, ya que las longitudes de los vectores $A\mathbf{v}_1,\dots, A\mathbf{v}_n$ son los valores singulares de A, y puesto que hay r valores singulares diferentes de cero, $A\mathbf{v}_i \ne \mathbf{0}$ si y solo si $1 \le i \le r$. Así que $A\mathbf{v}_1,\dots, A\mathbf{v}_r$ son vectores linealmente independientes, y están en Col A. Finalmente, para cualquier $\mathbf{y}$ en Col A, por ejemplo, $\mathbf{y} = A\mathbf{x}$, se puede escribir $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, y

$$\begin{aligned}\mathbf{y} = A\mathbf{x} &= c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + c_{r+1}A\mathbf{v}_{r+1} + \cdots + c_n A\mathbf{v}_n \\ &= c_1 A\mathbf{v}_1 + \cdots + c_r A\mathbf{v}_r + 0 + \cdots + 0\end{aligned}$$

Así, $\mathbf{y}$ está en Gen $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$, lo que demuestra que $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ es una base (ortogonal) para Col A. Por lo tanto, rango A = dim Col $A = r$. ■

NOTA NUMÉRICA

En algunos casos, el rango de A puede ser muy sensible a pequeños cambios en las entradas de A. El método evidente de contar el número de columnas pivote en A no funciona muy bien si A se reduce por filas con una computadora. Con frecuencia, el error por redondeo genera una forma escalonada con rango completo.

En la práctica, la manera más confiable de estimar el rango de una matriz A grande consiste en contar el número de valores singulares diferentes de cero. En este caso, se supone que los valores singulares distintos de cero y extremadamente pequeños son iguales a cero para todos los fines prácticos, y el *rango efectivo* de la matriz es el número que se obtiene al contar los valores singulares restantes diferentes de cero.[1]

Descomposición en valores singulares

La descomposición de A implica a una matriz Σ "diagonal" de $m \times n$ de la forma

$$\Sigma = \begin{array}{c} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{array}{l} \\ \leftarrow m - r \text{ filas} \end{array} \\ \uparrow\, n - r \text{ columnas} \end{array} \tag{3}$$

donde D es una matriz diagonal de $r \times r$ para alguna r que no exceda el mínimo de m y n. [Si r es igual a m o n o a ambos, entonces no aparecen algunas (o todas) las matrices cero].

TEOREMA 10 **Descomposición en valores singulares**

Sea A una matriz de $m \times n$ con rango r. Entonces existe una matriz Σ de $m \times n$ como la de la ecuación (3) para la cual las entradas diagonales en D son los primeros r valores singulares de A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, y existen una matriz ortogonal U de $m \times m$ y una matriz ortogonal V de $n \times n$ tales que

$$A = U\Sigma V^T$$

Cualquier factorización $A = U\Sigma V^T$, con U y V ortogonales, Σ como en la ecuación (3), y entradas diagonales positivas en D, se llama **descomposición en valores singulares** (o **DVS**) de A. Las matrices U y V no están determinadas de forma única por A, pero las entradas diagonales en D son necesariamente los valores singulares de A. Véase el ejercicio 19. Las columnas de U en esta descomposición se llaman **vectores singulares izquierdos** de A, y las columnas de V se llaman **vectores singulares derechos** de A.

[1] En general, no es sencillo estimar el rango. Para un análisis de los aspectos sutiles implicados, véase Philip E. Gill, Walter Murray y Margaret H. Wright, *Numerical Linear Algebra and Optimization*, vol.1 (Redwood City, CA: Addison-Wesley, 1991), sección 5.8.

DEMOSTRACIÓN Sean λ_i y $\mathbf{v}_i$ como en el teorema 9, así $\{A\mathbf{v}_1,\ldots, A\mathbf{v}_r\}$ es una base ortogonal para Col A. Normalice cada $A\mathbf{v}_i$ para obtener una base ortonormal $\{\mathbf{u}_1,\ldots, \mathbf{u}_r\}$, donde

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|}A\mathbf{v}_i = \frac{1}{\sigma_i}A\mathbf{v}_i$$

y

$$A\mathbf{v}_i = \sigma_i\mathbf{u}_i \qquad (1 \le i \le r) \tag{4}$$

Ahora se extiende $\{\mathbf{u}_1,\ldots, \mathbf{u}_r\}$ a una base ortonormal $\{\mathbf{u}_1,\ldots, \mathbf{u}_m\}$ de $\mathbb{R}^m$, y sean

$$U = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,] \quad \text{y} \quad V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\,]$$

Por construcción, U y V son matrices ortogonales. También, a partir de la ecuación (4),

$$AV = [\,A\mathbf{v}_1 \quad \cdots \quad A\mathbf{v}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] = [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,]$$

Sea D la matriz diagonal con entradas diagonales $\sigma_1,\ldots, \sigma_r$, con Σ como en la ecuación (3) anterior. Entonces,

$$U\Sigma = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,]\left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & 0 & & 0 \end{array}\right]$$

$$\begin{aligned} &= [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,] \\ &= AV \end{aligned}$$

Puesto que V es una matriz ortogonal, entonces $U\Sigma V^T = AVV^T = A$. ■

Los siguientes dos ejemplos se concentran en la estructura interna de una descomposición en valores singulares. Un algoritmo eficiente y numéricamente estable para esta descomposición utilizaría un enfoque diferente. Véase la nota numérica al final de la sección.

EJEMPLO 3 Utilice los resultados de los ejemplos 1 y 2 para construir una descomposición en valores singulares de $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

SOLUCIÓN Una construcción se puede dividir en tres pasos.

Paso 1. ***Encuentre una diagonalización ortogonal de*** $\boldsymbol{A^TA}$. Es decir, determine los valores propios de A^TA y un conjunto ortonormal correspondiente de vectores propios. Si A tuviera solo dos columnas, los cálculos podrían realizarse a mano. Matrices mucho más grandes, por lo general, requieren de un programa de matrices.[2] Sin embargo, para la matriz A en cuestión, los datos propios para A^TA se dan en el ejemplo 1.

Paso 2. ***Obtenga V y*** $\boldsymbol{\Sigma}$. Acomode los valores propios de A^TA en orden decreciente. En el ejemplo 1, los valores propios ya están listados en orden descendente: 360, 90 y 0. Los vectores propios unitarios correspondientes, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, son los vectores singulares derechos de A. Utilizando el ejemplo 1, se construye

$$V = [\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

[2] Véase la *Guía de estudio* si se desea consultar las instrucciones adecuadas para software y la calculadora graficadora. MATLAB, por ejemplo, puede dar los valores propios y los vectores propios mediante una sola instrucción, `eig`.

Las raíces cuadradas de los valores propios son los valores singulares:

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

Los valores singulares diferentes de cero son las entradas diagonales de D. La matriz Σ tiene el mismo tamaño que A, con D en su esquina superior izquierda y ceros en las entradas restantes.

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \qquad \Sigma = [\, D \quad 0 \,] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Paso 3. ***Construya U.*** Cuando A tiene rango r, las primeras r columnas de U son los vectores normalizados obtenidos de $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$. En este ejemplo, A tiene dos valores singulares diferentes de cero, de manera que rango $A = 2$. Recuerde de la ecuación (2) y del párrafo anterior al ejemplo 2 que $\|A\mathbf{v}_1\| = \sigma_1$ y $\|A\mathbf{v}_2\| = \sigma_2$. Por lo tanto,

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{6\sqrt{10}}\begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \frac{1}{3\sqrt{10}}\begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

Observe que $\{\mathbf{u}_1, \mathbf{u}_2\}$ ya es una base para $\mathbb{R}^2$, por lo que no se necesitan vectores adicionales para U, y $U = [\mathbf{u}_1\ \mathbf{u}_2]$. La descomposición en valores singulares de A es

$$A = \underbrace{\begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix}}_{\substack{\uparrow \\ U}} \underbrace{\begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}}_{\substack{\uparrow \\ \Sigma}} \underbrace{\begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}}_{\substack{\uparrow \\ V^T}}$$

■

EJEMPLO 4 Encuentre una descomposición en valores singulares de $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

SOLUCIÓN Primero, calcule $A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$. Los valores propios de A^TA son 18 y 0, con los vectores propios unitarios correspondientes

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Esos vectores unitarios forman las columnas de V:

$$V = [\, \mathbf{v}_1 \quad \mathbf{v}_2 \,] = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Los valores singulares son $\sigma_1 = \sqrt{18} = 3\sqrt{2}$ y $\sigma_2 = 0$. Como solo existe un valor singular distinto de cero, la "matriz" D se puede escribir como un número individual. Es decir, $D = 3\sqrt{2}$. La matriz Σ tiene el mismo tamaño que A, con D en su esquina superior izquierda:

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Para construir U, primero se obtienen $A\mathbf{v}_1$ y $A\mathbf{v}_2$:

$$A\mathbf{v}_1 = \begin{bmatrix} 2/\sqrt{2} \\ -4/\sqrt{2} \\ 4/\sqrt{2} \end{bmatrix}, \quad A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

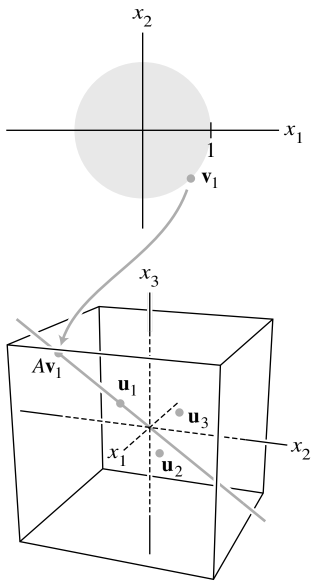


FIGURA 3

Para validar los cálculos, compruebe que $\|A\mathbf{v}_1\| = \sigma_1 = 3\sqrt{2}$. Desde luego, $A\mathbf{v}_2 = \mathbf{0}$ porque $\|A\mathbf{v}_2\| = \sigma_2 = 0$. Hasta ahora, la única columna encontrada para U es

$$\mathbf{u}_1 = \frac{1}{3\sqrt{2}} A\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

Las otras columnas de U se obtienen ampliando el conjunto $\{\mathbf{u}_1\}$ a una base ortonormal para $\mathbb{R}^3$. En este caso, se requieren dos vectores unitarios ortogonales $\mathbf{u}_2$ y $\mathbf{u}_3$ que sean ortogonales a $\mathbf{u}_1$. (Véase la figura 3). Cada vector debe satisfacer $\mathbf{u}_1^T\mathbf{x} = 0$, lo que es equivalente a la ecuación $x_1 - 2x_2 + 2x_3 = 0$. Una base para el conjunto solución de esta ecuación es

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

(Compruebe que $\mathbf{w}_1$ y $\mathbf{w}_2$ son ortogonales a $\mathbf{u}_1$). Aplique el proceso de Gram-Schmidt (con normalizaciones) a $\{\mathbf{w}_1, \mathbf{w}_2\}$, para obtener

$$\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

Finalmente, establezca $U = [\mathbf{u}_1\ \mathbf{u}_2\ \mathbf{u}_3]$, tome Σ y V^T de las ecuaciones anteriores, y escriba

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

■

Aplicaciones de la descomposición en valores singulares

Como se mencionó antes, la DVS se utiliza con frecuencia para estimar el rango de una matriz. Ahora se describirán brevemente otras aplicaciones numéricas, y en la sección 7.5 se expone una aplicación al procesamiento de imágenes.

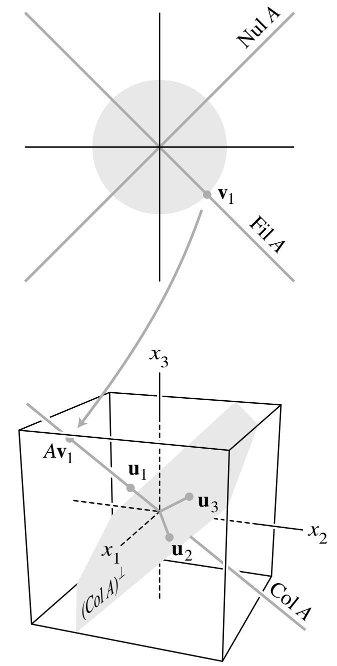
EJEMPLO 5 (El número de condición) La mayoría de los cálculos numéricos que implican una ecuación $A\mathbf{x} = \mathbf{b}$ son tan confiables como es posible cuando se utiliza la DVS de A. Las dos matrices ortogonales U y V no afectan las longitudes de vectores o los ángulos entre vectores (teorema 7 de la sección 6.2). Cualquier posible inestabilidad en los cálculos numéricos se identifica en Σ. Si los valores singulares de A son extremadamente grandes o pequeños, los errores por redondeo son casi inevitables, pero un análisis de errores se facilita si se conocen las entradas en Σ y V.

Si A es una matriz invertible de $n \times n$, entonces la razón σ_1/σ_n de los valores singulares más grande y más pequeño da el **número de condición** de A. Los ejercicios 41 a 43 de la sección 2.3 mostraron cómo el número de condición afecta la sensibilidad de una solución de $A\mathbf{x} = \mathbf{b}$ ante cambios (o errores) en las entradas de A. (En realidad, un "número de condición" de A se puede calcular en diversas formas, pero la definición que se presenta aquí se utiliza ampliamente para estudiar $A\mathbf{x} = \mathbf{b}$). ■

EJEMPLO 6 (Bases para subespacios fundamentales) Dada una DVS para una matriz A de $m \times n$, sean $\mathbf{u}_1, \ldots, \mathbf{u}_m$ los vectores singulares izquierdos, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ los vectores singulares derechos, $\sigma_1, \ldots, \sigma_n$ los valores singulares, y r el rango de A. De acuerdo con el teorema 9,

$$\{\mathbf{u}_1, \ldots, \mathbf{u}_r\} \tag{5}$$

es una base ortonormal para Col A.



Los subespacios fundamentales en el ejemplo 4.

Recuerde del teorema 3 de la sección 6.1 que $(\text{Col } A)^\perp = \text{Nul } A^T$. Por lo tanto,

$$\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\} \tag{6}$$

es una base ortonormal para Nul A^T.

Puesto que $\|A\mathbf{v}_i\| = \sigma_1$ para $1 \le i \le n$, y σ_i es 0 si y solo si $i > r$, por lo tanto los vectores $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ generan un subespacio de Nul A de dimensión $n - r$. De acuerdo con el teorema del rango, dim Nul $A = n -$ rango A. De ello se deduce que

$$\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\} \tag{7}$$

es una base ortonormal para Nul A, de acuerdo con el teorema de la base (en la sección 4.5).

A partir de (5) y (6), el complemento ortogonal de Nul A^T es Col A. Intercambiando A y A^T, observe que $(\text{Nul } A)^\perp = \text{Col } A^T = \text{Fil } A$. En consecuencia, a partir de (7),

$$\{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \tag{8}$$

es una base ortonormal para Fil A.

La figura 4 resume las ecuaciones (5) a (8), pero muestra la base ortogonal $\{\sigma_1\mathbf{u}_1, \ldots, \sigma_r\mathbf{u}_r\}$ para Col A en vez de la base normalizada, con la finalidad de recordarnos que $A\mathbf{v}_i = \sigma_i\mathbf{u}_i$ para $1 \le i \le r$. Las bases ortonormales explícitas para los cuatro subespacios fundamentales determinados por A son útiles en algunos cálculos, particularmente en problemas de optimización restringidos. ■

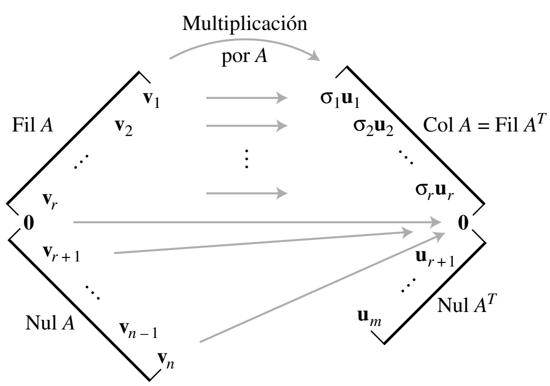


FIGURA 4 Los cuatro subespacios fundamentales y la acción de A.

Los cuatro subespacios fundamentales y el concepto de valores singulares proporcionan los enunciados finales del teorema de la matriz invertible. (Recuerde que se omitieron los enunciados del teorema acerca de A^T, ya que, de otra forma, el número de enunciados se duplicaría). Los otros enunciados se presentaron en las secciones 2.3, 2.9, 3.2, 4.6 y 5.2.

TEOREMA

Teorema de la matriz invertible (concluido)

Sea A una matriz de $n \times n$. Entonces, cada uno de los siguientes enunciados es equivalente al enunciado de que A es una matriz invertible.

u) $(\text{Col } A)^\perp = \{\mathbf{0}\}$.

v) $(\text{Nul } A)^\perp = \mathbb{R}^n$.

w) Fil $A = \mathbb{R}^n$.

x) A tiene n valores singulares diferentes de cero.

EJEMPLO 7 (DVS reducida y la seudoinversa de A) Cuando Σ contiene filas o columnas de ceros, es posible una descomposición más compacta de A. Utilizando la notación ya establecida, sea $r = \text{rango } A$, y U y V se particionan en submatrices cuyos primeros bloques contienen r columnas:

$$\begin{aligned} U &= [\, U_r \quad U_{m-r} \,], \quad \text{donde } U_r = [\, \mathbf{u}_1 \ \cdots \ \mathbf{u}_r \,] \\ V &= [\, V_r \quad V_{n-r} \,], \quad \text{donde } V_r = [\, \mathbf{v}_1 \ \cdots \ \mathbf{v}_r \,] \end{aligned}$$

Entonces U_r es de $m \times r$, y V_r es de $n \times r$. (Para simplificar la notación, considere U_{m-r} o V_{n-r} aun cuando tal vez alguna de ellas no tenga columnas). Por lo tanto, la multiplicación de matrices particionadas indica que

$$A = [\, U_r \quad U_{m-r} \,] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T \tag{9}$$

Esta factorización de A se llama **descomposición en valores singulares reducida** de A. Como las entradas diagonales en D son diferentes de cero, D es invertible. La siguiente matriz se llama la **seudoinversa** (o **inversa de Moore-Penrose**) de A:

$$A^+ = V_r D^{-1} U_r^T \tag{10}$$

Al final del capítulo, los ejercicios complementarios 12 a 14 exploran algunas de las propiedades de la descomposición en valores singulares reducida y de la seudoinversa. ■

EJEMPLO 8 (Solución por mínimos cuadrados) Dada la ecuación $A\mathbf{x} = \mathbf{b}$, utilice la seudoinversa de A en la ecuación (10) para definir

$$\hat{\mathbf{x}} = A^+\mathbf{b} = V_r D^{-1} U_r^T \mathbf{b}$$

Entonces, a partir de la DVS en la ecuación (9),

$$\begin{aligned} A\hat{\mathbf{x}} &= (U_r D V_r^T)(V_r D^{-1} U_r^T \mathbf{b}) \\ &= U_r D D^{-1} U_r^T \mathbf{b} \qquad \text{Porque } V_r^T V_r = I_r \\ &= U_r U_r^T \mathbf{b} \end{aligned}$$

De (5) se deduce que $U_r U_r^T \mathbf{b}$ es la proyección ortogonal $\hat{\mathbf{b}}$ de $\mathbf{b}$ sobre Col A. (Véase el teorema 10 de la sección 6.3). Por lo tanto, $\hat{\mathbf{x}}$ es una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. En efecto, esta $\hat{\mathbf{x}}$ tiene la longitud más pequeña en comparación con todas las soluciones por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. Véase el ejercicio complementario 14. ■

NOTA NUMÉRICA

Los ejemplos 1 a 4 y los ejercicios ilustran el concepto de valores singulares y sugieren cómo efectuar cálculos a mano. En la práctica, se debería evitar el cálculo de A^TA, porque cualquier error en las entradas de A se eleva al cuadrado en las entradas de A^TA. Existen métodos iterativos rápidos que producen los valores singulares y los vectores singulares de A con muchos decimales.

Lecturas adicionales

Horn, Roger A. y Charles R. Johnson, *Matrix Analysis* (Cambridge: Cambridge University Press, 1990).

Long, Cliff, "Visualization of Matrix Singular Value Decomposition". *Mathematics Magazine* **56** (1983), pp. 161-167.

Moler, C. B. y D. Morrison, "Singular Value Analysis of Cryptograms". *Amer. Math. Monthly* **90** (1983), pp. 78-87.

Strang, Gilbert, *Linear Algebra and Its Applications*, 4a. ed. (Belmont, CA: Brooks/Cole, 2005).

Watkins, David S., *Fundamentals of Matrix Computations* (Nueva York: Wiley, 1991), pp. 390-398, 409-421.

PROBLEMA DE PRÁCTICA

WEB

Dada una descomposición en valores singulares, $A = U\Sigma V^T$, encuentre una DVS de A^T. ¿Cómo están relacionados los valores singulares de A y A^T?

7.4 EJERCICIOS

En los ejercicios 1 a 4, obtenga los valores singulares de las matrices.

1. $\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$ **2.** $\begin{bmatrix} -5 & 0 \\ 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$ **4.** $\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$

En los ejercicios 5 a 12, encuentre una DVS para cada matriz. [*Sugerencia:* En el ejercicio 11, una elección para U es $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$. En el ejercicio 12, una columna de U puede ser $\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$].

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$ **6.** $\begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$

7. $\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$ **8.** $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

9. $\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$ **10.** $\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$

11. $\begin{bmatrix} -3 & 1 \\ 6 & -2 \\ 6 & -2 \end{bmatrix}$ **12.** $\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 1 \end{bmatrix}$

13. Encuentre la DVS de $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$. [*Sugerencia:* Trabaje con A^T].

14. En el ejercicio 7, obtenga un vector unitario $\mathbf{x}$ en el cual $A\mathbf{x}$ tiene longitud máxima.

15. Suponga que la siguiente factorización es una DVS de una matriz A, con las entradas en U y V redondeadas a dos decimales.

$$A = \begin{bmatrix} .40 & -.78 & .47 \\ .37 & -.33 & -.87 \\ -.84 & -.52 & -.16 \end{bmatrix} \begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .30 & -.51 & -.81 \\ .76 & .64 & -.12 \\ .58 & -.58 & .58 \end{bmatrix}$$

a) ¿Cuál es el rango de A?

b) Utilice esta descomposición de A, sin hacer cálculos, y escriba una base para Col A y una base para Nul A. [*Sugerencia:* Primero escriba las columnas de V].

16. Repita el ejercicio 15 para la siguiente DVS de una matriz A de 3×4:

$$A = \begin{bmatrix} -.86 & -.11 & -.50 \\ .31 & .68 & -.67 \\ .41 & -.73 & -.55 \end{bmatrix} \begin{bmatrix} 12.48 & 0 & 0 & 0 \\ 0 & 6.34 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .66 & -.03 & -.35 & .66 \\ -.13 & -.90 & -.39 & -.13 \\ .65 & .08 & -.16 & -.73 \\ -.34 & .42 & -.84 & -.08 \end{bmatrix}$$

En los ejercicios 17 a 24, A es una matriz de $m \times n$ con una descomposición en valores singulares $A = U\Sigma V^T$, donde U es una matriz ortogonal de $m \times m$, Σ es una matriz "diagonal" de $m \times n$ con r entradas positivas y sin entradas negativas, y V es una matriz ortogonal de $n \times n$. Justifique sus respuestas.

17. Suponga que A es cuadrada e invertible. Encuentre una descomposición en valores singulares de A^{-1}.

18. Demuestre que si A es cuadrada, entonces $|\det A|$ es el producto de los valores singulares de A.

19. Demuestre que las columnas de V son vectores propios de A^TA, que las columnas de U son vectores propios de AA^T, y que las entradas diagonales de Σ son los valores singulares de A. [*Sugerencia:* Utilice la DVS para calcular A^TA y AA^T].

20. Demuestre que si A es una matriz positiva definida de $n \times n$, entonces una diagonalización ortogonal $A = PDP^T$ es una descomposición en valores singulares de A.

21. Demuestre que si P es una matriz ortogonal de $m \times m$, entonces PA y A tienen los mismos valores singulares.

22. Justifique el enunciado del ejemplo 2 referente a que el segundo valor singular de una matriz A es el máximo de $\|A\mathbf{x}\|$ conforme $\mathbf{x}$ varía sobre todos los vectores unitarios ortogonales a $\mathbf{v}_1$, siendo $\mathbf{v}_1$ un vector singular derecho correspondiente al primer valor singular de A. [*Sugerencia:* Utilice el teorema 7 de la sección 7.3].

23. Sean $U = [\mathbf{u}_1 \cdots \mathbf{u}_m]$ y $V = [\mathbf{v}_1 \cdots \mathbf{v}_n]$, donde $\mathbf{u}_j$ y $\mathbf{v}_j$ son como en el teorema 10. Demuestre que

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T$$

24. Utilizando la notación del ejercicio 23, demuestre que $A^T\mathbf{u}_j = \sigma_j\mathbf{v}_j$ para $1 \le j \le r = \text{rango } A$.

25. Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Describa cómo encontrar una base $\mathcal{B}$ para $\mathbb{R}^n$ y una base $\mathcal{C}$ para $\mathbb{R}^m$ tal que la matriz para T respecto de $\mathcal{B}$ y $\mathcal{C}$ sea una matriz "diagonal" de $m \times n$.

[**M**] Calcule una DVS para cada matriz en los ejercicios 26 y 27. Informe las entradas matriciales finales con dos decimales. Utilice el método de los ejemplos 3 y 4.

26. $A = \begin{bmatrix} -18 & 13 & -4 & 4 \\ 2 & 19 & -4 & 12 \\ -14 & 11 & -12 & 8 \\ -2 & 21 & 4 & 8 \end{bmatrix}$

27. $A = \begin{bmatrix} 6 & -8 & -4 & 5 & -4 \\ 2 & 7 & -5 & -6 & 4 \\ 0 & -1 & -8 & 2 & 2 \\ -1 & -2 & 4 & 4 & -8 \end{bmatrix}$

28. [**M**] Determine los valores singulares de la matriz de 4×4 en el ejercicio 9 de la sección 2.3, y calcule el número de condición σ_1/σ_4.

29. [**M**] Calcule los valores singulares de la matriz de 5×5 en el ejercicio 10 de la sección 2.3, y determine el número de condición σ_1/σ_5.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Si $A = U\Sigma V^T$, donde Σ es de $m \times n$, entonces $A^T = (V^T)^T\Sigma^T U^T = V\Sigma^T U^T$. Esto es una DVS de A^T, ya que V y U son matrices ortogonales y Σ^T es una matriz "diagonal" de $n \times m$. Como Σ y Σ^T tienen las mismas entradas diagonales diferentes de cero, A y A^T tienen los mismos valores singulares distintos de cero. [*Nota:* Si A es de $2 \times n$, entonces AA^T es solo de 2×2 y sus valores propios pueden ser más fáciles de calcular (a mano) que los valores propios de A^TA].

7.5 APLICACIONES AL PROCESAMIENTO DE IMÁGENES Y ESTADÍSTICA

Las fotografías satelitales en la introducción del capítulo dan un ejemplo de datos mutidimensionales, o *multivariados*; es decir, se trata de información organizada de manera que cada dato en el conjunto de datos esté identificado con un punto (vector) en $\mathbb{R}^n$. El principal objetivo de esta sección es explicar una técnica, llamada *análisis de componentes principales*, empleada para analizar datos multivariados. Los cálculos ilustrarán el uso de la diagonalización ortogonal y la descomposición en valores singulares.

El análisis de componentes principales es aplicable a cualquier conjunto de datos consistente en listas de mediciones realizadas a una colección de objetos o individuos. Por ejemplo, considere un proceso químico que produce un material plástico. Para monitorizar el proceso, se toman 300 muestras del material producido, y cada muestra se somete a una batería de ocho pruebas, tales como punto de fusión, densidad, ductilidad, resistencia a la tensión, etcétera. El informe del laboratorio para cada muestra es un vector en $\mathbb{R}^8$, y el conjunto de tales vectores forma una matriz de 8×300, llamada **matriz de observaciones**.

En términos generales, se puede decir que los datos de control del proceso son octodimensionales. Los siguientes dos ejemplos describen datos que se pueden visualizar gráficamente.

EJEMPLO 1 Un ejemplo de datos bidimensionales está dado por un conjunto de pesos y alturas de N estudiantes de licenciatura. Sea $\mathbf{X}_j$ el **vector de observaciones** en $\mathbb{R}^2$ que lista el peso y la altura del j-ésimo estudiante. Si w denota el peso y h la altura, entonces la matriz de observaciones tiene la forma

$$\begin{array}{cccc} \left[\begin{matrix} w_1 \\ h_1 \end{matrix}\right. & \begin{matrix} w_2 \\ h_2 \end{matrix} & \begin{matrix} \cdots \\ \cdots \end{matrix} & \left.\begin{matrix} w_N \\ h_N \end{matrix}\right] \\ \uparrow & \uparrow & & \uparrow \\ \mathbf{X}_1 & \mathbf{X}_2 & & \mathbf{X}_N \end{array}$$

El conjunto de vectores de observación se puede visualizar como una *gráfica de dispersión* bidimensional. Véase la figura 1. ■

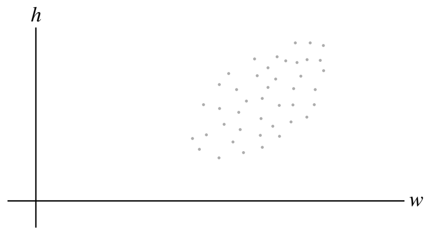


FIGURA 1 Una gráfica de dispersión de vectores de observaciones $\mathbf{X}_1, \ldots, \mathbf{X}_N$.

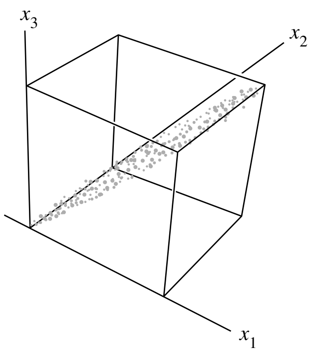


FIGURA 2 Gráfica de dispersión de datos espectrales para una imagen satelital.

EJEMPLO 2 Las primeras tres fotografías del Valle Railroad, Nevada, que se presentan en la introducción al capítulo, se pueden ver como *una* imagen de la región, con *tres componentes espectrales*, ya que se hicieron mediciones simultáneas de la región con tres distintas longitudes de onda. Cada fotografía da diferente información sobre la misma región física. Por ejemplo, el primer pixel en la esquina superior izquierda de cada fotografía corresponde al mismo lugar sobre el suelo (aproximadamente 30 metros por 30 metros). Cada pixel se asocia con un vector de observaciones en $\mathbb{R}^3$ que lista las intensidades de la señal para ese pixel en las tres bandas espectrales.

Por lo común, una imagen es de 2000×2000 pixeles, por lo que existen 4 millones de pixeles en la imagen. Los datos para la imagen forman una matriz con 3 filas y 4 millones de columnas (con las columnas arregladas en cualquier orden conveniente). En este caso, el carácter "multidimensional" de los datos se refiere a las tres dimensiones *espectrales* más que a las dos dimensiones *espaciales* que naturalmente pertenecen a cualquier fotografía. Los datos se pueden visualizar como un cúmulo de 4 millones de puntos en $\mathbb{R}^3$, quizá como en la figura 2. ■

Media y covarianza

Como preparación para el análisis de componentes principales, sea $[\mathbf{X}_1 \cdots \mathbf{X}_N]$ una matriz de observaciones de $p \times N$, tal como la que se describió anteriormente. La **media muestral**, $\mathbf{M}$, de los vectores de observaciones $\mathbf{X}_1, \ldots, \mathbf{X}_N$ está dada por

$$\mathbf{M} = \frac{1}{N}(\mathbf{X}_1 + \cdots + \mathbf{X}_N)$$

Para los datos de la figura 1, la media muestral es el punto en el "centro" de la gráfica de dispersión. Para $k = 1, \ldots, N$, sea

$$\hat{\mathbf{X}}_k = \mathbf{X}_k - \mathbf{M}$$

Las columnas de la matriz de $p \times N$

$$B = [\,\hat{\mathbf{X}}_1 \quad \hat{\mathbf{X}}_2 \quad \cdots \quad \hat{\mathbf{X}}_N\,]$$

tienen una media muestral cero, y se dice que B está en la **forma de desviación media.** Cuando la media muestral se resta de los datos de la figura 1, la gráfica de dispersión resultante tiene la forma que se observa en la figura 3.

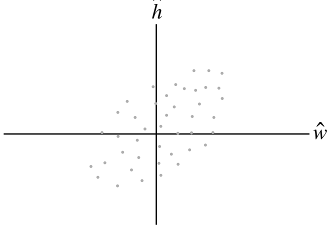


FIGURA 3 Datos de peso-altura en la forma de desviación media.

La **matriz de covarianza** (**muestral**) es la matriz S de $p \times p$ definida por

$$S = \frac{1}{N-1} BB^T$$

Como cualquier matriz de la forma BB^T es positiva semidefinida, entonces S lo es. (Véase el ejercicio 25 de la sección 7.2 con B y B^T intercambiadas).

EJEMPLO 3 Se realizan tres mediciones en cada uno de los cuatro individuos en una muestra aleatoria de una población. Los vectores de observaciones son

$$\mathbf{X}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{X}_2 = \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix}, \quad \mathbf{X}_3 = \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix}, \quad \mathbf{X}_4 = \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix}$$

Calcule la media muestral y la matriz de covarianza.

SOLUCIÓN La media muestral es

$$\mathbf{M} = \frac{1}{4}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \\ 13 \end{bmatrix} + \begin{bmatrix} 7 \\ 8 \\ 1 \end{bmatrix} + \begin{bmatrix} 8 \\ 4 \\ 5 \end{bmatrix} \right) = \frac{1}{4}\begin{bmatrix} 20 \\ 16 \\ 20 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 5 \end{bmatrix}$$

Restando la media muestral de $\mathbf{X}_1, \ldots, \mathbf{X}_4$ se obtiene

$$\hat{\mathbf{X}}_1 = \begin{bmatrix} -4 \\ -2 \\ -4 \end{bmatrix}, \quad \hat{\mathbf{X}}_2 = \begin{bmatrix} -1 \\ -2 \\ 8 \end{bmatrix}, \quad \hat{\mathbf{X}}_3 = \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix}, \quad \hat{\mathbf{X}}_4 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

y

$$B = \begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix}$$

La matriz de covarianza muestral es

$$\begin{aligned} S &= \frac{1}{3}\begin{bmatrix} -4 & -1 & 2 & 3 \\ -2 & -2 & 4 & 0 \\ -4 & 8 & -4 & 0 \end{bmatrix} \begin{bmatrix} -4 & -2 & -4 \\ -1 & -2 & 8 \\ 2 & 4 & -4 \\ 3 & 0 & 0 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 30 & 18 & 0 \\ 18 & 24 & -24 \\ 0 & -24 & 96 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 0 \\ 6 & 8 & -8 \\ 0 & -8 & 32 \end{bmatrix} \end{aligned}$$

■

Para analizar las entradas en $S = [s_{ij}]$, considere que $\mathbf{X}$ representa un vector que varía sobre el conjunto de vectores de observaciones y denote las coordenadas de $\mathbf{X}$ como $x_1, \ldots, x_p$. Entonces x_1, por ejemplo, es un escalar que varía sobre el conjunto de las primeras coordenadas de $\mathbf{X}_1, \ldots, \mathbf{X}_N$. Para $j = 1, \ldots, p$, la entrada diagonal s_{jj} en S es la **varianza** de x_j.

La varianza de x_j mide la dispersión de los valores de x_j. (Véase el ejercicio 13). En el ejemplo 3, la varianza de x_1 es 10 y la varianza de x_3 es 32. El hecho de que 32 sea mayor que 10 indica que el conjunto de las terceras entradas en los vectores de respuesta tiene una dispersión más amplia que el conjunto de las primeras entradas.

La **varianza total** de los datos es la suma de las varianzas en la diagonal de S. En general, la suma de las entradas diagonales de una matriz cuadrada S se denomina **traza** de la matriz, y se representa como $\text{tr}(S)$. Así,

$$\{\text{varianza total}\} = \text{tr}(S)$$

La entrada s_{ij} de S para $i \neq j$ es la **covarianza** de x_i y x_j. Observe que en el ejemplo 3, la covarianza entre x_1 y x_3 es 0 porque la entrada (1, 3) de S es 0. Los especialistas en estadística dicen que x_1 y x_3 **no están correlacionadas**. El análisis de los datos multivariados en $\mathbf{X}_1, \ldots, \mathbf{X}_N$ se simplifica notablemente cuando la mayoría de las variables (o todas) $x_1, \ldots, x_p$ no están correlacionadas, es decir, cuando la matriz de covarianza de $\mathbf{X}_1, \ldots, \mathbf{X}_N$ es diagonal o casi diagonal.

Análisis de componentes principales

Para simplificar, suponga que la matriz $[\mathbf{X}_1 \cdots \mathbf{X}_N]$ ya está en forma de desviación media. El objetivo del análisis de componentes principales es encontrar una matriz ortogonal de $p \times p$, $P = [\mathbf{u}_1 \cdots \mathbf{u}_p]$ que determine un cambio de variable, $\mathbf{X} = P\mathbf{Y}$, o

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}$$

con la propiedad de que las nuevas variables $y_1, \ldots, y_p$ no están correlacionadas y están en orden de varianza decreciente.

El cambio ortogonal de variable $\mathbf{X} = P\mathbf{Y}$ significa que cada vector de observación $\mathbf{X}_k$ recibe un "nuevo nombre", $\mathbf{Y}_k$, tal que $\mathbf{X}_k = P\mathbf{Y}_k$. Observe que $\mathbf{Y}_k$ es el vector de coordenadas de $\mathbf{X}_k$ con respecto a las columnas de P, y $\mathbf{Y}_k = P^{-1}\mathbf{X}_k = P^T\mathbf{X}_k$ para $k = 1, \ldots, N$.

No es difícil comprobar que para cualquier P ortogonal, la matriz de covarianza de $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ es P^TSP (ejercicio 11). Así, la matriz ortogonal P deseada es aquella que hace diagonal a P^TSP. Sea D una matriz diagonal con los valores propios $\lambda_1, \ldots, \lambda_p$ de S sobre la diagonal, ordenados de tal manera que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$, y sea P una matriz ortogonal cuyas columnas son los vectores propios unitarios correspondientes $\mathbf{u}_1, \ldots, \mathbf{u}_p$. De esta forma, $S = PDP^T$ y $P^TSP = D$.

Los vectores propios unitarios $\mathbf{u}_1, \ldots, \mathbf{u}_p$ de la matriz de covarianza S se llaman **componentes principales** de los datos (en la matriz de observaciones). La **primera componente principal** es el vector propio correspondiente al mayor valor propio de S, la **segunda componente principal** es el vector propio asociado al segundo mayor valor propio, y así sucesivamente.

La primera componente principal $\mathbf{u}_1$ determina la nueva variable y_1 de la siguiente manera. Sean $c_1, \ldots, c_p$ las entradas en $\mathbf{u}_1$. Como $\mathbf{u}_1^T$ es la primera fila de P^T, entonces la ecuación $\mathbf{Y} = P^T\mathbf{X}$ indica que

$$y_1 = \mathbf{u}_1^T\mathbf{X} = c_1x_1 + c_2x_2 + \cdots + c_px_p$$

Así, y_1 es una combinación lineal de las variables originales $x_1, \ldots, x_p$, utilizando como pesos las entradas en el vector propio $\mathbf{u}_1$. En forma similar, $\mathbf{u}_2$ determina la variable y_2, y así sucesivamente.

EJEMPLO 4 Los datos iniciales para la imagen multiespectral del Valle Railroad (ejemplo 2) consistían en 4 millones de vectores en $\mathbb{R}^3$. La matriz de covarianza asociada es[1]

$$S = \begin{bmatrix} 2382.78 & 2611.84 & 2136.20 \\ 2611.84 & 3106.47 & 2553.90 \\ 2136.20 & 2553.90 & 2650.71 \end{bmatrix}$$

[1] Los datos del ejemplo 4 y de los ejercicios 5 y 6 fueron proporcionados por Earth Satellite Corporation de Rockville, Maryland.

Encuentre las componentes principales de los datos, y liste la nueva variable determinada por la primera componente principal.

SOLUCIÓN Los valores propios de S y las componentes principales asociadas (los vectores propios unitarios) son

$$\lambda_1 = 7614.23 \qquad \lambda_2 = 427.63 \qquad \lambda_3 = 98.10$$
$$\mathbf{u}_1 = \begin{bmatrix} .5417 \\ .6295 \\ .5570 \end{bmatrix} \quad \mathbf{u}_2 = \begin{bmatrix} -.4894 \\ -.3026 \\ .8179 \end{bmatrix} \quad \mathbf{u}_3 = \begin{bmatrix} .6834 \\ -.7157 \\ .1441 \end{bmatrix}$$

Utilizando dos decimales para simplificar, la variable para la primera componente principal es

$$y_1 = .54x_1 + .63x_2 + .56x_3$$

Esta ecuación se utilizó para crear la fotografía *d*) en la introducción del capítulo. Las variables x_1, x_2, x_3 son las intensidades de la señal en las tres bandas espectrales. Los valores de x_1, convertidos a una escala de grises entre negro y blanco, produjeron la fotografía *a*). De manera similar, los valores de x_2 y x_3 generaron las fotografías *b*) y *c*), respectivamente. En cada pixel de la fotografía *d*), el valor en la escala de grises se calcula a partir de y_1, una combinación lineal ponderada de x_1, x_2, x_3. En este sentido, la fotografía *d*) "muestra" la primera componente principal de los datos. ■

En el ejemplo 4, la matriz de covarianza para los datos transformados, utilizando las variables y_1, y_2, y_3, es

$$D = \begin{bmatrix} 7614.23 & 0 & 0 \\ 0 & 427.63 & 0 \\ 0 & 0 & 98.10 \end{bmatrix}$$

Aunque D es evidentemente más sencilla que la matriz de covarianza original S, aún no es notorio el mérito de construir las nuevas variables. Sin embargo, las varianzas de las variable y_1, y_2, y_3 aparecen sobre la diagonal de D, y sin duda la primer varianza en D es mucho mayor que las otras dos. Como se verá, este hecho permitirá verificar que, en esencia, los datos son unidimensionales más que tridimensionales.

Reducción de la dimensión de datos multivariados

El análisis de componentes principales es potencialmente valioso en aplicaciones donde la mayor parte de la variación, o el rango dinámico, en los datos se debe a variaciones en *tan solo unas cuantas* de las nuevas variables, $y_1, \ldots, y_p$.

Es posible demostrar que un cambio ortogonal de variables, $\mathbf{X} = P\mathbf{Y}$, no modifica la varianza total de los datos. (A grandes rasgos, esto es cierto porque la multiplicación por la izquierda por P no altera las longitudes de los vectores ni los ángulos entre ellos. Véase el ejercicio 12). Esto significa que si $S = PDP^T$, entonces

$$\left\{\begin{matrix} \text{varianza total} \\ \text{de } x_1, \ldots, x_p \end{matrix}\right\} = \left\{\begin{matrix} \text{varianza total} \\ \text{de } y_1, \ldots, y_p \end{matrix}\right\} = \text{tr}(D) = \lambda_1 + \cdots + \lambda_p$$

La varianza de y_j es λ_j, y el cociente $\lambda_j/\text{tr}(S)$ mide la fracción de la varianza total que se "explica" o se "capta" con y_j.

EJEMPLO 5 Calcule los diversos porcentajes de varianza de los datos multiespectrales de Valle Railroad que se muestran en las fotografías de las componentes principales, *d*), *e*), *f*), presentadas en la introducción del capítulo.

SOLUCIÓN La varianza total de los datos es

$$\text{tr}(D) = 7614.23 + 427.63 + 98.10 = 8139.96$$

[Verifique que este número también sea igual a tr(S)]. Los porcentajes de la varianza total explicados por las componentes principales son

Primera componente	Segunda componente	Tercera componente
$\dfrac{7614.23}{8139.96} = 93.5\%$	$\dfrac{427.63}{8139.96} = 5.3\%$	$\dfrac{98.10}{8139.96} = 1.2\%$

En un sentido, el 93.5% de la información recopilada por Landsat para la región de Valle Railroad se presenta en la fotografía d), el 5.3% en e) y solamente el 1.2% restante en f). ■

Los cálculos del ejemplo 5 revelan que los datos prácticamente no tienen varianza en la tercera (nueva) coordenada. Todos los valores de y_3 son cercanos a cero. Geométricamente, los puntos de datos están cercanos al plano $y_3 = 0$, y sus ubicaciones se pueden determinar más o menos exactamente conociendo solo los valores de y_1 y y_2. De hecho, y_2 también tiene una varianza relativamente pequeña, lo que significa que los puntos están aproximadamente sobre una recta, y los datos, en esencia, son unidimensionales. Véase la figura 2, en la que los datos semejan un palo de paleta.

Caracterizaciones de las variables de componentes principales

Si $y_1,\ldots, y_p$ se originan de un análisis de componentes principales de una matriz de observaciones de $p \times N$, entonces la varianza de y_1 es tan grande como sea posible en el siguiente sentido: si $\mathbf{u}$ es cualquier vector unitario y si $y = \mathbf{u}^T\mathbf{X}$, entonces la varianza de los valores de y cuando $\mathbf{X}$ varía sobre los datos originales $\mathbf{X}_1,\ldots, \mathbf{X}_N$ es $\mathbf{u}^T S\mathbf{u}$. De acuerdo con el teorema 8 de la sección 7.3, el valor máximo de $\mathbf{u}^T S\mathbf{u}$, sobre todos los vectores unitarios $\mathbf{u}$, es el mayor valor propio λ_1 de S, y esta varianza se alcanza cuando $\mathbf{u}$ es el vector propio correspondiente $\mathbf{u}_1$. De manera similar, el teorema 8 indica que y_2 tiene la máxima varianza posible entre todas las variables $y = \mathbf{u}^T\mathbf{X}$ que *no están correlacionadas* con y_1. De igual forma, y_3 tiene la máxima varianza posible entre todas las variables no correlacionadas con y_1 y y_2, y así sucesivamente.

NOTA NUMÉRICA

La descomposición en valores singulares es la herramienta más importante para efectuar el análisis de componentes principales en aplicaciones prácticas. Si B es una matriz de observaciones de $p \times N$ en la forma de desviación media, y si $A = \left(1/\sqrt{N-1}\right)B^T$, entonces A^TA es la matriz de covarianza, S. Los cuadrados de los valores singulares de A son los p valores propios de S, y los vectores singulares derechos de A son las componentes principales de los datos.

Como se mencionó en la sección 7.4, el cálculo iterativo de la DVS de A es más rápido y preciso que una descomposición de valores propios de S. Esto es particularmente cierto, por ejemplo, en el procesamiento de imágenes hiperespectrales (con $p = 224$) que se mencionó en la introducción del capítulo. El análisis de componentes principales se completa en segundos en estaciones de trabajo especializadas.

Lectura adicional

Lillesand, Thomas M. y Ralph W. Kiefer, *Remote Sensing and Image Interpretation*, 4a. ed. (Nueva York: John Wiley, 2000).

PROBLEMAS DE PRÁCTICA

La siguiente tabla lista los pesos y las estaturas de cinco jóvenes:

Joven	#1	#2	#3	#4	#5
Peso (lb)	120	125	125	135	145
Altura (in.)	61	60	64	68	72

1. Encuentre la matriz de covarianza para los datos.
2. Realice un análisis de componentes principales de los datos para encontrar un solo *índice de tamaño* que explique la mayor parte de la variación en los datos.

7.5 EJERCICIOS

En los ejercicios 1 y 2, convierta la matriz de observaciones a la forma de desviación media y construya la matriz de covarianza de la muestra.

1. $\begin{bmatrix} 19 & 22 & 6 & 3 & 2 & 20 \\ 12 & 6 & 9 & 15 & 13 & 5 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 & 2 & 6 & 7 & 3 \\ 3 & 11 & 6 & 8 & 15 & 11 \end{bmatrix}$

3. Encuentre las componentes principales de los datos del ejercicio 1.

4. Encuentre las componentes principales de los datos del ejercicio 2.

5. [**M**] Se tomó una imagen Landsat con tres componentes espectrales de la base Homestead de la Fuerza Aérea, en Florida (después de que la base sufrió los estragos provocados por el huracán Andrew en 1992). La matriz de covarianza de los datos se muestra a continuación. Determine la primera componente principal de los datos, y calcule el porcentaje de la varianza total que está contenida en esa componente.

$$S = \begin{bmatrix} 164.12 & 32.73 & 81.04 \\ 32.73 & 539.44 & 249.13 \\ 81.04 & 249.13 & 189.11 \end{bmatrix}$$

6. [**M**] La siguiente matriz de covarianza se obtuvo de una imagen Landsat del río Columbia, Washington, utilizando datos de tres bandas espectrales. Sean x_1, x_2, x_3 las componentes espectrales de cada pixel en la imagen. Encuentre una nueva variable de la forma $y_1 = c_1x_1 + c_2x_2 + c_3x_3$ que tenga la máxima varianza posible, sujeta a la restricción $c_1^2 + c_2^2 + c_3^2 = 1$. ¿Qué porcentaje de la varianza total en los datos se explica mediante y_1?

$$S = \begin{bmatrix} 29.64 & 18.38 & 5.00 \\ 18.38 & 20.82 & 14.06 \\ 5.00 & 14.06 & 29.21 \end{bmatrix}$$

7. Sean x_1, x_2 las variables para los datos bidimensionales del ejercicio 1. Determine una nueva variable y_1 de la forma $y_1 = c_1x_1 + c_2x_2$, con $c_1^2 + c_2^2 = 1$, tal que y_1 tenga la máxima varianza posible sobre los datos proporcionados. ¿Qué parte de la varianza en los datos se explica por y_1?

8. Repita el ejercicio 7 para los datos del ejercicio 2.

9. Suponga que se aplican tres pruebas a una muestra aleatoria de estudiantes de licenciatura. Sean $\mathbf{X}_1, \ldots, \mathbf{X}_N$ los vectores de observaciones en $\mathbb{R}^3$ que listan las tres calificaciones de cada estudiante, y para $j = 1, 2, 3$, denote con x_j la calificación de un estudiante en el j-ésimo examen. Suponga que la matriz de covarianza de los datos es

$$S = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 7 \end{bmatrix}$$

Sea y un "índice" de desempeño de estudiantes, con $y = c_1x_1 + c_2x_2 + c_3x_3$ y $c_1^2 + c_2^2 + c_3^2 = 1$. Seleccione c_1, c_2, c_3 de manera que la varianza de y sobre el conjunto de datos sea lo más grande posible. [*Sugerencia:* Los valores propios de la matriz de covarianza de la muestra son $\lambda = 3, 6$ y 9].

10. [**M**] Repita el ejercicio 9 con $S = \begin{bmatrix} 5 & 4 & 2 \\ 4 & 11 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

11. A partir de los datos multivariados $\mathbf{X}_1, \ldots, \mathbf{X}_N$ (en $\mathbb{R}^p$) en la forma de desviación media, sea P una matriz de $p \times p$, y defina $\mathbf{Y}_k = P^T\mathbf{X}_k$ para $k = 1, \ldots, N$.
   a) Demuestre que $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ están en la forma de desviación media. [*Sugerencia:* Deje que $\mathbf{w}$ sea el vector en $\mathbb{R}^N$ con un 1 en cada entrada. Entonces $[\mathbf{X}_1 \cdots \mathbf{X}_N]\mathbf{w} = \mathbf{0}$ (el vector cero en $\mathbb{R}^p$)].
   b) Demuestre que si la matriz de covarianza de $\mathbf{X}_1, \ldots, \mathbf{X}_N$ es S, entonces la matriz de covarianza de $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ es P^TSP.

12. Sea $\mathbf{X}$ un vector que varía sobre las columnas de una matriz de observaciones de $p \times N$, en tanto que P es una matriz ortogonal de $p \times p$. Demuestre que el cambio de variable $\mathbf{X} = P\mathbf{Y}$ no cambia la varianza total de los datos. [*Sugerencia:* De acuerdo con el ejercicio 11, es suficiente demostrar que $\text{tr}(P^TSP) = \text{tr}(S)$. Utilice una propiedad de la traza mencionada en el ejercicio 25 de la sección 5.4].

13. La matriz de covarianza de la muestra es una generalización de una fórmula para la varianza de una muestra de N mediciones escalares, por ejemplo, $t_1, \ldots, t_N$. Si m es el promedio de $t_1, \ldots, t_N$, entonces la *varianza de la muestra* está dada por

$$\frac{1}{N-1}\sum_{k=1}^{n}(t_k - m)^2 \tag{1}$$

Muestre cómo la matriz de covarianza de la muestra, S, definida antes del ejemplo 3, se puede escribir en una forma similar a (1). [*Sugerencia:* Utilice multiplicación de matrices particionadas para escribir a S como $1/(N-1)$ por la suma de N matrices de tamaño $p \times p$. Para $1 \le k \le N$, escriba $\mathbf{X}_k - \mathbf{M}$ en vez de $\hat{\mathbf{X}}_k$].

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Primero arregle los datos en la forma de desviación media. Es fácil ver que el vector de la media muestral es $\mathbf{M} = \begin{bmatrix} 130 \\ 65 \end{bmatrix}$. Reste $\mathbf{M}$ de los vectores de observaciones (las columnas de la tabla) y obtenga

$$B = \begin{bmatrix} -10 & -5 & -5 & 5 & 15 \\ -4 & -5 & -1 & 3 & 7 \end{bmatrix}$$

Entonces la matriz de covarianza de la muestra es

$$S = \frac{1}{5-1}\begin{bmatrix} -10 & -5 & -5 & 5 & 15 \\ -4 & -5 & -1 & 3 & 7 \end{bmatrix}\begin{bmatrix} -10 & -4 \\ -5 & -5 \\ -5 & -1 \\ 5 & 3 \\ 15 & 7 \end{bmatrix}$$

$$= \frac{1}{4}\begin{bmatrix} 400 & 190 \\ 190 & 100 \end{bmatrix} = \begin{bmatrix} 100.0 & 47.5 \\ 47.5 & 25.0 \end{bmatrix}$$

2. Los valores propios de S son (con dos decimales)

$$\lambda_1 = 123.02 \quad \text{y} \quad \lambda_2 = 1.98$$

El vector propio unitario a λ_1 es $\mathbf{u} = \begin{bmatrix} .900 \\ .436 \end{bmatrix}$. (Como S es de 2×2, los cálculos se pueden hacer a mano si no se cuenta con un programa de matrices). Para el *índice de tamaño*, sea

$$y = .900\hat{w} + .436\hat{h}$$

donde $\hat{w}$ y $\hat{h}$ son el peso y la altura, respectivamente, en la forma de desviación media. La varianza de este índice sobre el conjunto de datos es 123.02. Puesto que la varianza total es $\operatorname{tr}(S) = 100 + 25 = 125$, el índice de tamaño explica prácticamente toda la varianza de los datos (el 98.4% de esta).

La figura 4 muestra los datos originales para el problema de práctica 1 y la recta determinada por la primera componente principal $\mathbf{u}$. (En forma vectorial paramétrica, la recta es $\mathbf{x} = \mathbf{M} + t\mathbf{u}$). Es posible demostrar que la recta es la mejor aproximación a los

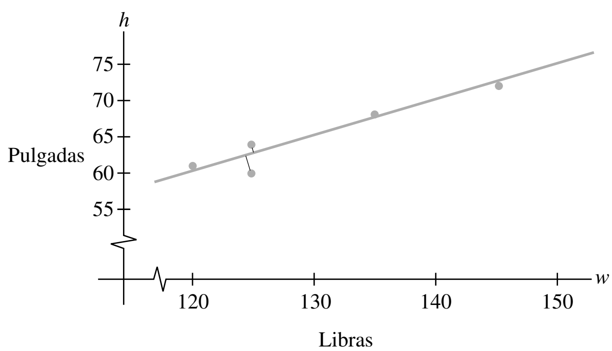


FIGURA 4 Una recta de regresión ortogonal determinada por la primera componente principal de los datos.

datos, en el sentido de que la suma de los cuadrados de las distancias *ortogonales* a la recta se minimiza. En efecto, el análisis de componentes principales es equivalente a la *regresión ortogonal*, pero eso es otra historia. Quizá se presente otra vez.

CAPÍTULO 7 EJERCICIOS COMPLEMENTARIOS

1. Marque cada enunciado como verdadero o falso. Justifique sus respuestas. En cada inciso, A representa una matriz de $n \times n$.

a) Si A es diagonalizable ortogonalmente, entonces A es simétrica.

b) Si A es una matriz ortogonal, entonces A es simétrica.

c) Si A es una matriz ortogonal, entonces $\|A\mathbf{x}\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

d) Los ejes principales de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$ pueden ser columnas de cualquier matriz P que diagonaliza a A.

e) Si P es una matriz de $n \times n$ con columnas ortogonales, entonces $P^T = P^{-1}$.

f) Si todo coeficiente en una forma cuadrática es positivo, entonces la forma cuadrática es definida positiva.

g) Si $\mathbf{x}^TA\mathbf{x} > 0$ para alguna $\mathbf{x}$, entonces la forma cuadrática $\mathbf{x}^TA\mathbf{x}$ es positiva definida.

h) Con un cambio adecuado de variable, cualquier forma cuadrática se puede convertir en una forma cuadrática sin productos cruzados.

i) El valor más grande de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$, para $\|\mathbf{x}\| = 1$, es la entrada más grande en la diagonal de A.

j) El valor máximo de una forma cuadrática positiva definida $\mathbf{x}^TA\mathbf{x}$ es el mayor valor propio de A.

k) Una forma cuadrática positiva definida se puede convertir en una forma negativa definida mediante un adecuado cambio de variable $\mathbf{x} = P\mathbf{u}$, para alguna matriz ortogonal P.

l) Una forma cuadrática indefinida es aquella cuyos valores propios no están definidos.

m) Si P es una matriz ortogonal de $n \times n$, entonces el cambio de variable $\mathbf{x} = P\mathbf{u}$ convierte $\mathbf{x}^TA\mathbf{x}$ en una forma cuadrática cuya matriz es $P^{-1}AP$.

n) Si U es de $m \times n$ con columnas ortogonales, entonces $UU^T\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre Col U.

o) Si B es de $m \times n$ y $\mathbf{x}$ es un vector unitario en $\mathbb{R}^n$, entonces $\|B\mathbf{x}\| \leq \sigma_1$, donde σ_1 es el primer valor singular de B.

p) Una descomposición en valores singulares de una matriz B de $m \times n$ se puede escribir como $B = P\Sigma Q$, donde P es una matriz ortogonal de $m \times m$, Q es una matriz ortogonal de $n \times n$, y Σ es una matriz "diagonal" de $m \times n$.

q) Si A es de $n \times n$, entonces A y A^TA tienen los mismos valores singulares.

2. Sean $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base ortonormal para $\mathbb{R}^n$, y $\lambda_1, \ldots, \lambda_n$ cualesquiera escalares reales. Defina

$$A = \lambda_1\mathbf{u}_1\mathbf{u}_1^T + \cdots + \lambda_n\mathbf{u}_n\mathbf{u}_n^T$$

a) Demuestre que A es simétrica.

b) Pruebe que $\lambda_1, \ldots, \lambda_n$ son los valores propios de A.

3. Sea A una matriz simétrica de $n \times n$ y rango r. Explique por qué la descomposición espectral de A representa a esta como la suma de r matrices de rango 1.

4. Sea A una matriz simétrica de $n \times n$.

a) Demuestre que $(\text{Col } A)^\perp = \text{Nul } A$. [*Sugerencia:* Véase la sección 6.1].

b) Demuestre que cada $\mathbf{y}$ en $\mathbb{R}^n$ se puede escribir en la forma $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, con $\hat{\mathbf{y}}$ en Col A y $\mathbf{z}$ en Nul A.

5. Demuestre que si $\mathbf{v}$ es un vector propio de una matriz A de $n \times n$, y $\mathbf{v}$ corresponde a un valor propio de A distinto de cero, entonces $\mathbf{v}$ está en Col A. [*Sugerencia:* Utilice la definición de un vector propio].

6. Sea A una matriz simétrica de $n \times n$. Aplique el ejercicio 5 y una base de vectores propios en $\mathbb{R}^n$ con la finalidad de dar una segunda demostración de la descomposición en el ejercicio 4*b*).

7. Demuestre que una matriz A de $n \times n$ es positiva definida si y solo si A admite una *factorización de Cholesky*, a saber, $A = R^TR$ para alguna matriz triangular superior invertible R cuyas entradas diagonales son todas positivas. [*Sugerencia:* Utilice una factorización QR y el ejercicio 26 de la sección 7.2].

8. Con base en el ejercicio 7, demuestre que si A es positiva definida, entonces A tiene una factorización LU, $A = LU$, donde U tiene pivotes positivos en su diagonal. (Lo contrario también es verdadero).

Si A es de $m \times n$, entonces la matriz $G = A^TA$ es la *matriz de Gram de* A. En este caso, las entradas de G son los productos interiores de las columnas de A. (Véase los ejercicios 9 y 10).

9. Demuestre que la matriz de Gram de cualquier matriz A es positiva semidefinida, con el mismo rango que A. (Véase los ejercicios de la sección 6.5).

10. Demuestre que si una matriz G de $n \times n$ es positiva semidefinida y tiene rango r, entonces G es la matriz de Gram de alguna matriz A de $r \times n$. Esto se llama *factorización reveladora del rango* de G. [*Sugerencia:* Considere la descomposición espectral de G, y primero escriba G como BB^T para una matriz B de $n \times r$].

11. Pruebe que cualquier matriz A de $n \times n$, admite una *descomposición polar* de la forma $A = PQ$, donde P es una matriz positiva semidefinida con el mismo rango que A, y donde Q es una matriz ortogonal de $n \times n$. [*Sugerencia:* Utilice la descomposición en valores singulares, $A = U\Sigma V^T$, y observe que $A = (U\Sigma U^T)(UV^T)$]. Esta descomposición se utiliza, por ejemplo, en ingeniería mecánica para modelar la deformación de un material. La matriz P describe la elongación o compresión del material (en las direcciones de los vectores propios de P), y Q describe la rotación del material en el espacio.

Los ejercicios 12 a 14 se refieren a una matriz A de $m \times n$ con una descomposición en valores singulares reducida, $A = U_r D V_r^T$, y la seudoinversa $A^+ = V_r D^{-1} U_r^T$.

12. Compruebe las propiedades de A^+:

a) Para toda $\mathbf{y}$ en $\mathbb{R}^m$, $AA^+\mathbf{y}$ es la proyección ortogonal de $\mathbf{y}$ sobre Col A.

b) Para toda $\mathbf{x}$ en $\mathbb{R}^n$, $A^+A\mathbf{x}$ es la proyección ortogonal de $\mathbf{x}$ sobre Fil A.

c) $AA^+A = A$ y $A^+AA^+ = A^+$.

13. Suponga que la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, y sea $\mathbf{x}^+ = A^+\mathbf{b}$. De acuerdo con el ejercicio 23 de la sección 6.3, existe exactamente un vector $\mathbf{p}$ en Fil A tal que $A\mathbf{p} = \mathbf{b}$. Los siguientes pasos prueban que $\mathbf{x}^+ = \mathbf{p}$ y $\mathbf{x}^+$ es la *solución de longitud mínima* de $A\mathbf{x} = \mathbf{b}$.

a) Demuestre que $\mathbf{x}^+$ está en Fil A. [*Sugerencia:* Escriba $\mathbf{b}$ como $A\mathbf{x}$ para alguna $\mathbf{x}$, y utilice el ejercicio 12].

b) Demuestre que $\mathbf{x}^+$ es una solución de $A\mathbf{x} = \mathbf{b}$.

c) Demuestre que si $\mathbf{u}$ es cualquier solución de $A\mathbf{x} = \mathbf{b}$, entonces $\|\mathbf{x}^+\| \leq \|\mathbf{u}\|$, con la igualdad solo si $\mathbf{u} = \mathbf{x}^+$.

14. Dada cualquier $\mathbf{b}$ en $\mathbb{R}^m$, adapte el ejercicio 13 para demostrar que $A^+\mathbf{b}$ es la *solución por mínimos cuadrados de longitud mínima*. [*Sugerencia:* Considere la ecuación $A\mathbf{x} = \hat{\mathbf{b}}$, donde $\hat{\mathbf{b}}$ es la proyección ortogonal de $\mathbf{b}$ sobre Col A].

[**M**] En los ejercicios 15 y 16, construya la seudoinversa de A. Comience aplicando un programa de matrices para obtener la DVS de A, o, si no tiene un programa de ese tipo, comience con una diagonalización ortogonal de A^TA. Utilice la seudoinversa para resolver $A\mathbf{x} = \mathbf{b}$, para $\mathbf{b} = (6, -1, -4, 6)$, y sea $\hat{\mathbf{x}}$ la solución. Haga un cálculo para comprobar que $\hat{\mathbf{x}}$ está en Fil A. Encuentre un vector $\mathbf{u}$ distinto de cero en Nul A, y compruebe que $\|\hat{\mathbf{x}}\| < \|\hat{\mathbf{x}} + \mathbf{u}\|$, lo que debe ser cierto de acuerdo con el ejercicio 13*c*).

15. $A = \begin{bmatrix} -3 & -3 & -6 & 6 & 1 \\ -1 & -1 & -1 & 1 & -2 \\ 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 & -1 \end{bmatrix}$

16. $A = \begin{bmatrix} 4 & 0 & -1 & -2 & 0 \\ -5 & 0 & 3 & 5 & 0 \\ 2 & 0 & -1 & -2 & 0 \\ 6 & 0 & -3 & -6 & 0 \end{bmatrix}$

8 Geometría de espacios vectoriales

EJEMPLO INTRODUCTORIO

Los sólidos platónicos

En la ciudad de Atenas, en el año 387 a. C., el filósofo griego Platón fundó una Academia, a la que muchos consideran como la primera universidad en el mundo. Aunque ahí se estudiaban ciencias como astronomía, biología, teoría política y filosofía, el tema fundamental de estudio fue la geometría. En efecto, a la entrada de la Academia se leía la inscripción: *No entre aquí quien no sepa geometría.*

Los griegos quedaron muy impresionados por los patrones geométricos, como los sólidos regulares. Un poliedro se llama regular si sus caras son polígonos regulares congruentes y todos los ángulos en los vértices son iguales. Cerca de 150 años antes de Euclides, los pitagóricos conocían al menos tres sólidos regulares: el tetraedro (4 caras triangulares), el cubo (6 caras cuadradas) y el octaedro (8 caras triangulares). (Véase la figura 1). Esas formas se presentan de manera natural en cristales de minerales comunes. Solo existen cinco de estos sólidos regulares; a los ya mencionados hay que añadir el dodecaedro (12 caras pentagonales) y el icosaedro (20 caras triangulares).

En el Libro XIII de sus *Elementos*, Platón analizó la teoría básica de esos cinco sólidos, y desde entonces se les conoce con el nombre de sólidos platónicos.

Durante siglos no hubo necesidad de considerar objetos geométricos de más de tres dimensiones. Pero en la actualidad es común que los matemáticos trabajen con objetos en espacios vectoriales de cuatro, cinco o incluso cientos de dimensiones. No necesariamente son claras las propiedades geométricas que pudieran atribuirse a esos objetos de grandes dimensiones.

Por ejemplo, ¿qué propiedades tienen las rectas en espacios bidimensionales y los planos en espacios tridimensionales? ¿Cuáles de esas propiedades podrían ser útiles en mayores dimensiones? Las secciones 8.1 y 8.4 presentan algunas respuestas. Los hiperplanos de la sección 8.4 serán importantes para entender la naturaleza multidimensional de los problemas de programación lineal.

¿A que "se parecería" el análogo de un poliedro en más de tres dimensiones? Se obtiene una respuesta parcial con las proyecciones bidimensionales del objeto tetradimensional, creadas de manera análoga a las proyecciones bidimensionales de un objeto tridimensional. La sección 8.5 ilustra esta idea para el "cubo" y el "simplejo" de cuatro dimensiones.

El estudio geométrico en grandes dimensiones no solo aporta nuevas formas de visualizar conceptos algebraicos abstractos, sino que también genera herramientas aplicables en $\mathbb{R}^3$. Por ejemplo, las secciones 8.2 y 8.6 incluyen aplicaciones a gráficos de computadora, y en la sección 8.5 se delinea una demostración (en el ejercicio 21) de que en $\mathbb{R}^3$ solo existen cinco poliedros regulares.

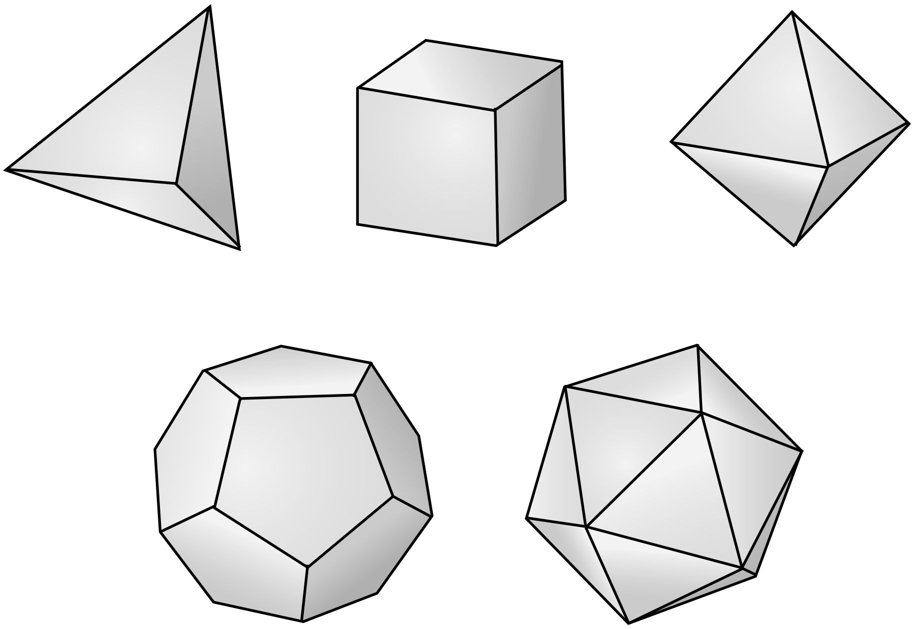


FIGURA 1 Los cinco sólidos platónicos.

En los capítulos anteriores, la mayoría de las aplicaciones implicaron cálculos algebraicos con subespacios y combinaciones lineales de vectores. Este capítulo estudia conjuntos de vectores que se pueden visualizar como objetos geométricos, segmentos de recta, polígonos y objetos sólidos. Los vectores individuales se ven como puntos. Los conceptos que se presentan en este capítulo se emplean en gráficos de computadora, en programación lineal y en otras áreas de matemáticas.[1]

A lo largo del capítulo, los conjuntos de vectores se describen mediante combinaciones lineales, pero con diferentes restricciones acerca de los pesos utilizados en dichas combinaciones. Por ejemplo, en la sección 8.1, la suma de los pesos es 1, mientras que en la sección 8.2, los pesos son positivos y suman 1. Desde luego, las visualizaciones son en $\mathbb{R}^2$ o $\mathbb{R}^3$, pero los conceptos también se aplican a $\mathbb{R}^n$ y a otros espacios vectoriales.

8.1 COMBINACIONES AFINES

Una combinación afín de vectores es un tipo especial de combinación lineal. Dados los vectores (o "puntos") $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ en $\mathbb{R}^n$ y los escalares $c_1, \dots, c_p$, una **combinación afín** de $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ es una combinación lineal

$$c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$$

tal que los pesos satisfacen $c_1 + \cdots + c_p = 1$.

[1] Véase Foley, Van Dam, Feiner y Hughes, *Computer Graphics-Principles and Practice*, 2a. edición (Boston: Addison-Wesley, 1996), pp. 1083-1112. Este material también analiza "espacios afines" sin coordenadas.

DEFINICIÓN

El conjunto de todas las combinaciones afines de puntos en un conjunto S es la **envolvente afín** (o el **afín generado**) de S, y se denota como aff S.

La envolvente afín de un solo punto $\mathbf{v}_1$ es justamente el conjunto $\{\mathbf{v}_1\}$, ya que tiene la forma $c_1\mathbf{v}_1$, donde $c_1 = 1$. Es frecuente que la envolvente afín de dos puntos distintos se escriba de manera especial. Suponga que $\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ con $c_1 + c_2 = 1$. Se escribe t en lugar de c_2, de manera que $c_1 = 1 - c_2 = 1 - t$. Entonces, la envolvente afín de $\{\mathbf{v}_1, \mathbf{v}_2\}$ es el conjunto

$$y = (1 - t)\mathbf{v}_1 + t\mathbf{v}_2, \quad \text{con } t \text{ en } \mathbb{R} \tag{1}$$

Este conjunto de puntos incluye a $\mathbf{v}_1$ (cuando $t = 0$) y a $\mathbf{v}_2$ (cuando $t = 1$). Si $\mathbf{v}_2 = \mathbf{v}_1$, entonces la ecuación (1) nuevamente describe solo un punto. En otras palabras, (1) describe la *recta* que pasa por $\mathbf{v}_1$ y $\mathbf{v}_2$. Para ver esto, se rescribe (1) en la forma:

$$y = \mathbf{v}_1 + t(\mathbf{v}_2 - \mathbf{v}_1) = \mathbf{p} + t\mathbf{u}, \quad \text{con } t \text{ en } \mathbb{R}$$

donde $\mathbf{p}$ es $\mathbf{v}_1$ y $\mathbf{u}$ es $\mathbf{v}_2 - \mathbf{v}_1$. El conjunto de todos los múltiplos de $\mathbf{u}$ es Gen $\{\mathbf{u}\}$, la recta que pasa por $\mathbf{u}$ y por el origen. El hecho de sumar $\mathbf{p}$ a cada punto sobre esta recta traslada Gen $\{\mathbf{u}\}$ a la recta que pasa por $\mathbf{p}$ paralela a la recta que pasa por $\mathbf{u}$ y por el origen. Véase la figura 1. (Compare esta figura con la número 5 de la sección 1.5).

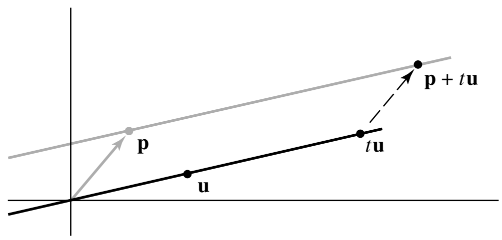


FIGURA 1

La figura 2 utiliza los puntos originales $\mathbf{v}_1$ y $\mathbf{v}_2$, y muestra aff$\{\mathbf{v}_1, \mathbf{v}_2\}$ como la recta que pasa por $\mathbf{v}_1$ y $\mathbf{v}_2$.

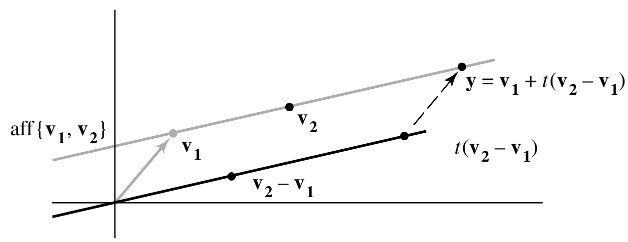


FIGURA 2

Observe que mientras que el punto $\mathbf{y}$ en la figura 2 es una combinación afín de $\mathbf{v}_1$ y $\mathbf{v}_2$, el punto $\mathbf{y} - \mathbf{v}_1$ es igual a $t(\mathbf{v}_2 - \mathbf{v}_1)$, que es una combinación lineal (de hecho, un múltiplo) de $\mathbf{v}_2 - \mathbf{v}_1$. Esta relación entre $\mathbf{y}$ y $\mathbf{y} - \mathbf{v}_1$ es válida para cualquier combinación afín de puntos, como lo indica el siguiente teorema.

TEOREMA 1

Un punto $\mathbf{y}$ en $\mathbb{R}^n$ es una combinación afín de $\mathbf{v}_1, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$ si y solo si $\mathbf{y} - \mathbf{v}_1$ es una combinación lineal de los puntos trasladados $\mathbf{v}_2 - \mathbf{v}_1, \ldots, \mathbf{v}_p - \mathbf{v}_1$.

DEMOSTRACIÓN Si $\mathbf{y} - \mathbf{v}_1$ es una combinación lineal de $\mathbf{v}_2 - \mathbf{v}_1, \ldots, \mathbf{v}_p - \mathbf{v}_1$, entonces existen pesos $c_2, \ldots, c_p$ tales que

$$\mathbf{y} - \mathbf{v}_1 = c_2(\mathbf{v}_2 - \mathbf{v}_1) + \cdots + c_p(\mathbf{v}_p - \mathbf{v}_1) \tag{2}$$

Luego,

$$\mathbf{y} = (1 - c_2 - \cdots - c_p)\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p \tag{3}$$

y los pesos en esta combinación lineal suman 1. Así, $\mathbf{y}$ es una combinación afín de $\mathbf{v}_1, \ldots, \mathbf{v}_p$. A la inversa, suponga que

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p \tag{4}$$

donde $c_1 + \cdots + c_p = 1$. Como $c_1 = 1 - c_2 - \cdots - c_p$, entonces la ecuación (4) se puede escribir como en la ecuación (3), y esto conduce a (2), lo que demuestra que $\mathbf{y} - \mathbf{v}_1$ es una combinación lineal de $\mathbf{v}_2 - \mathbf{v}_1, \ldots, \mathbf{v}_p - \mathbf{v}_1$. ■

En el enunciado del teorema 1, el punto $\mathbf{v}_1$ se podría remplazar por cualquier otro punto de la lista $\mathbf{v}_1, \ldots, \mathbf{v}_p$. En la demostración solo cambiaría la notación.

EJEMPLO 1 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ y $\mathbf{y} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Si es posible, escriba $\mathbf{y}$ como una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$.

SOLUCIÓN Se calculan los puntos trasladados

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_4 - \mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \end{bmatrix}, \quad \mathbf{y} - \mathbf{v}_1 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

Se encuentran los escalares c_2, c_3 y c_4 tales que

$$c_2(\mathbf{v}_2 - \mathbf{v}_1) + c_3(\mathbf{v}_3 - \mathbf{v}_1) + c_4(\mathbf{v}_4 - \mathbf{v}_1) = \mathbf{y} - \mathbf{v}_1 \tag{5}$$

ahora se reduce por filas la matriz aumentada teniendo estos puntos como columnas:

$$\begin{bmatrix} 1 & 0 & -3 & 3 \\ 3 & 1 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -3 & 3 \\ 0 & 1 & 9 & -10 \end{bmatrix}$$

Esto demuestra que la ecuación (5) es consistente, y la solución general es $c_2 = 3c_4 + 3$, $c_3 = -9c_4 - 10$, con c_4 libre. Cuando $c_4 = 0$,

$$\mathbf{y} - \mathbf{v}_1 = 3(\mathbf{v}_2 - \mathbf{v}_1) - 10(\mathbf{v}_3 - \mathbf{v}_1) + 0(\mathbf{v}_4 - \mathbf{v}_1)$$

y

$$\mathbf{y} = 8\mathbf{v}_1 + 3\mathbf{v}_2 - 10\mathbf{v}_3$$

Como otro ejemplo, se toma $c_4 = 1$. Entonces, $c_2 = 6$ y $c_3 = -19$, así que

$$\mathbf{y} - \mathbf{v}_1 = 6(\mathbf{v}_2 - \mathbf{v}_1) - 19(\mathbf{v}_3 - \mathbf{v}_1) + 1(\mathbf{v}_4 - \mathbf{v}_1)$$

y

$$\mathbf{y} = 13\mathbf{v}_1 + 6\mathbf{v}_2 - 19\mathbf{v}_3 + \mathbf{v}_4$$ ■

Mientras que el procedimiento del ejemplo 1 funciona para puntos arbitrarios $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ en $\mathbb{R}^n$, la pregunta se puede contestar más directamente si los puntos seleccionados $\mathbf{v}_i$ son una base para $\mathbb{R}^n$. Por ejemplo, sea $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ dicha base. Entonces cualquier $\mathbf{y}$ en $\mathbb{R}^n$ es una combinación *lineal* única de $\mathbf{b}_1, \ldots, \mathbf{b}_n$. Esta combinación es una combinación afín de las $\mathbf{b}$ si y solo si los pesos suman 1. (Estos pesos son justamente las $\mathcal{B}$-coordenadas de $\mathbf{y}$, como en la sección 4.4).

EJEMPLO 2 Sean $\mathbf{b}_1 = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}, \mathbf{p}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$ y $\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

El conjunto $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ es una base para $\mathbb{R}^3$. Determine si los puntos $\mathbf{p}_1$ y $\mathbf{p}_2$ son combinaciones afines de los puntos en $\mathcal{B}$.

SOLUCIÓN Encuentre las $\mathcal{B}$-coordenadas de $\mathbf{p}_1$ y $\mathbf{p}_2$. Estos dos cálculos se pueden combinar reduciendo por filas la matriz $[\mathbf{b}_1\ \ \mathbf{b}_2\ \ \mathbf{b}_3\ \ \mathbf{p}_1\ \ \mathbf{p}_2]$, con dos columnas aumentadas:

$$\begin{bmatrix} 4 & 0 & 5 & 2 & 1 \\ 0 & 4 & 2 & 0 & 2 \\ 3 & 2 & 4 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -2 & \frac{2}{3} \\ 0 & 1 & 0 & -1 & \frac{2}{3} \\ 0 & 0 & 1 & 2 & -\frac{1}{3} \end{bmatrix}$$

Lea la columna 4 para construir $\mathbf{p}_1$, y la columna 5 para construir $\mathbf{p}_2$:

$$\mathbf{p}_1 = -2\mathbf{b}_1 - \mathbf{b}_2 + 2\mathbf{b}_3 \quad \text{y} \quad \mathbf{p}_2 = \tfrac{2}{3}\mathbf{b}_1 + \tfrac{2}{3}\mathbf{b}_2 - \tfrac{1}{3}\mathbf{b}_3$$

La suma de los pesos en la combinación lineal para $\mathbf{p}_1$ es -1, no 1, de manera que $\mathbf{p}_1$ *no* es una combinación afín de las $\mathbf{b}$. Sin embargo, $\mathbf{p}_2$ *es* una combinación afín de las $\mathbf{b}$ porque la suma de los pesos para $\mathbf{p}_2$ es 1. ■

DEFINICIÓN

Un conjunto S es **afín** si $\mathbf{p}, \mathbf{q} \in S$ implica que $(1 - t)\mathbf{p} + t\mathbf{q} \in S$ para todo número real t.

Geométricamente, un conjunto es afín si siempre que dos puntos estén en el conjunto, la recta completa que pasa por esos puntos también está en el conjunto. (Si S solamente contiene un punto, $\mathbf{p}$, entonces la recta pasa por $\mathbf{p}$ y $\mathbf{p}$ es justamente un punto, una recta “degenerada”). Algebraicamente, para que un conjunto S sea afín, la definición requiere que cada combinación afín de dos puntos de S pertenezca a S. De manera notable, esto es equivalente a pedir que S contenga toda combinación afín de un número arbitrario de puntos de S.

TEOREMA 2

Un conjunto S es afín si y solo si cada combinación afín de puntos de S está en S. Es decir, S es afín si y solo si $S = \text{aff } S$.

DEMOSTRACIÓN Se supone que S es afín y se utiliza inducción sobre el número m de puntos de S que participan en una combinación afín. Cuando m es 1 o 2, una combinación afín de m puntos de S se encuentra en S, de acuerdo con la definición de conjunto afín. Ahora, se supone que cada combinación afín de k o menos puntos de S conducen a un punto en S, y se considera una combinación de $k + 1$ puntos. Tomemos $\mathbf{v}_i$ en S para $i = 1, \ldots, k + 1$, y sea $\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1}$, donde $c_1 + \cdots + c_{k+1} = 1$. Como las c_i suman 1, al menos una de ellas no debe ser igual a 1. Al reindexar las $\mathbf{v}_i$ y las c_i, si es necesario, se puede suponer que $c_{k+1} \neq 1$. Sea $t = c_1 + \cdots + c_k$. Entonces $t = 1 - c_{k+1} \neq 0$, y

$$\mathbf{y} = (1 - c_{k+1})\left(\frac{c_1}{t}\mathbf{v}_1 + \cdots + \frac{c_k}{t}\mathbf{v}_k\right) + c_{k+1}\mathbf{v}_{k+1} \tag{6}$$

De acuerdo con la hipótesis de inducción, el punto $\mathbf{z} = (c_1/t)\mathbf{v}_1 + \cdots + (c_k/t)\mathbf{v}_k$ está en S, ya que los coeficientes suman 1. De esta forma, (6) presenta a $\mathbf{y}$ como una combinación afín de dos puntos en S, y así $\mathbf{y} \in S$. De acuerdo con el principio de inducción, cada combinación afín de dichos puntos está en S. Es decir, aff $S \subset S$. Pero la inclusión inversa, $S \subset$ aff S, siempre se aplica. Por lo tanto, cuando S es afín, $S =$ aff S. A la inversa, si $S =$ aff S, entonces las combinaciones afines de dos (o más) puntos de S están en S, por consiguiente, S es afín. ■

La siguiente definición aporta terminología para conjuntos afines que pone de relieve su cercana relación con subespacios de $\mathbb{R}^n$.

DEFINICIÓN

Un traslado de un conjunto S en $\mathbb{R}^n$ por un vector $\mathbf{p}$ es el conjunto $S + \mathbf{p} = \{\mathbf{s} + \mathbf{p} : \mathbf{s} \in S\}$.[2] Un **plano afín** en $\mathbb{R}^n$ es un traslado de un subespacio de $\mathbb{R}^n$. Dos planos afines son **paralelos** si uno es el traslado del otro. La **dimensión de un plano** afín es la dimensión del subespacio paralelo correspondiente. La **dimensión de un conjunto** S, escrito como dim S, es la dimensión del plano afín más pequeño que contiene a S. Una **recta** en $\mathbb{R}^n$ es un plano afín de dimensión 1. Un **hiperplano** en $\mathbb{R}^n$ es un plano afín de dimensión $n - 1$.

En $\mathbb{R}^3$, los subespacios propios[3] consisten en el origen $\mathbf{0}$, el conjunto de todas las líneas que pasan por $\mathbf{0}$, y el conjunto de todos los planos que pasan por $\mathbf{0}$. Así, los planos afines propios en $\mathbb{R}^3$ son puntos (dimensión cero), rectas (unidimensionales) y planos (bidimensionales), los cuales pueden pasar o no por el origen.

El siguiente teorema indica que esas descripciones geométricas de rectas y planos en $\mathbb{R}^3$ (como traslados de subespacios) en realidad coinciden con sus anteriores descripciones algebraicas como conjuntos de todas las combinaciones afines de dos o tres puntos, respectivamente.

TEOREMA 3

Un conjunto S no vacío es afín si y solo si es un plano afín.

DEMOSTRACIÓN Supongamos que S es afín. Sea $\mathbf{p}$ cualquier punto fijo en S, y $W = S + (-\mathbf{p})$, de manera que $S = W + \mathbf{p}$. Para demostrar que S es un plano afín, es suficiente demostrar que W es un subespacio de $\mathbb{R}^n$. Puesto que $\mathbf{p}$ está en S, el vector cero está en W. Para demostrar que W es cerrado bajo sumas y múltiplos escalares, basta con probar que si $\mathbf{u}_1$ y $\mathbf{u}_2$ son elementos de W, entonces $\mathbf{u}_1 + t\mathbf{u}_2$ está en W para cada t real. Como $\mathbf{u}_1$ y $\mathbf{u}_2$ están en W, existen $\mathbf{s}_1$ y $\mathbf{s}_2$ en S tales que $\mathbf{u}_1 = \mathbf{s}_1 - \mathbf{p}$ y $\mathbf{u}_2 = \mathbf{s}_2 - \mathbf{p}$. Así, para cada t real,

$$\begin{aligned}\mathbf{u}_1 + t\mathbf{u}_2 &= (\mathbf{s}_1 - \mathbf{p}) + t(\mathbf{s}_2 - \mathbf{p}) \\ &= (1 - t)\mathbf{s}_1 + t(\mathbf{s}_1 + \mathbf{s}_2 - \mathbf{p}) - \mathbf{p}\end{aligned}$$

Sea $\mathbf{y} = \mathbf{s}_1 + \mathbf{s}_2 - \mathbf{p}$. Entonces $\mathbf{y}$ es una combinación afín de puntos en S. Como S es afín, $\mathbf{y}$ está en S (de acuerdo con el teorema 2). Pero entonces $(1 - t)\mathbf{s}_1 + t\mathbf{y}$ también está en S. Así, $\mathbf{u}_1 + t\mathbf{u}_2$ está en $-\mathbf{p} + S = W$. Esto demuestra que W es un subespacio de $\mathbb{R}^n$. Por lo tanto, S es un plano afín, porque $S = W + \mathbf{p}$.

A la inversa, suponga que S es un plano afín. Es decir, $S = W + \mathbf{p}$ para alguna $\mathbf{p} \in \mathbb{R}^n$ y algún subespacio W. Para probar que S es afín, es suficiente demostrar que para cualquier par de puntos $\mathbf{s}_1$ y $\mathbf{s}_2$ en S, la recta que pasa por $\mathbf{s}_1$ y $\mathbf{s}_2$ está en S. Por definición de W, existen $\mathbf{u}_1$ y $\mathbf{u}_2$ en W tales que $\mathbf{s}_1 = \mathbf{u}_1 + \mathbf{p}$ y $\mathbf{s}_2 = \mathbf{u}_2 + \mathbf{p}$. Así, para cada t real,

$$\begin{aligned}(1 - t)\mathbf{s}_1 + t\mathbf{s}_2 &= (1 - t)(\mathbf{u}_1 + \mathbf{p}) + t(\mathbf{u}_2 + \mathbf{p}) \\ &= (1 - t)\mathbf{u}_1 + t\mathbf{u}_2 + \mathbf{p}\end{aligned}$$

Como W es un subespacio, $(1 - t)\mathbf{u}_1 + t\mathbf{u}_2 \in W$ y así $(1 - t)\mathbf{s}_1 + t\mathbf{s}_2 \in W + \mathbf{p} = S$. Por lo tanto, S es afín. ■

[2] Si $\mathbf{p} = \mathbf{0}$, entonces el traslado es S mismo. Véase la figura 4 de la sección 1.5.

[3] Un subconjunto A de un conjunto B es un subconjunto **propio** de B si $A \neq B$. La misma condición se aplica a subespacios propios y planos afines propios en $\mathbb{R}^n$: no son iguales a $\mathbb{R}^n$.

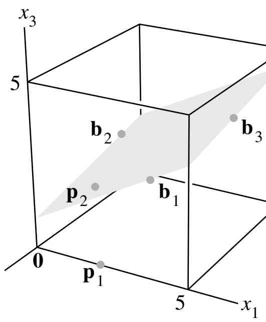


FIGURA 3

El teorema 3 ofrece un enfoque geométrico para la envolvente afín de un conjunto: es el plano afín que consiste en todas las combinaciones afines de puntos en el conjunto. Por ejemplo, la figura 3 muestra los puntos estudiados en el ejemplo 2. Aunque el conjunto de todas las combinaciones *lineales* de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_3$ es todo de $\mathbb{R}^3$, el conjunto de todas las combinaciones *afines* es solamente el plano a través de $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_3$. Observe que $\mathbf{p}_2$ (del ejemplo 2) está en el plano por $\mathbf{b}_1$, $\mathbf{b}_2$ y $\mathbf{b}_3$, mientras que $\mathbf{p}_1$ no se encuentra en ese plano. También, véase el ejercicio 14.

El siguiente ejemplo ofrece una mirada renovada a un conjunto familiar: el conjunto de todas soluciones de un sistema $A\mathbf{x} = \mathbf{b}$.

EJEMPLO 3 Suponga que las soluciones de una ecuación $A\mathbf{x} = \mathbf{b}$ son todas de la forma $\mathbf{x} = x_3\mathbf{u} + \mathbf{p}$, donde $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$. De la sección 1.5 recuerde que este conjunto es paralelo al conjunto solución de $A\mathbf{x} = \mathbf{0}$, que consiste en todos los puntos de la forma $x_3\mathbf{u}$. Encuentre puntos $\mathbf{v}_1$ y $\mathbf{v}_2$ tales que el conjunto solución de $A\mathbf{x} = \mathbf{b}$ sea aff $\{\mathbf{v}_1, \mathbf{v}_2\}$.

SOLUCIÓN El conjunto solución es una recta que pasa por $\mathbf{p}$ en la dirección de $\mathbf{u}$, como se muestra en la figura 1. Como aff$\{\mathbf{v}_1, \mathbf{v}_2\}$ es una recta que pasa por $\mathbf{v}_1$ y $\mathbf{v}_2$, identifique dos puntos sobre la recta $\mathbf{x} = x_3\mathbf{u} + \mathbf{p}$. Se presentan dos sencillas elecciones cuando $x_3 = 0$ y $x_3 = 1$. Es decir, tome $\mathbf{v}_1 = \mathbf{p}$ y $\mathbf{v}_2 = \mathbf{u} + \mathbf{p}$, de forma que

$$\mathbf{v}_2 = \mathbf{u} + \mathbf{p} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \\ -2 \end{bmatrix}$$

En este caso, el conjunto solución se describe como el conjunto de todas las combinaciones afines de la forma

$$\mathbf{x} = (1 - x_3)\begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix} + x_3\begin{bmatrix} 6 \\ -3 \\ -2 \end{bmatrix} \qquad \blacksquare$$

Anteriormente, el teorema 1 reveló una importante conexión entre combinaciones afines y combinaciones lineales. El siguiente teorema ofrece otra perspectiva de las combinaciones afines, que para $\mathbb{R}^2$ y $\mathbb{R}^3$ está cercanamente relacionada con aplicaciones a gráficos generados por computadora, como se explicará en la siguiente sección (y como se vio en la sección 2.7).

DEFINICIÓN

Para $\mathbf{v}$ en $\mathbb{R}^n$, la **forma homogénea** estándar de $\mathbf{v}$ es el punto $\tilde{\mathbf{v}} = \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix}$ en $\mathbb{R}^{n+1}$.

TEOREMA 4

Un punto $\mathbf{y}$ en $\mathbb{R}^n$ es una combinación afín de $\mathbf{v}_1,\ldots, \mathbf{v}_p$ en $\mathbb{R}^n$ si y solo si la forma homogénea de $\mathbf{y}$ está en Gen $\{\tilde{\mathbf{v}}_1,\ldots, \tilde{\mathbf{v}}_p\}$. De hecho, $\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, con $c_1 + \cdots + c_p = 1$, si y solo si $\tilde{\mathbf{y}} = c_1\tilde{\mathbf{v}}_1 + \cdots + c_p\tilde{\mathbf{v}}_p$.

DEMOSTRACIÓN Un punto $\mathbf{y}$ está en aff $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ si y solo si existen pesos $c_1,\ldots, c_p$ tales que

$$\begin{bmatrix} \mathbf{y} \\ 1 \end{bmatrix} = c_1\begin{bmatrix} \mathbf{v}_1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} \mathbf{v}_2 \\ 1 \end{bmatrix} + \cdots + c_p\begin{bmatrix} \mathbf{v}_p \\ 1 \end{bmatrix}$$

Esto ocurre si y solo si $\tilde{\mathbf{y}}$ se encuentra en Gen $\{\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2,\ldots, \tilde{\mathbf{v}}_p\}$. $\blacksquare$

EJEMPLO 4 Sean $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 7 \\ 1 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$. Utilice el teorema 4 para escribir $\mathbf{p}$ como una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, si es posible.

SOLUCIÓN Se reduce por filas la matriz aumentada para la ecuación

$$x_1\tilde{\mathbf{v}}_1 + x_2\tilde{\mathbf{v}}_2 + x_3\tilde{\mathbf{v}}_3 = \tilde{\mathbf{p}}$$

Para simplificar la aritmética, se mueve la cuarta fila de unos hasta arriba (lo que equivale a tres intercambios de filas). Después de esto, el número de operaciones aritméticas aquí es básicamente el mismo que el número que necesita el método del teorema 1.

$$[\,\tilde{\mathbf{v}}_1 \quad \tilde{\mathbf{v}}_2 \quad \tilde{\mathbf{v}}_3 \quad \tilde{\mathbf{p}}\,] \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 1 & 1 & 4 \\ 1 & 2 & 7 & 3 \\ 1 & 2 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & 1 \\ 0 & 1 & 6 & 2 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

$$\sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & 1.5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & .5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

De acuerdo con el teorema 4, $1.5\,\mathbf{v}_1 - \mathbf{v}_2 + .5\mathbf{v}_3 = \mathbf{p}$. Véase la figura 4, que muestra el plano que contiene a $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{p}$ (junto con los puntos sobre los ejes coordenados). ■

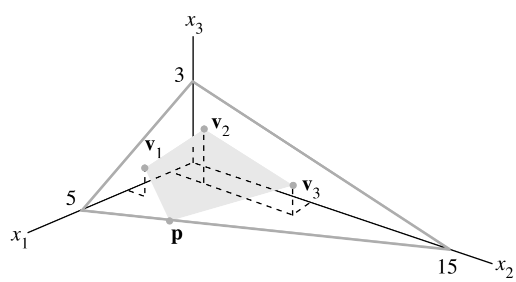


FIGURA 4

PROBLEMA DE PRÁCTICA

Dibuje los puntos $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ en un hoja de papel para gráficas, y explique por qué $\mathbf{p}$ *debe* ser una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Después, encuentre la combinación afín para $\mathbf{p}$. [*Sugerencia:* Pregúntese cuál es la dimensión de aff$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$].

8.1 EJERCICIOS

En los ejercicios 1 a 4, escriba $\mathbf{y}$ como una combinación afín de los otros puntos indicados, si es posible.

1. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$

3. $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 4 \\ -2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ -2 \\ 6 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 17 \\ 1 \\ 5 \end{bmatrix}$

4. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ -6 \\ 7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 3 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -3 \\ 4 \\ -4 \end{bmatrix}$

En los ejercicios 5 y 6, sean $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}$ y $S = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. Observe que S es una base ortogonal para $\mathbb{R}^3$. Escriba cada uno de los puntos dados como una combinación afín de los puntos del conjunto S, si es posible. [*Sugerencia:* Para encontrar los pesos, aplique el teorema 5 de la sección 6.2 en vez de la reducción por filas].

5. a. $\mathbf{p}_1 = \begin{bmatrix} 3 \\ 8 \\ 4 \end{bmatrix}$ b. $\mathbf{p}_2 = \begin{bmatrix} 6 \\ -3 \\ 3 \end{bmatrix}$ c. $\mathbf{p}_3 = \begin{bmatrix} 0 \\ -1 \\ -5 \end{bmatrix}$

6. a. $\mathbf{p}_1 = \begin{bmatrix} 0 \\ -19 \\ -5 \end{bmatrix}$ b. $\mathbf{p}_2 = \begin{bmatrix} 1.5 \\ -1.3 \\ -.5 \end{bmatrix}$ c. $\mathbf{p}_3 = \begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix}$

7. Sean

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 1 \end{bmatrix},$$

$$\mathbf{p}_1 = \begin{bmatrix} 5 \\ -3 \\ 5 \\ 3 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} -9 \\ 10 \\ 9 \\ -13 \end{bmatrix}, \quad \mathbf{p}_3 = \begin{bmatrix} 4 \\ 2 \\ 8 \\ 5 \end{bmatrix}$$

y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Es posible demostrar que S es linealmente independiente.

a) ¿Está $\mathbf{p}_1$ en Gen S? ¿Se encuentra $\mathbf{p}_1$ en aff S?

b) ¿Se encuentra $\mathbf{p}_2$ en Gen S? ¿Está $\mathbf{p}_2$ en aff S?

c) ¿Está $\mathbf{p}_3$ en Gen S? ¿Se encuentra $\mathbf{p}_3$ en aff S?

8. Repita el ejercicio 7 considerando

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 3 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 6 \\ -5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 12 \\ -6 \end{bmatrix},$$

$$\mathbf{p}_1 = \begin{bmatrix} 4 \\ -1 \\ 15 \\ -7 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} -5 \\ 3 \\ -8 \\ 6 \end{bmatrix} \quad \text{y} \quad \mathbf{p}_3 = \begin{bmatrix} 1 \\ 6 \\ -6 \\ -8 \end{bmatrix}$$

9. Suponga que las soluciones de una ecuación $A\mathbf{x} = \mathbf{b}$ son todas de la forma $\mathbf{x} = x_3\mathbf{u} + \mathbf{p}$, donde $\mathbf{u} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$.

Encuentre los puntos $\mathbf{v}_1$ y $\mathbf{v}_2$ tales que el conjunto solución de $A\mathbf{x} = \mathbf{b}$ sea aff$\{\mathbf{v}_1, \mathbf{v}_2\}$.

10. Suponga que las soluciones de una ecuación $A\mathbf{x} = \mathbf{b}$ son todas de la forma $\mathbf{x} = x_3\mathbf{u} + \mathbf{p}$, donde $\mathbf{u} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$.

Obtenga puntos $\mathbf{v}_1$ y $\mathbf{v}_2$ tales que el conjunto solución de $A\mathbf{x} = \mathbf{b}$ sea aff$\{\mathbf{v}_1, \mathbf{v}_2\}$.

En los ejercicios 11 y 12, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) El conjunto de todas las combinaciones afines de puntos en un conjunto S se denomina envolvente afín de S.

b) Si $\{\mathbf{b}_1,\dots, \mathbf{b}_k\}$ es un subconjunto linealmente independiente de $\mathbb{R}^n$ y si $\mathbf{p}$ es una combinación lineal de $\mathbf{b}_1,\dots, \mathbf{b}_k$, entonces $\mathbf{p}$ es una combinación afín de $\mathbf{b}_1,\dots, \mathbf{b}_k$.

c) La envolvente afín de dos puntos distintos se llama recta.

d) Un plano afín es un subespacio.

e) El plano en $\mathbb{R}^3$ es un hiperplano.

12. *a*) Si $S = \{\mathbf{x}\}$, entonces aff S es el conjunto vacío.

b) Un conjunto es afín si y solo si contiene su envolvente afín.

c) Un plano afín de dimensión 1 se llama recta.

d) Un plano afín de dimensión 2 se llama hiperplano.

e) Un plano afín que pasa por el origen es un subespacio.

13. Suponga que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$. Demuestre que Gen $\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$ es un plano en $\mathbb{R}^3$. [*Sugerencia:* Piense qué se puede decir sobre $\mathbf{u}$ y $\mathbf{v}$ cuando Gen $\{\mathbf{u}, \mathbf{v}\}$ es un plano].

14. Demuestre que si $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$, entonces aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es el plano a través de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

15. Sea A una matriz de $m \times n$ y, si $\mathbf{b}$ está en $\mathbb{R}^m$, demuestre que el conjunto S de todas las soluciones de $A\mathbf{x} = \mathbf{b}$ es un subconjunto afín de R^n.

16. Considere que $\mathbf{v} \in \mathbb{R}^n$ y $k \in \mathrm{R}$. Demuestre que $S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}\cdot\mathbf{v} = k\}$ es un subconjunto afín de $\mathbb{R}^n$.

17. Seleccione un conjunto S de tres puntos tales que aff S sea el plano en $\mathbb{R}^3$ cuya ecuación es $x_3 = 5$. Justifique su trabajo.

18. Seleccione un conjunto S de cuatro puntos distintos en $\mathbb{R}^3$ tal que aff S sea el plano $2x_1 + x_2 - 3x_3 = 12$. Justifique su trabajo.

19. Sea S un subconjunto afín de $\mathbb{R}^n$, y suponga que $f: \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal, y que $f(S)$ denota el conjunto de imágenes $\{f(\mathbf{x}): \mathbf{x} \in S\}$. Demuestre que $f(S)$ es un subconjunto afín de $\mathbb{R}^m$.

20. Sean $f : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal, T un subconjunto afín de R^m, y $S = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \in T\}$. Demuestre que S es un subconjunto afín de R^n.

En los ejercicios 21 a 26, demuestre el enunciado sobre subconjuntos A y B de $\mathbb{R}^n$, o dé el ejemplo requerido en $\mathbb{R}^2$. Una demostración para un ejercicio puede utilizar los resultados obtenidos en los ejercicios anteriores (así como los teoremas disponibles en el libro).

21. Si $A \subset B$ y B es afín, entonces aff $A \subset B$.

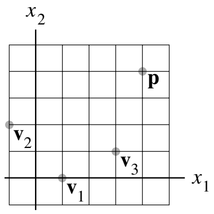
22. Si $A \subset B$, entonces aff $A \subset$ aff B.

23. $[(\text{aff } A) \cup (\text{aff } B)] \subset \text{aff } (A \cup B)$. [*Sugerencia:* Para demostrar que $D \cup E \subset F$, demuestre que $D \subset F$ y $E \subset F$].

24. Encuentre un ejemplo en $\mathbb{R}^2$ para demostrar que la igualdad no necesariamente es válida en el enunciado del ejercicio 23. [*Sugerencia:* Considere los conjuntos A y B, cada uno los cuales solamente contiene uno o dos puntos].

25. $\text{aff } (A \cap B) \subset (\text{aff } A \cap \text{aff } B)$

26. Encuentre un ejemplo en $\mathbb{R}^2$ para demostrar que la igualdad no necesita ser válida en el enunciado del ejercicio 25.



SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Como los puntos $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ no son colineales (es decir, no están sobre una recta), aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ no puede ser unidimensional. Por lo tanto, aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ debe ser igual a $\mathbb{R}^2$. Para encontrar los pesos reales utilizados para expresar $\mathbf{p}$ como una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, primero calcule

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{p} - \mathbf{v}_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Para escribir $\mathbf{p} - \mathbf{v}_1$ como una combinación lineal de $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$, se reduce por filas la matriz que tiene estos puntos como columnas:

$$\begin{bmatrix} -2 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 2 \end{bmatrix}$$

Así, $\mathbf{p} - \mathbf{v}_1 = \frac{1}{2}(\mathbf{v}_2 - \mathbf{v}_1) + 2(\mathbf{v}_3 - \mathbf{v}_1)$, lo que muestra que

$$\mathbf{p} = \left(1 - \tfrac{1}{2} - 2\right)\mathbf{v}_1 + \tfrac{1}{2}\mathbf{v}_2 + 2\mathbf{v}_3 = -\tfrac{3}{2}\mathbf{v}_1 + \tfrac{1}{2}\mathbf{v}_2 + 2\mathbf{v}_3$$

Esto expresa $\mathbf{p}$ como una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, porque los coeficientes suman 1.

Alternativamente, utilice el método del ejemplo 3 y reduzca por filas:

$$\begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{p} \\ 1 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 3 & 4 \\ 0 & 2 & 1 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -\frac{3}{2} \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Esto demuestra que $\mathbf{p} = -\frac{3}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 + 2\mathbf{v}_3$.

8.2 INDEPENDENCIA AFÍN

En esta sección se continuará explorando la relación entre conceptos lineales y conceptos afines. Primero considere un conjunto de tres vectores en $\mathbb{R}^3$, por ejemplo, $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Si S es linealmente dependiente, entonces uno de los vectores es una combinación lineal de los otros dos vectores. ¿Qué ocurre cuando uno de los vectores es una combinación *afín* de los otros? Por ejemplo, suponga que

$$\mathbf{v}_3 = (1 - t)\mathbf{v}_1 + t\mathbf{v}_2, \quad \text{para alguna } t \text{ en } \mathbb{R}.$$

Entonces,

$$(1 - t)\mathbf{v}_1 + t\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$$

Esta es una relación de dependencia lineal ya que no todos los pesos son cero. Pero hay algo más que es verdad: los pesos en la relación de dependencia suman cero:

$$(1 - t) + t + (-1) = 0$$

Esta es la propiedad adicional necesaria para definir *dependencia afín*.

DEFINICIÓN

Un conjunto indexado de puntos $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es **afínmente dependiente** si existen números reales $c_1, \ldots, c_p$, no todos cero, tales que

$$c_1 + \cdots + c_p = 0 \quad \text{y} \quad c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{1}$$

De lo contrario, será **afínmente independiente**.

Una combinación afín es un tipo especial de combinación lineal, y la dependencia afín es un tipo restringido de dependencia lineal. Por lo tanto, cada conjunto afínmente dependiente es, de manera automática, linealmente dependiente.

Un conjunto $\{\mathbf{v}_1\}$ de solo un punto (aun el vector cero) debe ser afínmente independiente porque las propiedades requeridas de los coeficientes c_i no se pueden satisfacer cuando solo existe un coeficiente. Para $\{\mathbf{v}_1\}$, la primera ecuación en (1) es justamente $c_1 = 0$, y al menos un coeficiente (el único) debe ser distinto de cero.

En el ejercicio 13 se le pide demostrar que un conjunto indexado $\{\mathbf{v}_1, \mathbf{v}_2\}$ es afínmente dependiente si y solo si $\mathbf{v}_1 = \mathbf{v}_2$. El siguiente teorema trata el caso general y muestra cómo el concepto de dependencia afín es análogo al de dependencia lineal. Los incisos *c*) y *d*) brindan métodos útiles para determinar si un conjunto es afínmente dependiente. De la sección 8.1, recuerde que si $\mathbf{v}$ está en $\mathbb{R}^n$, entonces el vector $\tilde{\mathbf{v}}$ en $\mathbb{R}^{n+1}$ denota la forma homogénea de $\mathbf{v}$.

TEOREMA 5

Dado un conjunto indexado $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$, con $p \geq 2$, los siguientes enunciados son lógicamente equivalentes. Es decir, todos ellos son verdaderos o todos son falsos.

a) S es afínmente dependiente.

b) Uno de los puntos en S es una combinación afín de los demás puntos en S.

c) El conjunto $\{\mathbf{v}_2 - \mathbf{v}_1, \ldots, \mathbf{v}_p - \mathbf{v}_1\}$ en $\mathbb{R}^n$ es linealmente dependiente.

d) El conjunto $\{\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_p\}$ de formas homogéneas en $\mathbb{R}^{n+1}$ es linealmente dependiente.

DEMOSTRACIÓN Suponga que el enunciado *a*) es verdadero, y sea que $c_1, \ldots, c_p$ satisfagan (1). Al renombrar los puntos en caso necesario, se puede suponer que $c_1 \neq 0$ y dividir ambas ecuaciones en (1) entre c_1, de manera que $1 + (c_2/c_1) + \cdots + (c_p/c_1) = 0$ y

$$\mathbf{v}_1 = (-c_2/c_1)\mathbf{v}_2 + \cdots + (-c_p/c_1)\mathbf{v}_p \tag{2}$$

Observe que los coeficientes en el lado derecho de (2) suman 1. Así, *a*) implica *b*). Ahora, suponga que *b*) es verdadero. Al renombrar los puntos en caso necesario, se puede suponer que $\mathbf{v}_1 = c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$, donde $c_2 + \cdots + c_p = 1$. Entonces,

$$(c_2 + \cdots + c_p)\mathbf{v}_1 = c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p \tag{3}$$

y

$$c_2(\mathbf{v}_2 - \mathbf{v}_1) + \cdots + c_p(\mathbf{v}_p - \mathbf{v}_1) = \mathbf{0} \tag{4}$$

No todos los coeficientes $c_2, \ldots, c_p$ pueden ser cero porque suman 1. Así, *b*) implica *c*).

Ahora si *c*) es verdadero, entonces existen pesos $c_2, \ldots, c_p$, no todos cero, tales que la ecuación (4) se cumple. Se rescribe (4) como (3) y se establece $c_1 = -(c_2 + \cdots + c_p)$. Entonces, $c_1 + \cdots + c_p = 0$. Así, (3) muestra que (1) es verdadera. Por lo que *c*) implica *a*), lo que prueba que *a*), *b*) y *c*) son lógicamente equivalentes. Por último, *d*) es equivalente a *a*) porque las dos ecuaciones en (1) son equivalentes a la siguiente ecuación que implica las formas homogéneas de los puntos en S:

$$c_1\begin{bmatrix}\mathbf{v}_1\\1\end{bmatrix} + \cdots + c_p\begin{bmatrix}\mathbf{v}_p\\1\end{bmatrix} = \begin{bmatrix}\mathbf{0}\\0\end{bmatrix} \qquad \blacksquare$$

En el enunciado *c*) del teorema 5, $\mathbf{v}_1$ se podría remplazar por cualquiera de los otros puntos en la lista $\mathbf{v}_1, \ldots, \mathbf{v}_p$. En la demostración, solo cambiaría la notación. Así, para determinar si un conjunto es afínmente dependiente, se resta un punto en el conjunto de los demás puntos, y se comprueba si el conjunto trasladado de $p - 1$ puntos es linealmente dependiente.

EJEMPLO 1 La envolvente afín de dos distintos puntos $\mathbf{p}$ y $\mathbf{q}$ es una recta. Si un tercer punto $\mathbf{r}$ está sobre la recta, entonces $\{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$ es un conjunto afínmente dependiente. Si un punto $\mathbf{s}$ no está sobre la recta que pasa por $\mathbf{p}$ y $\mathbf{q}$, entonces esos tres puntos no son colineales y $\{\mathbf{p}, \mathbf{q}, \mathbf{s}\}$ es un conjunto afínmente independiente. Véase la figura 1. ■

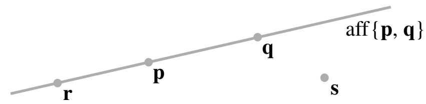


FIGURA 1 $\{\mathbf{p}, \mathbf{q}, \mathbf{r}\}$ es afínmente dependiente.

EJEMPLO 2 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 7 \\ 6.5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \\ 7 \end{bmatrix}$ y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Determine si S es afínmente independiente.

SOLUCIÓN Se calcula $\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \\ -.5 \end{bmatrix}$ y $\mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$. Esos dos puntos no son múltiplos, así que forman un conjunto linealmente independiente, S'. De manera que son falsos todos los enunciados del teorema 5, y S es afínmente independiente. La figura 2 muestra a S y al conjunto trasladado S'. Observe que Gen S' es un plano que pasa por el origen y aff S es un plano paralelo que pasa por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. (Desde luego, aquí solo se muestra una parte de cada plano). ■

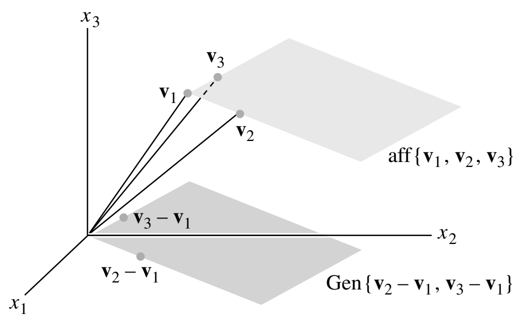


FIGURA 2 Un conjunto $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ afínmente independiente.

EJEMPLO 3 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 7 \\ 6.5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \\ 7 \end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix} 0 \\ 14 \\ 6 \end{bmatrix}$, y $S = \{\mathbf{v}_1, \dots, \mathbf{v}_4\}$. ¿$S$ es afínmente dependiente?

SOLUCIÓN Se calcula $\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \\ -.5 \end{bmatrix}$, $\mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ y $\mathbf{v}_4 - \mathbf{v}_1 = \begin{bmatrix} -1 \\ 11 \\ -1 \end{bmatrix}$, y se reduce por filas la matriz:

$$\begin{bmatrix} 1 & -1 & -1 \\ 4 & 1 & 11 \\ -.5 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 5 & 15 \\ 0 & -.5 & -1.5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 5 & 15 \\ 0 & 0 & 0 \end{bmatrix}$$

Recuerde, de la sección 4.6 (o de la sección 2.8) que las columnas son linealmente dependientes porque no toda columna es una columna pivote; así que $\mathbf{v}_2 - \mathbf{v}_1$, $\mathbf{v}_3 - \mathbf{v}_1$ y $\mathbf{v}_4 - \mathbf{v}_1$

son linealmente dependientes. De acuerdo con el enunciado *c*) del teorema 5, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ es afínmente dependiente. Esta dependencia también se puede establecer empleando el inciso *d*) del teorema 5 en vez del *c*). ■

Los cálculos del ejemplo 3 muestran que $\mathbf{v}_4 - \mathbf{v}_1$ es una combinación lineal de $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$, lo que significa que $\mathbf{v}_4 - \mathbf{v}_1$ está en Gen $\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$. De acuerdo con el teorema 1 de la sección 8.1, $\mathbf{v}_4$ se encuentra en aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. De hecho, si en el ejemplo 3 se completara la reducción por filas de la matriz se mostraría que

$$\mathbf{v}_4 - \mathbf{v}_1 = 2(\mathbf{v}_2 - \mathbf{v}_1) + 3(\mathbf{v}_3 - \mathbf{v}_1) \tag{5}$$

$$\mathbf{v}_4 = -4\mathbf{v}_1 + 2\mathbf{v}_2 + 3\mathbf{v}_3 \tag{6}$$

Véase la figura 3.

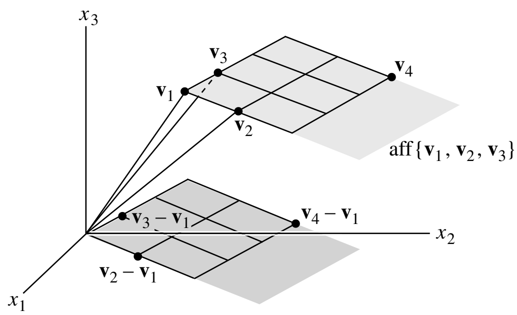


FIGURA 3 $\mathbf{v}_4$ está en el plano aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

La figura 3 muestra rejillas sobre Gen$\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$ y aff$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. La rejilla sobre aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ se basa en la ecuación (5). Otro "sistema de coordenadas" se puede basar en la ecuación (6), en la cual los coeficientes -4, 2 y 3 son las coordenadas *afines* o *baricéntricas* de $\mathbf{v}_4$.

Coordenadas baricéntricas

La definición de coordenadas baricéntricas depende de la siguiente versión afín del teorema de representación única de la sección 4.4. Para la demostración, véase el ejercicio 17 de esta sección.

TEOREMA 6

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto afínmente independiente en $\mathbb{R}^n$. Entonces, cada $\mathbf{p}$ en aff S tiene una única representación como una combinación afín de $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Es decir, para cada $\mathbf{p}$ existe un único conjunto de escalares $c_1, \ldots, c_k$ tales que

$$\mathbf{p} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \quad \text{y} \quad c_1 + \cdots + c_k = 1 \tag{7}$$

DEFINICIÓN

Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto afínmente independiente. Así, para cada punto $\mathbf{p}$ en aff S, los coeficientes $c_1, \ldots, c_p$ en la representación única (7) de $\mathbf{p}$ son las **coordenadas baricéntricas** (que, en ocasiones, también se conocen como **coordenadas afines**) de $\mathbf{p}$.

Observe que (7) es equivalente a la ecuación

$$\begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} \mathbf{v}_1 \\ 1 \end{bmatrix} + \cdots + c_k \begin{bmatrix} \mathbf{v}_k \\ 1 \end{bmatrix} \tag{8}$$

que implica las formas homogéneas de los puntos. La reducción por filas de la matriz aumentada $\begin{bmatrix} \tilde{\mathbf{v}}_1 & \cdots & \tilde{\mathbf{v}}_k & \tilde{\mathbf{p}} \end{bmatrix}$ para (8) produce las coordenadas baricéntricas de $\mathbf{p}$.

EJEMPLO 4 Sean $\mathbf{a} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 9 \\ 3 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$. Encuentre las coordenadas baricéntricas de $\mathbf{p}$ determinadas por el conjunto afínmente independiente $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

SOLUCIÓN Reduzca por filas la matriz aumentada de puntos en forma homogénea, moviendo la última fila de unos hacia la parte superior para así simplificar la aritmética:

$$\begin{bmatrix} \tilde{\mathbf{a}} & \tilde{\mathbf{b}} & \tilde{\mathbf{c}} & \tilde{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} 1 & 3 & 9 & 5 \\ 7 & 0 & 3 & 3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 9 & 5 \\ 7 & 0 & 3 & 3 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{3} \\ 0 & 0 & 1 & \frac{5}{12} \end{bmatrix}$$

Las coordenadas son $\frac{1}{4}$, $\frac{1}{3}$ y $\frac{5}{12}$ de manera que $\mathbf{p} = \frac{1}{4}\mathbf{a} + \frac{1}{3}\mathbf{b} + \frac{5}{12}\mathbf{c}$. ■

Las coordenadas baricéntricas tienen interpretaciones físicas y geométricas. Originalmente, A. F. Moebius las definió en 1827 para un punto $\mathbf{p}$ dentro de una región triangular con vértices $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$. Moebius escribió que las coordenadas baricéntricas de $\mathbf{p}$ son tres números no negativos $m_{\mathbf{a}}$, $m_{\mathbf{b}}$ y $m_{\mathbf{c}}$ tales que $\mathbf{p}$ es el centro de masa de un sistema que consiste en el triángulo (sin masa) y las masas $m_{\mathbf{a}}$, $m_{\mathbf{b}}$ y $m_{\mathbf{c}}$ en los vértices correspondientes. Las masas están unívocamente determinadas al requerir que su suma sea 1. Este punto de vista es útil en la física actual.[1]

La figura 4 da una interpretación geométrica a las coordenadas baricéntricas del ejemplo 4, al mostrar el triángulo $\Delta\mathbf{abc}$ y tres pequeños triángulos $\Delta\mathbf{pbc}$, $\Delta\mathbf{apc}$ y $\Delta\mathbf{abp}$. Las áreas de los pequeños triángulos son proporcionales a las coordenadas baricéntricas de $\mathbf{p}$. En efecto,

$$\begin{aligned} \text{área}(\Delta\mathbf{pbc}) &= \frac{1}{4} \cdot \text{área}(\Delta\mathbf{abc}) \\ \text{área}(\Delta\mathbf{apc}) &= \frac{1}{3} \cdot \text{área}(\Delta\mathbf{abc}) \\ \text{área}(\Delta\mathbf{abp}) &= \frac{5}{12} \cdot \text{área}(\Delta\mathbf{abc}) \end{aligned} \tag{9}$$

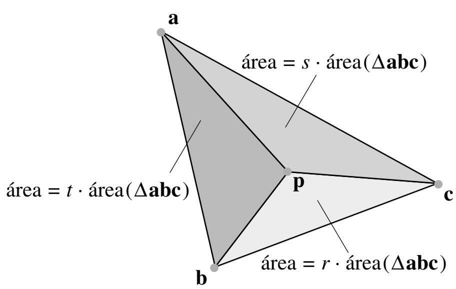


FIGURA 4 $\mathbf{p} = r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$. Aquí, $r = \frac{1}{4}$, $s = \frac{1}{3}$, $t = \frac{5}{12}$.

En los ejercicios 21 a 23 se comprueban las fórmulas de la figura 4. Las igualdades análogas para volúmenes de tetraedros son válidas para el caso en que $\mathbf{p}$ es un punto dentro de un tetraedro en $\mathbb{R}^3$, con vértices en $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ y $\mathbf{d}$.

[1] Véase el ejercicio 29 de la sección 1.3. En astronomía, sin embargo, el término "coordenadas baricéntricas" por lo general se refiere a las coordenadas ordinarias de puntos en $\mathbb{R}^3$ y que ahora se conoce como *Sistema de referencia celeste internacional*, un sistema de coordenadas cartesianas para el espacio exterior, con el origen en el centro de masa (el baricentro) del Sistema Solar.

Cuando un punto no está dentro del triángulo (o del tetraedro), algunas o todas las coordenadas baricéntricas serán negativas. En la figura 5 se ilustra el caso de un triángulo, para los vértices **a**, **b**, **c**, y los valores coordenados anteriores r, s, t. Por ejemplo, los puntos sobre la recta que pasa por **b** y **c** tienen $r = 0$ porque son combinaciones afines solamente de **b** y **c**. La recta paralela que pasa por **a** identifica los puntos con $r = 1$.

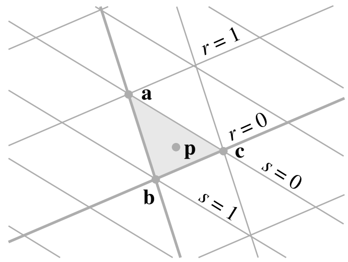


FIGURA 5 Coordenadas baricéntricas para puntos en aff $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

Coordenadas baricéntricas en gráficos generados por computadora

Cuando un diseñador trabaja con objetos geométricos en un programa de generación de gráficos, puede utilizar una aproximación de "malla de alambre" para un objeto en ciertos puntos clave durante el proceso con la finalidad de crear una imagen final realista.[2] Por ejemplo, si la superficie de una parte del objeto consiste en pequeñas superficies triangulares planas, entonces el programa de gráficos fácilmente puede agregar color, iluminación y sombreado a cada pequeña superficie cuando solo se conoce esa información en los vértices. Las coordenadas baricéntricas ofrecen la herramienta para interpolar suavemente la información de los vértices hacia el interior de un triángulo. La interpolación en un punto es simplemente la combinación lineal de los valores en los vértices utilizando como pesos las coordenadas baricéntricas.

Con frecuencia, los colores en un monitor de computadora se describen mediante las coordenadas RGB. Una terna (r, g, b) indica la cantidad de cada color (rojo, verde y azul) con parámetros que varían de 0 a 1. Por ejemplo, rojo puro es (1, 0, 0), blanco es (1, 1, 1) y negro es (0, 0, 0).

EJEMPLO 5 Sean $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ 3 \\ 4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 5 \\ 1 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 3 \\ 3 \\ 3.5 \end{bmatrix}$. Los colores en los vértices $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ de un triángulo son magenta (1, 0, 1), magenta claro (1, .4, 1) y púrpura (.6, 0, 1), respectivamente. Encuentre el color interpolado en **p**. Véase la figura 6.

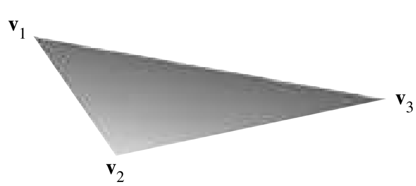


FIGURA 6 Colores interpolados.

[2] El ejemplo introductorio del capítulo 2 menciona un modelo de alambre para un avión Boeing 777, empleado para visualizar el flujo de aire sobre la superficie de la aeronave.

SOLUCIÓN Primero, determine las coordenadas baricéntricas de $\mathbf{p}$. Aquí se presenta el cálculo utilizando las formas homogéneas de los puntos; el primer paso consiste en mover la fila 4 a la fila 1:

$$\begin{bmatrix} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 & \tilde{\mathbf{p}} \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 3 & 4 & 1 & 3 \\ 1 & 3 & 5 & 3 \\ 5 & 4 & 1 & 3.5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & .25 \\ 0 & 1 & 0 & .50 \\ 0 & 0 & 1 & .25 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Así, $\mathbf{p} = .25\mathbf{v}_1 + .5\mathbf{v}_2 + .25\mathbf{v}_3$. Utilice las coordenadas baricéntricas de $\mathbf{p}$ para hacer una combinación lineal de los datos de color. Los valores RGB para $\mathbf{p}$ son

$$.25\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + .50\begin{bmatrix} 1 \\ .4 \\ 1 \end{bmatrix} + .25\begin{bmatrix} .6 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} .9 \\ .2 \\ 1 \end{bmatrix} \begin{matrix} \text{rojo} \\ \text{verde} \\ \text{azul} \end{matrix} \qquad \blacksquare$$

Uno de los últimos pasos en la preparación de la escena gráfica para mostrarla en el monitor consiste en eliminar las "superficies ocultas" que no deberían ser visibles en la pantalla. Imagine que la pantalla de visualización está formada por un millón de pixeles, y considere un rayo o una "línea de visión" desde el ojo del observador hacia un pixel y la colección de objetos que forman la escena 3D. El color y el resto de la información que se muestra en el pixel sobre el monitor deberían provenir del primer objeto sobre el que incide el rayo. Véase la figura 7. Cuando los objetos en la escena gráfica se aproximan por mallas de alambre con parches triangulares, el problema de la superficie oculta se puede resolver empleando coordenadas baricéntricas.

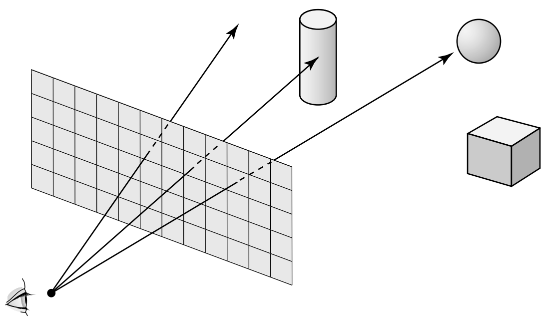


FIGURA 7 Un rayo proveniente del ojo cruza la pantalla hacia el objeto más cercano.

También es posible utilizar matemáticas para encontrar las intersecciones del triángulo de rayos para realizar sombreados extremadamente realistas de objetos. Actualmente, este método de *trazado de rayos* es muy lento para trabajos en tiempo real, pero esto podría cambiar en el futuro gracias a recientes avances en implementaciones de hardware.[3]

EJEMPLO 6 Sean

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 8 \\ 1 \\ -4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ 11 \\ -2 \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} .7 \\ .4 \\ -3 \end{bmatrix},$$

y $\mathbf{x}(t) = \mathbf{a} + t\mathbf{b}$ para $t \geq 0$. Determine el punto donde el rayo $\mathbf{x}(t)$ se interseca con el plano que contiene el triángulo con vértices $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. ¿Este punto se encuentra dentro del triángulo?

[3] Véase Joshua Fender y Jonathan Rose, "A High-Speed Ray Tracing Engine Built on a Field-Programmable System", en *Proc. Int. Conf. on Field-Programmable Technology*, IEEE (2003). (Un solo procesador puede calcular 600 millones de intersecciones de triángulos de rayos por segundo).

SOLUCIÓN El plano es aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Un punto típico en este plano se puede escribir como $(1 - c_2 - c_3)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ para algunos c_2 y c_3. (En esta combinación los pesos suman 1). El rayo $\mathbf{x}(t)$ hace intersección con el plano cuando c_2, c_3 y t satisfacen

$$(1 - c_2 - c_3)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{a} + t\mathbf{b}$$

Reacomode esto como $c_2(\mathbf{v}_2 - \mathbf{v}_1) + c_3(\mathbf{v}_3 - \mathbf{v}_1) + t(-\mathbf{b}) = \mathbf{a} - \mathbf{v}_1$. En forma matricial,

$$\begin{bmatrix} \mathbf{v}_2 - \mathbf{v}_1 & \mathbf{v}_3 - \mathbf{v}_1 & -\mathbf{b} \end{bmatrix} \begin{bmatrix} c_2 \\ c_3 \\ t \end{bmatrix} = \mathbf{a} - \mathbf{v}_1$$

Para los puntos específicos dados aquí,

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 7 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} 4 \\ 10 \\ 4 \end{bmatrix}, \quad \mathbf{a} - \mathbf{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 16 \end{bmatrix}$$

La reducción por filas de la matriz aumentada anterior produce

$$\begin{bmatrix} 7 & 4 & -.7 & -1 \\ 0 & 10 & -.4 & -1 \\ 2 & 4 & 3 & 16 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & .3 \\ 0 & 1 & 0 & .1 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

Así, $c_2 = .3$, $c_3 = .1$, y $t = 5$. Por lo tanto, el punto de intersección es

$$\mathbf{x}(5) = \mathbf{a} + 5\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 10 \end{bmatrix} + 5\begin{bmatrix} .7 \\ .4 \\ -3 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 2.0 \\ -5.0 \end{bmatrix}$$

Además,

$$\begin{aligned} \mathbf{x}(5) &= (1 - .3 - .1)\mathbf{v}_1 + .3\mathbf{v}_2 + .1\mathbf{v}_3 \\ &= .6\begin{bmatrix} 1 \\ 1 \\ -6 \end{bmatrix} + .3\begin{bmatrix} 8 \\ 1 \\ -4 \end{bmatrix} + .1\begin{bmatrix} 5 \\ 11 \\ -2 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 2.0 \\ -5.0 \end{bmatrix} \end{aligned}$$

El punto de intersección está dentro del triángulo porque todos los pesos baricéntricos para $\mathbf{x}(5)$ son positivos. ■

PROBLEMAS DE PRÁCTICA

1. Describa un método rápido para determinar cuándo tres puntos son colineales.

2. Los puntos $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$ y $\mathbf{v}_4 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ forman un conjunto afínmente dependiente. Encuentre los pesos $c_1, \ldots, c_4$ que produzcan una **relación de dependencia afín** $c_1\mathbf{v}_1 + \cdots + c_4\mathbf{v}_4 = \mathbf{0}$, donde $c_1 + \cdots + c_4 = 0$, pero no todas las c_i son ceros. [*Sugerencia:* Véase el final de la demostración del teorema 5].

8.2 EJERCICIOS

En los ejercicios 1 a 6, determine si el conjunto de puntos es afínmente dependiente. (Véase el problema de práctica 2). Si es así, entonces construya una relación de dependencia afín para los puntos.

1. $\begin{bmatrix} 3 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ **2.** $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 11 \end{bmatrix}, \begin{bmatrix} 0 \\ 15 \\ -9 \end{bmatrix}$

4. $\begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -6 \end{bmatrix}, \begin{bmatrix} -2 \\ 7 \\ -3 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ -3 \end{bmatrix}$

6. $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 0 \end{bmatrix}$

En los ejercicios 7 y 8, encuentre las coordenadas baricéntricas de **p** con respecto al conjunto de puntos afínmente independiente que lo precede.

7. $\begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \\ 0 \end{bmatrix}, \mathbf{p} = \begin{bmatrix} 5 \\ 4 \\ -2 \\ 2 \end{bmatrix}$

8. $\begin{bmatrix} 0 \\ 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -6 \\ 5 \end{bmatrix}, \mathbf{p} = \begin{bmatrix} -1 \\ 1 \\ -4 \\ 0 \end{bmatrix}$

En los ejercicios 9 y 10, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

9. *a*) Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en $\mathbb{R}^n$ y si el conjunto $\{\mathbf{v}_1 - \mathbf{v}_2, \mathbf{v}_3 - \mathbf{v}_2, \ldots, \mathbf{v}_p - \mathbf{v}_2\}$ es linealmente dependiente, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es afínmente dependiente. (Lea esto con sumo cuidado).

b) Si $\mathbf{v}_1, \ldots, \mathbf{v}_p$ están en $\mathbb{R}^n$ y si el conjunto de formas homogéneas $\{\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_p\}$ en $\mathbb{R}^{n+1}$ es linealmente independiente, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ es afínmente dependiente.

c) Un conjunto finito de puntos $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es afínmente dependiente si existen números reales $c_1, \ldots, c_k$, no todos ceros, tales que $c_1 + \cdots + c_k = 1$ y $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.

d) Si $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es afínmente independiente y si **p** en $\mathbb{R}^n$ tiene una coordenada baricéntrica negativa determinada por S, entonces **p** no está en aff S.

e) Si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, **a** y **b** están en $\mathbb{R}^3$ y si un rayo $\mathbf{a} + t\mathbf{b}$ para $t \geq 0$ se interseca con el triángulo de vértices $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, entonces todas las coordenadas baricéntricas del punto de intersección son no negativas.

10. *a*) Si $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es un conjunto afínmente dependiente, entonces el conjunto $\{\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_p\}$ en $\mathbb{R}^{n+1}$ de formas homogéneas puede ser linealmente independiente.

b) Si $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$ están en $\mathbb{R}^3$ y si el conjunto $\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1, \mathbf{v}_4 - \mathbf{v}_1\}$ es linealmente independiente, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ es afínmente independiente.

c) Dado $S = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ en $\mathbb{R}^n$, cada **p** en aff S tiene representación única como una combinación afín de $\mathbf{b}_1, \ldots, \mathbf{b}_k$.

d) Cuando una información de color se especifica en cada vértice $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ de un triángulo en $\mathbb{R}^3$, entonces el color se puede interpolar en un punto **p** en aff $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ empleando las coordenadas baricéntricas de **p**.

e) Si T es un triángulo en $\mathbb{R}^2$ y si un punto **p** está sobre una arista del triángulo, entonces no todas las coordenadas baricéntricas de **p** (para este triángulo) son positivas.

11. Explique por qué cualquier conjunto de cinco o más puntos en $\mathbb{R}^3$ debe tener dependencia afín.

12. Demuestre que un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ es afínmente dependiente cuando $p \geq n + 2$.

13. Utilice sólo la definición de dependencia afín para demostrar que un conjunto indexado $\{\mathbf{v}_1, \mathbf{v}_2\}$ en $\mathbb{R}^n$ es afínmente dependiente si y solo si $\mathbf{v}_1 = \mathbf{v}_2$.

14. Las condiciones para dependencia afín son más estrictas que para dependencia lineal, así que un conjunto afínmente dependiente, de manera automática, es linealmente dependiente. Además, un conjunto linealmente independiente no puede ser afínmente dependiente y, por lo tanto, debe ser afínmente independiente. Construya dos conjuntos indexados linealmente dependientes S_1 y S_2 en $\mathbb{R}^2$ tales que S_1 y S_2 tengan dependencia e independencia afines, respectivamente. En cada caso, el conjunto deberá contener uno, dos o tres puntos diferentes de cero.

15. Sean $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ y sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

a) Demuestre que el conjunto S es afínmente independiente.

b) Encuentre las coordenadas baricéntricas de $\mathbf{p}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{p}_3 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\mathbf{p}_4 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{p}_5 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, con respecto a S.

c) Sea T el triángulo con vértices $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Cuando se extienden los lados de T, las rectas dividen $\mathbb{R}^2$ en siete regiones. Véase la figura 8. Observe los signos de las coordenadas baricéntricas de los puntos en cada región. Por ejemplo, $\mathbf{p}_5$ está dentro del triángulo T y todas sus coordenadas baricéntricas son positivas. El punto $\mathbf{p}_1$ tiene coordenadas $(-, +, +)$. Su tercera coordenada es positiva porque $\mathbf{p}_1$ está sobre el lado $\mathbf{v}_3$ de la recta que pasa por $\mathbf{v}_1$ y $\mathbf{v}_2$. Su primera coordenada es negativa porque $\mathbf{p}_1$ es opuesto al lado $\mathbf{v}_1$ de la recta que pasa por $\mathbf{v}_2$ y $\mathbf{v}_3$. El punto $\mathbf{p}_2$ está sobre la arista $\mathbf{v}_2\mathbf{v}_3$ de T. Sus coordenadas son $(0, +, +)$. Sin calcular los valores reales, determine los signos de las coordenadas baricéntricas de los puntos $\mathbf{p}_6$, $\mathbf{p}_7$ y $\mathbf{p}_8$ que se muestran en la figura 8.

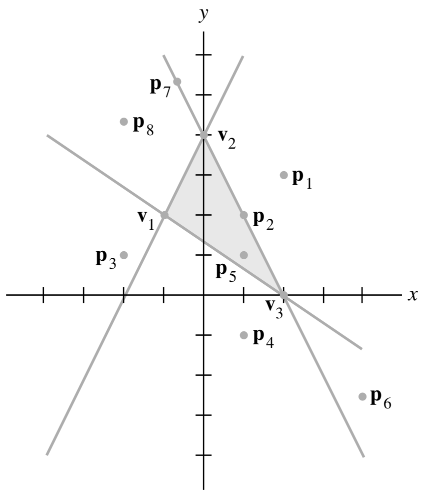


FIGURA 8

16. Sean $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\mathbf{p}_3 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{p}_4 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\mathbf{p}_5 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$, $\mathbf{p}_6 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{p}_7 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$ y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

a) Demuestre que el conjunto S es afínmente independiente.

b) Encuentre las coordenadas baricéntricas de $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ con respecto a S.

c) En papel milimétrico trace el triángulo T con vértices $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, extienda los lados como en la figura 5, y trace los puntos $\mathbf{p}_4$, $\mathbf{p}_5$, $\mathbf{p}_6$ y $\mathbf{p}_7$. Sin calcular los valores reales, determine los signos de las coordenadas baricéntricas de los puntos $\mathbf{p}_4$, $\mathbf{p}_5$, $\mathbf{p}_6$ y $\mathbf{p}_7$.

17. Demuestre el teorema 6 para un conjunto afínmente independiente $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ en $\mathbb{R}^n$. [*Sugerencia:* Un método es imitar la demostración del teorema 7 de la sección 4.4].

18. Sea T un tetraedro en posición "estándar", con tres aristas a lo largo de los tres ejes coordenados positivos en $\mathbb{R}^3$, y suponga que los vértices son $a\mathbf{e}_1$, $b\mathbf{e}_2$, $c\mathbf{e}_3$ y $\mathbf{0}$, donde $[\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3] = I_3$. Obtenga las fórmulas para las coordenadas baricéntricas de un punto arbitrario $\mathbf{p}$ en $\mathbb{R}^3$.

19. Sean $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ un conjunto de puntos en $\mathbb{R}^n$ afínmente dependiente y $f: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Demuestre que $\{f(\mathbf{p}_1), f(\mathbf{p}_2), f(\mathbf{p}_3)\}$ es afínmente dependiente en $\mathbb{R}^m$.

20. Suponga que $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es un conjunto afínmente independiente en $\mathbb{R}^n$ y que $\mathbf{q}$ es un punto arbitrario en $\mathbb{R}^n$. Demuestre que el conjunto trasladado $\{\mathbf{p}_1 + \mathbf{q}, \mathbf{p}_2 + \mathbf{q}, \mathbf{p}_3 + \mathbf{q}\}$ también es afínmente independiente.

En los ejercicios 21 a 24, **a**, **b** y **c** son puntos no colineales en $\mathbb{R}^2$ y **p** es cualquier otro punto en $\mathbb{R}^2$. Sea que $\Delta\mathbf{abc}$ denote la región triangular cerrada determinada por **a**, **b** y **c**, y sea $\Delta\mathbf{pbc}$ la región determinada por **p**, **b** y **c**. Por conveniencia, suponga que **a**, **b** y **c** se arreglan de tal forma que $\det[\,\tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{c}}\,]$ es positivo, donde $\tilde{\mathbf{a}}$, $\tilde{\mathbf{b}}$ y $\tilde{\mathbf{c}}$ son las formas homogéneas estándar de los puntos.

21. Demuestre que el área de $\Delta\mathbf{abc}$ es $\det[\,\tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{c}}\,]/2$. [*Sugerencia:* Consulte las secciones 3.2 y 3.3, incluyendo los ejercicios].

22. Sea **p** un punto sobre la recta que pasa por **a** y **b**. Demuestre que $\det[\,\tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{p}}\,] = 0$.

23. Sea **p** cualquier punto interior de $\Delta\mathbf{abc}$, con coordenadas baricéntricas (r, s, t), de manera que

$$[\,\tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{c}}\,]\begin{bmatrix} r \\ s \\ t \end{bmatrix} = \tilde{\mathbf{p}}$$

Utilice el ejercicio 19 y un hecho relacionado con determinantes (capítulo 3) para demostrar que

$$r = (\text{área de } \Delta\mathbf{pbc})/(\text{área de } \Delta\mathbf{abc})$$
$$s = (\text{área de } \Delta\mathbf{apc})/(\text{área de } \Delta\mathbf{abc})$$
$$t = (\text{área de } \Delta\mathbf{abp})/(\text{área de } \Delta\mathbf{abc})$$

24. Tome a **q** sobre el segmento de recta de **b** a **c** y considere la recta que pasa por **q** y **a**, que se puede representar como $\mathbf{p} = (1 - x)\mathbf{q} + x\mathbf{a}$ para toda x real. Pruebe que, para cada x, $\det[\,\tilde{\mathbf{p}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{c}}\,] = x \cdot \det[\,\tilde{\mathbf{a}}\ \ \tilde{\mathbf{b}}\ \ \tilde{\mathbf{c}}\,]$. A partir de esto y de un trabajo anterior, concluya que el parámetro x es la primera coordenada baricéntrica de **p**. Sin embargo, por construcción, el parámetro x también determina la distancia relativa entre **p** y **q** sobre el segmento de **q** a **a**. (Cuando $x = 1$, $\mathbf{p} = \mathbf{a}$). Cuando este resultado se aplica al ejemplo 5, demuestra que los colores en el vértice **a** y el punto **q** se interpolan suavemente conforme **p** se mueve a lo largo de la recta entre **a** y **q**.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Del ejemplo 1, el problema es determinar si los puntos son afínmente dependientes. Utilice el método del ejemplo 2 y reste un punto de los otros dos. Si uno de esos dos nuevos puntos es múltiplo del otro, entonces los tres puntos originales se encuentran sobre una recta.

2. En esencia, la demostración del teorema 5 señala que una relación de dependencia afín entre puntos corresponde a una relación de dependencia lineal entre las formas homogéneas de los puntos, utilizando los *mismos* pesos. Así, por reducción de filas:

$$\begin{bmatrix} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 & \tilde{\mathbf{v}}_4 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 5 & 1 \\ 1 & 0 & 4 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 4 & 1 & 5 & 1 \\ 1 & 0 & 4 & 2 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1.25 \\ 0 & 0 & 1 & .75 \end{bmatrix}$$

Vea esta matriz como la matriz de coeficientes de $A\mathbf{x} = \mathbf{0}$ con cuatro variables. Entonces x_4 es libre, $x_1 = x_4$, $x_2 = -1.25x_4$, y $x_3 = -.75x_4$. Una solución es $x_1 = x_4 = 4$, $x_2 = -5$ y $x_3 = -3$. Una dependencia lineal entre las formas homogéneas es $4\tilde{\mathbf{v}}_1 - 5\tilde{\mathbf{v}}_2 - 3\tilde{\mathbf{v}}_3 + 4\tilde{\mathbf{v}}_4 = \mathbf{0}$. De esta forma, $4\mathbf{v}_1 - 5\mathbf{v}_2 - 3\mathbf{v}_3 + 4\mathbf{v}_4 = \mathbf{0}$.

Otro método de solución es trasladar el problema al origen restando $\mathbf{v}_1$ de los otros puntos, encontrar una relación de dependencia lineal entre los puntos trasladados, y después reacomodar los términos. La cantidad de aritmética implicada es casi la misma que en el enfoque que se acaba de mostrar.

8.3 COMBINACIONES CONVEXAS

La sección 8.1 consideró combinaciones lineales especiales de la forma

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \quad \text{donde} \quad c_1 + c_2 + \cdots + c_k = 1$$

Esta sección, además, impone la restricción de que los pesos sean no negativos.

DEFINICIÓN Una **combinación convexa** de puntos $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ en $\mathbb{R}^n$ es una combinación lineal de la forma

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

tal que $c_1 + c_2 + \cdots + c_k = 1$ y $c_i \geq 0$ para toda i. El conjunto de todas las combinaciones convexas de puntos en un conjunto S es la **envolvente convexa** de S, y se denota como conv S.

La envolvente convexa de un solo punto $\mathbf{v}_1$ es justamente el conjunto $\{\mathbf{v}_1\}$, igual que la envolvente afín. En otros casos, la envolvente convexa está adecuadamente contenida en la envolvente afín. Recuerde que la envolvente afín de puntos distintos $\mathbf{v}_1$ y $\mathbf{v}_2$ es la recta

$$\mathbf{y} = (1 - t)\mathbf{v}_1 + t\mathbf{v}_2, \quad \text{con } t \text{ en } \mathbb{R}$$

Como los pesos en una combinación convexa son no negativos, los puntos en conv $\{\mathbf{v}_1, \mathbf{v}_2\}$ se pueden escribir en la forma

$$\mathbf{y} = (1 - t)\mathbf{v}_1 + t\mathbf{v}_2, \quad \text{con } 0 \leq t \leq 1$$

que es el **segmento de recta** entre $\mathbf{v}_1$ y $\mathbf{v}_2$, y de aquí en adelante se denotará como $\overline{\mathbf{v}_1\mathbf{v}_2}$.

Si un conjunto S es afínmente independiente y si $\mathbf{p} \in$ aff S, entonces $\mathbf{p} \in$ conv S si y solo si las coordenadas baricéntricas de $\mathbf{p}$ son no negativas. El ejemplo 1 muestra una situación especial en la cual S es mucho más que solo afínmente independiente.

EJEMPLO 1 Sean

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 6 \\ -3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -6 \\ 3 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 6 \\ 0 \\ 3 \end{bmatrix}, \quad \mathbf{p}_1 = \begin{bmatrix} 0 \\ 3 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} -10 \\ 5 \\ 11 \\ -4 \end{bmatrix},$$

y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Observe que S es un conjunto ortogonal. Determine si $\mathbf{p}_1$ está en Gen S, aff S y conv S. Y después haga lo mismo para $\mathbf{p}_2$.

SOLUCIÓN Si $\mathbf{p}_1$, es al menos, una combinación *lineal* de los puntos en S, entonces es fácil calcular los pesos porque S es un conjunto ortogonal. Sea W el subespacio generado por S. Un cálculo como en la sección 6.3 indica que la proyección ortogonal de $\mathbf{p}_1$ sobre W es el mismo $\mathbf{p}_1$:

$$\begin{aligned}\text{proy}_W &= \frac{\mathbf{p}_1\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{p}_1\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 + \frac{\mathbf{p}_1\cdot\mathbf{v}_3}{\mathbf{v}_3\cdot\mathbf{v}_3}\mathbf{v}_3 \\ &= \frac{18}{54}\mathbf{v}_1 + \frac{18}{54}\mathbf{v}_2 + \frac{18}{54}\mathbf{v}_3 \\ &= \frac{1}{3}\begin{bmatrix}3\\0\\6\\-3\end{bmatrix} + \frac{1}{3}\begin{bmatrix}-6\\3\\3\\0\end{bmatrix} + \frac{1}{3}\begin{bmatrix}3\\6\\0\\3\end{bmatrix} = \begin{bmatrix}0\\3\\3\\0\end{bmatrix} = \mathbf{p}_1\end{aligned}$$

Esto muestra que $\mathbf{p}_1$ está *en* Gen S. Además, puesto que los coeficientes suman 1, $\mathbf{p}_1$ está en aff S. De hecho, $\mathbf{p}_1$ está en conv S porque los coeficientes también son no negativos.

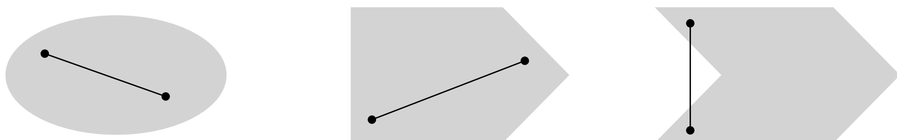
Para $\mathbf{p}_2$, un cálculo similar indica que proy$_W$ $\mathbf{p}_2 \neq \mathbf{p}_2$. Como proy$_W$ $\mathbf{p}_2$ es el punto en Gen S más cercano a $\mathbf{p}_2$, entonces el punto $\mathbf{p}_2$ no está en Gen S. En particular, $\mathbf{p}_2$ no puede estar en aff S o conv S. ■

Recuerde que un conjunto S es afín si contiene todas las rectas determinadas por pares de puntos en S. Cuando la atención se restringe a combinaciones convexas, la condición adecuada implica segmentos de recta en vez de rectas.

DEFINICIÓN

Un conjunto S es **convexo** si para cada $\mathbf{p}$, $\mathbf{q} \in S$, el segmento de recta $\overline{\mathbf{pq}}$ está contenido en S.

Intuitivamente, un conjunto S es convexo si todo par de puntos en el conjunto se puede "ver" entre sí sin que la línea de visión salga del conjunto. La figura 1 ilustra esta idea.



Convexo Convexo No convexo

FIGURA 1

El siguiente resultado es análogo al teorema 2 para conjuntos afines.

TEOREMA 7

Un conjunto S es convexo si y solo si cada combinación convexa de puntos de S se encuentra en S. Es decir, S es convexo si y solo si S = conv S.

DEMOSTRACIÓN El argumento es similar al utilizado en la demostración del teorema 2. La única diferencia está en el paso de inducción. Al tomar una combinación convexa de $k + 1$ puntos, considere $\mathbf{y} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1}$, donde $c_1 + \cdots + c_{k+1} = 1$ y

$0 \leq c_i \leq 1$ para toda i. Si $c_{k+1} = 1$, entonces $\mathbf{y} = \mathbf{v}_{k+1}$, que pertenece a S, y no hay más que probar. Si $c_{k+1} < 1$, sea $t = c_1 + \cdots + c_k$. Entonces, $t = 1 - c_{k+1} > 0$ y

$$\mathbf{y} = (1 - c_{k+1})\left(\frac{c_1}{t}\mathbf{v}_1 + \cdots + \frac{c_k}{t}\mathbf{v}_k\right) + c_{k+1}\mathbf{v}_{k+1} \tag{1}$$

Según la hipótesis de inducción, el punto $\mathbf{z} = (c_1/t)\mathbf{v}_1 + \cdots + (c_k/t)\mathbf{v}_k$ está en S porque los coeficientes no negativos suman 1. Así, la ecuación (1) presenta a $\mathbf{y}$ como una combinación convexa de dos puntos en S. De acuerdo con el principio de inducción, cada combinación convexa de tales puntos se encuentra en S. ■

El teorema 9 que se expone a continuación ofrece una caracterización más geométrica de la envolvente convexa de un conjunto. Se requiere un resultado preliminar de intersecciones de conjuntos. Recuerde de la sección 4.1 (ejercicio 32) que la intersección de dos subespacios es, en sí misma, un subespacio. De hecho, la intersección de cualquier colección de subespacios es, en sí misma, un subespacio. Un resultado similar es válido para conjuntos afines y conjuntos convexos.

TEOREMA 8

Sea $\{S_\alpha : \alpha \in \mathcal{A}\}$ cualquier colección de conjuntos convexos. Entonces, $\cap_{\alpha\in\mathcal{A}} S_\alpha$ es convexo. Si $\{T_\beta : \beta \in \mathcal{B}\}$ es cualquier colección de conjuntos afines, entonces $\cap_{\beta\in\mathcal{B}} T_\beta$ es afín.

DEMOSTRACIÓN Si $\mathbf{p}$ y $\mathbf{q}$ están en $\cap S_\alpha$, entonces $\mathbf{p}$ y $\mathbf{q}$ están en cada S_α. Como cada S_α es convexo, el segmento de recta entre $\mathbf{p}$ y $\mathbf{q}$ está en S_α para toda α y, por lo tanto, ese segmento está contenido en $\cap S_\alpha$. La demostración para el caso afín es similar. ■

TEOREMA 9

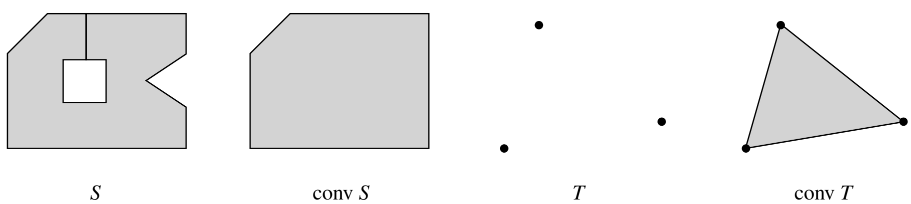
Para cualquier conjunto S, la envolvente convexa de S es la intersección de todos los conjuntos convexos que contienen a S.

DEMOSTRACIÓN Sea T la intersección de todos los conjuntos convexos que contienen a S. Como conv S es un conjunto convexo que contiene a S, se deduce que $T \subset$ conv S. Por otro lado, sea C cualquier conjunto convexo que contiene a S. Entonces, C contiene a toda combinación convexa de puntos de C (teorema 7) y, por consiguiente, también contiene a toda combinación convexa de puntos del subconjunto S. Es decir, conv $S \subset C$. Puesto que esto es verdad para todo conjunto convexo C que contiene a S, esto también es válido para la intersección de todos ellos. Es decir, conv $S \subset T$. ■

El teorema 9 señala que conv S es, en un sentido natural, el "menor" conjunto convexo que contiene a S. Por ejemplo, considere un conjunto S que está dentro de algún gran rectángulo en $\mathbb{R}^2$, e imagine estirar una banda de hule alrededor del exterior de S. Cuando la banda se contrae alrededor de S, indica la frontera de la envolvente convexa de S. O bien, para usar otra analogía, la envolvente convexa de S *llena* todos los huecos en el interior de S y *rellena* todas las abolladuras en la frontera de S.

EJEMPLO 2

a) A continuación se muestran las envolventes convexas de los conjuntos S y T en $\mathbb{R}^2$.



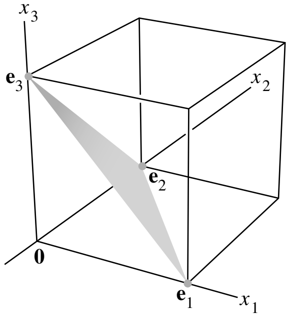


FIGURA 2

b) Sea S el conjunto que consta de la base estándar para $\mathbb{R}^3$, $S = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Entonces, conv S es una superficie triangular en $\mathbb{R}^3$, con vértices $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$. Véase la figura 2. ■

EJEMPLO 3 Sea $S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0 \text{ y } y = x^2 \right\}$. Demuestre que la envolvente convexa de S es la unión del origen y $\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x > 0 \text{ y } y \geq x^2 \right\}$. Véase la figura 3.

SOLUCIÓN Cada punto en conv S debe estar en un segmento de recta que conecte dos puntos de S. La línea discontinua de la figura 3 indica que, excepto por el origen, el eje y positivo no está en conv S, porque el origen es el único punto de S sobre el eje y. Parecería razonable que la figura 3 muestre conv S, pero, ¿cómo se podría estar seguro de que el punto $(10^{-2}, 10^4)$, por ejemplo, está sobre un segmento de recta del origen a un punto sobre la curva en S?

Considere cualquier punto $\mathbf{p}$ en la región sombreada de la figura 3, por ejemplo,

$$\mathbf{p} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{con } a > 0 \text{ y } b \geq a^2$$

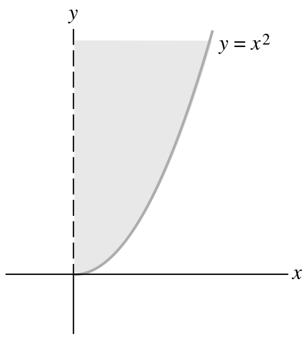


FIGURA 3

La recta que pasa por $\mathbf{0}$ y $\mathbf{p}$ tiene la ecuación $y = (b/a)t$ para t real. Esta recta se interseca con S donde t satisface $(b/a)t = t^2$, es decir, cuando $t = b/a$. Así, $\mathbf{p}$ está sobre el segmento de recta de $\mathbf{0}$ a $\begin{bmatrix} b/a \\ b^2/a^2 \end{bmatrix}$, lo que indica que la figura 3 es correcta. ■

El siguiente teorema es básico en el estudio de conjuntos convexos. Constantin Caratheodory fue el primero en probarlo en 1907. Si $\mathbf{p}$ está en la envolvente convexa de S, entonces, por definición, $\mathbf{p}$ debe ser una combinación convexa de puntos de S. Pero la definición no establece cuántos puntos de S se requieren para construir la combinación. El notable teorema de Caratheodory dice que en un espacio n-dimensional, el número de puntos de S en la combinación convexa nunca es mayor que $n + 1$.

TEOREMA 10

(Caratheodory) Si S es un subconjunto no vacío de $\mathbb{R}^n$, entonces cada punto en conv S se puede expresar como una combinación convexa de $n + 1$ o menos puntos de S.

DEMOSTRACIÓN Dado $\mathbf{p}$ en conv S, se puede escribir $\mathbf{p} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$, donde $\mathbf{v}_i \in S$, $c_1 + \cdots + c_k = 1$ y $c_i \geq 0$, para alguna k e $i = 1, \ldots, k$. El objetivo es demostrar que tal expresión existe para $\mathbf{p}$ con $k \leq n + 1$.

Si $k > n + 1$, entonces $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es afínmente dependiente, de acuerdo con el ejercicio 12 de la sección 8.2. Así, existen escalares $d_1, \ldots, d_k$, no todos cero, tales que

$$\sum_{i=1}^{k} d_i\mathbf{v}_i = \mathbf{0} \quad \text{y} \quad \sum_{i=1}^{k} d_i = 0$$

Considere las dos ecuaciones

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{p}$$

y

$$d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k = \mathbf{0}$$

Restando un múltiplo adecuado de la segunda ecuación a la primera, se elimina uno de los términos $\mathbf{v}_i$ y se obtiene una combinación convexa de menos que k elementos de S que es igual a $\mathbf{p}$.

Como no todos los coeficientes d_i son cero, podemos suponer (reacomodando subíndices, si es necesario) que $d_k > 0$ y que $c_k/d_k \leq c_i/d_i$ para todas aquellas i para las que $d_i > 0$. Para $i = 1,\ldots, k$, sea $b_i = c_i - (c_k/d_k)d_i$. Entonces, $b_k = 0$ y

$$\sum_{i=1}^{k} b_i = \sum_{i=1}^{k} c_i - \frac{c_k}{d_k}\sum_{i=1}^{k} d_i = 1 - 0 = 1$$

Además, cada $b_i \geq 0$. En realidad, si $d_i \leq 0$, entonces $b_i \geq c_i \geq 0$. Si $d_i > 0$, entonces $b_i = d_i(c_i/d_i - c_k/d_k) \geq 0$. Por construcción,

$$\begin{aligned}\sum_{i=1}^{k-1} b_i\mathbf{v}_i &= \sum_{i=1}^{k} b_i\mathbf{v}_i = \sum_{i=1}^{k}\left(c_i - \frac{c_k}{d_k}d_i\right)\mathbf{v}_i \\ &= \sum_{i=1}^{k} c_i\mathbf{v}_i - \frac{c_k}{d_k}\sum_{i=1}^{k} d_i\mathbf{v}_i = \sum_{i=1}^{k} c_i\mathbf{v}_i = \mathbf{p}\end{aligned}$$

De esta forma, $\mathbf{p}$ ahora es una combinación convexa de $k - 1$ de los puntos $\mathbf{v}_1,\ldots, \mathbf{v}_k$. Este proceso se puede repetir hasta que $\mathbf{p}$ se exprese como una combinación convexa de, a lo sumo, $n + 1$ de los puntos de S. ■

El siguiente ejemplo ilustra los cálculos de la demostración anterior.

EJEMPLO 4 Sean

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \quad \mathbf{p} = \begin{bmatrix} \frac{10}{3} \\ \frac{5}{2} \end{bmatrix}$$

y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. Luego,

$$\tfrac{1}{4}\mathbf{v}_1 + \tfrac{1}{6}\mathbf{v}_2 + \tfrac{1}{2}\mathbf{v}_3 + \tfrac{1}{12}\mathbf{v}_4 = \mathbf{p} \tag{2}$$

Utilice el procedimiento de la demostración del teorema de Caratheodory para expresar $\mathbf{p}$ como una combinación convexa de tres puntos de S.

SOLUCIÓN El conjunto S es afínmente dependiente. Utilice las técnicas de la sección 8.2 para obtener una relación de dependencia afín

$$-5\mathbf{v}_1 + 4\mathbf{v}_2 - 3\mathbf{v}_3 + 4\mathbf{v}_4 = \mathbf{0} \tag{3}$$

Después, seleccione los puntos $\mathbf{v}_2$ y $\mathbf{v}_4$ en (3), cuyos coeficientes sean positivos. Para cada punto, calcule la razón de los cocientes en las ecuaciones (2) y (3). La razón para $\mathbf{v}_2$ es $\frac{1}{6} \div 4 = \frac{1}{24}$, y para $\mathbf{v}_4$ es $\frac{1}{12} \div 4 = \frac{1}{48}$. La razón para $\mathbf{v}_4$ es menor, así que reste $\frac{1}{48}$ multiplicado por la ecuación (3) de la ecuación (2) para eliminar $\mathbf{v}_4$:

$$\begin{gathered}\left(\tfrac{1}{4} + \tfrac{5}{48}\right)\mathbf{v}_1 + \left(\tfrac{1}{6} - \tfrac{4}{48}\right)\mathbf{v}_2 + \left(\tfrac{1}{2} + \tfrac{3}{48}\right)\mathbf{v}_3 + \left(\tfrac{1}{12} - \tfrac{4}{48}\right)\mathbf{v}_4 = \mathbf{p} \\ \tfrac{17}{48}\mathbf{v}_1 + \tfrac{4}{48}\mathbf{v}_2 + \tfrac{27}{48}\mathbf{v}_3 = \mathbf{p}\end{gathered}$$ ■

Por lo general, este resultado no se puede mejorar al disminuir el número requerido de puntos. En realidad, dados cualesquiera tres puntos no colineales en $\mathbb{R}^2$, el centroide del triángulo formado por ellos está en la envolvente convexa de los tres, pero no está en la envolvente convexa de dos cualesquiera de ellos.

PROBLEMAS DE PRÁCTICA

1. Sean $\mathbf{v}_1 = \begin{bmatrix} 6 \\ 2 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 7 \\ 1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -2 \\ 4 \\ -1 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ y $\mathbf{p}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$, y $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Determine si $\mathbf{p}_1$ y $\mathbf{p}_2$ están en conv S.
2. Sea S el conjunto de puntos sobre la curva $y = 1/x$ para $x > 0$. Explique geométricamente por qué conv S consiste en todos los puntos sobre la curva S y por arriba de ella.

8.3 EJERCICIOS

1. En $\mathbb{R}^2$, sea $S = \left\{\begin{bmatrix} 0 \\ y \end{bmatrix} : 0 \le y < 1\right\} \cup \left\{\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right\}$. Describa (o bosqueje) la envolvente convexa de S.

2. Describa la envolvente convexa del conjunto S de puntos $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ que satisfacen las condiciones indicadas. Justifique sus respuestas. (Demuestre que un punto arbitrario $\mathbf{p}$ en S pertenece a conv S).
   a) $y = 1/x$ y $x \ge 1/2$
   b) $y = \operatorname{sen} x$
   c) $y = x^{1/2}$ y $x \ge 0$

3. Considere los puntos del ejercicio 5 de la sección 8.1. ¿Cuáles de $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ están en conv S?

4. Considere los puntos del ejercicio 6 de la sección 8.1. ¿Cuáles de $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ están en conv S?

5. Sean

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ -3 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix},$$
$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix},$$

y sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. Determine si $\mathbf{p}_1$ y $\mathbf{p}_2$ están en conv S.

6. Sean $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \mathbf{p}_1 = \begin{bmatrix} -1 \\ 2 \\ -\frac{3}{2} \\ \frac{5}{2} \end{bmatrix},$

$\mathbf{p}_2 = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{4} \\ \frac{7}{4} \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 6 \\ -4 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{p}_4 = \begin{bmatrix} -1 \\ -2 \\ 0 \\ 4 \end{bmatrix}$, y S el conjunto ortogonal $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Determine si cada $\mathbf{p}_i$ está en Gen S, aff S o conv S.
   a) $\mathbf{p}_1$ *b*) $\mathbf{p}_2$ *c*) $\mathbf{p}_3$ *d*) $\mathbf{p}_4$

Los ejercicios 7 a 10 utilizan la terminología de la sección 8.2.

7. *a*) Sea $T = \left\{\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}\right\}$, y sean
   $$\mathbf{p}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \text{ y } \mathbf{p}_4 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$
   Encuentre las coordenadas baricéntricas de $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ y $\mathbf{p}_4$ con respecto a T.
   b) Utilice sus respuestas del inciso *a*) para determinar si cada uno de $\mathbf{p}_1, \ldots, \mathbf{p}_4$ en el inciso *a*) está dentro, fuera o en la arista de conv T, una región triangular.

8. Repita el ejercicio 7 para $T = \left\{\begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$ y
   $$\mathbf{p}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{p}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{p}_3 = \begin{bmatrix} 1 \\ \frac{1}{3} \end{bmatrix} \text{ y } \mathbf{p}_4 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

9. Sea $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ un conjunto afínmente independiente. Considere los puntos $\mathbf{p}_1, \ldots, \mathbf{p}_5$, cuyas coordenadas baricéntricas con respecto a S están dadas por $(2, 0, 0, -1)$, $\left(0, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ $\left(\frac{1}{2}, 0, \frac{3}{2}, -1\right)$, $\left(\frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{6}\right)$ y $\left(\frac{1}{3}, 0, \frac{2}{3}, 0\right)$, respectivamente. Determine si cada uno de $\mathbf{p}_1, \ldots, \mathbf{p}_5$ está dentro, fuera o sobre la superficie de conv S, un tetraedro. ¿Algunos de esos puntos están en la arista de conv S?

10. Repita el ejercicio 9 para los puntos $\mathbf{q}_1, \ldots, \mathbf{q}_5$ cuyas coordenadas baricéntricas con respecto a S están dadas por $\left(\frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{2}\right)$, $\left(\frac{3}{4}, -\frac{1}{4}, 0, \frac{1}{2}\right)$, $\left(0, \frac{3}{4}, \frac{1}{4}, 0\right)$, $(0, -2, 0, 3)$ y $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right)$, respectivamente.

En los ejercicios 11 y 12, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) Si $\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ y $c_1 + c_2 + c_3 = 1$, entonces $\mathbf{y}$ es una combinación convexa de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.
    b) Si S es un conjunto no vacío, entonces conv S contiene algunos puntos que no están en S.
    c) Si S y T son conjuntos convexos, entonces $S \cup T$ también es convexo.

12. *a*) Un conjunto es convexo si $\mathbf{x}, \mathbf{y} \in S$ implica que el segmento de recta entre $\mathbf{x}$ y $\mathbf{y}$ está contenido en S.
    b) Si S y T son conjuntos convexos, entonces $S \cap T$ también es convexo.

c) Si S es un subconjunto no vacío de $\mathbb{R}^5$ y $\mathbf{y} \in \text{conv } S$, entonces existen distintos puntos $\mathbf{v}_1, \ldots, \mathbf{v}_6$ en S tales que $\mathbf{y}$ es una combinación convexa de $\mathbf{v}_1, \ldots, \mathbf{v}_6$.

13. Sea S un subconjunto convexo de $\mathbb{R}^n$ y suponga que $f: \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal. Demuestre que el conjunto $f(S) = \{f(\mathbf{x}): \mathbf{x} \in S\}$ es un subconjunto convexo de $\mathbb{R}^m$.

14. Sean $f: \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal y T un subconjunto convexo de $\mathbb{R}^m$. Demuestre que el conjunto $S = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \in T\}$ es un subconjunto convexo de $\mathbb{R}^n$.

15. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$ y

$\mathbf{p} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Confirme que

$$\mathbf{p} = \tfrac{1}{3}\mathbf{v}_1 + \tfrac{1}{3}\mathbf{v}_2 + \tfrac{1}{6}\mathbf{v}_3 + \tfrac{1}{6}\mathbf{v}_4 \text{ y } \mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 - \mathbf{v}_4 = \mathbf{0}$$

Utilice el procedimiento de la demostración del teorema de Caratheodory para expresar $\mathbf{p}$ como una combinación convexa de tres de las $\mathbf{v}_i$. Haga esto de *dos* formas.

16. Repita el ejercicio 9 para los puntos $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

$\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{v}_4 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, considerando que

$$\mathbf{p} = \tfrac{1}{121}\mathbf{v}_1 + \tfrac{72}{121}\mathbf{v}_2 + \tfrac{37}{121}\mathbf{v}_3 + \tfrac{1}{11}\mathbf{v}_4$$

y

$$10\mathbf{v}_1 - 6\mathbf{v}_2 + 7\mathbf{v}_3 - 11\mathbf{v}_4 = \mathbf{0}$$

En los ejercicios 17 a 20, demuestre el enunciado en cuestión respecto de los subconjuntos A y B de $\mathbb{R}^n$. Una demostración para un ejercicio puede emplear resultados de ejercicios anteriores.

17. Si $A \subset B$ y B es convexo, entonces $\text{conv } A \subset B$.

18. Si $A \subset B$, entonces $\text{conv } A \subset \text{conv } B$.

19. *a*) $[(\text{conv } A) \cup (\text{conv } B)] \subset \text{conv } (A \cup B)$

b) Encuentre un ejemplo en $\mathbb{R}^2$ para demostrar que la igualdad no necesariamente es válida en el inciso *a*).

20. *a*) $\text{conv } (A \cap B) \subset [(\text{conv } A) \cap (\text{conv } B)]$

b) Encuentre un ejemplo en $\mathbb{R}^2$ para probar que la igualdad no necesita ser válida en el inciso *a*).

21. Considere que $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$ son puntos en $\mathbb{R}^n$, y defina $\mathbf{f}_0(t) = (1-t)\mathbf{p}_0 + t\mathbf{p}_1$, $\mathbf{f}_1(t) = (1-t)\mathbf{p}_1 + t\mathbf{p}_2$ y $\mathbf{g}(t) = (1-t)\mathbf{f}_0(t) + t\mathbf{f}_1(t)$ para $0 \le t \le 1$. Para los puntos que se muestran en la siguiente figura, realice un esquema que muestre $\mathbf{f}_0\left(\frac{1}{2}\right)$, $\mathbf{f}_1\left(\frac{1}{2}\right)$ y $\mathbf{g}\left(\frac{1}{2}\right)$.

22. Repita el ejercicio 21 para $\mathbf{f}_0\left(\frac{3}{4}\right)$, $\mathbf{f}_1\left(\frac{3}{4}\right)$ y $\mathbf{g}\left(\frac{3}{4}\right)$.

23. Sea $\mathbf{g}(t)$ como se define en el ejercicio 21. Su gráfica se llama *curva cuadrática de Bézier*, y se utiliza en algunos diseños de gráficos computacionales. Los puntos $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$ se llaman *puntos de control* de la curva. Calcule una fórmula para $\mathbf{g}(t)$ que implique solamente a $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$. Después, demuestre que $\mathbf{g}(t)$ está en $\text{conv } \{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$ para $0 \le t \le 1$.

24. Dados los puntos de control $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ en $\mathbb{R}^n$, sea $\mathbf{g}_1(t)$ para $0 \le t \le 1$ la curva cuadrática de Bézier del ejercicio 23 determinada por $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$, y sea $\mathbf{g}_2(t)$ definida de manera similar para $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Para $0 \le t \le 1$, defina $\mathbf{h}(t) = (1-t)\mathbf{g}_1(t) + t\mathbf{g}_2(t)$. Demuestre que la gráfica de $\mathbf{h}(t)$ está en la envolvente convexa de los cuatro puntos de control. A esta curva se le llama *curva cúbica de Bézier*, y su definición aquí es un paso en un algoritmo para construir curvas de Bézier (analizadas en la sección 8.6). Una curva de Bézier de grado k se determina por $k+1$ puntos de control, y su gráfica está en la envolvente convexa de esos puntos de control.

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. Los puntos $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ no son ortogonales, así que se calcula

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} -8 \\ 2 \\ -3 \end{bmatrix}, \mathbf{p}_1 - \mathbf{v}_1 = \begin{bmatrix} -5 \\ 1 \\ -1 \end{bmatrix} \text{ y } \mathbf{p}_2 - \mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \\ -1 \end{bmatrix}$$

Aumente la matriz $[\mathbf{v}_2 - \mathbf{v}_1 \quad \mathbf{v}_3 - \mathbf{v}_1]$ con $\mathbf{p}_1 - \mathbf{v}_1$ y $\mathbf{p}_2 - \mathbf{v}_1$ y reduzca por filas:

$$\begin{bmatrix} 1 & -8 & -5 & -3 \\ -1 & 2 & 1 & 0 \\ 3 & -3 & -1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & \frac{1}{3} & 1 \\ 0 & 1 & \frac{2}{3} & \frac{1}{2} \\ 0 & 0 & 0 & -\frac{5}{2} \end{bmatrix}$$

La tercera columna indica que $\mathbf{p}_1 - \mathbf{v}_1 = \frac{1}{3}(\mathbf{v}_2 - \mathbf{v}_1) + \frac{2}{3}(\mathbf{v}_3 - \mathbf{v}_1)$, lo que conduce a $\mathbf{p}_1 = 0\mathbf{v}_1 + \frac{1}{3}\mathbf{v}_2 + \frac{2}{3}\mathbf{v}_3$. Así, $\mathbf{p}_1$ está en $\text{conv } S$. De hecho, $\mathbf{p}_1$ está en $\text{conv } \{\mathbf{v}_2, \mathbf{v}_3\}$.

La última columna de la matriz indica que $\mathbf{p}_2 - \mathbf{v}_1$ no es una combinación lineal de $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$. Así, $\mathbf{p}_2$ no es una combinación afín de $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, por lo que $\mathbf{p}_2$ quizá no esté en conv S.

Un método de solución alternativo consiste en reducir por filas la matriz aumentada de las formas homogéneas:

$$\begin{bmatrix} \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 & \tilde{\mathbf{p}}_1 & \tilde{\mathbf{p}}_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

2. Si $\mathbf{p}$ es un punto arriba de S, entonces la recta que pasa por $\mathbf{p}$ con pendiente -1 se intersecará con S en dos puntos antes de llegar a los ejes x y y positivos.

8.4 HIPERPLANOS

Los hiperplanos desempeñan un papel especial en la geometría de $\mathbb{R}^n$ porque dividen el espacio en dos porciones, justo como un plano separa a $\mathbb{R}^3$ en dos partes y una recta corta a $\mathbb{R}^2$. La clave al trabajar con hiperplanos es utilizar simples descripciones *implícitas*, en vez de las representaciones *explícitas* o paramétricas de rectas y planos empleadas en el estudio anterior de conjuntos afines.[1]

Una ecuación implícita de una recta en $\mathbb{R}^2$ tiene la forma $ax + by = d$. Una ecuación implícita de un plano en $\mathbb{R}^3$ tiene la forma $ax + by + cz = d$. Ambas ecuaciones describen la recta o el plano como el conjunto de todos los puntos en los cuales una expresión lineal (también llamada *funcional lineal*) tiene un valor fijo, d.

DEFINICIÓN

Una **funcional lineal** en $\mathbb{R}^n$ es una transformación lineal f de $\mathbb{R}^n$ en $\mathbb{R}$. Para todo escalar d en $\mathbb{R}$, el símbolo $[f : d]$ denota el conjunto de todas las $\mathbf{x}$ en $\mathbb{R}^n$ para las cuales el valor de f es d. Es decir,

$$[f : d] \text{ es el conjunto } \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = d\}$$

La *funcional* **cero** es la transformación tal que $f(\mathbf{x}) = 0$ para toda $\mathbf{x}$ en $\mathbb{R}^n$. Todas las demás funcionales lineales en $\mathbb{R}^n$ son **diferentes de cero**.

EJEMPLO 1 En $\mathbb{R}^2$, la recta $x - 4y = 13$ es un hiperplano en $\mathbb{R}^2$, y es el conjunto de puntos en los que la funcional lineal $f(x, y) = x - 4y$ tiene valor 13. Es decir, la recta es el conjunto $[f : 13]$. ■

EJEMPLO 2 En $\mathbb{R}^3$, el plano $5x - 2y + 3z = 21$ es un hiperplano, el conjunto de puntos en los que la funcional lineal $g(x, y, z) = 5x - 2y + 3z$ tiene valor 21. Este hiperplano es el conjunto $[g : 21]$. ■

Si f es una funcional lineal en $\mathbb{R}^n$, entonces la matriz estándar de esta transformación lineal f es una matriz A de $1 \times n$, por ejemplo, $A = [a_1 \quad a_2 \quad \cdots \quad a_n]$. De manera que

$$[f : 0] \quad \text{es igual que} \quad \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 0\} = \text{Nul } A \tag{1}$$

[1] Las representaciones paramétricas se estudiaron en la sección 1.5.

Si f es una funcional distinta de cero, entonces rango de $A = 1$, y dim Nul $A = n - 1$, de acuerdo con el teorema del rango.[2] Así, el subespacio $[f : 0]$ tiene dimensión $n - 1$ y, por lo tanto, es un hiperplano. Además, si d es cualquier número en $\mathbb{R}$, entonces

$$[f : d] \quad \text{es igual que} \quad \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = d\} \tag{2}$$

Recuerde del teorema 6 de la sección 1.5 que el conjunto de soluciones de $A\mathbf{x} = \mathbf{b}$ se obtiene al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{0}$, empleando cualquier solución particular $\mathbf{p}$ de $A\mathbf{x} = \mathbf{b}$. Cuando A es la matriz estándar de la transformación f, este teorema dice que

$$[f : d] = [f : 0] + \mathbf{p} \quad \text{para cualquier } \mathbf{p} \text{ en } [f : d] \tag{3}$$

Por lo tanto, los conjuntos $[f : d]$ son hiperplanos paralelos a $[f : 0]$. Véase la figura 1.

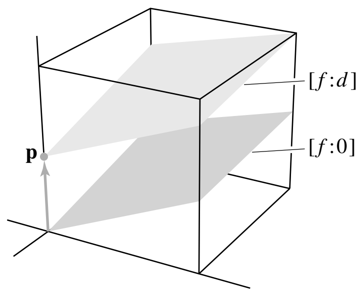


FIGURA 1 Hiperplanos paralelos, con $f(\mathbf{p}) = d$.

Cuando A es una matriz de $1 \times n$, la ecuación $A\mathbf{x} = d$ se puede escribir con un producto interno $\mathbf{n}\cdot\mathbf{x}$, empleando $\mathbf{n}$ en $\mathbb{R}^n$ con las mismas entradas que A. Así, a partir de (2),

$$[f : d] \quad \text{es igual que} \quad \{\mathbf{x} \in \mathbb{R}^n : \mathbf{n}\cdot\mathbf{x} = d\} \tag{4}$$

Entonces, $[f : 0] = \{\mathbf{x} \in \mathbb{R}^n\!: \mathbf{n}\cdot\mathbf{x} = 0\}$, lo que muestra que $[f : 0]$ es el complemento ortogonal del subespacio generado por $\mathbf{n}$. En la terminología de cálculo y geometría en $\mathbb{R}^3$, $\mathbf{n}$ se llama un vector **normal** a $[f : 0]$. (Un vector "normal" en este sentido no necesita tener longitud unitaria). Además, se dice que $\mathbf{n}$ es **normal** a cada hiperplano paralelo $[f : d]$, aunque $\mathbf{n}\cdot\mathbf{x}$ no es cero cuando $d \neq 0$.

Otro nombre para $[f : d]$ es *conjunto nivel* de f, y $\mathbf{n}$ algunas veces se llama *gradiente* de f cuando $f(\mathbf{x}) = \mathbf{n}\cdot\mathbf{x}$ para cada $\mathbf{x}$.

EJEMPLO 3 Sean $\mathbf{n} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ -6 \end{bmatrix}$, y sea $H = \{\mathbf{x} : \mathbf{n}\cdot\mathbf{x} = 12\}$, de manera que $H = [f : 12]$, donde $f(x, y) = 3x + 4y$. Por lo tanto, H es la recta $3x + 4y = 12$. Encuentre una descripción implícita del hiperplano paralelo (recta) $H_1 = H + \mathbf{v}$.

SOLUCIÓN Primero, encuentre un punto $\mathbf{p}$ en H_1. Para hacer esto, determine un punto en H y súmele $\mathbf{v}$. Por ejemplo, $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$ está en H, así que $\mathbf{p} = \begin{bmatrix} 1 \\ -6 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ está en H_1. Ahora, calcule $\mathbf{n}\cdot\mathbf{p} = -9$. Esto indica que $H_1 = [f : -9]$. Véase la figura 2, que también muestra el subespacio $H_0 = \{\mathbf{x} : \mathbf{n}\cdot\mathbf{x} = 0\}$. ■

Los siguientes tres ejemplos revelan conexiones entre descripciones implícitas y explícitas de hiperplanos. El ejemplo 4 inicia con una forma implícita.

[2] Véase el teorema 14 de la sección 2.9 o el teorema 14 de la sección 4.6.

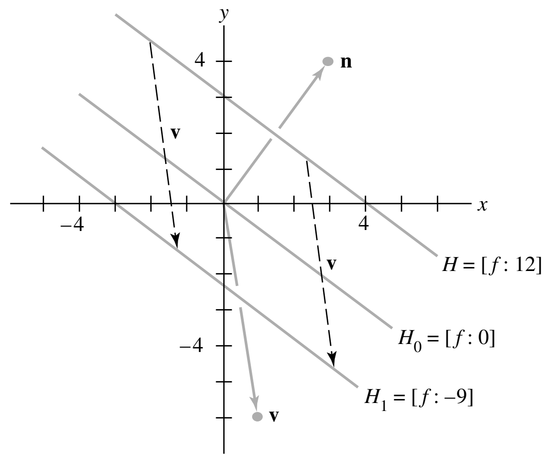


FIGURA 2

EJEMPLO 4 En $\mathbb{R}^2$, dé una descripción explícita de la recta $x - 4y = 13$ en forma vectorial paramétrica.

SOLUCIÓN Esto equivale a resolver una ecuación no homogénea $A\mathbf{x} = \mathbf{b}$, donde $A = [1 \quad -4]$ y $\mathbf{b}$ es el número 13 en $\mathbb{R}$. Escriba $x = 13 + 4y$, donde y es una variable libre. En forma paramétrica, la solución es

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 13 + 4y \\ y \end{bmatrix} = \begin{bmatrix} 13 \\ 0 \end{bmatrix} + y\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \mathbf{p} + y\mathbf{q}, \quad y \in \mathbb{R}$$ ■

Es más complicado convertir una descripción explícita de una recta a la forma implícita. La idea básica es construir $[f : 0]$ y después encontrar d para $[f : d]$.

EJEMPLO 5 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$, y L_1 la recta que pasa por $\mathbf{v}_1$ y $\mathbf{v}_2$. Encuentre una funcional lineal f y una constante d tales que $L_1 = [f : d]$.

SOLUCIÓN La recta L_1 es paralela a la recta trasladada L_0 a través de $\mathbf{v}_2 - \mathbf{v}_1$ y el origen. La ecuación que define a L_0 tiene la forma

$$[a \quad b]\begin{bmatrix} x \\ y \end{bmatrix} = 0 \quad \text{o} \quad \mathbf{n}\cdot\mathbf{x} = 0 \quad \text{donde} \quad \mathbf{n} = \begin{bmatrix} a \\ b \end{bmatrix} \tag{5}$$

Como $\mathbf{n}$ es ortogonal al subespacio L_0, que contiene a $\mathbf{v}_2 - \mathrm{v}_1$, calcule

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 6 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

y resuelva

$$\begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} 5 \\ -2 \end{bmatrix} = 0$$

Por inspección, una solución es $[a \quad b] = [2 \quad 5]$. Sea $f(x, y) = 2x + 5y$. A partir de (5), $L_0 = [f : 0]$, y $L_1 = [f : d]$ para alguna d. Como $\mathbf{v}_1$ está sobre la recta L_1, $d = f(\mathbf{v}_1) = 2(1) + 5(2) = 12$. Así, la ecuación para L_1 es $2x + 5y = 12$. Como comprobación, observe que $f(\mathbf{v}_2) = f(6, 0) = 2(6) + 5(0) = 12$, de manera que $\mathbf{v}_2$ también está sobre L_1. ■

EJEMPLO 6 Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$. Encuentre una descripción implícita $[f : d]$ del plano H_1 que pasa por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

SOLUCIÓN H_1 es paralelo a un plano H_0 que pasa por el origen y que contiene los puntos trasladados

$$\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

Como esos dos puntos son linealmente independientes, $H_0 = \text{Gen}\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$. Sea $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ la normal a H_0. Entonces, $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$ son ortogonales a $\mathbf{n}$. Es decir, $(\mathbf{v}_2 - \mathbf{v}_1)\cdot\mathbf{n} = 0$ y $(\mathbf{v}_3 - \mathbf{v}_1)\cdot\mathbf{n} = 0$. Esas dos ecuaciones forman un sistema cuya matriz aumentada se puede reducir por filas:

$$\begin{bmatrix} 1 & -2 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0, \quad \begin{bmatrix} 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0, \quad \begin{bmatrix} 1 & -2 & 3 & 0 \\ 2 & 0 & 1 & 0 \end{bmatrix}$$

Las operaciones de fila dan $a = (-\frac{2}{4})c$, $b = (\frac{5}{4})c$, con c libre. Establezca $c = 4$, por ejemplo. Entonces, $\mathbf{n} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}$ y $H_0 = [f : 0]$, donde $f(\mathbf{x}) = -2x_1 + 5x_2 + 4x_3$.

El hiperplano paralelo H_1 es $[f : d]$. Para encontrar d, considere el hecho de que $\mathbf{v}_1$ está en H_1, y calcule $d = f(\mathbf{v}_1) = f(1, 1, 1) = -2(1) + 5(1) + 4(1) = 7$. Como comprobación, calcule $f(\mathbf{v}_2) = f(2, -1, 4) = -2(2) + 5(-1) + 4(4) = 16 - 9 = 7$. ■

El procedimiento del ejemplo 6 se generaliza a mayores dimensiones. Sin embargo, para el caso especial de $\mathbb{R}^3$, también se puede utilizar la fórmula del **producto cruz** para determinar $\mathbf{n}$, aplicando un determinante simbólico como una regla mnemotécnica:

$$\begin{aligned}
\mathbf{n} &= (\mathbf{v}_2 - \mathbf{v}_1) \times (\mathbf{v}_3 - \mathbf{v}_1) \\
&= \begin{vmatrix} 1 & 2 & \mathbf{i} \\ -2 & 0 & \mathbf{j} \\ 3 & 1 & \mathbf{k} \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ 3 & 1 \end{vmatrix}\mathbf{i} - \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}\mathbf{j} + \begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix}\mathbf{k} \\
&= -2\mathbf{i} + 5\mathbf{j} + 4\mathbf{k} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}
\end{aligned}$$

Si solo se necesita la fórmula para f, el cálculo del producto cruz se puede escribir como un determinante ordinario

$$\begin{aligned}
f(x_1, x_2, x_3) &= \begin{vmatrix} 1 & 2 & x_1 \\ -2 & 0 & x_2 \\ 3 & 1 & x_3 \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ 3 & 1 \end{vmatrix}x_1 - \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix}x_2 + \begin{vmatrix} 1 & 2 \\ -2 & 0 \end{vmatrix}x_3 \\
&= -2x_1 + 5x_2 + 4x_3
\end{aligned}$$

Hasta el momento, cada hiperplano examinado se ha descrito como $[f : d]$ para alguna funcional lineal f y alguna d en $\mathbb{R}$, o bien, de manera equivalente como $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{n}\cdot\mathbf{x} = d\}$ para alguna $\mathbf{n}$ en $\mathbb{R}^n$. El siguiente teorema indica que *todo* hiperplano tiene estas descripciones equivalentes.

TEOREMA 11

Un subconjunto H de $\mathbb{R}^n$ es un hiperplano si y solo si $H = [f : d]$ para alguna funcional lineal f diferente de cero y algún escalar d en $\mathbb{R}$. De esta forma, si H es un hiperplano, existe un vector $\mathbf{n}$ diferente de cero y un número real d tales que $H = \{\mathbf{x} : \mathbf{n}\cdot\mathbf{x} = d\}$.

DEMOSTRACIÓN Suponga que H es un hiperplano, tome $\mathbf{p} \in H$, y sea $H_0 = H - \mathbf{p}$. Entonces, H_0 está en un subespacio de dimensión $(n - 1)$. Después, tome cualquier punto $\mathbf{y}$ que no esté en H_0. De acuerdo con el teorema de descomposición ortogonal de la sección 6.3,

$$\mathbf{y} = \mathbf{y}_1 + \mathbf{n}$$

donde $\mathbf{y}_1$ es un vector en H_0 y $\mathbf{n}$ es ortogonal a todo vector en H_0. La función f definida por

$$f(\mathbf{x}) = \mathbf{n}\cdot\mathbf{x} \quad \text{para } \mathbf{x} \in \mathbb{R}^n$$

es una funcional lineal, por las propiedades del producto interno. Ahora, $[f : 0]$ es un hiperplano que contiene a H_0, por construcción de $\mathbf{n}$. Se deduce que

$$H_0 = [f : 0]$$

[Argumento: H_0 contiene una base S de $n - 1$ vectores, y como S está en el subespacio $(n - 1)$-dimensional $[f : 0]$, entonces S también debe ser una base para $[f : 0]$, de acuerdo con el teorema de la base]. Finalmente, sea $d = f(\mathbf{p}) = \mathbf{n}\cdot\mathbf{p}$. Por lo tanto, como se muestra en la ecuación (3) anterior,

$$[f : d] = [f : 0] + \mathbf{p} = H_0 + \mathbf{p} = H$$

El enunciado inverso de que $[f : d]$ es un hiperplano se deduce de las ecuaciones (1) y (3) anteriores. ■

Muchas importantes aplicaciones de hiperplanos dependen de la posibilidad de "separar" dos conjuntos mediante un hiperplano. Intuitivamente, esto significa que uno de los conjuntos está sobre un lado del hiperplano, y el otro conjunto está sobre el lado contrario. La terminología y la notación que se presentan a continuación ayudarán a hacer más precisa esta idea.

TOPOLOGÍA EN $\mathbb{R}^n$: TÉRMINOS Y HECHOS

Para cualquier punto $\mathbf{p}$ en $\mathbb{R}^n$ y cualquier real $\delta > 0$, la **bola abierta** $B(\mathbf{p}, \delta)$ con centro en $\mathbf{p}$ y radio δ está dada por

$$B(\mathbf{p}, \delta) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{p}\| < \delta\}$$

Dado un conjunto S en $\mathbb{R}^n$, un punto $\mathbf{p}$ es un **punto interior** de S si existe un $\delta > 0$ tal que $B(\mathbf{p}, \delta) \subset S$. Si toda bola abierta centrada en $\mathbf{p}$ se interseca con S y con el complemento de S, entonces $\mathbf{p}$ se llama **punto frontera** de S. Un conjunto es **abierto** si no contiene a ninguno de sus puntos frontera. (Esto es equivalente a decir que todos sus puntos son puntos interiores). Un conjunto es **cerrado** si contiene todos sus puntos frontera. (Si S contiene solamente algunos de sus puntos frontera, entonces S no es abierto ni cerrado). Un conjunto S es **acotado** si existe una $\delta > 0$ tal que $S \subset B(\mathbf{0}, \delta)$. Un conjunto en $\mathbb{R}^n$ es **compacto** si es cerrado y acotado.

Teorema: La envolvente convexa de un conjunto abierto es abierta, y la envolvente convexa de un conjunto compacto es compacta. (La envolvente convexa de un conjunto cerrado no necesita ser cerrada. Véase el ejercicio 27).

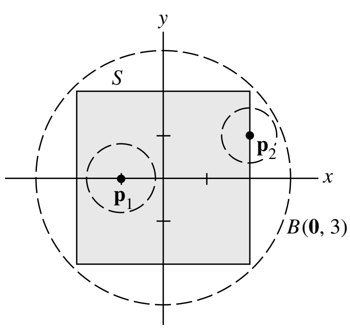


FIGURA 3
El conjunto S es cerrado y acotado.

EJEMPLO 7 Sean

$$S = \text{conv}\left\{\begin{bmatrix}-2\\2\end{bmatrix}, \begin{bmatrix}-2\\-2\end{bmatrix}, \begin{bmatrix}2\\-2\end{bmatrix}, \begin{bmatrix}2\\2\end{bmatrix}\right\}, \quad \mathbf{p}_1 = \begin{bmatrix}-1\\0\end{bmatrix} \quad \text{y} \quad \mathbf{p}_2 = \begin{bmatrix}2\\1\end{bmatrix},$$

como se muestra en la figura 3. Entonces, $\mathbf{p}_1$ es un punto interior porque $B\left(\mathbf{p}, \frac{3}{4}\right) \subset S$. El punto $\mathbf{p}_2$ es un punto frontera porque cada bola abierta centrada en $\mathbf{p}_2$ se interseca con S y el complemento de S. El conjunto S es cerrado porque contiene todos sus puntos frontera. El conjunto S es acotado porque $S \subset B(\mathbf{0}, 3)$. Por lo tanto, S también es compacto. ■

Notación: Si f es una funcional lineal, entonces $f(A) \leq d$ significa que $f(\mathbf{x}) \leq d$ para cada $\mathbf{x} \in A$. Las notaciones correspondientes se emplearán cuando las desigualdades estén invertidas o cuando sean estrictas.

DEFINICIÓN

El hiperplano $H = [f : d]$ **separa** a dos conjuntos A y B si es válida una de las siguientes situaciones:

i. $f(A) \leq d$ y $f(B) \geq d$, o
ii. $f(A) \geq d$ y $f(B) \leq d$.

Si en las condiciones anteriores todas las desigualdades débiles se remplazan por desigualdades estrictas, entonces se dice que H **separa estrictamente** A y B.

Observe que la separación estricta requiere que los dos conjuntos sean disjuntos, lo que no ocurre en la separación simple. Es claro que si dos círculos en el plano son tangentes externamente, entonces su recta tangente común los separa (pero no estrictamente).

Aunque es necesario que los dos conjuntos sean disjuntos para separarlos estrictamente, esta condición no es suficiente, ni siquiera para conjuntos convexos cerrados. Por ejemplo, sean

$$A = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq \frac{1}{2} \text{ y } \frac{1}{x} \leq y \leq 2 \right\} \quad \text{y} \quad B = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0 \text{ y } y = 0 \right\}$$

Entonces, A y B son conjuntos convexos cerrados disjuntos, pero no se pueden separar estrictamente mediante un hiperplano (recta en $\mathbb{R}^2$). Véase la figura 4. Así, el problema de separar (o de separar estrictamente) dos conjuntos con un hiperplano es más complicado de lo que parece.

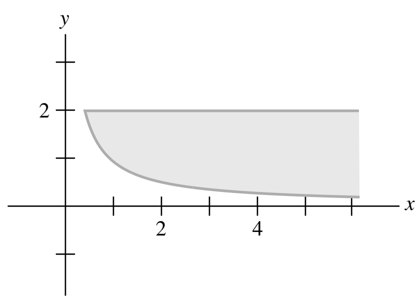


FIGURA 4 Conjuntos convexos cerrados disjuntos.

Hay muchas condiciones interesantes sobre los conjuntos A y B que implican la existencia de un hiperplano de separación, pero los siguientes dos teoremas son suficientes para esta sección. La demostración del primer teorema requiere de un poco de material preliminar,[3] pero el segundo es consecuencia inmediata del primero.

TEOREMA 12

Suponga que A y B son conjuntos convexos no vacíos tales que A es compacto y B es cerrado. Entonces, existe un hiperplano H que separa estrictamente A y B si y solo si $A \cap B = \varnothing$.

TEOREMA 13

Suponga que A y B son conjuntos compactos no vacíos. Entonces, existe un hiperplano que separa estrictamente A y B si y solo si $(\text{conv } A) \cap (\text{conv } B) = \varnothing$.

[3] Una demostración del teorema 12 se presenta en Steven R. Lay, *Convex Sets and Their Applications* (Nueva York: John Wiley & Sons, 1982; Mineola, NY: Dover Publications, 2007), pp. 34-39.

DEMOSTRACIÓN Suponga que (conv A) $\cap$ (conv B) $= \varnothing$. Como la envolvente convexa de un conjunto compacto es compacta, entonces el teorema 12 asegura que existe un hiperplano H que separa estrictamente conv A y conv B. Es claro que H también separa estrictamente a los conjuntos más pequeños A y B.

A la inversa, suponga que el hiperplano $H = [f : d]$ separa estrictamente A y B. Sin pérdida de generalidad, suponga que $f(A) < d$ y $f(B) > d$. Sea $\mathbf{x} = c_1\mathbf{x}_1 + \cdots + c_k\mathbf{x}_k$ cualquier combinación convexa de elementos de A. Entonces,

$$f(\mathbf{x}) = c_1 f(\mathbf{x}_1) + \cdots + c_k f(\mathbf{x}_k) < c_1 d + \cdots + c_k d = d$$

ya que $c_1 + \cdots + c_k = 1$. Por lo tanto, $f(\text{conv } A) < d$. Asimismo, $f(\text{conv } B) > d$, por lo que $H = [f : d]$ separa estrictamente conv A y conv B. De acuerdo con el teorema 12, conv A y conv B deben ser disjuntos. ■

EJEMPLO 8 Sean

$$\mathbf{a}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \quad \text{y } \mathbf{b}_2 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix},$$

y sean $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $B = \{\mathbf{b}_1, \mathbf{b}_2\}$. Demuestre que el hiperplano $H = [f : 5]$, donde $f(x_1, x_2, x_3) = 2x_1 - 3x_2 + x_3$, no separa A y B. ¿Existe un hiperplano paralelo a H que separe A y B? ¿Se intersecan las envolventes convexas de A y B?

SOLUCIÓN Evalúe la funcional lineal f en cada uno de los puntos en A y B:

$$f(\mathbf{a}_1) = 2, \quad f(\mathbf{a}_2) = -11, \quad f(\mathbf{a}_3) = -6, \quad f(\mathbf{b}_1) = 4 \quad \text{y} \quad f(\mathbf{b}_2) = 12$$

Como $f(\mathbf{b}_1) = 4$ es menor que 5 y $f(\mathbf{b}_2) = 12$ es mayor que 5, los puntos de B están en ambos lados de $H = [f : 5]$ y así H no separa A y B.

Puesto que $f(A) < 3$ y $f(B) > 3$, el hiperplano paralelo $[f : 3]$ separa estrictamente A y B. Según el teorema 13, (conv A) $\cap$ (conv B) $= \varnothing$.

Precaución: Si no hay hiperplanos paralelos a H que separen estrictamente A y B, esto *no* implicaría necesariamente la intersección de sus envolventes convexas. Tal vez algún otro hiperplano no paralelo a H los separara estrictamente. ■

PROBLEMA DE PRÁCTICA

Sean $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{n}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{n}_2 = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$, y considere que H_1 es el hiperplano (plano) en $\mathbb{R}^3$ que pasa por el punto $\mathbf{p}_1$ y que tiene vector normal $\mathbf{n}_1$, y H_2 es el hiperplano que pasa por el punto $\mathbf{p}_2$ con vector normal $\mathbf{n}_2$. Dé una descripción explícita de $H_1 \cap H_2$ mediante una fórmula que muestre cómo generar todos los puntos en $H_1 \cap H_2$.

8.4 EJERCICIOS

1. Sea L la recta en $\mathbb{R}^2$ que pasa por los puntos $\begin{bmatrix} -1 \\ 4 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$. Encuentre una funcional lineal f y un número real d tales que $L = [f : d]$.

2. Sea L la recta en $\mathbb{R}^2$ que pasa por los puntos $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ y $\begin{bmatrix} -2 \\ -1 \end{bmatrix}$. Obtenga una funcional lineal f y un número real d tales que $L = [f : d]$.

En los ejercicios 3 y 4, determine si cada conjunto es abierto o cerrado, o bien, ni abierto ni cerrado.

3. *a*) $\{(x, y) : y > 0\}$

b) $\{(x, y) : x = 2 \text{ y } 1 \leq y \leq 3\}$

c) $\{(x, y) : x = 2 \text{ y } 1 < y < 3\}$

d) $\{(x, y) : xy = 1 \text{ y } x > 0\}$

e) $\{(x, y) : xy \geq 1 \text{ y } x > 0\}$

4. *a*) $\{(x, y) : x^2 + y^2 = 1\}$

b) $\{(x, y) : x^2 + y^2 > 1\}$

c) $\{(x, y) : x^2 + y^2 \leq 1 \text{ y } y > 0\}$

d) $\{(x, y) : y \geq x^2\}$

e) $\{(x, y) : y < x^2\}$

En los ejercicios 5 y 6, determine si cada conjunto es compacto o no, y si es convexo o no.

5. Utilice los conjuntos del ejercicio 3.

6. Utilice los conjuntos del ejercicio 4.

En los ejercicios 7 a 10, sea H el hiperplano que pasa por los puntos mencionados. *a*) Encuentre un vector $\mathbf{n}$ que sea normal al hiperplano. *b*) Obtenga una funcional lineal f y un número real d tales que $H = [f : d]$.

7. $\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 5 \end{bmatrix}$ **8.** $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 7 \\ -4 \\ 4 \end{bmatrix}$

9. $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

10. $\begin{bmatrix} 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -1 \\ -1 \end{bmatrix}$

11. Sean $\mathbf{p} = \begin{bmatrix} 1 \\ -3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 3 \end{bmatrix}$ y $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 4 \\ 0 \\ 4 \end{bmatrix}$, y H el hiperplano en $\mathbb{R}^4$ con normal $\mathbf{n}$ y que pasa por $\mathbf{p}$. ¿Cuáles de los puntos $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ están sobre un mismo lado de H como el origen, y cuáles no?

12. Sean $\mathbf{a}_1 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{a}_3 = \begin{bmatrix} -1 \\ 6 \\ 0 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$ y $\mathbf{n} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$, y sean $A = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ y $B = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. Encuentre un hiperplano H con normal $\mathbf{n}$ que separa A y B. ¿Existe un hiperplano paralelo a H que separe estrictamente A y B?

13. Sean $\mathbf{p}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 3 \end{bmatrix}$, $\mathbf{n}_1 = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 2 \end{bmatrix}$ y $\mathbf{n}_2 = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 5 \end{bmatrix}$, sean H_1 el hiperplano en $\mathbb{R}^4$ que pasa por $\mathbf{p}_1$ con normal $\mathbf{n}_1$, y H_2 el hiperplano que pasa por $\mathbf{p}_2$ con normal $\mathbf{n}_2$. Dé una descripción explícita de $H_1 \cap H_2$. [*Sugerencia:* Obtenga un punto $\mathbf{p}$ en $H_1 \cap H_2$ y dos vectores linealmente independientes $\mathbf{v}_1$ y $\mathbf{v}_2$ que generen un subespacio paralelo al plano afín bidimensional $H_1 \cap H_2$].

14. Sean F_1 y F_2 planos tetradimensionales en $\mathbb{R}^6$ y suponga que $F_1 \cap F_2 \neq \varnothing$. ¿Cuáles son las posibles dimensiones de $F_1 \cap F_2$?

En los ejercicios 15 a 20, escriba una fórmula para una funcional lineal f y especifique un número d, tal que $[f : d]$ sea el hiperplano descrito en el ejercicio.

15. Sea A la matriz $[1 \ \ -3 \ \ 4 \ \ -2]$ de 1×4, y sea $b = 5$. Además, $H = \{\mathbf{x} \text{ en } \mathbb{R}^4 : A\mathbf{x} = b\}$.

16. Sea A la matriz de $[2 \ \ 5 \ \ -3 \ \ 0 \ \ 6]$ de 1×5. Observe que Nul A está en $\mathbb{R}^5$. Sea $H = \text{Nul } A$.

17. Sea H el plano en $\mathbb{R}^3$ generado por las filas de $B = \begin{bmatrix} 1 & 3 & 5 \\ 0 & 2 & 4 \end{bmatrix}$. Es decir, $H = \text{Fil } B$. [*Sugerencia:* Indague cuál es la relación entre H y Nul B. Véase la sección 6.1].

18. Sea H el plano en $\mathbb{R}^3$ generado por las filas de $B = \begin{bmatrix} 1 & 4 & -5 \\ 0 & -2 & 8 \end{bmatrix}$. Es decir, $H = \text{Fil } B$.

19. Sea H el espacio columna de la matriz $B = \begin{bmatrix} 1 & 0 \\ 4 & 2 \\ -7 & -6 \end{bmatrix}$. Es decir, $H = \text{Col } B$. [*Sugerencia:* Indague cuál es relación entre Col B y Nul B^T. Véase la sección 6.1].

20. Sea H el espacio columna de la matriz $B = \begin{bmatrix} 1 & 0 \\ 5 & 2 \\ -4 & -4 \end{bmatrix}$. Es decir, $H = \text{Col } B$.

En los ejercicios 21 y 22, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

21. *a*) Una transformación lineal de $\mathbb{R}$ a $\mathbb{R}^n$ se llama una funcional lineal.

b) Si f es una funcional lineal definida en $\mathbb{R}^n$, entonces existe un número real k tal que $f(\mathbf{x}) = k\mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

c) Si un hiperplano separa estrictamente a los conjuntos A y B, entonces $A \cap B = \varnothing$.

d) Si A y B son conjuntos convexos cerrados y $A \cap B = \varnothing$, entonces existe un hiperplano que separa estrictamente A y B.

22. *a*) Si d es un número real y f es una funcional lineal diferente de cero definida en $\mathbb{R}^n$, entonces $[f : d]$ es un hiperplano en $\mathbb{R}^n$.

b) Dado cualquier vector $\mathbf{n}$ y cualquier número real d, el conjunto $\{\mathbf{x} : \mathbf{n}\cdot\mathbf{x} = d\}$ es un hiperplano.

c) Si A y B son conjuntos disjuntos no vacíos tales que A es compacto y B es cerrado, entonces existe un hiperplano que separa estrictamente A y B.

d) Si existe un hiperplano H tal que H no separa estrictamente a dos conjuntos A y B, entonces $(\text{conv } A) \cap (\text{conv } B) \neq \varnothing$.

23. Sean $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Encuentre un hiperplano $[f : d]$ (en este caso, una recta) que separe estrictamente $\mathbf{p}$ de conv $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

24. Repita el ejercicio 23 para $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

25. Sea $\mathbf{p} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Encuentre un hiperplano $[f : d]$ que separe estrictamente $B(\mathbf{0}, 3)$ y $B(\mathbf{p}, 1)$. [*Sugerencia:* Después de encontrar f, demuestre que el punto $\mathbf{v} = (1 - .75)\mathbf{0} + .75\mathbf{p}$ no está en $B(\mathbf{0}, 3)$ ni en $B(\mathbf{p}, 1)$].

26. Sean $\mathbf{q} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{p} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$. Encuentre un hiperplano $[f : d]$ que separe estrictamente $B(\mathbf{q}, 3)$ y $B(\mathbf{p}, 1)$.

27. Dé un ejemplo de un subconjunto cerrado S de $\mathbb{R}^2$ tal que conv S no sea cerrado.

28. Dé un ejemplo de un conjunto compacto A y un conjunto cerrado B en $\mathbb{R}^2$ tales que $(\text{conv } A) \cap (\text{conv } B) = \varnothing$, pero considerando que A y B no se puedan separar estrictamente con un hiperplano.

29. Pruebe que la bola abierta $B(\mathbf{p}, \delta) = \{\mathbf{x} : \|\mathbf{x} - \mathbf{p}\| < \delta\}$ es un conjunto convexo. [*Sugerencia:* Utilice la desigualdad del triángulo].

30. Demuestre que la envolvente convexa de un conjunto acotado también está acotada.

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

Primero, calcule $\mathbf{n}_1\cdot\mathbf{p}_1 = -3$ y $\mathbf{n}_2\cdot\mathbf{p}_2 = 7$. El hiperplano H_1 es el conjunto solución de la ecuación $x_1 + x_2 - 2x_3 = -3$, y H_2 es el conjunto solución de la ecuación $-2x_1 + x_2 + 3x_3 = 7$. Entonces,

$$H_1 \cap H_2 = \{\mathbf{x} : x_1 + x_2 - 2x_3 = -3 \text{ y } -2x_1 + x_2 + 3x_3 = 7\}$$

Esta es una descripción implícita de $H_1 \cap H_2$. Para encontrar una descripción explícita, resuelva el sistema de ecuaciones por reducción de filas:

$$\begin{bmatrix} 1 & 1 & -2 & -3 \\ -2 & 1 & 3 & 7 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -\frac{5}{3} & -\frac{10}{3} \\ 0 & 1 & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Por lo tanto, $x_1 = -\frac{10}{3} + \frac{5}{3}x_3, x_2 = \frac{1}{3} + \frac{1}{3}x_3, x_3 = x_3$. Sean $\mathbf{p} = \begin{bmatrix} -\frac{10}{3} \\ \frac{1}{3} \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ 1 \end{bmatrix}$.

La solución general se puede escribir como $\mathbf{x} = \mathbf{p} + x_3\mathbf{v}$. Así, $H_1 \cap H_2$ es la recta que pasa por $\mathbf{p}$ en la dirección de $\mathbf{v}$. Observe que $\mathbf{v}$ es ortogonal tanto a $\mathbf{n}_1$ como a $\mathbf{n}_2$.

8.5 POLÍTOPOS

Esta sección estudia propiedades geométricas de una importante clase de conjuntos convexos compactos llamados polítopos. Esos conjuntos se presentan en todo género de aplicaciones, incluyendo teoría de juegos, programación lineal y problemas de optimización más generales, como el diseño de controles de retroalimentación para sistemas de ingeniería.

Un **polítopo** en $\mathbb{R}^n$ es la envolvente convexa de un conjunto finito de puntos. En $\mathbb{R}^2$, un polítopo es simplemente un polígono. En $\mathbb{R}^3$, un polítopo se llama poliedro. Características importantes de un poliedro son sus caras, aristas y vértices. Por ejemplo, el cubo tiene 6 caras cuadradas, 12 aristas y 8 vértices. Las siguientes definiciones aportan terminología para mayores dimensiones, así como para $\mathbb{R}^2$ y $\mathbb{R}^3$. Recuerde que la dimensión de un conjunto en $\mathbb{R}^n$ es la dimensión del plano afín más pequeño que lo contiene. Además, observe que un polítopo es un tipo especial de conjunto convexo compacto, porque un conjunto finito en $\mathbb{R}^n$ es compacto y la envolvente convexa de este conjunto también es compacta, de acuerdo con el teorema incluido en el recuadro de términos y hechos en topología, de la sección 8.4.

DEFINICIÓN

Sea S un subconjunto convexo compacto de $\mathbb{R}^n$. Un subconjunto F no vacío de S es una **cara** (propia) de S si $F \neq S$ y existe un hiperplano $H = [f : d]$ tal que $F = S \cap H$ y $f(S) \leq d$ o $f(S) \geq d$. El hiperplano H es un **hiperplano de soporte** para S. Si la dimensión de F es k, entonces F es una ***k*-cara** de S.

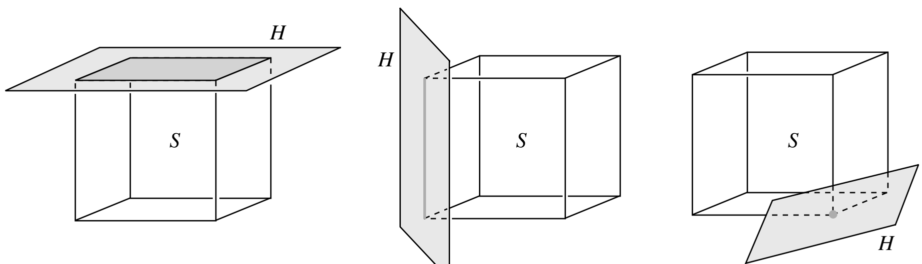
Si P es un polítopo de dimensión k, entonces P se llama un ***k*-polítopo**. Una 0-cara de P se llama **vértice**, una 1-cara es una **arista**, y una $(k - 1)$-dimensional cara es una **faceta** de S.

EJEMPLO 1 Suponga que S es un cubo en $\mathbb{R}^3$. Cuando un plano H se traslada en $\mathbb{R}^3$ hasta que apenas toca (soporta) al cubo, pero no corta el interior de este, entonces hay tres posibilidades para $H \cap S$, dependiendo de la orientación de H. (Véase la figura 1).

$H \cap S$ puede ser una cara cuadrada bidimensional (faceta) del cubo.

$H \cap S$ puede ser una arista unidimensional del cubo.

$H \cap S$ puede ser un vértice de dimensión 0 del cubo. ■



$H \cap S$ es bidimensional. $H \cap S$ es unidimensional. $H \cap S$ es de dimensión 0.

FIGURA 1

La mayoría de las aplicaciones de polítopos de alguna manera implican a los vértices, porque estos tienen una propiedad especial que se identifica en la siguiente definición.

DEFINICIÓN

Sea S un conjunto convexo. Un punto $\mathbf{p}$ en S es un **punto extremo** de S si $\mathbf{p}$ no está en el interior de algún segmento de recta contenido en S. Más precisamente, si $\mathbf{x}$, $\mathbf{y} \in S$ y $\mathbf{p} \in \overline{\mathbf{xy}}$, entonces $\mathbf{p} = \mathbf{x}$ o $\mathbf{p} = \mathbf{y}$. El conjunto de todos los puntos extremos de S se llama **perfil** de S.

Un vértice de cualquier conjunto convexo S compacto es automáticamente un punto extremo de S. Este hecho queda asentado en la demostración del teorema 14, que se presenta a continuación. Al trabajar con un polítopo, por ejemplo, $P = \text{conv}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ para $\mathbf{v}_1,\ldots,\mathbf{v}_k$ en $\mathbb{R}^n$, por lo general es útil saber que $\mathbf{v}_1,\ldots,\mathbf{v}_k$ son los puntos extremos de P. Sin embargo, esta lista podría contener puntos extraños. Es decir, por ejemplo, algún vector $\mathbf{v}_i$ podría ser el punto medio de una arista del polítopo. Desde luego, en este caso $\mathbf{v}_i$ realmente no se necesita para generar la envolvente convexa. La siguiente definición describe la propiedad de los vértices que hace que sean puntos extremos.

DEFINICIÓN

El conjunto $\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ es una **representación mínima** del polítopo P si $P = \text{conv}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ y para cada $i = 1,\ldots,k$, $\mathbf{v}_i \notin \text{conv}\,\{\mathbf{v}_j : j \neq i\}$.

Cada polítopo tiene una representación mínima. Si $P = \text{conv}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ y si alguna $\mathbf{v}_i$ es una combinación convexa de los otros puntos, entonces es posible eliminar $\mathbf{v}_i$ del conjunto de puntos sin cambiar la envolvente convexa. Este proceso se puede repetir hasta obtener la representación mínima. Es factible demostrar que la representación mínima es única.

TEOREMA 14

Suponga que $M = \{\mathbf{v}_1,\ldots,\mathbf{v}_k\}$ es la representación mínima del polítopo P. Entonces los siguientes tres enunciados son equivalentes:

a) $\mathbf{p} \in M$.

b) $\mathbf{p}$ es un vértice de P.

c) $\mathbf{p}$ es un punto extremo de P.

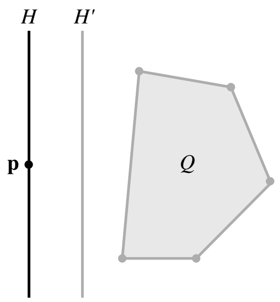


FIGURA 2

DEMOSTRACIÓN *a*) ⇒ *b*). Suponga que $\mathbf{p} \in M$ y sea $Q = \text{conv}\,\{\mathbf{v} : \mathbf{v} \in M \text{ y } \mathbf{v} \neq \mathbf{p}\}$. De la definición de M se deduce que $\mathbf{p} \notin Q$, y como Q es compacto, el teorema 13 implica la existencia de un hiperplano H' que separa estrictamente $\{\mathbf{p}\}$ y Q. Sea H el hiperplano que pasa por $\mathbf{p}$ paralelo a H'. Véase la figura 2.

Entonces Q está en uno de los semiespacios cerrados H^+ acotado por H, y $P \subseteq H^+$. Por lo tanto, H es soporte para P en $\mathbf{p}$. Además, $\mathbf{p}$ es el único punto de P que puede estar en H, de manera que $H \cap P = \{\mathbf{p}\}$, y $\mathbf{p}$ es un vértice de P.

b) ⇒ *c*). Sea $\mathbf{p}$ un vértice en P. Entonces existe un hiperplano $H = [f : d]$ tal que $H \cap P = \{\mathbf{p}\}$ y $f(P) \geq d$. Si $\mathbf{p}$ no fuera un punto extremo, entonces se tendrían puntos $\mathbf{x}$ y $\mathbf{y}$ en P tales que $\mathbf{p} = (1 - c)\mathbf{x} + c\mathbf{y}$ con $0 < c < 1$. Es decir,

$$c\mathbf{y} = \mathbf{p} - (1 - c)\mathbf{x} \quad \text{y} \quad \mathbf{y} = \left(\frac{1}{c}\right)(\mathbf{p}) - \left(\frac{1}{c} - 1\right)(\mathbf{x})$$

Se deduce que $f(\mathbf{y}) = \dfrac{1}{c} f(\mathbf{p}) - \left(\dfrac{1}{c} - 1\right) f(\mathbf{x})$. Pero, $f(\mathbf{p}) = d$ y $f(\mathbf{x}) \geq d$, por lo que

$$f(\mathbf{y}) \leq \left(\frac{1}{c}\right)(d) - \left(\frac{1}{c} - 1\right)(d) = d$$

Por otro lado, $\mathbf{y} \in P$, así que $f(\mathbf{y}) \geq d$. Se deduce que $f(\mathbf{y}) = d$ y que $\mathbf{y} \in H \cap P$. Esto contradice el hecho de que $\mathbf{p}$ es un vértice. Entonces, $\mathbf{p}$ debe ser un punto extremo. (Observe que esta parte de la demostración no depende de que P sea un polítopo. Es válida para cualquier conjunto convexo compacto).

c) ⇒ *a*). Es claro que cualquier punto extremo de P debe ser un miembro de M. ■

EJEMPLO 2 Recuerde que el perfil de un conjunto S es el conjunto de puntos extremos de S. El teorema 14 muestra que el perfil de un polígono en $\mathbb{R}^2$ es el conjunto de vértices. (Véase la figura 3). El perfil de una bola cerrada es su frontera. Un conjunto abierto no tiene puntos extremos, de manera que su perfil es vacío. Un semiespacio cerrado no tiene puntos extremos, por lo que su perfil es vacío. ■

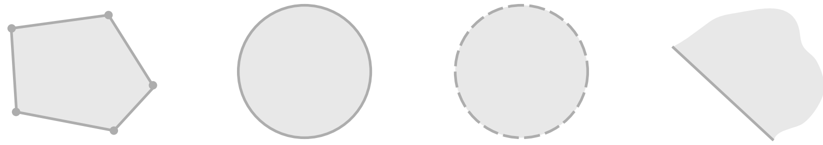


FIGURA 3

El ejercicio 18 le pide demostrar que un punto $\mathbf{p}$ en un conjunto convexo S es un punto extremo de S si, y solo si, cuando $\mathbf{p}$ se elimina de S, los puntos restantes aún siguen formando un conjunto convexo. Se deduce que si S^* es cualquier subconjunto de S tal que conv S^* es igual a S, entonces S^* debe contener el perfil de S. Los conjuntos del ejemplo 2 muestran que, en general, S^* puede ser más grande que el perfil de S. Sin embargo, es cierto que cuando S es compacto, S^* se puede considerar como el perfil de S, como lo demostrará el teorema 15. Así, cada conjunto S compacto no vacío tiene un punto extremo, y el conjunto de todos los puntos extremos es el subconjunto más pequeño de S cuya envolvente convexa es igual a S.

TEOREMA 15

Sea S un conjunto convexo compacto no vacío. Entonces S es la envolvente convexa de su perfil (el conjunto de puntos extremos de S).

DEMOSTRACIÓN La demostración es por inducción sobre la dimensión del conjunto S.[1] ■

Una importante aplicación del teorema 15 es el siguiente teorema. Es uno de los resultados teóricos clave en el desarrollo de la programación lineal. Las funcionales lineales son continuas, y las funciones continuas siempre alcanzan sus máximos y mínimos sobre un conjunto compacto. El significado del teorema 16 es que para conjuntos convexos compactos, el máximo (y el mínimo) realmente se obtiene en un punto extremo de S.

TEOREMA 16

Sea f una funcional lineal definida sobre un conjunto convexo S compacto y no vacío. Entonces, existen puntos extremos $\hat{\mathbf{v}}$ y $\hat{\mathbf{w}}$ de S tales que

$$f(\hat{\mathbf{v}}) = \max_{\mathbf{v}\in S} f(\mathbf{v}) \quad \text{y} \quad f(\hat{\mathbf{w}}) = \min_{\mathbf{v}\in S} f(\mathbf{v})$$

DEMOSTRACIÓN Suponga que f alcanza su máximo m sobre S en algún punto $\mathbf{v}'$ en S. Es decir, $f(\mathbf{v}') = m$. Se desea demostrar que existe un punto extremo en S con la misma propiedad. De acuerdo con el teorema 15, $\mathbf{v}'$ es una combinación convexa de los puntos extremos de S. Es decir, existen puntos extremos $\mathbf{v}_1,\ldots, \mathbf{v}_k$ de S y no negativos $c_1,\ldots, c_k$ tales que

$$\mathbf{v}' = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \quad \text{con } c_1 + \cdots + c_k = 1$$

Si ninguno de los puntos extremos de S satisface $f(\mathbf{v}) = m$, entonces

$$f(\mathbf{v}_i) < m \quad \text{para } i = 1, \ldots, k$$

[1] Los detalles se pueden encontrar en Steven R. Lay, *Convex Sets and Their Applications* (Nueva York: John Wiley & Sons, 1982; Mineola, NY: Dover Publications, 2007, p. 43.

porque m es el máximo de f sobre S. Pero entonces, como f es lineal,

$$\begin{aligned} m = f(\mathbf{v}') &= f(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \\ &= c_1 f(\mathbf{v}_1) + \cdots + c_k f(\mathbf{v}_k) \\ &< c_1 m + \cdots + c_k m = m(c_1 + \cdots + c_k) = m \end{aligned}$$

Esta contradicción implica que algún punto extremo $\hat{\mathbf{v}}$ de S debe satisfacer $f(\hat{\mathbf{v}}) = m$.

La demostración para $\hat{\mathbf{w}}$ es similar. ∎

EJEMPLO 3 Dados los puntos $\mathbf{p}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{p}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ en $\mathbb{R}^2$, sea $S =$ conv $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. Para cada funcional lineal f, encuentre el valor máximo m de f sobre el conjunto S, y obtenga todos los puntos $\mathbf{x}$ en S en los que $f(\mathbf{x}) = m$.

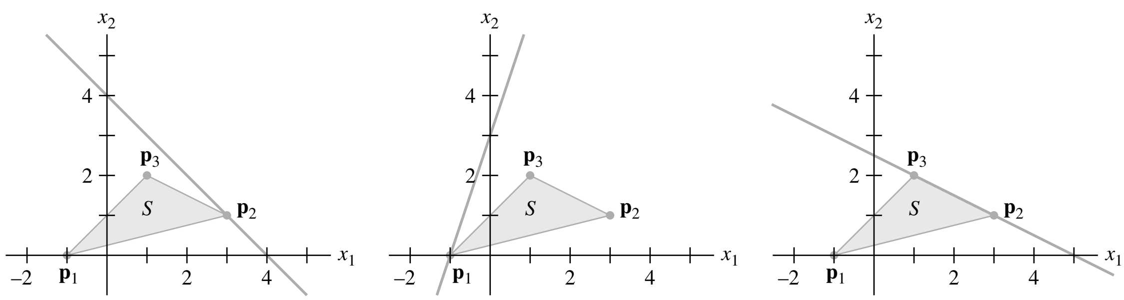
a) $f_1(x_1, x_2) = x_1 + x_2$ *b)* $f_2(x_1, x_2) = -3x_1 + x_2$ *c)* $f_3(x_1, x_2) = x_1 + 2x_2$

SOLUCIÓN De acuerdo con el teorema 16, el valor máximo se alcanza en uno de los puntos extremos de S. Así que para encontrar m, se evalúa f en cada punto extremo y se selecciona el valor más grande.

a) $f_1(\mathbf{p}_1) = -1$, $f_1(\mathbf{p}_2) = 4$ y $f_1(\mathbf{p}_3) = 3$, de manera que $m_1 = 4$. Trace la gráfica de la recta $f_1(x_1, x_2) = m_1$, es decir, $x_1 + x_2 = 4$, y observe que $\mathbf{x} = \mathbf{p}_2$ es el único punto en S en el cual $f_1(\mathbf{x}) = 4$. Véase la figura 4*a*).

b) $f_2(\mathbf{p}_1) = 3$, $f_2(\mathbf{p}_2) = -8$ y $f_2(\mathbf{p}_3) = -1$, por lo que $m_1 = 3$. Trace la gráfica de la recta $f_2(x_1, x_2) = m_2$, es decir, $-3x_1 + x_2 = 3$, y observe que $\mathbf{x} = \mathbf{p}_1$ es el único punto en S en el cual $f_2(\mathbf{x}) = 3$. Véase la figura 4*b*).

c) $f_3(\mathbf{p}_1) = -1$, $f_3(\mathbf{p}_2) = 5$ y $f_3(\mathbf{p}_3) = 5$, de manera que $m_1 = 5$. Trace la gráfica de la recta $f_3(x_1, x_2) = m_3$, es decir, $x_1 + 2x_2 = 5$. Aquí, f_3 alcanza su valor máximo en $\mathbf{p}_2$, en $\mathbf{p}_3$ y en cada punto de la envolvente convexa de $\mathbf{p}_2$ y $\mathbf{p}_3$. Véase la figura 4*c*). ∎



a) $x_1 + x_2 = 4$ *b)* $-3x_1 + x_2 = 3$ *c)* $x_1 + 2x_2 = 5$

FIGURA 4

La situación ilustrada en el ejemplo 3 para $\mathbb{R}^2$ también se aplica en mayores dimensiones. El valor máximo de una funcional lineal f en un polítopo P está en la intersección de un hiperplano soporte y P. Esta intersección es un solo punto extremo de P, o bien, la envolvente convexa de 2 o más puntos extremos de P. En cualquier caso, la intersección es un polítopo, y sus puntos extremos forman un subconjunto de los puntos extremos de P.

Por definición, un polítopo es la envolvente convexa de un conjunto finito de puntos. Esta es una representación explícita del polítopo porque identifica puntos en el conjunto. Un polítopo también se puede representar implícitamente como la intersección de un número finito de semiespacios cerrados. El ejemplo 4 ilustra esto en $\mathbb{R}^2$.

EJEMPLO 4 Sean

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{p}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{p}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

en $\mathbb{R}^2$, y $S = \text{conv}\,\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. Álgebra sencilla indica que la recta que pasa por $\mathbf{p}_1$ y $\mathbf{p}_2$ está dada por $x_1 + x_2 = 1$, y S está en el lado de esta recta donde

$$x_1 + x_2 \geq 1 \quad \text{o, de manera equivalente,} \quad -x_1 - x_2 \leq -1.$$

De forma similar, la recta que pasa por $\mathbf{p}_2$ y $\mathbf{p}_3$ es $x_1 - x_2 = 1$, y S está sobre el lado donde

$$x_1 - x_2 \leq 1$$

Además, la recta que pasa por $\mathbf{p}_3$ y $\mathbf{p}_1$ es $-x_1 + 3x_2 = 3$, y S se encuentra en el lado donde

$$-x_1 + 3x_2 \leq 3$$

Véase la figura 5. Se deduce que S se puede describir como el conjunto solución del sistema de desigualdades lineales

$$\begin{aligned} -x_1 - x_2 &\leq -1 \\ x_1 - x_2 &\leq 1 \\ -x_1 + 3x_2 &\leq 3 \end{aligned}$$

Este sistema se puede escribir como $A\mathbf{x} \leq \mathbf{b}$, donde

$$A = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}.$$

Observe que una desigualdad entre dos vectores, tales como $A\mathbf{x}$ y $\mathbf{b}$, se aplica a cada una de las coordenadas correspondientes en esos vectores. ■

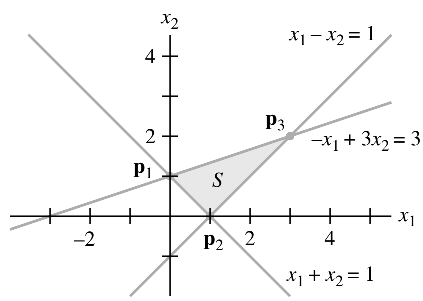


FIGURA 5

En ocasiones será necesario remplazar una descripción implícita de un polítopo por una representación mínima del polítopo, listando todos los puntos extremos del polítopo. En casos sencillos, es factible una solución gráfica. El siguiente ejemplo muestra cómo manejar la situación cuando varios puntos de interés están muy cercanos para identificarlos fácilmente en una gráfica.

EJEMPLO 5 Sea P el conjunto de puntos en $\mathbb{R}^2$ que satisfacen $A\mathbf{x} \leq \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 3 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 18 \\ 8 \\ 21 \end{bmatrix}$$

y $\mathbf{x} \geq \mathbf{0}$. Encuentre la representación mínima de P.

SOLUCIÓN La condición $\mathbf{x} \geq \mathbf{0}$ coloca a P en el primer cuadrante de $\mathbb{R}^2$, una condición típica en problemas de programación lineal. Las tres desigualdades en $A\mathbf{x} \leq \mathbf{b}$ implican tres rectas frontera:

$$(1)\ \ x_1 + 3x_2 = 18 \quad (2)\ \ x_1 + x_2 = 8 \quad (3)\ \ 3x_1 + 2x_2 = 21$$

Las tres rectas tienen pendientes negativas, por lo que una idea general de la forma de P es fácil de visualizar. Incluso un burdo esquema de las gráficas de esas rectas revelará que (0, 0), (7, 0) y (0, 6) son vértices del polítopo P.

¿Qué ocurre con las intersecciones de las rectas (1), (2) y (3)? A partir de la gráfica, algunas veces es claro cuáles intersecciones incluir. Si no es así, el siguiente procedimiento algebraico funcionará bien:

> Cuando se encuentra un punto de intersección que corresponde a dos desigualdades, pruébelo en las otras desigualdades para ver si el punto está en el polítopo.

La intersección de (1) y (2) es $\mathbf{p}_{12} = (3, 5)$. Ambas coordenadas son no negativas, así que $\mathbf{p}_{12}$ satisface todas las desigualdades, excepto quizá la tercera. Pruebe esto:

$$3(3) + 2(5) = 19 < 21$$

Este punto de intersección satisface la desigualdad para (3), de manera que $\mathbf{p}_{12}$ está en el polítopo.

La intersección de (2) y (3) es $\mathbf{p}_{23} = (5, 3)$. Esto satisface todas las desigualdades, excepto quizá la desigualdad para (1). Pruebe esto:

$$1(5) + 3(3) = 14 < 18$$

Esto demuestra que $\mathbf{p}_{23}$ está en el polítopo.

Por último, la intersección de (1) y (3) es $\mathbf{p}_{13} = \left(\frac{27}{7}, \frac{33}{7}\right)$. Pruebe esto en la desigualdad para (2):

$$1\left(\tfrac{27}{7}\right) + 1\left(\tfrac{33}{7}\right) = \tfrac{60}{7} \approx 8.6 > 8$$

Por lo tanto, $\mathbf{p}_{13}$ **no** satisface la segunda desigualdad, lo que demuestra que $\mathbf{p}_{13}$ **no** está en P. En conclusión, la representación mínima del polítopo P es

$$\left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 6 \end{bmatrix}\right\}$$ ■

El resto de esta sección analiza la construcción de dos polítopos básicos en $\mathbb{R}^3$ (y mayores dimensiones). El primero se presenta en problemas de programación lineal. Ambos polítopos brindan oportunidades para visualizar $\mathbb{R}^4$ en una forma notable.

Simplejo

Un **simplejo** es la envolvente convexa de un conjunto finito de vectores afínmente independientes. Para construir un simplejo k-dimensional (o k-simplejo), se procede como sigue:

0-simplejo S^0: un solo punto $\{\mathbf{v}_1\}$
1-simplejo S^1: conv $(S^0 \cup \{\mathbf{v}_2\})$, y $\mathbf{v}_2$ no está en aff S^0
2-simplejo S^2: conv $(S^1 \cup \{\mathbf{v}_3\})$, y $\mathbf{v}_3$ no se encuentra en aff S^1
⋮
k-simplejo S^k: conv $(S^{k-1} \cup \{\mathbf{v}_{k+1}\})$, y $\mathbf{v}_{k+1}$ no está en aff S^{k-1}

El simplejo S^1 es un segmento de recta. El triángulo S^2 proviene de seleccionar un punto $\mathbf{v}_3$ que no está en la recta que contiene a S_1 y después formar la envolvente convexa con S_2.

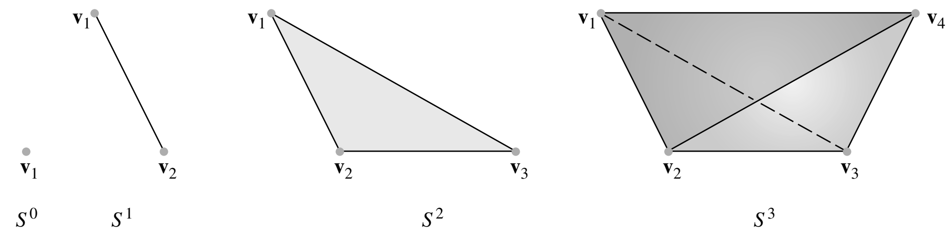


FIGURA 6

(Véase la figura 6). El tetraedro S_3 se produce al elegir un punto $\mathbf{v}_4$ que no esté en el plano de S^2 y después formar la envolvente convexa con S^2.

Antes de continuar, considere algunos de los patrones que se están presentando. El triángulo S^2 tiene tres aristas. Cada una de esas aristas es un segmento de recta semejante a S^1. ¿De dónde provienen esos tres segmentos de recta? Uno de ellos es S^1. Otro viene de la unión del punto final $\mathbf{v}_2$ con el nuevo punto $\mathbf{v}_3$. El tercero proviene de unir el otro punto final $\mathbf{v}_1$ con $\mathbf{v}_3$. Se podría decir que cada punto extremo en S^1 se estira dentro de un segmento de recta en S^2.

El tetraedro S^3 de la figura 6 tiene cuatro caras triangulares. Una de ellas es el triángulo original S^2, y las otras tres se obtienen al estirar las aristas de S^2 al nuevo punto $\mathbf{v}_4$. Además, observe que los vértices de S^2 se estiran para formar las aristas de S^3. Las otras aristas de S^3 provienen de las aristas de S^2. Esto sugiere cómo "visualizar" el S^4 de cuatro dimensiones.

La construcción de S^4, llamado pentatopo, implica formar la envolvente convexa de S^3 con un punto $\mathbf{v}_5$ que no está en el 3-espacio S^3. Desde luego, es imposible un esquema completo, pero la figura 7 es sugestiva: S^4 tiene cinco vértices, y cualesquiera cuatro de los vértices determinan una faceta en la forma de un tetraedro. Por ejemplo, la figura enfatiza la faceta con vértices $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_4$ y $\mathbf{v}_5$ y la faceta con vértices $\mathbf{v}_2$, $\mathbf{v}_3$, $\mathbf{v}_4$ y $\mathbf{v}_5$. Hay cinco de estas

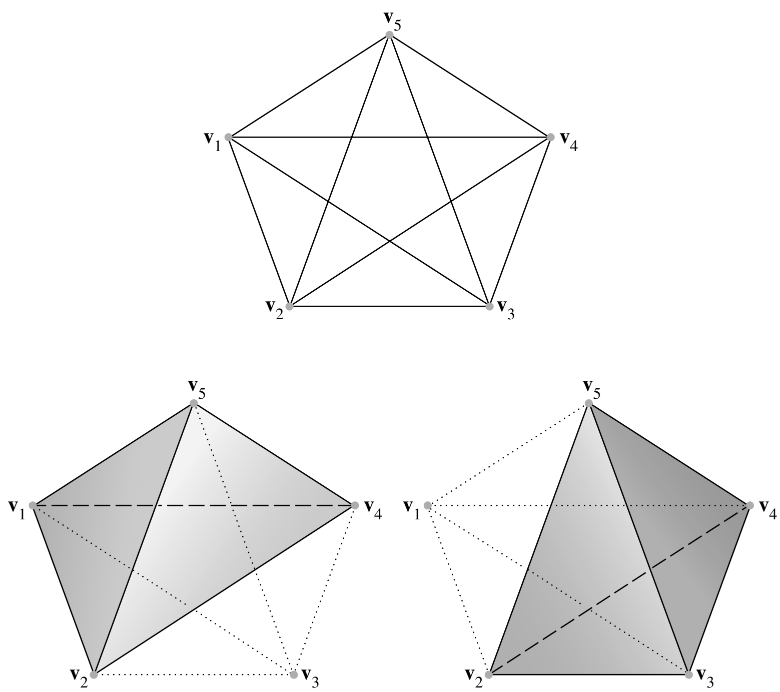


FIGURA 7 El simplejo de cuatro dimensiones S^4 proyectado sobre $\mathbb{R}^2$, con dos facetas tetraédricas resaltadas.

facetas. La figura 7 identifica las 10 aristas de S^4, las cuales se pueden emplear para visualizar las 10 caras triangulares.

La figura 8 muestra otra representación del simplejo de cuatro dimensiones S^4. Esta vez el quinto vértice se presenta "dentro" del tetraedro S^3. Las facetas tetraédricas resaltadas también parecen estar "dentro" de S^3.

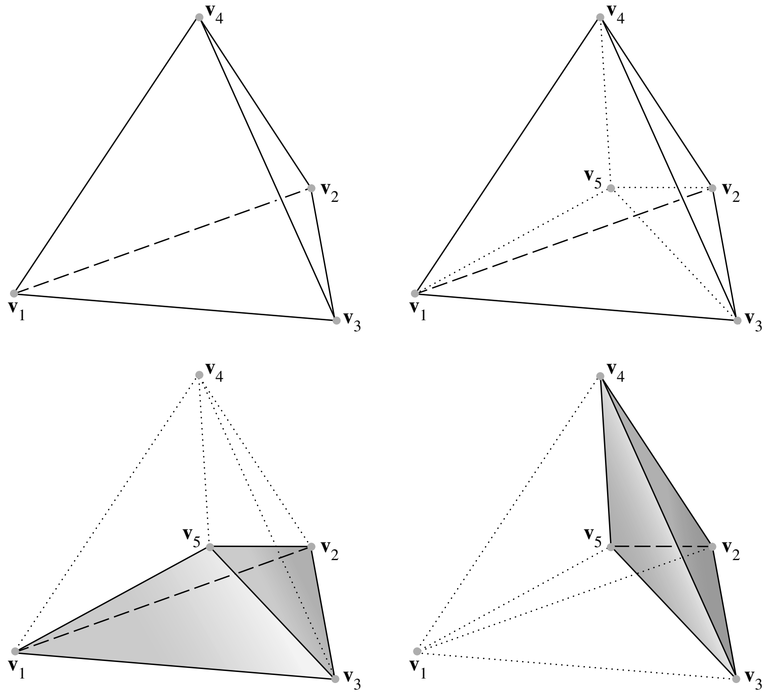


FIGURA 8 El quinto vértice de S^4 está "dentro" de S^3.

Hipercubo

Sea $I_i = \overline{\mathbf{0}\mathbf{e}_i}$ el segmento de recta del origen $\mathbf{0}$ al vector $\mathbf{e}_i$ de la base estándar en $\mathbb{R}^n$. Entonces, para k tal que $1 \leq k \leq n$, el vector suma[2]

$$C^k = I_1 + I_2 + \cdots + I_k$$

se llama el **hipercubo** k-dimensional.

Para visualizar la construcción de C^k, inicie con los casos sencillos. El hipercubo C^1 es el segmento de recta I_1. Si C^1 se traslada mediante $\mathbf{e}_2$, la envolvente convexa de sus posiciones inicial y final describe un cuadrado C^2. (Véase la figura 9 de la página 478). Al trasladar C^2 mediante $\mathbf{e}_3$ se produce el cubo C^3. Una traslación similar de C^3 mediante el vector $\mathbf{e}_4$ produce el hipercubo de cuatro dimensiones C^4.

Una vez más, esto es difícil de visualizar, pero la figura 10 muestra una proyección bidimensional de C^4. Cada una de las aristas de C^3 se estira para formar una cara cuadrada de C^4. Y cada cara cuadrada de C^3 se amplía para formar una cara cúbica de C^4. La figura 11 muestra tres facetas de C^4. El inciso *a*) resalta el cubo proveniente de la cara cuadrada izquierda de C^3. El inciso *b*) muestra el cubo que proviene de la cara cuadrada frontal de C^3. Y el inciso *c*) destaca el cubo proveniente de la cara cuadrada superior de C^3.

[2] El vector suma de dos conjuntos A y B se define mediante $A + B = \{\mathbf{c} : \mathbf{c} = \mathbf{a} + \mathbf{b}$ para algunas $\mathbf{a} \in A$ y $\mathbf{b} \in B\}$.

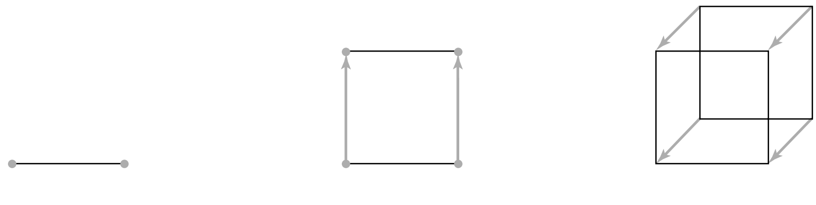


FIGURA 9 Construcción del cubo C^3.

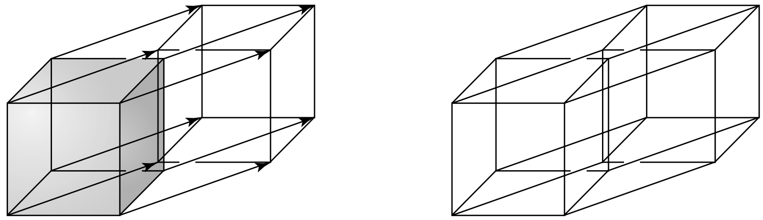


FIGURA 10 C^4 proyectado sobre $\mathbb{R}^2$.

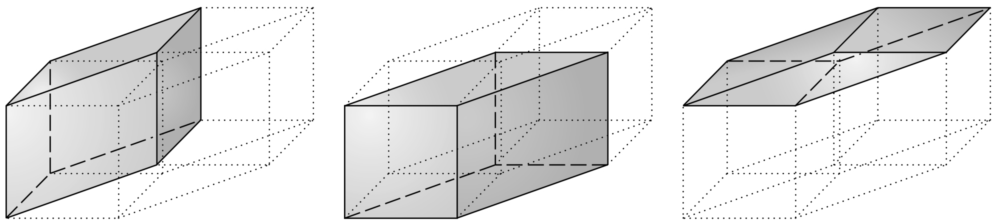


FIGURA 11 Tres de las facetas cúbicas de C^4.

La figura 12 muestra otra representación de C^4 en la que el cubo trasladado se coloca “dentro” de C^3. Esto hace más fácil visualizar las facetas cúbicas de C^4, porque existe menos distorsión.

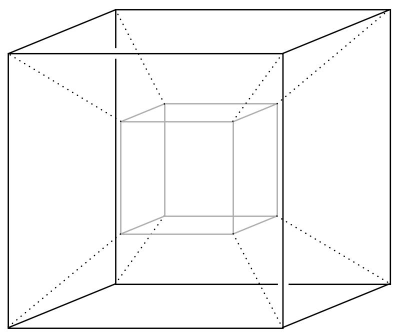


FIGURA 12 La imagen trasladada de C^3 se coloca “dentro” de C^3 para obtener C^4.

En total, el cubo de cuatro dimensiones C^4 tiene ocho caras cúbicas. Dos de ellas provienen de las imágenes originales y trasladadas de C^3, y seis provienen de las caras cuadradas de C^3 que se amplían para formar cubos. Las caras cuadradas bidimensionales de C^4 provienen de las caras cuadradas de C^3 y de su traslado, y las aristas de C^3 que se estiran para formar cuadrados. Por lo tanto, hay $2 \times 6 + 12 = 24$ caras cuadradas. Para contar las aristas, multiplique por dos el número de aristas en C^3 y sume el número de vértices en C^3. Esto da $2 \times 12 + 8 = 32$ aristas en C^4. Todos los vértices en C^4 provienen de C^3 y su traslado, de manera que hay $2 \times 8 = 16$ vértices.

Uno de los resultados verdaderamente notables en el estudio de polítopos es la siguiente fórmula; Leonard Euler (1707-1783) fue el primero en probarla. La fórmula establece una relación sencilla entre el número de caras de diferentes dimensiones en un polítopo. Para simplificar el enunciado de la fórmula, sea $f_k(P)$ el número de caras k-dimensionales de un polítopo n-dimensional P.[3]

$$\text{Fórmula de Euler:}\quad \sum_{k=0}^{n-1}(-1)^k f_k(P) = 1 + (-1)^{n-1}$$

En particular, cuando $n = 3$, $v - e + f = 2$, donde v, e y f denotan el número de vértices, aristas y facetas, respectivamente, de P.

PROBLEMA DE PRÁCTICA

1. Encuentre la representación mínima del polítopo P definido por las desigualdades $A\mathbf{x} \le \mathbf{b}$ y $\mathbf{x} \ge \mathbf{0}$, cuando $A = \begin{bmatrix} 1 & 3 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 12 \\ 9 \\ 12 \end{bmatrix}$.

8.5 EJERCICIOS

1. A partir de los puntos $\mathbf{p}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{p}_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ en $\mathbb{R}^2$, sea $S = \text{conv}\,\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. Para cada funcional lineal f, encuentre el valor máximo m de f en el conjunto S, y obtenga todos los puntos $\mathbf{x}$ en S en los cuales $f(\mathbf{x}) = m$.

   a) $f(x_1, x_2) = x_1 - x_2$ *b*) $f(x_1, x_2) = x_1 + x_2$

   c) $f(x_1, x_2) = -3x_1 + x_2$

2. Dados los puntos $\mathbf{p}_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{p}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ en $\mathbb{R}^2$, sea $S = \text{conv}\,\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. Para cada funcional lineal f, encuentre el valor máximo m de f en el conjunto S, y determine todos los puntos $\mathbf{x}$ en S en los cuales $f(\mathbf{x}) = m$.

   a) $f(x_1, x_2) = x_1 + x_2$ *b*) $f(x_1, x_2) = x_1 - x_2$

   c) $f(x_1, x_2) = -2x_1 + x_2$

3. Repita el ejercicio 1 considerando que m es el valor *mínimo* de f en S y no el valor máximo.

4. Repita el ejercicio 2 considerando que m es el valor *mínimo* de f en S y no el valor máximo.

En los ejercicios 5 a 8, encuentre la representación mínima del polítopo definido por las desigualdades $A\mathbf{x} \le \mathbf{b}$ y $\mathbf{x} \ge \mathbf{0}$.

5. $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 10 \\ 15 \end{bmatrix}$

6. $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 18 \\ 16 \end{bmatrix}$

7. $A = \begin{bmatrix} 1 & 3 \\ 1 & 1 \\ 4 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 18 \\ 10 \\ 28 \end{bmatrix}$

8. $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 8 \\ 6 \\ 7 \end{bmatrix}$

9. Sea $S = \{(x, y) : x^2 + (y - 1)^2 \le 1\} \le \{(3, 0)\}$. ¿El origen es un punto extremo de conv S? ¿Es el origen un vértice de conv S?

10. Encuentre un ejemplo de un conjunto convexo S cerrado en $\mathbb{R}^2$ tal que su perfil P sea no vacío, pero conv $P \neq S$.

11. Encuentre un ejemplo de un conjunto convexo S acotado en $\mathbb{R}^2$ tal que su perfil P sea no vacío, pero conv $P \neq S$.

12. *a*) Determine el número de k-caras del simplejo de cinco dimensiones S^5 para $k = 0, 1,\ldots, 4$. Verifique que su respuesta satisfaga la fórmula de Euler.

    b) Realice una tabla de los valores de $f_k(S^n)$ para $n = 1,\ldots, 5$ y $k = 0, 1,\ldots, 4$. ¿Identifica un patrón? Proponga una fórmula general para $f_k(S^n)$.

---

[3] Se presenta una demostración en Steven R. Lay, *Convex Sets and Their Applications* (Nueva York: John Wiley & Sons, 1982; Mineola, NY: Dover Publications, 2007), p. 131.

13. *a*) Determine el número de k-caras del hipercubo de cinco dimensiones C^5 para $k = 0, 1, \ldots, 4$. Compruebe que su respuesta satisfaga la fórmula de Euler.

b) Construya una tabla de los valores de $f_k(C^n)$ para $n = 1, \ldots, 5$ y $k = 0, 1, \ldots, 4$. ¿Identifica un patrón? Proponga una fórmula general para $f_k(C^n)$.

14. Suponga que $\mathbf{v}_1, \ldots, \mathbf{v}_k$ son vectores linealmente independientes en $\mathbb{R}^n$ $(1 \leq k \leq n)$. Entonces el conjunto $X^k = \text{conv}\,\{\pm\mathbf{v}_1, \ldots, \pm\mathbf{v}_k\}$ es un ***k*-politopo de cruce** (u **ortoplex**).

a) Elabore un bosquejo de X^1 y X^2.

b) Determine el número de k-caras del politopo de cruce tridimensional X^3 para $k = 0, 1, 2$. ¿Qué otro nombre tiene X^3?

c) Obtenga el número de k-caras del politopo de cruce de cuatro dimensiones X^4 para $k = 0, 1, 2, 3$. Compruebe que su respuesta satisfaga la fórmula de Euler.

d) Encuentre una fórmula para $f_k(X^n)$, el número de k-caras de X^n, para $0 \leq k \leq n - 1$.

15. Una ***k*-pirámide** P^k es la envolvente convexa de un $(k - 1)$-politopo Q y un punto $\mathbf{x} \notin \text{aff}\, Q$. Encuentre una fórmula para cada uno de los siguientes casos en términos de $f_j(Q)$, $j = 0, \ldots, n - 1$.

a) El número de vértices de $P^n : f_0(P^n)$.

b) El número de k-caras de $P^n : f_k(P^n)$, para $1 \leq k \leq n - 2$.

c) El número de facetas $(n - 1)$-dimensionales de P^n: $f_{n-1}(P^n)$.

En los ejercicios 16 y 17, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

16. *a*) Un politopo es la envolvente convexa de un conjunto finito de puntos.

b) Sea $\mathbf{p}$ un punto extremo de un conjunto convexo S. Si $\mathbf{u}, \mathbf{v} \in S$, $\mathbf{p} \in \overline{\mathbf{uv}}$ y $\mathbf{p} \neq \mathbf{u}$, entonces $\mathbf{p} = \mathbf{v}$.

c) Si S es un subconjunto convexo no vacío de $\mathbb{R}^n$, entonces S es la envolvente convexa de su perfil.

d) El simplejo de cuatro dimensiones S^4 tiene exactamente cinco facetas, cada una de las cuales es un tetraedro tridimensional.

17. *a*) Un cubo en $\mathbb{R}^3$ tiene cinco facetas.

b) Un punto $\mathbf{p}$ es un punto extremo de un politopo P si y solo si $\mathbf{p}$ es un vértice de P.

c) Si S es un conjunto convexo, compacto y no vacío, y una funcional lineal alcanza su máximo en un punto $\mathbf{p}$, entonces $\mathbf{p}$ es un punto extremo de S.

d) Un politopo bidimensional siempre tiene el mismo número de vértices y aristas.

18. Sea $\mathbf{v}$ un elemento del conjunto convexo S. Pruebe que $\mathbf{v}$ es un punto extremo de S si y solo si el conjunto $\{\mathbf{x} \in S : \mathbf{x} \neq \mathbf{v}\}$ es convexo.

19. Si $c \in \mathbb{R}$ y S es un conjunto, defina $cS = \{c\mathbf{x} : \mathbf{x} \in S\}$. Sea S un conjunto convexo y suponga que $c > 0$ y $d > 0$. Demuestre que $cS + dS = (c + d)S$.

20. Encuentre un ejemplo para demostrar que la convexidad de S es necesaria en el ejercicio 19.

21. Si A y B son conjuntos convexos, demuestre que $A + B$ es convexo.

22. Un poliedro (3-politopo) se llama **regular** si todas sus facetas son polígonos regulares congruentes y todos los ángulos en los vértices son iguales. En la siguiente demostración aporte los detalles de que solamente existen cinco poliedros regulares.

a) Suponga que un poliedro regular tiene r facetas, cada una de las cuales es un polígono regular de k lados, y que s aristas concurren en cada vértice. Considerando que v y e denotan los números de vértices y aristas en el poliedro, explique por qué $kr = 2e$ y $sv = 2e$.

b) Utilice la fórmula de Euler para demostrar que $\dfrac{1}{s} + \dfrac{1}{k} = \dfrac{1}{2} + \dfrac{1}{e}$.

c) Obtenga todas las soluciones integrales de la ecuación en el inciso *b*) que satisfacen las restricciones geométricas del problema. (¿Qué tan pequeños pueden ser k y s?).

Para su información, los cinco poliedros regulares son el tetraedro (4, 6, 4), el cubo (8, 12, 16), el octaedro (6, 12, 8), el dodecaedro (20, 30, 12) y el icosaedro (12, 30, 20). (Los números entre paréntesis indican los números de vértices, aristas y caras, respectivamente).

SOLUCIÓN AL PROBLEMA DE PRÁCTICA

1. La desigualdad matricial $A\mathbf{x} \leq \mathbf{b}$ conduce al siguiente sistema de desigualdades:

a) $x_1 + 3x_2 \leq 12$

b) $x_1 + 2x_2 \leq 9$

c) $2x_1 + x_2 \leq 12$

La condición $\mathbf{x} \geq \mathbf{0}$ coloca al politopo en el primer cuadrante del plano. Un vértice es (0, 0). Las x_1 de intersección de las tres rectas (cuando $x_2 = 0$) son 12, 9 y 6, por lo que (6, 0) es un vértice. Las x_2 de intersección de las tres rectas (cuando $x_1 = 0$) son 4, 4.5, y 12, de manera que (0, 4) es un vértice.

¿Cómo hacer que las tres rectas frontera intercepten en valores positivos de x_1 y x_2? La intersección de *a*) y *b*) está en $\mathbf{p}_{ab} = (3, 3)$. Al probar $\mathbf{p}_{ab}$ en *c*) se obtiene $2(3) + 1(3) = 9 < 12$, por lo que $\mathbf{p}_{ab}$ está en *P*. La intersección de *b*) y *c*) está en $\mathbf{p}_{bc} = (5, 2)$. Al probar $\mathbf{p}_{bc}$ en *a*) se obtiene $1(5) + 3(2) = 11 < 12$, de manera que $\mathbf{p}_{bc}$ está en *P*. La intersección de *a*) y *c*) está en $\mathbf{p}_{ac} = (4.8, 2.4)$. Al probar $\mathbf{p}_{ac}$ en *b*) se obtiene $1(4.8) + 2(2.4) = 9.6 > 9$. Así, $\mathbf{p}_{ac}$ no está en *P*.

Por último, los tres vértices (puntos extremos) de los polítopos son (0, 0), (6, 0), (5, 2), (3, 3) y (0, 4). Esos puntos forman la representación mínima de *P*. Esto se muestra gráficamente en la figura 13.

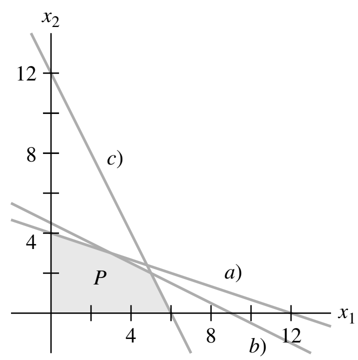


FIGURA 13

8.6 CURVAS Y SUPERFICIES

Durante miles de años, los constructores utilizaron largas tiras delgadas de madera para formar el casco de las embarcaciones. En épocas más recientes, los diseñadores emplearon tiras metálicas largas y flexibles para formar las superficies de autos y aviones. Pesos y clavijas moldearon las tiras en suaves curvas llamadas *splines cúbicos naturales*. La curva entre dos puntos de control sucesivos (clavijas o pesos) tiene una representación paramétrica usando polinomios cúbicos. Por desgracia, dichas curvas tienen la propiedad de que el movimiento de un punto de control afecta la forma de toda la curva, debido a las fuerzas físicas que las clavijas y los pesos ejercen sobre la tira. Los ingenieros diseñadores habían deseado por mucho tiempo tener un gran control local de la curva, de manera que el movimiento de un punto de control afectara solamente una pequeña parte de la curva. En 1962, un ingeniero francés especialista en automotores, Pierre Bézier, resolvió este problema agregando puntos de control y utilizando una clase especial de curvas, las cuales ahora llevan el nombre de su creador.

Curvas de Bézier

Las curvas que se describen a continuación desempeñan un papel importante en los gráficos generados por computadora y en ingeniería. Por ejemplo, se emplean en Adobe Illustrator y en Macromedia Freehand, y en lenguajes de programación como OpenGL. Esas curvas le permiten a un programa almacenar información exacta sobre segmentos curvos y superficies en un número relativamente pequeño de puntos de control. Todos los comandos de gráficos para los segmentos y las superficies solo se tienen que calcular para los puntos de control. La estructura especial de esas curvas también acelera otros cálculos en la “tubería de gráficos” que crea la imagen final en la pantalla.

Los ejercicios de la sección 8.3 introdujeron las curvas cuadráticas de Bézier y mostraron un método para construir curvas de Bézier de grado superior. Aquí el análisis se concentra en las curvas de Bézier cuadráticas y cúbicas, que se determinan con tres o cuatro puntos de control, denotados como $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Esos puntos pueden estar en $\mathbb{R}^2$ o $\mathbb{R}^3$, o bien, representarse mediante formas homogéneas en $\mathbb{R}^3$ o $\mathbb{R}^4$. Las descripciones paramétricas estándares de esas curvas, para $0 \leq t \leq 1$, son

$$\mathbf{w}(t) = (1-t)^2\mathbf{p}_0 + 2t(1-t)\mathbf{p}_1 + t^2\mathbf{p}_2 \tag{1}$$

$$\mathbf{x}(t) = (1-t)^3\mathbf{p}_0 + 3t(1-t)^2\mathbf{p}_1 + 3t^2(1-t)\mathbf{p}_2 + t^3\mathbf{p}_3 \tag{2}$$

La figura 1 muestra dos curvas típicas. Por lo regular, las curvas pasan solo a través de los puntos de control inicial y final, pero una curva de Bézier siempre está en la envolvente convexa de sus puntos de control. (Véase los ejercicios 21 a 24 de la sección 8.3).

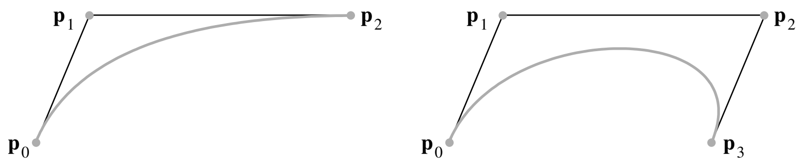


FIGURA 1 Curvas de Bézier cuadráticas y cúbicas.

Las curvas de Bézier son útiles en los gráficos generados por computadora porque sus propiedades esenciales se preservan bajo la acción de transformaciones lineales y traslaciones. Por ejemplo, si A es una matriz de tamaño adecuado, entonces de la linealidad de la multiplicación de matrices, para $0 \leq t \leq 1$,

$$\begin{aligned} A\mathbf{x}(t) &= A[(1-t)^3\mathbf{p}_0 + 3t(1-t)^2\mathbf{p}_1 + 3t^2(1-t)\mathbf{p}_2 + t^3\mathbf{p}_3] \\ &= (1-t)^3A\mathbf{p}_0 + 3t(1-t)^2A\mathbf{p}_1 + 3t^2(1-t)A\mathbf{p}_2 + t^3A\mathbf{p}_3 \end{aligned}$$

Los nuevos puntos de control son $A\mathbf{p}_0, \ldots, A\mathbf{p}_3$. En el ejercicio 1 se consideran traslaciones de curvas de Bézier.

Las curvas de la figura 1 sugieren que los puntos de control determinan las rectas tangentes a las curvas en los puntos de control inicial y terminal. De sus clases de cálculo, recuerde que para cualquier curva paramétrica, por ejemplo $\mathbf{y}(t)$, la dirección de la recta tangente a la curva en un punto $\mathbf{y}(t)$ está dada por la derivada $\mathbf{y}'(t)$, llamada **vector tangente** de la curva. (Esta derivada se calcula entrada por entrada).

EJEMPLO 1 Determine cómo el vector tangente de la curva cuadrática de Bézier $\mathbf{w}(t)$ está relacionado con los puntos de control de la curva, en $t = 0$ y $t = 1$.

SOLUCIÓN Escriba los pesos en la ecuación (1) como simples polinomios

$$\mathbf{w}(t) = (1 - 2t + t^2)\mathbf{p}_0 + (2t - 2t^2)\mathbf{p}_1 + t^2\mathbf{p}_2$$

Entonces, ya que la derivación es una transformación lineal sobre funciones,

$$\mathbf{w}'(t) = (-2 + 2t)\mathbf{p}_0 + (2 - 4t)\mathbf{p}_1 + 2t\mathbf{p}_2$$

Así,

$$\begin{aligned} \mathbf{w}'(0) &= -2\mathbf{p}_0 + 2\mathbf{p}_1 = 2(\mathbf{p}_1 - \mathbf{p}_0) \\ \mathbf{w}'(1) &= -2\mathbf{p}_1 + 2\mathbf{p}_2 = 2(\mathbf{p}_2 - \mathbf{p}_1) \end{aligned}$$

El vector tangente en $\mathbf{p}_0$, por ejemplo, apunta de $\mathbf{p}_0$ a $\mathbf{p}_1$, pero es dos veces más largo que el segmento de $\mathbf{p}_0$ a $\mathbf{p}_1$. Observe que $\mathbf{w}'(0) = \mathbf{0}$ cuando $\mathbf{p}_1 = \mathbf{p}_0$. En este caso, $\mathbf{w}(t) = (1 - t^2)\mathbf{p}_1 + t^2\mathbf{p}_2$, y la gráfica de $\mathbf{w}(t)$ es el segmento de recta de $\mathbf{p}_1$ a $\mathbf{p}_2$. ■

Conexión de dos curvas de Bézier

Dos curvas de Bézier básicas se pueden unir extremo con extremo, con el punto terminal de la primera curva $\mathbf{x}(t)$ como el punto inicial $\mathbf{p}_2$ de la segunda curva $\mathbf{y}(t)$. Se dice que la curva combinada tiene *continuidad geométrica* G^0 (en $\mathbf{p}_2$) porque los dos segmentos se juntan en $\mathbf{p}_2$. Si la recta tangente a la curva 1 en $\mathbf{p}_2$ tiene diferente dirección que la recta tangente a la curva 2, entonces una "esquina", o un cambio abrupto de dirección, puede ser evidente en $\mathbf{p}_2$. Véase la figura 2.

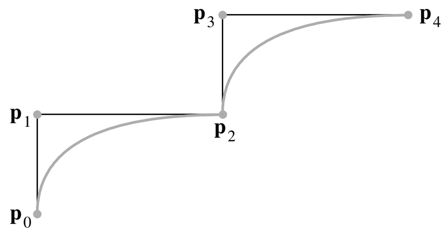


FIGURA 2 Continuidad G^0 en $\mathbf{p}_2$.

Para evitar un doblez pronunciado, es suficiente ajustar las curvas para que tengan lo que se llama *continuidad geométrica* G^1, donde ambos vectores tangentes en $\mathbf{p}_2$ apuntan en la misma dirección. Es decir, las derivadas $\mathbf{x}'(1)$ y $\mathbf{y}'(0)$ apuntan en la misma dirección, aunque pueden diferir en magnitud. Cuando los vectores tangentes son realmente iguales en $\mathbf{p}_2$, el vector tangente es continuo en $\mathbf{p}_2$, y se dice que la curva combinada tiene continuidad C^1, o *continuidad paramétrica* C^1. La figura 3 muestra continuidad G^1 en *a*) y continuidad C^1 en *b*).

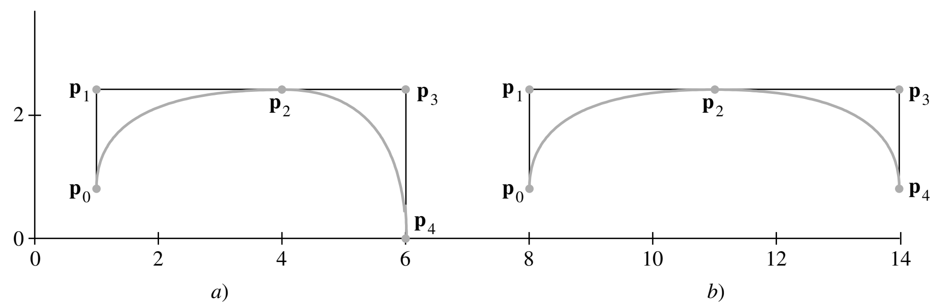


FIGURA 3 *a*) Continuidad G^1 y *b*) continuidad C^1.

EJEMPLO 2 Considere que $\mathbf{x}(t)$ y $\mathbf{y}(t)$ determinan dos curvas cuadráticas de Bézier, con puntos de control $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$ y $\{\mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4\}$, respectivamente. Las curvas se unen en $\mathbf{p}_2 = \mathbf{x}(1) = \mathbf{y}(0)$.

a) Suponga que la curva combinada tiene continuidad G^1 (en $\mathbf{p}_2$). ¿Qué restricción algebraica hace que esta condición se imponga sobre los puntos de control? Exprese esta restricción en lenguaje geométrico.

b) Repita el inciso *a*) para continuidad C^1.

SOLUCIÓN

a) A partir del ejemplo 1, $\mathbf{x}'(t) = 2(\mathbf{p}_2 - \mathbf{p}_1)$. Además, utilizando los puntos de control para $\mathbf{y}(t)$ en lugar de $\mathbf{w}(t)$, el ejemplo 1 indica que $\mathbf{y}'(0) = 2(\mathbf{p}_3 - \mathbf{p}_2)$. La continuidad G^1 significa que $\mathbf{y}'(0) = k\mathbf{x}'(1)$ para alguna constante positiva k. De forma equivalente,

$$\mathbf{p}_3 - \mathbf{p}_2 = k(\mathbf{p}_2 - \mathbf{p}_1), \quad \text{con } k > 0 \tag{3}$$

Geométricamente, (3) implica que $\mathbf{p}_2$ está sobre el segmento de recta de $\mathbf{p}_1$ a $\mathbf{p}_3$. Para probar esto, sea $t = (k + 1)^{-1}$, y observe que $0 < t < 1$. Despeje k, para obtener $k = (1 - t)/t$. Cuando esta expresión para k se emplea en la ecuación (3), un reordenamiento indica que $\mathbf{p}_2 = (1 - t)\mathbf{p}_1 + t\,\mathbf{p}_3$, lo que comprueba la afirmación acerca de $\mathbf{p}_2$.

b) La continuidad C^1 significa que $\mathbf{y}'(0) = \mathbf{x}'(1)$. Por lo tanto, $2(\mathbf{p}_3 - \mathbf{p}_2) = 2(\mathbf{p}_2 - \mathbf{p}_1)$, de manera que $\mathbf{p}_3 - \mathbf{p}_2 = \mathbf{p}_2 - \mathbf{p}_1$, y $\mathbf{p}_2 = (\mathbf{p}_1 + \mathbf{p}_3)/2$. Geométricamente, $\mathbf{p}_2$ es el punto medio del segmento de recta de $\mathbf{p}_1$ a $\mathbf{p}_3$. Véase la figura 3. ■

La figura 4 muestra la continuidad C^1 para dos curvas cúbicas de Bézier. Observe cómo el punto que une los dos segmentos de recta está a la mitad del segmento de recta entre los puntos de control adyacentes.

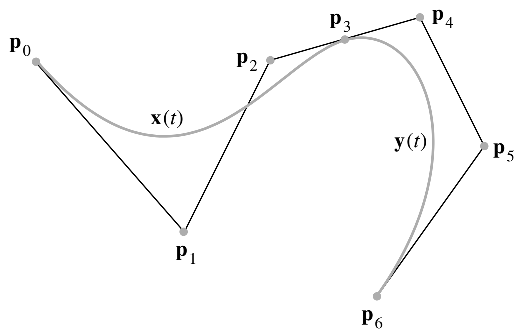


FIGURA 4 Dos curvas cúbicas de Bézier.

Dos curvas tienen continuidad (paramétrica) C^2 cuando tienen continuidad C^1 y las *segundas* derivadas $\mathbf{x}''(1)$ y $\mathbf{y}''(0)$ son iguales. Esto es posible para curvas cúbicas de Bézier, pero limita severamente las posiciones de los puntos de control. Otra clase de curvas cúbicas, llamadas *B-splines*, siempre tienen continuidad C^2 porque cada par de curvas comparte tres puntos de control en lugar de uno. Los gráficos que utilizan B-splines tienen más puntos de control y, en consecuencia, requieren más cálculos. Algunos ejercicios de esta sección examinan esas curvas.

De manera sorprendente, si $\mathbf{x}(t)$ y $\mathbf{y}(t)$ se unen en $\mathbf{p}_3$, la suavidad aparente de la curva en $\mathbf{p}_3$, por lo general, es la misma para continuidad G^1 y continuidad C^1. Esto se debe a que la magnitud de $\mathbf{x}'(t)$ no está relacionada con la forma física de la curva. La magnitud solo refleja la parametrización matemática de la curva. Por ejemplo, si una nueva función vectorial $\mathbf{z}(t)$ es igual a $\mathbf{x}(2t)$, entonces el punto $\mathbf{z}(t)$ atraviesa la curva de $\mathbf{p}_0$ a $\mathbf{p}_3$ dos veces más rápido que la versión original, porque $2t$ alcanza 1 cuando t es .5. Pero, de acuerdo con la regla de la cadena de cálculo, $\mathbf{z}'(t) = 2\cdot\mathbf{x}'(2t)$, por lo que el vector tangente a $\mathbf{z}(t)$ en $\mathbf{p}_3$ es el doble del vector tangente a $\mathbf{x}(t)$ en $\mathbf{p}_3$.

En la práctica, muchas curvas de Bézier sencillas se unen para crear objetos gráficos. Los programas de composición tipográfica constituyen una importante aplicación, porque muchas letras en un tipo de fuente implican segmentos curvos. Cada letra en letra PostScript®, por ejemplo, se almacena como un conjunto de puntos de control, junto con información sobre cómo construir el "contorno" de la letra utilizando segmentos de recta y curvas de Bézier. El aumento de dicha letra básicamente requiere multiplicar las coordenadas de cada punto de control por un factor de escala constante. Una vez que se ha calculado el contorno de la letra, se rellenan las partes sólidas adecuadas de la letra. La figura 5 ilustra esto para un carácter de una letra PostScript. Observe los puntos de control.

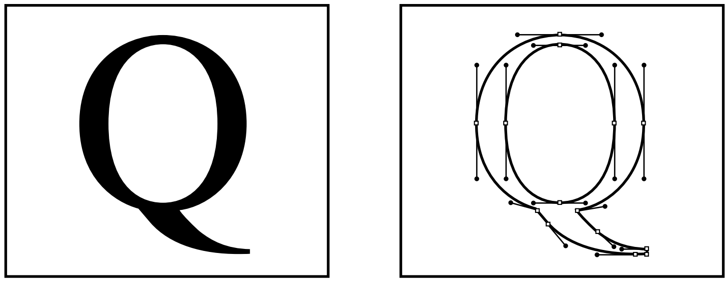


FIGURA 5 Un carácter PostScript.

Ecuaciones matriciales para curvas de Bézier

Como una curva de Bézier es una combinación lineal de puntos de control utilizando polinomios como pesos, la fórmula para $\mathbf{x}(t)$ se puede escribir como

$$\begin{aligned}\mathbf{x}(t) &= \begin{bmatrix}\mathbf{p}_0 & \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3\end{bmatrix}\begin{bmatrix}(1-t)^3\\ 3t(1-t)^2\\ 3t^2(1-t)\\ t^3\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{p}_0 & \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3\end{bmatrix}\begin{bmatrix}1-3t+3t^2-t^3\\ 3t-6t^2+3t^3\\ 3t^2-3t^3\\ t^3\end{bmatrix}\\ &= \begin{bmatrix}\mathbf{p}_0 & \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3\end{bmatrix}\begin{bmatrix}1 & -3 & 3 & -1\\ 0 & 3 & -6 & 3\\ 0 & 0 & 3 & -3\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\ t\\ t^2\\ t^3\end{bmatrix}\end{aligned}$$

La matriz cuyas columnas son los cuatro puntos de control es la **matriz de geometría**, G. La matriz de 4×4 de coeficientes polinomiales es la **matriz base de Bézier**, M_B. Si $\mathbf{u}(t)$ es el vector columna de potencias de t, entonces la curva de Bézier está dada por

$$\mathbf{x}(t) = GM_B\mathbf{u}(t) \tag{4}$$

Otras curvas cúbicas paramétricas en gráficos de computadora también se escriben en esta forma. Por ejemplo, si las entradas en la matriz M_B se cambian adecuadamente, las curvas resultantes son B-splines. Estas son “más suaves” que las curvas de Bézier, pero no pasan por ninguno de los puntos de control. Una curva cúbica de **Hermite** se presenta cuando la matriz M_B se remplaza por una matriz base de Hermite. En este caso, las columnas de la matriz de geometría consisten en los puntos inicial y final de las curvas y los vectores tangentes a las curvas en esos puntos.[1]

La curva de Bézier en la ecuación (4) se puede “factorizar” de otra manera, para utilizarse en el análisis de superficies de Bézier. Por conveniencia más adelante, el parámetro t

[1] El término *matriz base* viene de las filas de la matriz que listan los coeficientes de los polinomios de *mezclado* utilizados para definir la curva. Para una curva de Bézier cúbica, los cuatro polinomios son $(1-t)^3$, $3t(1-t)^2$, $3t^2(1-t)$, y t^3. Estos forman una base para el espacio $\mathbb{P}_3$ de polinomios de grado 3 o menor. Cada entrada en el vector $\mathbf{x}(t)$ es una combinación lineal de esos polinomios. Los pesos provienen de las filas de la matriz de geometría G en la ecuación (4).

se remplaza por un parámetro s:

$$\mathbf{x}(s) = \mathbf{u}(s)^T M_B^T \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \begin{bmatrix} 1 & s & s^2 & s^3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}$$

$$= \begin{bmatrix} (1-s)^3 & 3s(1-s)^2 & 3s^2(1-s) & s^3 \end{bmatrix} \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} \tag{5}$$

Esta fórmula no coincide plenamente con la transpuesta del producto de la derecha en (4), porque $\mathbf{x}(s)$ y los puntos de control se presentan en la ecuación (5) sin los símbolos de transpuesta. La matriz de los puntos de control en la ecuación (5) se llama el **vector de geometría**. Esto se debería ver como un bloque matricial de 4×1 (particionado) cuyas entradas son vectores columnas. La matriz a la izquierda del vector de geometría, en la segunda parte de la ecuación (5), también se puede considerar como un bloque matricial, con un escalar en cada bloque. La multiplicación de matrices particionadas tiene sentido, porque cada entrada (vector) en el vector de geometría se puede multiplicar por la izquierda por un escalar, así como por una matriz. Por lo tanto, el vector columna $\mathbf{x}(s)$ se representa mediante la ecuación (5).

Superficies de Bézier

Un parche superficial bicúbico en 3D se puede construir a partir de un conjunto de cuatro curvas de Bézier. Considere las cuatro matrices de geometría

$$\begin{bmatrix} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{13} & \mathbf{p}_{14} \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{p}_{21} & \mathbf{p}_{22} & \mathbf{p}_{23} & \mathbf{p}_{24} \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{p}_{31} & \mathbf{p}_{32} & \mathbf{p}_{33} & \mathbf{p}_{34} \end{bmatrix}$$
$$\begin{bmatrix} \mathbf{p}_{41} & \mathbf{p}_{42} & \mathbf{p}_{43} & \mathbf{p}_{44} \end{bmatrix}$$

y recuerde de la ecuación (4) que una curva de Bézier se produce cuando cualquiera de esas matrices se multiplica por la derecha por el siguiente vector de pesos:

$$M_B \mathbf{u}(t) = \begin{bmatrix} (1-t)^3 \\ 3t(1-t)^2 \\ 3t^2(1-t) \\ t^3 \end{bmatrix}$$

Sea G el bloque matricial (particionado) de 4×4 cuyas entradas son los puntos de control $\mathbf{p}_{ij}$ presentados antes. Entonces, el siguiente producto es un bloque matricial de 4×1, y cada entrada es una curva de Bézier:

$$GM_B \mathbf{u}(t) = \begin{bmatrix} \mathbf{p}_{11} & \mathbf{p}_{12} & \mathbf{p}_{13} & \mathbf{p}_{14} \\ \mathbf{p}_{21} & \mathbf{p}_{22} & \mathbf{p}_{23} & \mathbf{p}_{24} \\ \mathbf{p}_{31} & \mathbf{p}_{32} & \mathbf{p}_{33} & \mathbf{p}_{34} \\ \mathbf{p}_{41} & \mathbf{p}_{42} & \mathbf{p}_{43} & \mathbf{p}_{44} \end{bmatrix} \begin{bmatrix} (1-t)^3 \\ 3t(1-t)^2 \\ 3t^2(1-t) \\ t^3 \end{bmatrix}$$

De hecho,

$$GM_B \mathbf{u}(t) = \begin{bmatrix} (1-t)^3\mathbf{p}_{11} + 3t(1-t)^2\mathbf{p}_{12} + 3t^2(1-t)\mathbf{p}_{13} + t^3\mathbf{p}_{14} \\ (1-t)^3\mathbf{p}_{21} + 3t(1-t)^2\mathbf{p}_{22} + 3t^2(1-t)\mathbf{p}_{23} + t^3\mathbf{p}_{24} \\ (1-t)^3\mathbf{p}_{31} + 3t(1-t)^2\mathbf{p}_{32} + 3t^2(1-t)\mathbf{p}_{33} + t^3\mathbf{p}_{34} \\ (1-t)^3\mathbf{p}_{41} + 3t(1-t)^2\mathbf{p}_{42} + 3t^2(1-t)\mathbf{p}_{43} + t^3\mathbf{p}_{44} \end{bmatrix}$$

Ahora fije t. Así, $GM_B\mathbf{u}(t)$ es un vector columna que se puede utilizar como un vector de geometría en la ecuación (5) para una curva de Bézier en otra variable s. Esta observación produce la **superficie bicúbica de Bézier**:

$$\mathbf{x}(s, t) = \mathbf{u}(s)^T M_B^T GM_B\mathbf{u}(t), \quad \text{donde } 0 \leq s, t \leq 1 \tag{6}$$

La fórmula para $\mathbf{x}(s, t)$ es una combinación lineal de los 16 puntos de control. Si uno imagina que esos puntos de control están acomodados en un arreglo rectangular más o menos uniforme, como en la figura 6, entonces la superficie de Bézier se controla con una red de ocho curvas de Bézier, cuatro en la "dirección s" y cuatro en la "dirección t". La superficie realmente pasa por los cuatro puntos de control en sus "esquinas". Cuando está a la mitad de una superficie más grande, la superficie de 16 puntos comparte con sus vecinos sus 12 puntos de control frontera.

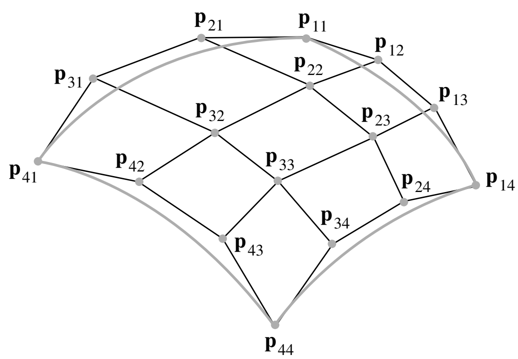


FIGURA 6 Dieciséis puntos de control para un parche superficial bicúbico de Bézier.

Aproximaciones a curvas y superficies

En programas CAD y en los que se utilizan para crear juegos computacionales realistas, el diseñador con frecuencia labora en una estación de trabajo gráfica para crear una "escena" que implica varias estructuras geométricas. Este proceso requiere interacción entre el diseñador y los objetos geométricos. Cada leve reposicionamiento de un objeto requiere nuevos cálculos matemáticos por el programa de gráficos. Las curvas y superficies de Bézier son útiles en este proceso porque implican mucho menos puntos de control que objetos aproximados por bastantes polígonos. Esto reduce notablemente el tiempo de cálculo y acelera el trabajo del diseñador.

Sin embargo, después de la composición de la escena, la preparación de la imagen final tiene diferentes demandas computacionales que se logran con mayor facilidad mediante objetos que consisten en superficies planas y aristas rectas, como los poliedros. El diseñador necesita *generar* la escena, introduciendo iluminación, agregando color y textura a las superficies, y simulando reflexiones en las superficies.

Calcular la dirección de la luz reflejada en un punto $\mathbf{p}$ sobre una superficie, por ejemplo, requiere saber las direcciones de la luz incidente y de la *normal superficial*, esto es, el vector perpendicular al plano tangente en $\mathbf{p}$. Determinar estos vectores normales es mucho más fácil sobre una superficie compuesta de, por ejemplo, pequeños polígonos planos que sobre una superficie curva cuyo vector normal cambia continuamente conforme se mueve $\mathbf{p}$. Si $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ son vértices adyacentes de un polígono plano, entonces la normal superficial es más o menos el producto cruz $(\mathbf{p}_2 - \mathbf{p}_1) \times (\mathbf{p}_2 - \mathbf{p}_3)$. Cuando el polígono es pequeño, solo se necesita un vector normal para reproducir el polígono entero. Además, dos rutinas de sombreado ampliamente utilizadas, el sombreado Gouraud y el sombreado Phong, requieren una superficie para definirse mediante polígonos.

Como resultado de esa necesidad de superficies planas, ahora las curvas y superficies de Bézier de la etapa de composición de escenas generalmente se aproximan mediante segmentos de líneas rectas y superficies poliédricas. La idea básica para aproximar una curva o superficie de Bézier es dividir la curva o superficie en partes más pequeñas, con más y más puntos de control.

Subdivisión recursiva de curvas y superficies de Bézier

La figura 7 muestra los cuatro puntos de control $\mathbf{p}_0, \ldots, \mathbf{p}_3$ para una curva de Bézier, junto con puntos de control para dos nuevas curvas, cada una de las cuales coincide con la mitad de la curva original. La curva "izquierda" inicia en $\mathbf{q}_0 = \mathbf{p}_0$ y termina en $\mathbf{q}_3$, en el punto medio de la curva original. La curva "derecha" comienza en $\mathbf{r}_0 = \mathbf{q}_3$ y termina en $\mathbf{r}_3 = \mathbf{p}_3$.

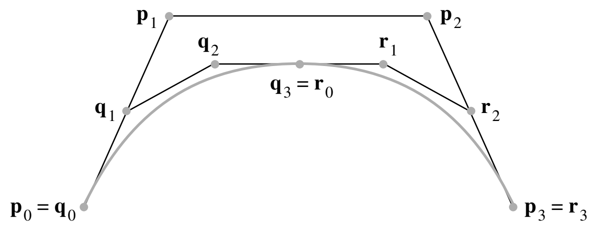


FIGURA 7 Subdivisión de una curva de Bézier.

La figura 8 muestra cómo los nuevos puntos de control encierran regiones que son "más delgadas" que la región encerrada por los puntos de control originales. Conforme las distancias entre los puntos de control disminuyen, los puntos de control de cada segmento de curva también se hacen más cercanos a un segmento de recta. Esta *propiedad de variación-disminución* de las curvas de Bézier depende del hecho de que una curva de Bézier siempre está en la envolvente convexa de los puntos de control.

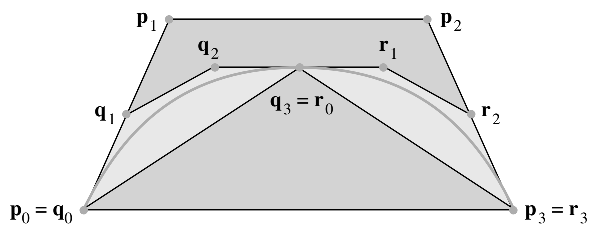


FIGURA 8 Envolventes convexas de los puntos de control.

Los nuevos puntos de control están relacionados con los puntos de control originales mediante fórmulas sencillas. Desde luego, $\mathbf{q}_0 = \mathbf{p}_0$ y $\mathbf{r}_3 = \mathbf{p}_3$. El punto medio de la curva original $\mathbf{x}(t)$ se encuentra en $\mathbf{x}(.5)$ cuando $\mathbf{x}(t)$ tiene la parametrización estándar.

$$\mathbf{x}(t) = (1 - 3t + 3t^2 - t^3)\mathbf{p}_0 + (3t - 6t^2 + 3t^3)\mathbf{p}_1 + (3t^2 - 3t^3)\mathbf{p}_2 + t^3\mathbf{p}_3 \tag{7}$$

para $0 \leq t \leq 1$. Así, los nuevos puntos de control $\mathbf{q}_3$ y $\mathbf{r}_0$ están dados por

$$\mathbf{q}_3 = \mathbf{r}_0 = \mathbf{x}(.5) = \tfrac{1}{8}(\mathbf{p}_0 + 3\mathbf{p}_1 + 3\mathbf{p}_2 + \mathbf{p}_3) \tag{8}$$

También son sencillas las fórmulas para los restantes puntos de control "interiores", pero la deducción de las fórmulas requiere algún trabajo que implica vectores tangentes de las curvas. Por definición, el vector tangente a una curva parametrizada $\mathbf{x}(t)$ es la derivada $\mathbf{x}'(t)$. Este vector muestra la dirección de la recta tangente a la curva en $\mathbf{x}(t)$. Para la curva de Bézier en (7),

$$\mathbf{x}'(t) = (-3 + 6t - 3t^2)\mathbf{p}_0 + (3 - 12t + 9t^2)\mathbf{p}_1 + (6t - 9t^2)\mathbf{p}_2 + 3t^2\mathbf{p}_3$$

para $0 \leq t \leq 1$. En particular,

$$\mathbf{x}'(0) = 3(\mathbf{p}_1 - \mathbf{p}_0) \quad \text{y} \quad \mathbf{x}'(1) = 3(\mathbf{p}_3 - \mathbf{p}_2) \tag{9}$$

Geométricamente, $\mathbf{p}_1$ está sobre la recta tangente a la curva en $\mathbf{p}_0$, y $\mathbf{p}_2$ está sobre la recta tangente a la curva en $\mathbf{p}_3$. Véase la figura 8. Además, a partir de $\mathbf{x}'(t)$, calcule

$$\mathbf{x}'(.5) = \tfrac{3}{4}(-\mathbf{p}_0 - \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \tag{10}$$

Sean $\mathbf{y}(t)$ la curva de Bézier determinada por $\mathbf{q}_0,\dots,\mathbf{q}_3$, y $\mathbf{z}(t)$ la curva de Bézier determinada por $\mathbf{r}_0,\dots,\mathbf{r}_3$. Como $\mathbf{y}(t)$ recorre la misma trayectoria que $\mathbf{x}(t)$, pero solo llega hasta $\mathbf{x}(.5)$ conforme t va de 0 a 1, entonces $\mathbf{y}(t) = \mathbf{x}(.5t)$ para $0 \le t \le 1$. De manera similar, ya que $\mathbf{z}(t)$ inicia en $\mathbf{x}(.5)$ cuando $t = 0$, entonces $\mathbf{z}(t) = \mathbf{x}(.5 + .5t)$ para $0 \le t \le 1$. De acuerdo con la regla de la cadena para derivadas,

$$\mathbf{y}'(t) = .5\mathbf{x}'(.5t) \quad \text{y} \quad \mathbf{z}'(t) = .5\mathbf{x}'(.5 + .5t) \quad \text{para } 0 \le t \le 1 \tag{11}$$

A partir de (9) con $\mathbf{y}'(0)$ en vez de $\mathbf{x}'(0)$, de (11) con $t = 0$, y de (9), los puntos de control para $\mathbf{y}(t)$ satisfacen

$$3(\mathbf{q}_1 - \mathbf{q}_0) = \mathbf{y}'(0) = .5\mathbf{x}'(0) = \tfrac{3}{2}(\mathbf{p}_1 - \mathbf{p}_0) \tag{12}$$

A partir de (9) con $\mathbf{y}'(1)$ en vez de $\mathbf{x}'(1)$, de (11) con $t = 1$, y de (10),

$$3(\mathbf{q}_3 - \mathbf{q}_2) = \mathbf{y}'(1) = .5\mathbf{x}'(.5) = \tfrac{3}{8}(-\mathbf{p}_0 - \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3) \tag{13}$$

Las ecuaciones (8), (9), (10), (12) y (13) se pueden resolver y así generar las fórmulas para $\mathbf{q}_0,\dots,\mathbf{q}_3$ que se muestran en el ejercicio 13. Geométricamente, las fórmulas se presentan en la figura 9. Los puntos de control interiores $\mathbf{q}_1$ y $\mathbf{r}_2$ son los puntos medios, respectivamente, del segmento de $\mathbf{p}_0$ a $\mathbf{p}_1$ y el segmento de $\mathbf{p}_2$ a $\mathbf{p}_3$. Cuando el punto medio del segmento de $\mathbf{p}_1$ a $\mathbf{p}_2$ se conecta a $\mathbf{q}_1$, ¡el segmento de recta resultante tiene a $\mathbf{q}_2$ como punto medio!

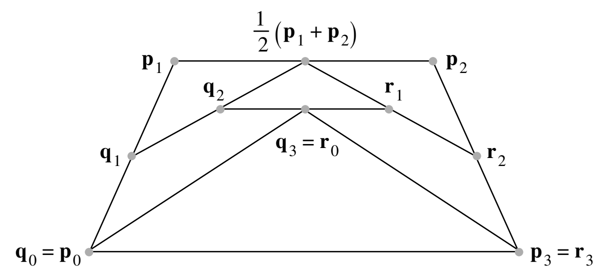


FIGURA 9 Estructura geométrica de los nuevos puntos de control.

Esto completa un paso del proceso de subdivisión. La "recursión" inicia, y las nuevas curvas se subdividen. La recursión continúa a una profundidad en la cual todas las curvas son lo suficientemente rectas. De manera alternativa, en cada paso la recursión puede ser "adaptiva" y no subdividir una de las dos nuevas curvas si esa curva es suficientemente recta. Una vez que la subdivisión se detiene por completo, los puntos extremos de cada curva se unen mediante segmentos de recta, y la escena está lista para el siguiente paso en la preparación de la imagen final.

Una superficie bicúbica de Bézier tiene la misma propiedad de variación-disminución que las curvas de Bézier que forman cada sección transversal de la superficie, de manera que el proceso ya descrito se puede aplicar en cada sección transversal. Omitiendo los detalles, a continuación se presenta la estrategia básica. Considere las cuatro curvas de Bézier "paralelas" cuyo parámetro es s, y aplique el proceso de subdivisión a cada una de ellas. Esto produce cuatro conjuntos de ocho puntos de control; cada conjunto determina una curva conforme s varía de 0 a 1. Sin embargo, al variar t, existen ocho curvas, cada una con cuatro puntos de control. Aplique el proceso de subdivisión a cada uno de esos conjuntos de cuatro puntos, para crear un total de 64 puntos de control. La recursión adaptiva también es posible en este caso, pero hay algunas sutilezas implicadas.[2]

[2] Véase Foley, Van Dam, Feiner y Hughes, *Computer Graphics–Principles and Practice*, 2a. ed. (Boston: Addison-Wesley, 1996), pp. 527-528.

PROBLEMAS DE PRÁCTICA

El término *spline* generalmente se refiere a una curva que pasa por puntos especificados. Sin embargo, una B-spline, normalmente no pasa a través de sus puntos de control. Un solo segmento tiene la forma paramétrica

$$\begin{aligned}\mathbf{x}(t) = \tfrac{1}{6}\big[&(1-t)^3\mathbf{p}_0 + (3t^3 - 6t^2 + 4)\mathbf{p}_1 \\ &+ (-3t^3 + 3t^2 + 3t + 1)\mathbf{p}_2 + t^3\mathbf{p}_3\big]\end{aligned} \tag{14}$$

para $0 \leq t \leq 1$, donde $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ son los puntos de control. Cuando t varía de 0 a 1, $\mathbf{x}(t)$ crea un curva corta cercana a $\overline{\mathbf{p}_1\mathbf{p}_2}$. Álgebra básica indica que la fórmula de B-spline también se puede escribir como

$$\begin{aligned}\mathbf{x}(t) = \tfrac{1}{6}\big[&(1-t)^3\mathbf{p}_0 + (3t(1-t)^2 - 3t + 4)\mathbf{p}_1 \\ &+ (3t^2(1-t) + 3t + 1)\mathbf{p}_2 + t^3\mathbf{p}_3\big]\end{aligned} \tag{15}$$

Esto revela la similitud con la curva de Bézier. Excepto por el factor $1/6$ al frente, los términos $\mathbf{p}_0$ y $\mathbf{p}_3$ son iguales. La componente $\mathbf{p}_1$ se ha incrementado por $-3t + 4$, y la componente $\mathbf{p}_2$ se incrementó por $3t + 1$. Esas componentes hacen más cercana la curva a $\overline{\mathbf{p}_1\mathbf{p}_2}$ que la curva de Bézier. El factor $1/6$ es necesario para mantener la suma de los coeficientes igual a 1. La figura 10 compara una B-spline con una curva de Bézier que tiene los mismos puntos de control.

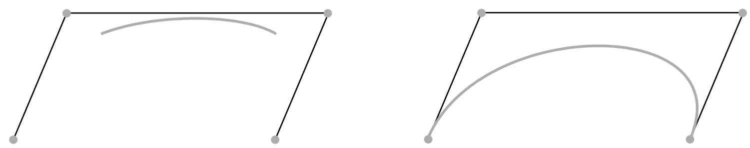


FIGURA 10 Un segmento B-spline y una curva de Bézier.

1. Demuestre que la B-spline no inicia en $\mathbf{p}_0$, pero $\mathbf{x}(0)$ está en conv $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$. Suponiendo que $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$ son afínmente independientes, encuentre las coordenadas afines de $\mathbf{x}(0)$ con respecto a $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$.
2. Demuestre que la B-spline no termina en $\mathbf{p}_3$, pero $\mathbf{x}(1)$ está en conv $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. Suponiendo que $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ son afínmente independientes, encuentre las coordenadas afines de $\mathbf{x}(1)$ con respecto a $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

8.6 EJERCICIOS

1. Suponga que una curva de Bézier se traslada a $\mathbf{x}(t) + \mathbf{b}$. Es decir, para $0 \leq t \leq 1$, la nueva curva es

$$\begin{aligned}\mathbf{x}(t) = (1-t)^3\mathbf{p}_0 &+ 3t(1-t)^2\mathbf{p}_1 \\ &+ 3t^2(1-t)\mathbf{p}_2 + t^3\mathbf{p}_3 + \mathbf{b}\end{aligned}$$

Demuestre que esta nueva curva es una vez más una curva de Bézier. [*Sugerencia:* ¿Dónde están los nuevos puntos de control?]

2. La forma vectorial paramétrica de una curva B-spline se definió en los problemas de práctica como

$$\begin{aligned}\mathbf{x}(t) = \tfrac{1}{6}\big[&(1-t)^3\mathbf{p}_0 + (3t(1-t) - 3t + 4)\mathbf{p}_1 \\ &+(3t^2(1-t) + 3t + 1)\mathbf{p}_2 + t^3\mathbf{p}_3\big] \quad \text{para } 0 \leq t \leq 1,\end{aligned}$$

donde $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$ son los puntos de control.

a) Demuestre que para $0 \leq t \leq 1$, $\mathbf{x}(t)$ está en la envolvente convexa de los puntos de control.

b) Suponga que una curva B-spline $\mathbf{x}(t)$ se traslada a $\mathbf{x}(t) + \mathbf{b}$ (como en el ejercicio 1). Demuestre que esta nueva curva es una vez más una B-spline.

3. Sea $\mathbf{x}(t)$ una curva de Bézier cúbica determinada por los puntos $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$.

a) Calcule el vector *tangente* $\mathbf{x}'(t)$. Determine cómo $\mathbf{x}'(0)$ y $\mathbf{x}'(1)$ se relacionan con los puntos de control, y realice descripciones geométricas de las *direcciones* de esos vectores tangentes. ¿Es posible tener $\mathbf{x}'(1) = \mathbf{0}$?

b) Calcule la segunda derivada $\mathbf{x}''(t)$ y determine cómo $\mathbf{x}''(0)$ y $\mathbf{x}''(1)$ se relacionan con los puntos de control. Dibuje un esquema basado en la figura 10 y construya un segmento de recta que apunte en la dirección de $\mathbf{x}''(0)$. [*Sugerencia:* Utilice $\mathbf{p}_1$ como el origen del sistema de coordenadas].

4. Sea $\mathbf{x}(t)$ la B-spline del ejercicio 2, con puntos de control $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$.

a) Calcule el vector tangente $\mathbf{x}'(t)$ y determine cómo las derivadas $\mathbf{x}'(0)$ y $\mathbf{x}'(1)$ se relacionan con los puntos de control. Realice descripciones geométricas de las *direcciones* de esos vectores tangentes. Explore qué pasa cuando $\mathbf{x}'(0)$ y $\mathbf{x}'(1)$ son iguales a $\mathbf{0}$. Justifique sus afirmaciones.

b) Calcule la segunda derivada $\mathbf{x}''(t)$ y encuentre cómo $\mathbf{x}''(0)$ y $\mathbf{x}''(1)$ se relacionan con los puntos de control. Dibuje un esquema basado en la figura 10, y construya un segmento de recta que apunte en la dirección de $\mathbf{x}''(1)$. [*Sugerencia:* Utilice $\mathbf{p}_2$ como el origen del sistema de coordenadas].

5. Sean $\mathbf{x}(t)$ y $\mathbf{y}(t)$ curvas de Bézier cúbicas con puntos de control $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ y $\{\mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5, \mathbf{p}_6\}$, respectivamente, de manera que $\mathbf{x}(t)$ y $\mathbf{y}(t)$ se unen en $\mathbf{p}_3$. Las siguientes preguntas se refieren a la curva que consiste en $\mathbf{x}(t)$ seguida de $\mathbf{y}(t)$. Para simplificar, suponga que la curva está en $\mathbb{R}^2$.

a) ¿Qué condición sobre los puntos de control garantizará que la curva tenga continuidad C^1 en $\mathbf{p}_3$? Justifique su respuesta.

b) ¿Qué ocurre cuando $\mathbf{x}'(1)$ y $\mathbf{y}'(0)$ son ambos el vector cero?

6. Una B-spline se construye mediante segmentos B-spline, descritos en el ejercicio 2. Sean $\mathbf{p}_0,\ldots, \mathbf{p}_4$ puntos de control. Para $0 \leq t \leq 1$, sean $\mathbf{x}(t)$ y $\mathbf{y}(t)$ determinadas por las matrices de geometría $[\mathbf{p}_0\ \mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]$ y $[\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3\ \mathbf{p}_4]$, respectivamente. Observe cómo los dos segmentos comparten tres puntos de control. No hay traslape entre los dos segmentos, sin embargo, se unen en un punto extremo común, cercano a $\mathbf{p}_2$.

a) Demuestre que la curva combinada tiene continuidad G^0, es decir, $\mathbf{x}(1) = \mathbf{y}(0)$.

b) Demuestre que la curva tiene continuidad C^1 en el punto de unión, $\mathbf{x}(1)$. Es decir, demuestre que $\mathbf{x}'(1) = \mathbf{y}'(0)$.

7. Sean $\mathbf{x}(t)$ y $\mathbf{y}(t)$ las curvas de Bézier del ejercicio 5, y suponga que la curva combinada tiene continuidad C^2 (que incluye la continuidad C^1) en $\mathbf{p}_3$. Sea $\mathbf{x}''(1) = \mathbf{y}''(0)$ y pruebe que $\mathbf{p}_5$ está completamente determinado por $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Así, los puntos $\mathbf{p}_0,\ldots, \mathbf{p}_3$ y la condición C^2 determinan todos, excepto uno, los puntos de control para $\mathbf{y}(t)$.

8. Sean $\mathbf{x}(t)$ y $\mathbf{y}(t)$ segmentos de una B-spline como en el ejercicio 6. Demuestre que la curva tiene continuidad C^2 (como también continuidad C^1) en $\mathbf{x}(1)$. Es decir, demuestre que $\mathbf{x}''(1) = \mathbf{y}''(0)$. Esta continuidad de alto orden es deseable en aplicaciones CAD, como en el diseño de carrocerías de automóviles, porque las curvas y superficies aparecen mucho más lisas. Sin embargo, las B-splines requieren tres veces más cálculos que las curvas de Bézier, para curvas de longitud comparable. Para superficies, las B-splines necesitan nueve veces el cálculo de las superficies de Bézier. Con frecuencia, los programadores eligen superficies de Bézier en las aplicaciones (como en los simuladores de vuelo) que requieren generación en tiempo real.

9. Una curva de Bézier cuártica se determina mediante cinco puntos de control, $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ y $\mathbf{p}_4$:

$$\begin{aligned}\mathbf{x}(t) = (1-t)^4\mathbf{p}_0 &+ 4t(1-t)^3\mathbf{p}_1 + 6t^2(1-t)^2\mathbf{p}_2\\ &+ 4t^3(1-t)\mathbf{p}_3 + t^4\mathbf{p}_4 \quad \text{for } 0 \leq t \leq 1\end{aligned}$$

Construya la matriz base cuártica M_B para $\mathbf{x}(t)$.

10. La "B" en B-spline se refiere al hecho de que un segmento $\mathbf{x}(t)$ se puede escribir en términos de una matriz base, M_S, en una forma similar a una curva de Bézier. Es decir,

$$\mathbf{x}(t) = GM_S\mathbf{u}(t) \quad \text{para } 0 \leq t \leq 1$$

donde G es la matriz de geometría $[\mathbf{p}_0\ \mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3]$ y $\mathbf{u}(t)$ es el vector columna $(1, t, t^2, t^3)$. En una B-spline *uniforme*, cada segmento utiliza la misma matriz base, pero cambia la matriz de geometría. Construya la matriz base M_S para $\mathbf{x}(t)$.

En los ejercicios 11 y 12, marque cada enunciado como verdadero o falso. Justifique sus respuestas.

11. *a*) La curva cúbica de Bézier está basada en cuatro puntos de control.

b) Dada una curva de Bézier cuadrática $\mathbf{x}(t)$ con puntos de control $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$, el segmento de recta dirigido $\mathbf{p}_1 - \mathbf{p}_0$ (de $\mathbf{p}_0$ a $\mathbf{p}_1$) es el vector tangente a la curva en $\mathbf{p}_0$.

c) Cuando dos curvas de Bézier cuadráticas con puntos de control $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$ y $\{\mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4\}$ se unen en $\mathbf{p}_2$, la curva de Bézier combinada tendrá continuidad C^1 en $\mathbf{p}_2$ si $\mathbf{p}_2$ es el punto medio del segmento de recta entre $\mathbf{p}_1$ y $\mathbf{p}_3$.

12. *a*) Las propiedades esenciales de la curvas de Bézier se preservan bajo la acción de transformaciones lineales, pero no traslaciones.

b) Cuando dos curvas de Bézier $\mathbf{x}(t)$ y $\mathbf{y}(t)$ se unen en el punto donde $\mathbf{x}(1) = \mathbf{y}(0)$, entonces la curva combinada tiene continuidad G^0 en ese punto.

c) La matriz base de Bézier es una matriz cuyas columnas son los puntos de control de la curva.

Los ejercicios 13 a 15 se refieren a la subdivisión de la curva de Bézier que se muestra en la figura 7. Sea $\mathbf{x}(t)$ la curva de Bézier, con puntos de control $\mathbf{p}_0,\ldots, \mathbf{p}_3$, y sean $\mathbf{y}(t)$ y $\mathbf{z}(t)$ las curvas de Bézier de subdivisión como en el libro, con puntos de control $\mathbf{q}_0,\ldots, \mathbf{q}_3$ y $\mathbf{r}_0,\ldots, \mathbf{r}_3$, respectivamente.

13. *a*) Utilice la ecuación (12) para demostrar que $\mathbf{q}_1$ es el punto medio del segmento de $\mathbf{p}_0$ a $\mathbf{p}_1$.

b) Aplique la ecuación (13) para demostrar que

$$8\mathbf{q}_2 = 8\mathbf{q}_3 + \mathbf{p}_0 + \mathbf{p}_1 - \mathbf{p}_2 - \mathbf{p}_3$$

c) Utilice el inciso *b*), la ecuación (8) y el inciso *a*) para demostrar que $\mathbf{q}_2$ es el punto medio del segmento de $\mathbf{q}_1$ al punto medio del segmento de $\mathbf{p}_1$ a $\mathbf{p}_2$. Es decir, $\mathbf{q}_2 = \frac{1}{2}[\mathbf{q}_1 + \frac{1}{2}(\mathbf{p}_1 + \mathbf{p}_2)]$.

14. *a*) Justifique cada signo de igualdad:

$$3(\mathbf{r}_3 - \mathbf{r}_2) = \mathbf{z}'(1) = .5\mathbf{x}'(1) = \tfrac{3}{2}(\mathbf{p}_3 - \mathbf{p}_2)$$

b) Demuestre que $\mathbf{r}_2$ es el punto medio del segmento de $\mathbf{p}_2$ a $\mathbf{p}_3$.

c) Justifique cada signo de igualdad: $3(\mathbf{r}_1 - \mathbf{r}_0) = \mathbf{z}'(0) = .5\mathbf{x}'(.5)$.

d) Utilice el inciso *c*) para demostrar que $8\mathbf{r}_1 = -\mathbf{p}_0 - \mathbf{p}_1 + \mathbf{p}_2 + \mathbf{p}_3 + 8\mathbf{r}_0$.

e) Use el inciso *d*), la ecuación (8) y el inciso *a*) para demostrar que $\mathbf{r}_1$ es el punto medio del segmento de $\mathbf{r}_2$ al punto medio del segmento de $\mathbf{p}_1$ a $\mathbf{p}_2$. Es decir, $\mathbf{r}_1 = \frac{1}{2}[\mathbf{r}_2 + \frac{1}{2}(\mathbf{p}_1 + \mathbf{p}_2)]$.

15. Algunas veces solo se necesita la mitad de una curva de Bézier para subdividir más. Por ejemplo, la subdivisión del lado "izquierdo" se logra con los incisos *a*) y *c*) del ejercicio 13 y la ecuación (8). Cuando se dividen ambas mitades de la curva $\mathbf{x}(t)$, es posible organizar los cálculos de manera eficiente para determinar al mismo tiempo los puntos de control izquierdos y derechos, sin emplear la ecuación (8) directamente.

a) Demuestre que los vectores tangentes $\mathbf{y}'(1)$ y $\mathbf{z}'(0)$ son iguales.

b) Utilice el inciso *a*) para demostrar que $\mathbf{q}_3$ (que es igual a $\mathbf{r}_0$) es el punto medio del segmento de $\mathbf{q}_2$ a $\mathbf{r}_1$.

c) Empleando el inciso *b*) y los resultados de los ejercicios 13 y 14, escriba un algoritmo que calcule los puntos de control de $\mathbf{y}(t)$ y $\mathbf{z}(t)$ de manera eficiente. Solo se necesitan sumas y división entre 2.

16. Explique por qué una curva cúbica de Bézier está determinada plenamente por $\mathbf{x}(0)$, $\mathbf{x}'(0)$, $\mathbf{x}(1)$ y $\mathbf{x}'(1)$.

17. La simbología TrueType®, creada por Apple Computer y Adobe Systems, utiliza curvas cuadráticas de Bézier, mientras que PostScript®, creado por Microsoft, emplea curvas cúbicas de Bézier. Las curvas cúbicas dan más flexibilidad para el diseño de los tipos de letra, pero es importante para Microsoft que cada letra que utiliza curvas cuadráticas se pueda transformar en una que use curvas cúbicas. Suponga que $\mathbf{w}(t)$ es una curva cuadrática, con puntos de control $\mathbf{p}_0$, $\mathbf{p}_1$ y $\mathbf{p}_2$.

a) Encuentre puntos de control $\mathbf{r}_0$, $\mathbf{r}_1$, $\mathbf{r}_2$ y $\mathbf{r}_3$ tales que la curva cúbica de Bézier $\mathbf{x}(t)$ con esos puntos de control y con la propiedad de que $\mathbf{x}(t)$ y $\mathbf{w}(t)$ tengan los mismos puntos inicial y terminal, así como los mismos vectores tangentes en $t = 0$ y $t = 1$. (Véase el ejercicio 16).

b) Demuestre que si $\mathbf{x}(t)$ se construye como en el inciso *a*), entonces $\mathbf{x}(t) = \mathbf{w}(t)$ para $0 \le t \le 1$.

18. Utilice la multiplicación matricial particionada para calcular el siguiente producto matricial, que se presenta en la fórmula alternativa (5) para una curva de Bézier:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix}$$

SOLUCIONES A LOS PROBLEMAS DE PRÁCTICA

1. De la ecuación (14) con $t = 0$, $\mathbf{x}(0) \neq \mathbf{p}_0$ porque

$$\mathbf{x}(0) = \tfrac{1}{6}[\mathbf{p}_0 + 4\mathbf{p}_1 + \mathbf{p}_2] = \tfrac{1}{6}\mathbf{p}_0 + \tfrac{2}{3}\mathbf{p}_1 + \tfrac{1}{6}\mathbf{p}_2$$

Los coeficientes son no negativos y suman 1, por lo que $\mathbf{x}(0)$ está en conv $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$, y las coordenadas afines con respecto a $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$ son $\left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right)$.

2. A partir de la ecuación (14) con $t = 1$, $\mathbf{x}(1) \neq \mathbf{p}_3$ porque

$$\mathbf{x}(1) = \tfrac{1}{6}[\mathbf{p}_1 + 4\mathbf{p}_2 + \mathbf{p}_3] = \tfrac{1}{6}\mathbf{p}_1 + \tfrac{2}{3}\mathbf{p}_2 + \tfrac{1}{6}\mathbf{p}_3$$

Los coeficientes son no negativos y suman 1, por lo que $\mathbf{x}(1)$ está en conv $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$, y las coordenadas afines con respecto a $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ son $\left(\frac{1}{6}, \frac{2}{3}, \frac{1}{6}\right)$.

APÉNDICE

A

Unicidad de la forma escalonada reducida

TEOREMA

Unicidad de la forma escalonada reducida

Cada matriz A de $m \times n$ es equivalente por filas a una única matriz escalonada reducida U.

DEMOSTRACIÓN La demostración se basa en la idea de la sección 4.3 de que las columnas de matrices equivalentes por filas tienen exactamente las mismas relaciones de dependencia lineal.

El algoritmo de reducción por filas indica que al menos existe una matriz U del tipo mencionado. Suponga que A es equivalente por filas a las matrices U y V en forma escalonada reducida. La entrada del extremo izquierdo, distinta de cero, en una fila de U es un "1 principal". La ubicación de este 1 principal se conoce como posición pivote, y la columna correspondiente que lo contiene es una columna pivote. (Esta definición solo emplea la naturaleza escalonada de U y V, y no supone la unicidad de la forma escalonada reducida).

Las columnas pivote de U y V son precisamente las columnas distintas de cero que no son linealmente dependientes de las columnas a su izquierda. (Esta condición se satisface automáticamente por una primera columna, si esta es distinta de cero). Como U y V son equivalentes por filas (ambas son equivalentes por filas a A), sus columnas tienen las mismas relaciones de dependencia lineal. Así, las columnas pivote de U y V se presentan en idénticas ubicaciones. Si existen r de tales columnas, y puesto que U y V están en forma escalonada reducida, entonces sus columnas pivote son las primeras r columnas de la matriz identidad de $m \times m$. En consecuencia, *son iguales las columnas pivote correspondientes de U y V*.

Por último, considere cualquier columna de U que no sea pivote, por ejemplo, la columna j. Esta columna es cero o una combinación lineal de las columnas pivote a su izquierda (porque esas columnas pivote son una base para el espacio generado por las columnas a la izquierda de la columna j). Ambos casos se pueden expresar mediante $U\mathbf{x} = \mathbf{0}$ para alguna $\mathbf{x}$ cuya j-ésima entrada es 1. Entonces, también $V\mathbf{x} = \mathbf{0}$, lo que significa que la columna j de V es cero o la misma combinación lineal de las columnas pivote de V a su izquierda. Como las columnas pivote correspondientes de U y V son iguales, entonces también son iguales las columnas j de U y V. Esto es válido para todas las columnas que no son pivote, de manera que $V = U$, lo que prueba la unicidad de U.

APÉNDICE

B

Números complejos

Un **número complejo** es un número que se escribe en la forma

$$z = a + bi$$

donde a y b son números reales, e i es un símbolo formal que satisface la relación $i^2 = -1$. El número a es la **parte real** de z, que se denota como Re z, y b es la **parte imaginaria** de z, que se denota como Im z. Dos números complejos se consideran iguales si y solo si sus partes real e imaginaria son iguales. Por ejemplo, si $z = 5 + (-2)i$, entonces Re $z = 5$ e Im $z = -2$. Para simplificar, se escribe $z = 5 - 2i$.

Un número real a se considera como un tipo especial de número complejo, al identificar a con $a + 0i$. Además, las operaciones aritméticas de los números reales se pueden extender al conjunto de números complejos.

El **sistema de números complejos**, que se denota con $\mathbb{C}$, es el conjunto de todos los números complejos, junto con las siguientes operaciones de suma y multiplicación:

$$(a + bi) + (c + di) = (a + c) + (b + d)i \tag{1}$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i \tag{2}$$

Esas reglas se reducen a la suma y multiplicación ordinarias de números reales cuando b y d son cero en las ecuaciones (1) y (2). Es fácil comprobar que las leyes usuales de la aritmética para $\mathbb{R}$ también son válidas para $\mathbb{C}$. Por esta razón, la multiplicación por lo general se calcula mediante desarrollos algebraicos, como se muestra en el siguiente ejemplo.

EJEMPLO 1

$$\begin{aligned}(5 - 2i)(3 + 4i) &= 15 + 20i - 6i - 8i^2 \\ &= 15 + 14i - 8(-1) \\ &= 23 + 14i\end{aligned}$$

Es decir, multiplique cada término de $5 - 2i$ por cada término de $3 + 4i$, utilice $i^2 = -1$, y escriba el resultado en la forma $a + bi$. ■

La resta de los números complejos z_1 y z_2 se define por

$$z_1 - z_2 = z_1 + (-1)z_2$$

En particular, se escribe $-z$ en vez de $(-1)z$.

El **conjugado** de $z = a + bi$ es el número complejo $\overline{z}$ (que se lee "z barra"), definido por

$$\boxed{\overline{z} = a - bi}$$

Para obtener $\overline{z}$ de z basta con cambiar el signo de la parte imaginaria.

EJEMPLO 2 El conjugado de $-3 + 4i$ es $-3 - 4i$; escriba $\overline{-3 + 4i} = -3 - 4i$. ■

Observe que si $z = a + bi$, entonces,

$$z\overline{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2 i^2 = a^2 + b^2 \tag{3}$$

Como $z\overline{z}$ es real y no negativo, entonces tiene una raíz cuadrada real. El **valor absoluto** (o **módulo**) de z es el número real $|z|$ definido por

$$\boxed{|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}}$$

Si z es un número real, entonces $z = a + 0i$, y $|z| = \sqrt{a^2}$, que es igual al valor absoluto ordinario de a.

A continuación se listan algunas propiedades útiles de conjugados y del valor absoluto; w y z denotan números complejos.

1. $\overline{z} = z$ si y solo si z es un número real.

2. $\overline{w + z} = \overline{w} + \overline{z}$.

3. $\overline{wz} = \overline{w}\,\overline{z}$; en particular, $\overline{rz} = r\overline{z}$ si r es un número real.

4. $z\overline{z} = |z|^2 \geq 0$.

5. $|wz| = |w||z|$.

6. $|w + z| \leq |w| + |z|$.

Si $z \neq 0$, entonces $|z| > 0$ y z tiene un inverso multiplicativo, que se denota como $1/z$ o z^{-1} y está dado por

$$\boxed{\frac{1}{z} = z^{-1} = \frac{\overline{z}}{|z|^2}}$$

Desde luego, un cociente w/z simplemente significa $w \cdot (1/z)$.

EJEMPLO 3 Sean $w = 3 + 4i$ y $z = 5 - 2i$. Calcule $z\overline{z}$, $|z|$ y w/z.

SOLUCIÓN A partir de la ecuación (3),

$$z\overline{z} = 5^2 + (-2)^2 = 25 + 4 = 29$$

Para el valor absoluto, $|z| = \sqrt{z\overline{z}} = \sqrt{29}$. Para calcular w/z, primero multiplique el numerador y el denominador por $\overline{z}$, el conjugado del denominador. Utilizando la ecuación (3), se elimina la i en el denominador:

$$\begin{aligned}\frac{w}{z} &= \frac{3+4i}{5-2i} \\ &= \frac{3+4i}{5-2i}\cdot\frac{5+2i}{5+2i} \\ &= \frac{15+6i+20i-8}{5^2+(-2)^2} \\ &= \frac{7+26i}{29} \\ &= \frac{7}{29}+\frac{26}{29}i\end{aligned}$$

■

Interpretación geométrica

Cada número complejo $z = a + bi$ corresponde a un punto (a, b) en el plano $\mathbb{R}^2$, como en la figura 1. El eje horizontal es el **eje real** ya que los puntos $(a, 0)$ en este corresponden a los números reales. El eje vertical es el **eje imaginario** porque los puntos $(0, b)$ sobre este corresponden a los **números imaginarios puros** de la forma $0 + bi$, o simplemente bi. El conjugado de z es la imagen de espejo de z en el eje real. El valor absoluto de z es la distancia de (a, b) al origen.

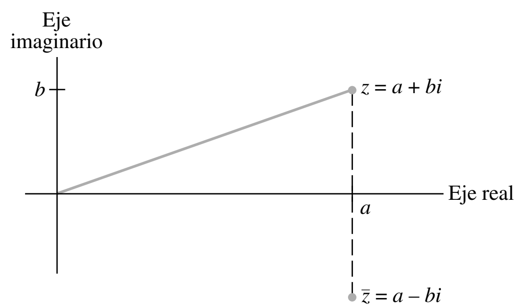


FIGURA 1 El complejo conjugado es una imagen de espejo.

La suma de los números complejos $z = a + bi$ y $w = c + di$ corresponde a la suma vectorial de (a, b) y (c, d) en $\mathbb{R}^2$, como se muestra en la figura 2.

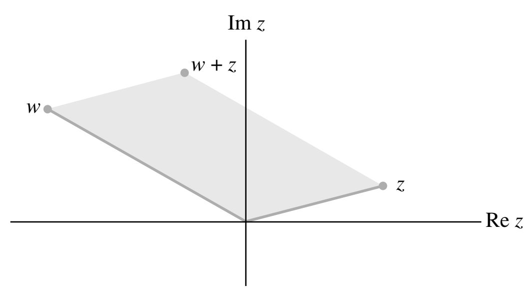


FIGURA 2 Suma de números complejos.

Para dar una representación gráfica de la multiplicación compleja, se utilizan **coordenadas polares** en $\mathbb{R}^2$. Dado un número complejo diferente de cero $z = a + bi$, sea φ el ángulo entre el eje real positivo y el punto (a, b), como se muestra en la figura 3, donde $-\pi < \varphi \leq \pi$. El ángulo φ se llama el **argumento** de z; y se escribe $\varphi = \arg z$. De trigonometría,

$$a = |z| \cos \varphi, \qquad b = |z| \operatorname{sen} \varphi$$

y así,

$$z = a + bi = |z|(\cos \varphi + i \operatorname{sen} \varphi)$$

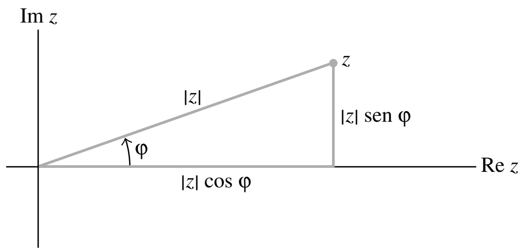


FIGURA 3 Coordenadas polares de z.

Si w es otro número complejo distinto de cero, por ejemplo,

$$w = |w| (\cos \vartheta + i \operatorname{sen} \vartheta)$$

entonces, empleando identidades trigonométricas estándares para el seno y coseno de la suma de dos ángulos, es posible comprobar que

$$wz = |w|\,|z|\,[\cos(\vartheta + \varphi) + i \operatorname{sen}(\vartheta + \varphi)] \tag{4}$$

Véase la figura 4. Para cocientes, se puede escribir una fórmula similar en forma polar. Las expresiones para productos y cocientes se enuncian como sigue.

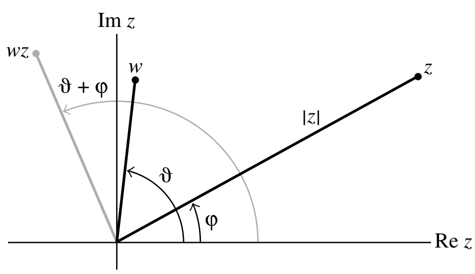
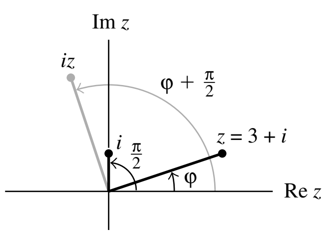


FIGURA 4 Multiplicación con coordenadas polares.

El producto de dos números complejos distintos de cero está dado, en forma polar, por el producto de sus valores absolutos y la suma de sus argumentos. El cociente de dos números complejos distintos de cero está dado por el cociente de sus valores absolutos y la diferencia de sus argumentos.



Multiplicación por i.

EJEMPLO 4

a) Si w tiene valor absoluto 1, entonces $w = \cos \vartheta + i \operatorname{sen} \vartheta$, donde ϑ es el argumento de w. La multiplicación de cualquier número z diferente de cero por w simplemente hace girar a z un ángulo ϑ.

b) El argumento de i es $\pi/2$ radianes, de manera que la multiplicación de z por i hace girar a z un ángulo de $\pi/2$ radianes. Por ejemplo, $3 + i$ gira $(3 + i)i = -1 + 3i$. ■

Potencias de un número complejo

La fórmula (4) se aplica cuando $z = w = r(\cos\varphi + i \operatorname{sen}\varphi)$. En este caso,

$$z^2 = r^2(\cos 2\varphi + i \operatorname{sen} 2\varphi)$$

y

$$\begin{aligned} z^3 &= z \cdot z^2 \\ &= r(\cos\varphi + i \operatorname{sen}\varphi) \cdot r^2(\cos 2\varphi + i \operatorname{sen} 2\varphi) \\ &= r^3(\cos 3\varphi + i \operatorname{sen} 3\varphi) \end{aligned}$$

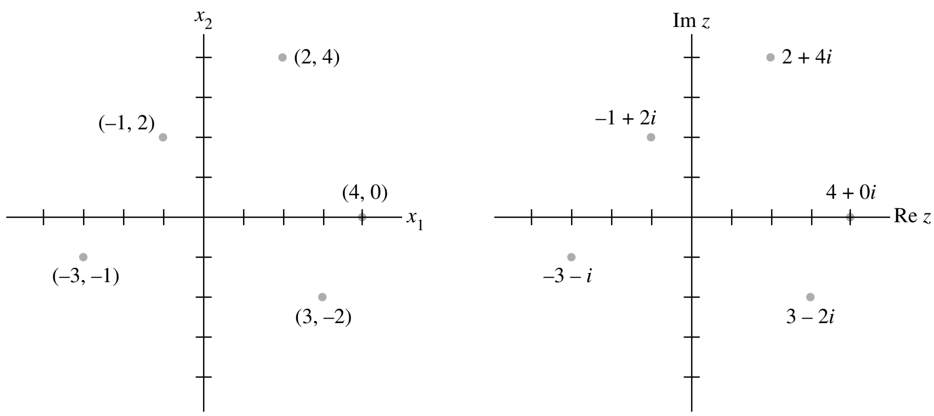
En general, para cualquier entero positivo k,

$$z^k = r^k(\cos k\varphi + i \operatorname{sen} k\varphi)$$

Este resultado se conoce como *teorema de De Moivre*.

Números complejos y $\mathbb{R}^2$

No obstante que los elementos de $\mathbb{R}^2$ y $\mathbb{C}$ están en correspondencia uno a uno, y las operaciones de suma son esencialmente las mismas, existe una diferencia lógica entre $\mathbb{R}^2$ y $\mathbb{C}$. En $\mathbb{R}^2$ solo se puede multiplicar un vector por un escalar real, mientras que en $\mathbb{C}$ es posible multiplicar dos números complejos cualesquiera para obtener un tercer número complejo. (El producto punto en $\mathbb{R}^2$ no cuenta, porque produce un escalar, no un elemento de $\mathbb{R}^2$.) Se utiliza notación escalar para los elementos de $\mathbb{C}$ con la finalidad de poner de relieve esta diferencia.



El plano real $\mathbb{R}^2$. El plano complejo $\mathbb{C}$.

Glosario

A

adjunta clásica: La matriz adj A formada de una matriz cuadrada A al remplazar la entrada (i, j) de A por el cofactor (i, j), para toda i y j, y después transponer la matriz resultante.

algoritmo de reducción por filas: Un método sistemático que emplea operaciones elementales de fila que reduce una matriz a la forma escalonada o a la forma escalonada reducida.

análisis de tendencia: El uso de polinomios ortogonales para ajustar datos, con el producto interior dado por la evaluación en un conjunto finito de puntos.

ángulo (entre vectores diferentes de cero $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^2$ o $\mathbb{R}^3$): El ángulo ϑ entre los dos segmentos de recta dirigidos del origen a los puntos $\mathbf{u}$ y $\mathbf{v}$. Relaciona al producto escalar mediante

$$\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta$$

aproximación de Fourier (de orden n): El punto más cercano, en el subespacio de polinomios trigonométricos de n-ésimo orden, a una función dada en $C[0, 2\pi]$.

aritmética de punto flotante: Aritmética con números representados como decimales $\pm\,.d_1 \cdots d_p \times 10^r$, donde r es un entero y el número p de dígitos a la derecha del punto decimal generalmente está entre 8 y 16.

atractor (de un sistema dinámico en $\mathbb{R}^2$): El origen cuando todas las trayectorias tienden a $\mathbf{0}$.

B

base (para un subespacio no trivial H de un espacio vectorial V): Un conjunto indexado $\mathcal{B} = \{\mathbf{v}_1,\dots, \mathbf{v}_p\}$ en V tal que: **i.** $\mathcal{B}$ es un conjunto linealmente independiente y **ii.** El subespacio generado por $\mathcal{B}$ coincide con H, es decir, $H = \text{Gen}\,\{\mathbf{v}_1,\dots, \mathbf{v}_p\}$.

base de vectores propios: Una base que consiste totalmente en vectores propios de una matriz dada.

base estándar: La base $\mathcal{E} = \{\mathbf{e}_1,\dots, \mathbf{e}_n\}$ para $\mathbb{R}^n$ que consiste en las columnas de la matriz identidad de $n \times n$, o la base $\{1, t,\dots, t^n\}$ para $\mathbb{P}_n$.

base ortogonal: Una base que también es un conjunto ortogonal.

base ortonormal: Una base que es un conjunto ortogonal de vectores unitarios.

$\mathcal{B}$-coordenadas de x: *Véase* coordenadas de $\mathbf{x}$ respecto de la base $\mathcal{B}$.

$\mathcal{B}$-matriz (para T): Una matriz $[T]_{\mathcal{B}}$ para una transformación lineal $T : V \to V$ respecto de una base $\mathcal{B}$ para V, con la propiedad de que $[T(\mathbf{x})]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[\mathbf{x}]_{\mathcal{B}}$ para toda $\mathbf{x}$ en V.

bola abierta B(p, δ) en $\mathbb{R}^n$: El conjunto $\{\mathbf{x} : \|\mathbf{x} - \mathbf{p}\| < \delta\}$ en $\mathbb{R}^n$, donde $\delta > 0$.

bola cerrada (en $\mathbb{R}^n$): Un conjunto $\{\mathbf{x} : \|\mathbf{x} - \mathbf{p}\| < \delta\}$ en $\mathbb{R}^n$, donde $\mathbf{p}$ está en $\mathbb{R}^n$ y $\delta > 0$.

C

cadena de Markov: Una secuencia de vectores de probabilidad $\mathbf{x}_0$, $\mathbf{x}_1, \mathbf{x}_2,\dots$, junto con una matriz estocástica P tal que $\mathbf{x}_{k+1} = P\mathbf{x}_k$ para $k = 0, 1, 2,\dots$

cambio de base: *Véase* matriz de cambio de coordenadas.

cambio relativo o **error relativo** (en $\mathbf{b}$): La cantidad $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ cuando $\mathbf{b}$ se cambia a $\mathbf{b} + \Delta\mathbf{b}$.

cociente de Rayleigh: $R(\mathbf{x}) = (\mathbf{x}^TA\mathbf{x})/(\mathbf{x}^T\mathbf{x})$. Una estimación de un valor propio de A (por lo general, una matriz simétrica).

codominio (de una transformación $T : \mathbb{R}^n \to \mathbb{R}^m$): El conjunto $\mathbb{R}^m$ que contiene al rango de T. En general, si T mapea un espacio vectorial V en un espacio vectorial W, entonces W es el codominio de T.

coeficientes de Fourier: Los pesos empleados para hacer que un polinomio trigonométrico sea una aproximación de Fourier a una función.

coeficientes de regresión: Los coeficientes β_0 y β_1 en la recta de mínimos cuadrados $y = \beta_0 + \beta_1 x$.

cofactor: Un número $C_{ij} = (-1)^{i+j}\det A_{ij}$, llamado el cofactor (i, j) de A, donde A_{ij} es la submatriz formada al eliminar la i-ésima fila y la j-ésima columna de A.

cofactores: *Véase* desarrollo por cofactores.

columna pivote: Una columna que contiene una posición pivote.

combinación afín: Una combinación lineal de vectores (puntos en $\mathbb{R}^n$) en la que la suma de los pesos implicados es 1.

combinación convexa (de puntos $\mathbf{v}_1,\dots, \mathbf{v}_k$ en $\mathbb{R}^n$): Una combinación lineal de vectores (puntos) en la que los pesos en la combinación son positivos y la suma de los pesos es 1.

combinación lineal: Una suma de múltiplos escalares de vectores. Los escalares se llaman *pesos*.

combinación positiva (de puntos $\mathbf{v}_1,\dots, \mathbf{v}_m$ en $\mathbb{R}^n$): Una combinación lineal $c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$, donde toda $c_i \geq 0$.

complemento de Schur: Una cierta matriz formada a partir de los bloques de una matriz particionada de 2×2, $A = [A_{ij}]$. Si A_{11} es invertible, su complemento de Schur está dado por $A_{22} - A_{21}A_{11}^{-1}A_{12}$. Si A_{22} es invertible, su complemento de Schur está dado por $A_{11} - A_{12}A_{22}^{-1}A_{21}$.

complemento ortogonal (de W): El conjunto $W^{\perp}$ de todos los vectores ortogonales a W.

componente de y ortogonal a u (para $\mathbf{u} \neq \mathbf{0}$): El vector $\mathbf{y} - \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$.

componentes principales (de los datos en una matriz B de observaciones): Los vectores propios unitarios de una matriz de covarianza muestral S para B, con los vectores propios arreglados de tal manera que los valores propios correspondientes de S disminuyan en magnitud. Si B es una forma de desviación media, entonces las componentes principales son los vectores singulares derechos en una descomposición en valores singulares de B^T.

composición de transformaciones lineales: Un mapeo generado al aplicar dos o más transformaciones lineales sucesivas. Si las transformaciones son transformaciones matriciales, por ejemplo, multiplicación por la izquierda de B seguida por multiplicación por la izquierda de A, entonces la composición es el mapeo $\mathbf{x} \mapsto A(B\mathbf{x})$.

conjunto abierto S en $\mathbb{R}^n$: Un conjunto que no contiene a ninguno de sus puntos frontera. (De manera equivalente, S es abierto si cada punto de S es un punto interior).

conjunto acotado en $\mathbb{R}^n$: Un conjunto que está contenido en una bola abierta $B(\mathbf{0}, \delta)$ para alguna $\delta > 0$.

conjunto afín (o **subconjunto afín**)**:** Un conjunto S de puntos tales que si $\mathbf{p}$ y $\mathbf{q}$ están en S, entonces, $(1 - t)\mathbf{p} + t\mathbf{q} \in S$ para cada número real t.

conjunto afínmente dependiente: Un conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ tal que existen números reales $c_1,\ldots, c_p$, no todos cero, tales que $c_1 + \cdots + c_p = 0$ y $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$.

conjunto afínmente independiente: Un conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$ que no tiene dependencia afín.

conjunto cerrado (en $\mathbb{R}^n$): Un conjunto que contiene todos sus puntos frontera.

conjunto compacto (en $\mathbb{R}^n$): Un conjunto en $\mathbb{R}^n$ que es cerrado y acotado.

conjunto convexo: Un conjunto S con la propiedad de que para cada $\mathbf{p}$ y $\mathbf{q}$ en S, el segmento de recta $\overline{\mathbf{pq}}$ está contenido en S.

conjunto fundamental de soluciones: Una base para el conjunto de todas las soluciones de una ecuación diferencial o de una ecuación en diferencias lineal homogénea.

conjunto generado por $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$: El conjunto Gen $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$.

conjunto generador (para un subespacio H): Cualquier conjunto $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ en H tal que $H =$ Gen $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$.

conjunto generador mínimo (para un subespacio H): Un conjunto $\mathcal{B}$ que genera a H y tiene la propiedad de que si uno de los elementos de $\mathcal{B}$ se elimina de $\mathcal{B}$, entonces el nuevo conjunto no genera a H.

conjunto linealmente independiente máximo (en V): Un conjunto linealmente independiente $\mathcal{B}$ en V tal que si un vector $\mathbf{v}$, que está en V pero no en $\mathcal{B}$, se agrega a $\mathcal{B}$ entonces el nuevo conjunto es linealmente dependiente.

conjunto nivel (o **gradiente**) de una funcional lineal f sobre $\mathbb{R}^n$: Un conjunto $[f : d] = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = d\}$

conjunto ortogonal: Un conjunto S de vectores tales que $\mathbf{u}\cdot\mathbf{v} = 0$ para cada par distinto $\mathbf{u}$, $\mathbf{v}$ en S.

conjunto ortonormal: Un conjunto ortogonal de vectores unitarios.

conjunto solución: El conjunto de todas las posibles soluciones de un sistema lineal. El conjunto solución es vacío cuando el sistema lineal es inconsistente.

contracción: Un mapeo $\mathbf{x} \mapsto r\mathbf{x}$ para algún escalar r, con $0 \le r \le 1$.

controlable (par de matrices): Un par de matrices (A, B) donde A es de $n \times n$, B tiene n filas, y

$$\text{rango}\begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix} = n$$

Relacionado con un modelo de espacio de estados de un sistema de control y la ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k$ $(k = 0, 1,\ldots)$.

convergente (secuencia de vectores): Una secuencia $\{\mathbf{x}_k\}$ tal que las entradas en $\mathbf{x}_k$ se pueden hacer tan cercanas como se quiera a las entradas de algún vector fijo para toda k suficientemente grande.

coordenadas baricéntricas (de un punto $\mathbf{p}$ respecto de un conjunto $S = \{\mathbf{v}_1,\ldots, \mathbf{v}_k\}$ afínmente independiente): El conjunto (único) de pesos $c_1,\ldots, c_k$ tales que $\mathbf{p} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ y $c_1 + \cdots + c_k = 1$. (Algunas veces también se llaman coordenadas afines de $\mathbf{p}$ respecto de S.)

coordenadas de x respecto de la base $\mathcal{B} = \{\mathbf{b}_1,\ldots, \mathbf{b}_n\}$: Los pesos $c_1,\ldots, c_n$ en la ecuación $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

coordenadas homogéneas: En $\mathbb{R}^3$, la representación de (x, y, z) como (X, Y, Z, H) para cualquier $H \neq 0$, donde $x = X/H$, $y = Y/H$, y $z = Z/H$. En $\mathbb{R}^2$, generalmente se toma $H = 1$, y las coordenadas homogéneas de (x, y) se escriben como $(x, y, 1)$.

corriente de circuito: La cantidad de corriente eléctrica que fluye por un circuito que hace la suma algebraica de las caídas de voltaje RI, alrededor del circuito, sea igual a la suma algebraica de las fuentes de voltaje en el circuito.

covarianza (de variables x_i y x_j, para $i \neq j$): La entrada s_{ij} en la matriz de covarianza S para una matriz de observaciones, donde x_i y x_j varían sobre las coordenadas i-ésima y j-ésima, respectivamente, de los vectores de observaciones.

cubo: Un objeto sólido tridimensional acotado por seis caras cuadradas, con tres caras que se encuentran en cada vértice.

curva de Bézier cuadrática: Una curva cuya descripción se puede describir en la forma $\mathbf{g}(t) = (1 - t)\mathbf{f}_0(t) + t\mathbf{f}_1(t)$ para $0 \le t \le 1$, donde $\mathbf{f}_0(t) = (1 - t)\mathbf{p}_0 + t\mathbf{p}_1$ y $\mathbf{f}_1(t) = (1 - t)\mathbf{p}_1 + t\mathbf{p}_2$. Los puntos $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$ son los *puntos de control* para la curva.

D

demandas intermedias: Demandas de bienes o servicios que se consumirán en el proceso de elaborar otros bienes y servicios para los consumidores. Si $\mathbf{x}$ es el nivel de producción y C es la matriz de consumo, entonces $C\mathbf{x}$ lista las demandas intermedias.

desarrollo columna-fila: La expresión de un producto AB como una suma de productos exteriores: $\text{col}_1(A)\ \text{fila}_1(B) + \cdots + \text{col}_n(A)\ \text{fila}_n(B)$, donde n es el número de columnas de A.

desarrollo por cofactores: Una fórmula para det A empleando los cofactores asociados con una fila o una columna, por ejemplo, para la fila 1:

$$\det A = a_{11}C_{11} + \cdots + a_{1n}C_{1n}$$

descomposición de vectores propios (de $\mathbf{x}$): Una ecuación, $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$, que expresa a $\mathbf{x}$ como una combinación lineal de vectores propios de una matriz.

descomposición en valores singulares (de una matriz A de $m \times n$): $A = U\Sigma V^T$, donde U es una matriz ortogonal de $m \times m$, V es una matriz ortogonal de $n \times n$, y Σ es una matriz de $m \times n$ con entradas no negativas sobre la diagonal principal (arregladas en orden decreciente de magnitud) y ceros en las demás entradas. Si rango $A = r$, entonces Σ tiene exactamente r entradas positivas (los valores singulares distintos de cero de A) sobre la diagonal.

descomposición en valores singulares reducida: Una factorización $A = UDV^T$, para una matriz A de $m \times n$, con rango r, donde U es de $m \times r$ con columnas ortonormales, D es una matriz diagonal de $r \times r$ con r valores singulares de A distintos de cero en su diagonal, y V es de $n \times r$ con columnas ortonormales.

descomposición espectral (de A): Una representación

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

donde $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es una base ortonormal de vectores propios de A, y $\lambda_1, \ldots, \lambda_n$ son los valores propios correspondientes de A.

descomposición ortogonal: La representación de un vector $\mathbf{y}$ como la suma de dos vectores, uno en un subespacio W especificado y el otro en $W^\perp$. En general, una descomposición $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$, donde $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ es una base ortogonal para un subespacio que contiene a $\mathbf{y}$.

descomposición polar (de A): Una factorización $A = PQ$, donde P es una matriz positiva semidefinida de $n \times n$ con el mismo rango que A, y Q es una matriz ortogonal de $n \times n$.

descripción explícita (de un subespacio W de $\mathbb{R}^n$): Una representación paramétrica de W como el conjunto de todas las combinaciones lineales de un conjunto de vectores específicos.

descripción implícita (de un subespacio W de $\mathbb{R}^n$): Un conjunto de una o más ecuaciones homogéneas que caracterizan a los puntos de W.

desigualdad de Cauchy-Schwarz: $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|u\|\cdot\|v\|$ para toda $\mathbf{u}$, $\mathbf{v}$.

desigualdad del triángulo: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ para toda $\mathbf{u}$, $\mathbf{v}$.

determinante (de una matriz cuadrada A): El número $\det A$ definido inductivamente mediante un desarrollo por cofactores sobre la primera fila de A. Además, $(-1)^r$ veces el producto de las entradas diagonales en cualquier forma escalonada U obtenida de A por remplazos de fila y r intercambios de fila (pero no por operaciones de escalamiento).

diagonal por bloques (matriz): Una matriz particionada $A = [A_{ij}]$ tal que cada bloque A_{ij} es una matriz cero para $i \neq j$.

diagonal principal (de una matriz): Las entradas con igual índice de fila y columna.

diferente de cero (matriz o vector): Una matriz (con posiblemente solo una fila o columna) que contiene al menos una entrada distinta de cero.

dilatación: Un mapeo $\mathbf{x} \mapsto r\mathbf{x}$ para algún escalar r, con $1 < r$.

dimensión:

de un conjunto S: La dimensión del plano afín más pequeño que contiene a S.

de un espacio vectorial V: El número de vectores en una base para V, que se escribe como $\dim V$. La dimensión del espacio cero es 0.

de un plano afín S: La dimensión del subespacio paralelo correspondiente.

de un subespacio S: El número de vectores en una base para S, que se escribe como $\dim S$.

distancia a un subespacio: La distancia de un punto dado (vector) $\mathbf{v}$ al punto más cercano en el subespacio.

distancia entre u y v: La longitud del vector $\mathbf{u} - \mathbf{v}$, que se denota como $\text{dist}(\mathbf{u}, \mathbf{v})$.

dominio (de una transformación T): El conjunto de todos los vectores $\mathbf{x}$ para los cuales $T(\mathbf{x})$ está definido.

E

ecuación auxiliar: Una ecuación polinomial en una variable r, creada con los coeficientes de una ecuación en diferencias homogénea.

ecuación característica (de A): $\det(A - \lambda I) = 0$.

ecuación en diferencias (o **relación de recurrencia lineal**): Una ecuación de la forma $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ cuya solución es una secuencia de vectores, $\mathbf{x}_0, \mathbf{x}_1, \ldots$

ecuación homogénea: Una ecuación de la forma $A\mathbf{x} = \mathbf{0}$, posiblemente escrita como una ecuación vectorial o como un sistema de ecuaciones lineales.

ecuación lineal (en las variables $x_1, \ldots, x_n$): Una ecuación que se puede escribir en la forma $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, donde b y los coeficientes $a_1, \ldots, a_n$ son números reales o complejos.

ecuación matricial: Una ecuación que implica al menos una matriz; por ejemplo, $A\mathbf{x} = \mathbf{b}$.

ecuación no homogénea: Una ecuación de la forma $A\mathbf{x} = \mathbf{b}$ con $\mathbf{b} \neq \mathbf{0}$, posiblemente escrita como una ecuación vectorial o como un sistema de ecuaciones lineales.

ecuación paramétrica de un plano: Una ecuación de la forma $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ (s, t en $\mathbb{R}$), con $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes.

ecuación paramétrica de una recta: Una ecuación de la forma $\mathbf{x} = \mathbf{p} + t\mathbf{v}$ (t en $\mathbb{R}$).

ecuación vectorial: Una ecuación que implica una combinación lineal de vectores con pesos indeterminados.

ecuaciones normales: El sistema de ecuaciones que se representa por $A^TA\mathrm{x} = \mathrm{A}^T\mathbf{b}$, cuya solución conduce a las soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. En estadística, una notación común es $X^TX\beta = X^T\mathbf{y}$.

ejes principales (de una forma cuadrática $\mathbf{x}^TA\mathbf{x}$): Las columnas ortonormales de una matriz ortogonal P tal que $P^{-1}AP$ es diagonal. (Esas columnas son vectores propios unitarios de A.) Por lo general, las columnas de P se ordenan de tal manera que los valores propios correspondientes de A se acomodan en orden decreciente de magnitud.

eliminación gaussiana: *Véase* algoritmo de reducción por filas.

entrada principal: La entrada distinta de cero en el extremo izquierdo en una fila de una matriz.

entradas diagonales (en una matriz): Entradas con iguales índices de fila y columna.

envolvente afín (o **afín generado**) de un conjunto S: El conjunto de todas las combinaciones afines de puntos en S, que se denota como $\text{aff}\, S$.

envolvente convexa (de un conjunto S): El conjunto de todas las combinaciones convexas de puntos en S, que se denota como: $\text{conv}\, S$.

envolvente positiva (de un conjunto S): El conjunto de todas las combinaciones positivas de puntos en S, que se denota como $\text{pos}\, S$.

error cuadrático medio: El error de una aproximación en un espacio con producto interior, donde el producto interior está dado por una integral definida.

error de mínimos cuadrados: La distancia $\|\mathbf{b} - A\hat{\mathbf{x}}\|$ de $\mathbf{b}$ a $A\hat{\mathbf{x}}$, cuando $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

error por redondeo: Error en la aritmética de punto flotante causado cuando el resultado de un cálculo se redondea (se trunca) al número de dígitos de punto flotante almacenado. Además, el error que resulta cuando la representación decimal de un número tal como 1/3 se aproxima por medio de un número de punto flotante con una cantidad finita de dígitos.

escalamiento (de un vector): Operación que consiste en multiplicar un vector (o una fila o columna de una matriz) por un escalar distinto de cero.

escalar: Un número (real) empleado para multiplicar un vector o una matriz.

espacio columna (de una matriz A de $m \times n$): El conjunto Col A de todas las combinaciones lineales de las columnas de A. Si $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, entonces Col A = Gen $\{\mathbf{a}_1,\ldots, \mathbf{a}_n\}$. De manera equivalente,

$$\text{Col } A = \{\mathbf{y} : \mathbf{y} = A\mathbf{x} \text{ para alguna } \mathbf{x} \text{ en } \mathbb{R}^n\}$$

espacio con producto interior: Un espacio vectorial sobre el cual se define un producto interior.

espacio fila (de una matriz A): El conjunto Fil A de todas las combinaciones lineales de los vectores formados con las filas de A; también se denota como Col A^T.

espacio nulo (de una matriz A de $m \times n$): El conjunto Nul A de todas las soluciones de la ecuación homogénea $A\mathbf{x} = \mathbf{0}$. Nul $A = \{\mathbf{x} : \mathbf{x}$ está en $\mathbb{R}^n$ y $A\mathbf{x} = \mathbf{0}\}$.

espacio propio (de A correspondiente a λ): El conjunto de *todas* las soluciones de $A\mathbf{x} = \lambda\mathbf{x}$, donde λ es un valor propio de A. Consiste en el vector cero y todos los vectores propios asociados a λ.

espacio vectorial: Un conjunto de objetos, llamados vectores, sobre el cual están definidas dos operaciones, llamadas suma y multiplicación por escalares. Deben satisfacerse 10 axiomas. Véase la primera definición en la sección 4.1.

espacio vectorial de dimensión finita: Un espacio vectorial que se genera por un conjunto finito de vectores.

espacio vectorial de dimensión infinita: Un espacio vectorial V diferente de cero que no tiene base finita.

espacios vectoriales isomorfos: Dos espacios vectoriales V y W para los cuales existe una transformación lineal T uno a uno que mapea V sobre W.

F

factorización (de A): Una ecuación que expresa A como un producto de dos o más matrices.

factorización de Cholesky: Una factorización $A = R^TR$, donde R es una matriz triangular superior invertible cuyas entradas diagonales son todas positivas.

factorización de Schur (de A, para escalares reales): Una factorización $A = URU^T$ de una matriz A de $n \times n$, que tiene n valor propios reales, donde U es una matriz ortogonal de $n \times n$, y R es una matriz triangular superior.

factorización LU: La representación de una matriz A en la forma $A = LU$ donde L es una matriz triangular inferior cuadrada con unos sobre la diagonal (una matriz triangular inferior unitaria) y U es una forma escalonada de A.

factorización LU permutada: La representación de una matriz A en la forma $A = LU$ donde L es una matriz cuadrada tal que una permutación de sus filas formará una matriz triangular inferior unitaria, y U es una forma escalonada de A.

factorización QR: Una factorización de una matriz A de $m \times n$ con columnas linealmente independientes, $A = QR$, donde Q es una matriz de $m \times n$ cuyas columnas forman una base ortonormal para Col A, y R es una matriz invertible triangular superior de $n \times n$ con entradas positivas sobre su diagonal.

fase progresiva (de reducción por filas): La primera parte del algoritmo que reduce una matriz a la forma escalonada.

fase regresiva (de reducción por filas): La última parte del algoritmo que reduce una matriz en forma escalonada a una forma escalonada reducida.

filtro lineal: Una ecuación en diferencias lineal empleada para transformar señales discretas en el tiempo.

flop: Una operación aritmética ($+$, $-$, $*$, $/$) que se efectúa sobre dos números reales de punto flotante.

forma cuadrática: Una función Q definida para $\mathbf{x}$ en $\mathbb{R}^n$ por $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, donde A es una matriz simétrica de $n \times n$ (llamada la **matriz de la forma cuadrática**).

forma cuadrática indefinida: Una forma cuadrática Q tal que $Q(\mathbf{x})$ toma valores positivos y negativos.

forma cuadrática negativa definida: Una forma cuadrática Q tal que $Q(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

forma cuadrática negativa semidefinida: Una forma cuadrática Q tal que $Q(\mathbf{x}) \leq 0$ para toda $\mathbf{x}$.

forma cuadrática positiva definida: Una forma cuadrática Q tal que $Q(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

forma cuadrática positiva semidefinida: Una forma cuadrática Q tal que $Q(\mathbf{x}) \geq 0$ para toda $\mathbf{x}$.

forma de desviación media (de un vector): Un vector cuyas entradas suman cero.

forma de desviación media (de una matriz de observaciones): Una matriz cuyos vectores fila están en forma de desviación media. Para cada fila, las entradas suman cero.

forma escalonada (o **forma escalonada por filas** de una matriz): Una matriz escalonada que es equivalente por filas a la matriz dada.

forma escalonada reducida (o **forma escalonada reducida por filas**): Una matriz escalonada reducida que es equivalente por filas a una matriz dada.

forma homogénea de (un vector) $\mathbf{v}$ en $\mathbb{R}^n$: El punto $\tilde{\mathbf{v}} = \begin{bmatrix} \mathbf{v} \\ 1 \end{bmatrix}$ en $\mathbb{R}^{n+1}$.

funcional lineal (en $\mathbb{R}^n$): Una transformación lineal f de $\mathbb{R}^n$ a $\mathbb{R}$.

funciones propias (de una ecuación diferencial $\mathbf{x}'(t) = A\mathbf{x}(t)$): Una función $\mathbf{x}(t) = \mathbf{v}e^{\lambda t}$, donde $\mathbf{v}$ es un vector propio de A, y λ es el valor propio correspondiente.

G

Gen $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$: El conjunto de todas las combinaciones lineales de $\mathbf{v}_1,\ldots, \mathbf{v}_p$. También, el *subespacio generado* por $\mathbf{v}_1,\ldots, \mathbf{v}_p$.

H

hiperplano (en $\mathbb{R}^n$): Un plano afín en $\mathbb{R}^n$ de dimensión $n - 1$. También: un traslado de un subespacio de dimensión $n - 1$.

hiperplano de soporte (a un conjunto convexo compacto S en $\mathbb{R}^n$): Un hiperplano $H = [f : d]$ tal que $H \cap S \neq \varnothing$ y $f(x) \leq d$ para toda x en S o $f(x) \geq d$ para toda x en S.

I

Im x: El vector en $\mathbb{R}^n$ formado con las partes imaginarias de las entradas de un vector **x** en $\mathbb{C}^n$.

imagen (de un vector **x** bajo una transformación T): El vector $T(\mathbf{x})$ asignado a **x** por T.

inversa (de uma matriz A de $n \times n$): Uma matriz A^{-1} de $n \times n$, tal que $AA^{-1} = A^{-1} = I_n$.

inversa de Moore-Penrose: *Véase* seudoinversa.

inversa por la derecha (de A): Cualquier matriz rectangular C tal que $AC = I$.

inversa por la izquierda (de A): Cualquier matriz rectangular C tal que $CA = I$.

isomorfismo: Un mapeo lineal uno a uno de un espacio vectorial sobre otro.

L

ley asociativa de la multiplicación: $A(BC) = (AB)C$, para toda A, B, C.

leyes de Kirchhoff: 1. (**ley de voltaje**) La suma algebraica de las caídas de voltaje RI en una dirección alrededor de un circuito es igual a la suma algebraica de las fuentes de voltaje en la misma dirección alrededor del circuito. **2.** (**ley de corriente**) La corriente en una rama es la suma algebraica de las corrientes de circuito que fluyen por esa rama.

leyes distributivas: (izquierda) $A(B + C) = AB + AC$, y (derecha) $(B + C)A = BA + CA$, para toda A, B, C.

linealmente dependientes (vectores): Un conjunto indexado $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ con la propiedad de que existen pesos $c_1,\ldots, c_p$, no todos cero, tales que $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$. Es decir, la ecuación vectorial $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ tiene una solución *no trivial*.

longitud (o **norma**, de **v**): El escalar $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.

M

magnitud (de un vector): *Véase* norma.

mapeo: *Véase* transformación.

mapeo de coordenadas (determinado por una base ordenada $\mathcal{B}$ en un espacio vectorial V): Un mapeo que a cada **x** en V le asocia su vector coordenado $[\mathbf{x}]_{\mathcal{B}}$.

mapeo sobre: Un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que cada **b** en R^m es la imagen de *al menos* una **x** en $\mathbb{R}^n$.

mapeo uno a uno: Un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ tal que cada **b** en $\mathbb{R}^m$ es la imagen de *a lo sumo* una **x** en R^n.

matrices conformadas para multiplicación por bloques: Dos matrices particionadas A y B tales que el producto por bloques AB está definido: la partición columna de A debe ajustar con la partición fila de B.

matrices equivalentes por filas: Dos matrices para las cuales existe una secuencia (finita) de operaciones de fila que transforma una matriz en la otra.

matrices que conmutan: Dos matrices A y B tales que $AB = BA$.

matrices semejantes: Matrices A y B tales que $P^{-1}AP = B$, o bien, de manera equivalente, $A = PBP^{-1}$, para alguna matriz invertible P.

matriz: Un arreglo rectangular de números.

matriz aumentada: Una matriz construida mediante una matriz de coeficientes para un sistema lineal y una o más columnas a la derecha. Cada columna adicional contiene las constantes del lado derecho de un sistema con la matriz de coeficientes dada.

matriz banda: Una matriz cuyas entradas diferentes de cero están dentro de una banda a lo largo de la diagonal principal.

matriz bidiagonal: Una matriz cuyas entradas diferentes de cero están sobre la diagonal principal y sobre una diagonal adyacente a la diagonal principal.

matriz casi singular: Una matriz mal condicionada.

matriz compañera: Una matriz de forma especial cuyo polinomio característico es $(-1)^n p(\lambda)$ cuando $p(\lambda)$ es un polinomio especificado cuyo término principal es λ^n.

matriz de cambio de coordenadas (de una base $\mathcal{B}$ a una base $\mathcal{C}$): Una matriz $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ que transforma vectores de $\mathcal{B}$-coordenadas a vectores de $\mathcal{C}$-coordenadas: $[\mathbf{x}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{x}]_{\mathcal{B}}$. Si $\mathcal{C}$ es la base estándar para $\mathbb{R}^n$, entonces algunas veces $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ se escribe como $P_{\mathcal{B}}$.

matriz de coeficientes: Una matriz cuyas entradas son los coeficientes de un sistema de ecuaciones lineales.

matriz de consumo: Una matriz en el modelo de entrada y salida de Leontief cuyas columnas son los vectores de consumo unitarios para los diversos sectores de una economía.

matriz de covarianza (o **matriz de covarianza muestral**): La matriz S de $p \times p$ definida por $S = (N - 1)^{-1}BB^T$, donde B, es una matriz de observaciones de $p \times N$ en forma de desviación de la media.

matriz de diseño: La matriz X en el modelo lineal $\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\epsilon}$, donde las columnas de X están determinadas de alguna forma por los valores observados de ciertas variables independientes.

matriz de entrada-salida: *Véase* matriz de consumo.

matriz de flexibilidad: Una matriz cuya j-ésima columna da las deflexiones de una viga elástica en puntos específicos cuando se aplica una unidad de fuerza en el j-ésimo punto sobre la viga.

matriz de Gram (de A): La matriz A^TA.

matriz de $\boldsymbol{m \times n}$: Una matriz con m filas y n columnas.

matriz de migración: Una matriz que indica el movimiento porcentual entre dos lugares diferentes, de un periodo al siguiente.

matriz de observaciones: Una matriz de $p \times N$ cuyas columnas son vectores de observaciones, cada columna lista p mediciones realizadas sobre un individuo u objeto en una población dada o en un conjunto.

matriz de proyección (o **matriz de proyección ortogonal**): Una matriz simétrica B tal que $B^2 = B$. Un ejemplo sencillo es $B = \mathbf{v}\mathbf{v}^T$, donde **v** es un vector unitario.

matriz de rango completo: Una matriz de $m \times n$ cuyo rango es el menor de m y n.

matriz de rigidez: La inversa de una matriz de flexibilidad. La j-ésima columna de una matriz de rigidez da las cargas que se deben aplicar en puntos específicos sobre la viga elástica para producir una deflexión unitaria en el j-ésimo punto sobre la viga.

matriz de transferencia: Una matriz A asociada con un circuito eléctrico que tiene terminales de entrada y salida, de tal manera que el vector de salida es igual a A por el vector de entrada.

matriz de Vandermonde: Una matriz V de $n \times n$ o su transpuesta, cuando V tiene la forma

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

matriz diagonal: Una matriz cuadrada cuyas entradas *fuera* de la diagonal principal son cero.

matriz diagonalizable: Una matriz que se puede escribir en la forma factorizada PDP^{-1}, donde D es una matriz diagonal y P es una matriz invertible.

matriz diagonalizable ortogonalmente: Una matriz A que admite una factorización, $A = PDP^{-1}$, con P una matriz ortogonal ($P^{-1} = P^T$) y D diagonal.

matriz elemental: Una matriz invertible que resulta al efectuar una operación elemental de fila sobre una matriz identidad.

matriz escalonada (o **matriz escalonada por filas**): Una matriz rectangular que tiene tres propiedades: **1.** Todas las filas distintas de cero están por arriba de cualquier fila constituida únicamente de ceros. **2.** Cada entrada principal de una fila está en una columna a la derecha de la entrada principal de la fila superior. **3.** En una columna, todas las entradas debajo de una entrada principal valen cero.

matriz escalonada reducida: Una matriz rectangular en forma escalonada que tiene estas propiedades adicionales: la entrada principal en cada fila distinta de cero es 1, y cada uno de esos 1 es la única entrada distinta de cero en su columna respectiva.

matriz estándar (para una transformación lineal T): La matriz A tal que $T(\mathbf{x}) = A\mathbf{x}$ para toda $\mathbf{x}$ en el dominio de T.

matriz estocástica: Una matriz cuadrada cuyas columnas son vectores de probabilidad.

matriz estocástica regular: Una matriz estocástica P tal que alguna potencia matricial P^k solo contiene entradas estrictamente positivas.

matriz identidad (denotada por I o I_n): Una matriz cuadrada con unos sobre la diagonal y ceros en las demás entradas.

matriz indefinida: Una matriz simétrica A tal que $\mathbf{x}^TA\mathbf{x}$ toma valores positivos y negativos.

matriz invertible: Una matriz cuadrada que tiene una inversa.

matriz mal condicionada: Una matriz cuadrada con un número de condición grande (o posiblemente infinito); una matriz que es singular o que puede convertirse en singular si alguna de sus entradas se modifica ligeramente.

matriz negativa definida: Una matriz simétrica A tal que $\mathbf{x}^TA\mathbf{x} < 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

matriz negativa semidefinida: Una matriz simétrica A tal que $\mathbf{x}^TA\mathbf{x} \leq 0$ para toda $\mathbf{x}$.

matriz no singular: Una matriz invertible.

matriz ortogonal: Una matriz cuadrada U invertible tal que $U^{-1} = U^T$.

matriz para T respecto de las bases $\mathcal{B}$ y $\mathcal{C}$: Una matriz M para una transformación lineal $T : V \to W$ con la propiedad de que $[T(\mathbf{x})]_{\mathcal{C}} = M[\mathbf{x}]_{\mathcal{B}}$ para todas las $\mathbf{x}$ en V, donde $\mathcal{B}$ es una base para V, y $\mathcal{C}$ es una base para W. Cuando $W = V$ y $\mathcal{C} = \mathcal{B}$, la matriz M se llama matriz $\mathcal{B}$ para T y se denota como $[T]_{\mathcal{B}}$.

matriz particionada (o **matriz por bloques**): Una matriz cuyas entradas son en sí mismas matrices de tamaños adecuados.

matriz por bloques: *Véase* matriz particionada.

matriz positiva definida: Una matriz simétrica A tal que $\mathbf{x}^TA\mathbf{x} > 0$ para toda $\mathbf{x} \neq \mathbf{0}$.

matriz positiva semidefinida: Una matriz simétrica A tal que $\mathbf{x}^TA\mathbf{x} \geq 0$ para toda $\mathbf{x}$.

matriz simétrica: Una matriz A tal que $A^T = A$.

matriz singular: Una matriz cuadrada que no tiene inversa.

matriz triangular: Una matriz A con ceros arriba o debajo de las entradas diagonales.

matriz triangular inferior: Una matriz con ceros arriba de la diagonal principal.

matriz triangular inferior permutada: Una matriz tal que una permutación de sus filas formará una matriz triangular inferior.

matriz triangular inferior unitaria: Una matriz triangular inferior cuadrada con unos sobre la diagonal principal.

matriz triangular superior: Una matriz U (no necesariamente cuadrada) con ceros debajo de las entradas diagonales $u_{11}, u_{22}, \ldots$

media muestral: El promedio M de un conjunto de vectores, $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N$, dado por $\boldsymbol{M} = (1/N)(\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_N)$.

mejor aproximación: El punto más cercano (en un subespacio) a un vector dado.

método de potencias: Un algoritmo para estimar un valor propio estrictamente dominante de una matriz cuadrada.

método de potencias inverso: Un algoritmo para estimar un valor propio λ de una matriz cuadrada, cuando está disponible una buena estimación inicial de λ.

mínimos cuadrados ponderados: Problemas de mínimos cuadrados con un producto interior ponderado tal como

$$\langle \mathbf{x}, \mathbf{y} \rangle = w_1^2 x_1 y_1 + \cdots + w_n^2 x_n y_n$$

misma dirección (como un vector $\mathbf{v}$): Un vector que es un múltiplo positivo de $\mathbf{v}$.

modelo de entrada-salida: *Véase* modelo de entrada y salida de Leontief.

modelo de intercambio: *Véase* modelo de intercambio de Leontief.

modelo de intercambio (o **cerrado**) **de Leontief:** Un modelo de una economía donde están fijas las entradas y las salidas, y donde se busca un conjunto de precios para los productos de los sectores de tal forma que el ingreso de cada sector iguale los gastos. Esta condición de "equilibrio" se expresa como un sistema de ecuaciones lineales, con los precios como las incógnitas.

modelo de Leontief de entrada y salida (o **ecuación de producción de Leontief**): La ecuación $\mathbf{x} = C\mathbf{x} + \mathbf{d}$, donde $\mathbf{x}$ es la producción, $\mathbf{d}$ es la demanda final, y C es la matriz de consumo (o entrada-salida). La j-ésima columna de C lista las entradas que el sector j consume por unidad de salida.

modelo lineal (en estadística): Cualquier ecuación de la forma $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, donde X y $\mathbf{y}$ son conocidas, y β se elige para minimizar la longitud del **vector residual**, $\boldsymbol{\epsilon}$.

modelo matricial por etapas: Una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde $\mathbf{x}_k$ lista el número de hembras en una población en el tiempo k, con las hembras clasificadas por varias etapas de desarrollo (juveniles, subadultos y adultos).

multiplicación matricial por bloques: La multiplicación fila-columna de matrices particionadas como si las entradas en bloques fueran escalares.

multiplicación por la derecha (por A): Multiplicación de una matriz por la derecha por A.

multiplicación por la izquierda (por A): Multiplicación por A de un vector o una matriz por la izquierda.

multiplicidad algebraica: La multiplicidad de un valor propio como una raíz de la ecuación característica.

múltiplo escalar de u por *c*: El vector $c\mathbf{u}$ que se obtiene al multiplicar cada entrada en $\mathbf{u}$ por c.

N

norma (o **longitud**, de $\mathbf{v}$): El escalar $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.

normalización (de un vector $\mathbf{v}$ diferente de cero): El proceso de crear un vector unitario $\mathbf{u}$ que es un múltiplo positivo de $\mathbf{v}$.

núcleo (de una transformación lineal $T : V \to W$): El conjunto de $\mathbf{x}$ en V tal que $T(\mathbf{x}) = \mathbf{0}$.

número de condición (de A): El cociente σ_1/σ_n, donde σ_1 es el valor singular más grande de A y σ_n es el valor singular más pequeño. El número de condición es $+\infty$ cuando σ_n es cero.

O

operaciones elementales de fila: 1. Remplazo de una fila por la suma de sí misma y un múltiplo de otra fila. **2.** Intercambio de dos filas. **3.** (Escalamiento) Multiplicación de todas las entradas en una fila por una constante distinta de cero.

optimización restringida: El problema de maximizar una cantidad como $\mathbf{x}^TA\mathbf{x}$ o $\|A\mathbf{x}\|$ cuando $\mathbf{x}$ está sujeta a una o más restricciones, tales como $\mathbf{x}^T\mathbf{x} = 1$ o $\mathbf{x}^T\mathbf{v} = 0$.

origen: El vector cero.

ortogonal a *W*: Ortogonal a cada vector en W.

P

parte triangular inferior (de A): Una matriz triangular inferior cuyas entradas sobre la diagonal principal y abajo de ella coinciden con las de A.

perfil (de un conjunto S en $\mathbb{R}^n$): El conjunto de puntos extremos de S.

pesos: Los escalares que se emplean en una combinación lineal.

pivote: Un número distinto de cero que se emplea en una posición pivote para crear ceros mediante operaciones de fila o se cambia en un 1 principal, el cual, a la vez, se utiliza para crear ceros.

plano afín (en $\mathbb{R}^n$): Un traslado de un subespacio de $\mathbb{R}^n$.

plano que pasa por u, v y el origen: Un conjunto cuya ecuación paramétrica es $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (s, t en $\mathbb{R}$), con $\mathbf{u}$ y $\mathbf{v}$ linealmente independientes.

planos afines paralelos: Dos o más planos afines tales que cada uno es un traslado de los otros planos afines.

poliedro: Un polítopo en $\mathbb{R}^3$.

polígono: Un polítopo en $\mathbb{R}^2$.

polinomio característico (de A): $\det(A - \lambda I)$ o, en algunos libros, $\det(\lambda I - A)$.

polinomio de interpolación: Un polinomio cuya gráfica pasa a través de cada punto en un conjunto de puntos de datos en $\mathbb{R}^2$.

polinomio trigonométrico: Una combinación lineal de la función constante 1 y funciones seno y coseno, como cos nt y sen nt.

polítopo: La envolvente convexa de un conjunto finito de puntos en $\mathbb{R}^n$ (un tipo especial de conjunto convexo compacto).

posición estándar: La posición de la gráfica de una ecuación $\mathbf{x}^TA\mathbf{x} = c$, cuando A es una matriz diagonal.

posición pivote: Una posición en una matriz A que corresponde a una entrada principal en cada forma escalonada de A.

precios de equilibrio: Un conjunto de precios para la producción total de los diversos sectores en una economía, de tal manera que el ingreso de cada sector logre equilibrar sus gastos.

pregunta de existencia: Pregunta: "¿Existe una solución al sistema?". Es decir, "¿el sistema es consistente?". Además, "¿existe solución de $A\mathbf{x} = \mathbf{b}$ para *todas* las posibles $\mathbf{b}$?".

pregunta de unicidad: Pregunta: "¿Si existe una solución de un sistema, esa solución es la única?".

problema de mínimos cuadrados general: A partir de una matriz A de $m \times n$, y un vector $\mathbf{b}$ en $\mathbb{R}^m$, se encuentra $\hat{\mathbf{x}}$ en $\mathbb{R}^n$ tal que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

proceso de Gram-Schmidt: Un algoritmo para producir una base ortogonal u ortonormal para un subespacio que se genera por un conjunto dado de vectores.

producto *A*x: La combinación lineal de las columnas de A empleando como pesos las entradas correspondientes en $\mathbf{x}$.

producto cruzado: Un término cx_ix_j en una forma cuadrática, con $i \neq j$.

producto exterior: Un producto matricial $\mathbf{u}\mathbf{v}^T$ donde $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$ vistos como matrices de $n \times 1$. (El símbolo de transpuesta está sobre el "exterior" de los símbolos $\mathbf{u}$ y $\mathbf{v}$).

producto interior: El escalar $\mathbf{u}^T\mathbf{v}$, generalmente escrito como $\mathbf{u}\cdot\mathbf{v}$, donde $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^n$ vistos como matrices de $n \times 1$. También se llama el **producto punto** de $\mathbf{u}$ y $\mathbf{v}$. En general, una función sobre un espacio vectorial que asigna a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ un número $\langle \mathbf{u}, \mathbf{v}\rangle$, sujeto a ciertos axiomas. Véase la sección 6.7.

producto punto: *Véase* producto interior.

proyección ortogonal de y sobre u (o sobre la recta que pasa por $\mathbf{u}$ y por el origen, para $\mathbf{u} \neq \mathbf{0}$): El vector $\hat{\mathbf{y}}$ definido por $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$.

proyección ortogonal de y sobre *W*: El vector único $\hat{\mathbf{y}}$ en W tal que $\mathbf{y} - \hat{\mathbf{y}}$ es ortogonal a W. Notación: $\hat{\mathbf{y}} = \mathrm{proy}_W\,\mathbf{y}$.

punto espiral (de un sistema dinámico en $\mathbb{R}^2$): El origen cuando las trayectorias caen en espiral hacia $\mathbf{0}$.

punto extremo (de un conjunto convexo S): Un punto $\mathbf{p}$ en S tal que $\mathbf{p}$ no está en el interior de cualquier segmento de recta contenido en S. (Es decir, si $\mathbf{x}$, $\mathbf{y}$ están en S, y $\mathbf{p}$ está sobre el segmento de recta $\overline{\mathbf{xy}}$, entonces $\mathbf{p} = \mathbf{x}$ o $\mathbf{p} = \mathbf{y}$).

punto frontera de un conjunto S en $\mathbb{R}^n$: Un punto $\mathbf{p}$ tal que cada bola abierta en $\mathbb{R}^n$ centrada en $\mathbf{p}$ se interseca tanto con S como con el complemento de S.

punto interior (de un conjunto S en $\mathbb{R}^n$): Un punto $\mathbf{p}$ en S tal que para alguna $\delta > 0$, la bola abierta $\mathbf{B}(\mathbf{p}, \delta)$ centrada en $\mathbf{p}$ está contenida en S.

punto silla (de un sistema dinámico en $\mathbb{R}^2$): El origen cuando algunas trayectorias son atraídas hacia $\mathbf{0}$ y otras trayectorias son repelidas desde $\mathbf{0}$.

R

rango (de una matriz A): La dimensión del espacio columna de A, que se denota como rango A.

rango (de una transformación lineal T): El conjunto de todos los vectores de la forma $T(\mathbf{x})$ para alguna $\mathbf{x}$ en el dominio de T.

Re x: El vector en $\mathbb{R}^n$ formado de las partes reales de las entradas de un vector $\mathbf{x}$ en $\mathbb{C}^n$.

recta de mínimos cuadrados: La recta $y = \hat{\beta}_0 + \hat{\beta}_1 x$ que minimiza el error de mínimos cuadrados en la ecuación $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$.

recta que pasa por p paralela a v: El conjunto $\{\mathbf{p} + t\mathbf{v}\colon t \text{ en } \mathbb{R}\}$.

red en escalera: Una red eléctrica ensamblada que conecta en serie dos o más circuitos eléctricos.

reflexión de Householder: Una transformación $\mathbf{x} \mapsto Q\mathbf{x}$, donde $Q = I - 2\mathbf{u}\mathbf{u}^T$ y $\mathbf{u}$ es un vector unitario ($\mathbf{u}^T\mathbf{u} = 1$).

regla de Cramer: Una fórmula para cada entrada en la solución $\mathbf{x}$ de la ecuación $A\mathbf{x} = \mathbf{b}$ cuando A es una matriz invertible.

regla del paralelogramo para la adición: Una interpretación geométrica de la suma de dos vectores $\mathbf{u}$, $\mathbf{v}$ como la diagonal del paralelogramo determinado por $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{0}$.

regla fila-columna: La regla para calcular un producto AB en el cual la entrada (i, j) de AB es la suma de los productos de las entradas correspondientes de la fila i de A y columna j de B.

regla fila-vector para calcular $A\mathbf{x}$: Regla para calcular un producto $A\mathbf{x}$ en el cual la i-ésima entrada de $A\mathbf{x}$ es la suma de los productos de las entradas correspondientes de la fila i de A y del vector $\mathbf{x}$.

regresión múltiple: Un modelo lineal que implica diversas variables independientes y una variable dependiente.

relación de dependencia afín: Una ecuación de la forma $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$, donde los pesos $c_1, \ldots, c_p$ no todos son cero y $c_1 + \cdots + c_p = 0$.

relación de dependencia lineal: Una ecuación vectorial homogénea donde se especifican todos los pesos y al menos un peso es distinto de cero.

relación de recurrencia: *Véase* ecuación en diferencias.

remplazo de filas: Una operación elemental de fila que sustituye una fila de una matriz por la suma de la fila y un múltiplo de otra fila.

repulsor (de un sistema dinámico en $\mathbb{R}^2$): El origen cuando todas las trayectorias excepto la secuencia cero constante tienden a alejarse de $\mathbf{0}$.

resta vectorial: Se calcula $\mathbf{u} + (-1)\mathbf{v}$ y se escribe el resultado como $\mathbf{u} - \mathbf{v}$.

rotación de Givens: Una transformación lineal de $\mathbb{R}^n$ a $\mathbb{R}^n$ empleada en programas computacionales para crear entradas cero en un vector (por lo general, una columna de una matriz).

S

señal (o **señal discreta de tiempo**): Una secuencia doblemente infinita de números, $\{y_k\}$; una función definida sobre los enteros; pertenece al espacio vectorial $\mathbb{S}$.

serie de Fourier: Una serie infinita que converge a una función en el espacio con producto interior $C[0, 2\pi]$, con el producto interior dado por una integral definida.

seudoinversa (de A): La matriz $VD^{-1}U^T$, cuando UDV^T es una descomposición en valores singulares reducida de A.

simplejo: La envolvente convexa de un conjunto finito de vectores en $\mathbb{R}^n$ afínmente independientes.

sistema de ecuaciones lineales (o **sistema lineal**): Una colección de una o más ecuaciones lineales que implican el mismo conjunto de variables, por ejemplo, $x_1, \ldots, x_n$.

sistema desacoplado: Una ecuación en diferencias $\mathbf{y}_{k+1} = A\mathbf{y}_k$, o una ecuación diferencial $\mathbf{y}'(t) = A\mathbf{y}(t)$, en la cual A es una matriz diagonal. La evolución discreta de cada entrada en $\mathbf{y}_k$ (como una función de k), o la evolución continua de cada entrada en la función vectorial $\mathbf{y}(t)$, no sufren alteración por lo que ocurre con las otras entradas conforme $k \to \infty$ o $t \to \infty$.

sistema dinámico: *Véase* sistema dinámico lineal discreto.

sistema dinámico lineal discreto: Una ecuación en diferencias de la forma $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que describe los cambios en un sistema (por lo general, un sistema físico) conforme pasa el tiempo. El sistema físico se mide en tiempos discretos, cuando $k = 0, 1, 2, \ldots$, y el **estado** del sistema al tiempo k es un vector $\mathbf{x}_k$ cuyas entradas indican ciertos hechos de interés sobre el sistema.

sistema lineal: Una colección de una o más ecuaciones lineales que implica las mismas variables, por ejemplo, $x_1, \ldots, x_n$.

sistema lineal consistente: Un sistema lineal con al menos una solución.

sistema lineal inconsistente: Un sistema lineal que carece de solución.

sistema sobredeterminado: Un sistema de ecuaciones con más ecuaciones que incógnitas.

sistema subdeterminado: Un sistema de ecuaciones con menos ecuaciones que incógnitas.

sistemas (lineales) equivalentes: Sistemas lineales con el mismo conjunto solución.

sólido regular: Uno de los cinco posibles poliedros regulares en $\mathbb{R}^3$: el tetraedro (4 caras triangulares iguales), el cubo (6 caras cuadradas), el octaedro (8 caras triangulares iguales), el dodecaedro (12 caras pentagonales iguales), y el icosaedro (20 caras triangulares iguales).

solución (de un sistema lineal que implica las variables $x_1, \dots, x_n$): Una lista $(s_1, s_2, \dots, s_n)$ de números que hacen que cada ecuación en el sistema sea un enunciado verdadero cuando los valores $s_1, \dots, s_n$ se sustituyen para $x_1, \dots, x_n$, respectivamente.

solución de mínimos cuadrados (de $A\mathbf{x} = \mathbf{b}$): Un vector $\hat{\mathbf{x}}$ tal que $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

solución general (de un sistema lineal): Una descripción paramétrica de un conjunto solución que expresa las variables básicas en términos de las variables libres (los parámetros), si las hay. Después de la sección 1.5, la descripción paramétrica se escribe en forma vectorial.

solución no trivial: Una solución distinta de cero de una ecuación homogénea o de un sistema de ecuaciones homogéneas.

solución trivial: La solución $\mathbf{x} = \mathbf{0}$ de una ecuación homogénea $A\mathbf{x} = \mathbf{0}$.

subconjunto propio de un conjunto *S*: Un subconjunto de S que no es igual a S mismo.

subespacio: Un subconjunto H de algún espacio vectorial V tal que H tiene estas propiedades: **1.** El vector cero de V está en H; **2.** H es cerrado bajo suma vectorial; y **3.** H es cerrado bajo multiplicación por escalares.

subespacio cero: El subespacio $\{\mathbf{0}\}$ que consta solamente del vector cero.

subespacio invariante (para A): Un subespacio H tal que $A\mathbf{x}$ está en H siempre que $\mathbf{x}$ se encuentre en H.

subespacio propio: Cualquier subespacio de un espacio vectorial V diferente de V.

subespacios fundamentales (determinados por A): El espacio nulo y el espacio columna de A, y el espacio nulo y el espacio columna de A^T, con Col A^T comúnmente llamado el espacio fila de A.

submatriz (de A): Cualquier matriz que se obtiene al eliminar algunas filas y/o columnas de A; también, la propia A.

suma de columna: La suma de las entradas en una columna de una matriz.

suma de fila: La suma de las entradas en una fila de una matriz.

suma vectorial: Operación que consiste en sumar vectores mediante la adición de las correspondientes entradas.

sustitución regresiva (con notación matricial): La fase regresiva de reducción por filas de una matriz aumentada que transforma una matriz escalonada en una matriz escalonada reducida; se utiliza para obtener la solución (o soluciones) de un sistema de ecuaciones lineales.

T

tamaño (de una matriz): Dos números, escritos en la forma $m \times n$, que especifican el número de filas (m) y de columnas (n) en la matriz.

tetraedro: Un objeto sólido tridimensional acotado por cuatro caras triangulares iguales, con tres caras que se encuentran en cada vértice.

transformación (o **función**, o **mapeo**) ***T* de $\mathbb{R}^n$ a $\mathbb{R}^m$:** Una regla que asigna a cada vector $\mathbf{x}$ en $\mathbb{R}^n$ un único vector $T(\mathbf{x})$ en $\mathbb{R}^m$. Notación: $T : \mathbb{R}^n \to \mathbb{R}^m$. Además, $T : \mathrm{V} \to W$ denota una regla que asigna a cada $\mathbf{x}$ en V un único vector $T(\mathbf{x})$ en W.

transformación afín: Un mapeo $T : \mathbb{R}^n \to \mathbb{R}^m$ de la forma $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, con una matriz A de $m \times n$ y $\mathbf{b}$ en $\mathbb{R}^m$.

transformación de semejanza: Una transformación que cambia A a la forma $P^{-1}AP$.

transformación lineal invertible: Una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que existe una función $S : \mathbb{R}^n \to \mathbb{R}^n$ que satisface tanto $T(S(\mathbf{x})) = \mathbf{x}$ como $S(T(\mathbf{x})) = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

transformación lineal *T* (de un espacio vectorial V a un espacio vectorial W): Una regla T que asigna a cada vector $\mathbf{x}$ en V un único vector $T(\mathbf{x})$ en W, tal que **i.** $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para toda $\mathbf{u}$, $\mathbf{v}$ en V, y **ii.** $T(c\mathbf{u}) = cT(\mathbf{u})$ para toda $\mathbf{u}$ en V y todos los escalares c. Notación: $T : V \to W$; también, $\mathbf{x} \mapsto A\mathbf{x}$ cuando $T : \mathbb{R}^n \to \mathbb{R}^m$ y A es la matriz estándar para T.

transformación matricial: Un mapeo $\mathbf{x} \mapsto A\mathbf{x}$, donde A es una matriz de $n \times n$ y $\mathbf{x}$ representa cualquier vector en $\mathbb{R}^n$.

transpuesta (de A): Una matriz A^T de $n \times m$, cuyas columnas son las filas correspondientes de la matriz A de $m \times n$.

traslación (por un vector $\mathbf{p}$): La operación de sumar $\mathbf{p}$ a un vector o a cada vector en un conjunto dado.

trayectoria: La gráfica de una solución $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ de un sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$; los puntos a menudo están conectados por una delgada curva, para facilitar la visualización de la trayectoria. Además, la gráfica $\mathbf{x}(t)$ para $t \geq 0$, cuando $\mathbf{x}(t)$ es una solución de una ecuación diferencial $\mathbf{x}'(t) = A\mathbf{x}(t)$.

traza (de una matriz cuadrada A): La suma de las entradas diagonales en A; se denota como tr A.

triangular superior en bloques (matriz): Una matriz particionada $A = [A_{ij}]$ tal que cada bloque A_{ij} es una matriz cero para $i > j$.

V

valor propio (de A): Un escalar λ tal que la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ tiene una solución para algún vector $\mathbf{x}$ distinto de cero.

valor propio complejo: Una raíz no real de la ecuación característica de una matriz de $n \times n$.

valor propio estrictamente dominante: Un valor propio λ_1 de una matriz A con la propiedad de que $|\lambda_1| > |\lambda_k|$ para todos los otros valor propios λ_k de A.

valores singulares (de A): Las raíces cuadradas (positivas) de los valores propios de A^TA, arregladas en orden decreciente de magnitud.

variable básica: Una variable en un sistema lineal que corresponde a una columna pivote en la matriz de coeficientes.

variable libre: Cualquier variable en un sistema lineal que no es una variable básica.

variables no correlacionadas: Cualesquiera dos variables x_i y x_j (con $i \neq j$) que incluyen las i-ésima y j-ésima coordenadas de los vectores de observación en una matriz de observación, de tal forma que la covarianza s_{ij} es cero.

varianza (de una variable x_j): La entrada diagonal s_{jj} en la matriz de covarianza S para una matriz de observaciones, donde x_j varía sobre la j-ésima coordenada de los vectores de observaciones.

varianza total: La traza de la matriz de covarianza S de una matriz de observaciones.

vector: Una lista de números; una matriz con una sola columna. En general, cualquier elemento de un espacio vectorial.

vector cero: El vector único, que se denota por $\mathbf{0}$, tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$ para toda $\mathbf{u}$. En $\mathbb{R}^n$, $\mathbf{0}$ es el vector cuyas entradas son todas iguales a cero.

vector columna: Una matriz con una sola columna, o una sola columna de una matriz que tiene varias columnas.

vector coordenado de x respecto de $\mathcal{B}$: El vector $[\mathbf{x}]_{\mathcal{B}}$ cuyas entradas son las coordenadas de $\mathbf{x}$ respecto de la base $\mathcal{B}$.

vector de consumo unitario: Un vector columna en el modelo de entrada y salida de Leontief, que lista las entradas que un sector necesita por cada unidad de su salida; una columna de la matriz de consumo.

vector de demanda final (o **cuenta de demandas finales**): El vector $\mathbf{d}$ en el modelo de entrada y salida de Leontief que lista los valores en dólares de los bienes y servicios requeridos por varios sectores de la parte no productiva de la economía. El vector $\mathbf{d}$ puede representar la demanda de los consumidores, consumo gubernamental, excedente de producción, exportaciones, u otras demandas externas.

vector de equilibrio: *Véase* vector de estado estable.

vector de estado: Un vector de probabilidad. En general, un vector que describe el "estado" de un sistema físico, frecuentemente en conexión con una ecuación en diferencias $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

vector de estado estable (para una matriz estocástica P): Un vector de probabilidad $\mathbf{q}$ tal que $P\mathbf{q} = \mathbf{q}$.

vector de observaciones: El vector $\mathbf{y}$ en el modelo lineal $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$, donde las entradas en $\mathbf{y}$ son los valores observados de una variable dependiente.

vector de parámetros: El vector incógnita β en el modelo lineal $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$.

vector de probabilidad: Un vector en $\mathbb{R}^n$ cuyas entradas son no negativas y suman uno.

vector de producción: El vector en el modelo entrada y salida de Leontief que lista las cantidades que producirán los diversos sectores de una economía.

vector fila: Una matriz con una sola fila, o bien, una fila de una matriz que tiene varias filas.

vector normal (a un subespacio V de $\mathbb{R}^n$): Un vector $\mathbf{n}$ en $\mathbb{R}^n$ tal que $\mathbf{n}\cdot\mathbf{x} = 0$ para toda $\mathbf{x}$ en V.

vector propio (de A): Un vector $\mathbf{x}$ *distinto de cero* tal que $A\mathbf{x} = \lambda\mathbf{x}$ para algún escalar λ.

vector propio complejo: Un vector distinto de cero $\mathbf{x}$ en $\mathbb{C}^n$ tal que $A\mathbf{x} = \lambda\mathbf{x}$, donde A es una matriz de $n \times n$, y λ es un valor propio complejo.

vector residual: La cantidad $\boldsymbol{\epsilon}$ que aparece en el modelo lineal general: $\mathbf{y} = X\beta + \boldsymbol{\epsilon}$; es decir, $\boldsymbol{\epsilon} = \mathbf{y} - X\beta$, la diferencia entre los valores observados y los valores predichos (de y).

vector unitario: Un vector $\mathbf{v}$ tal que $\|\mathbf{v}\| = 1$.

vectores iguales: Vectores en $\mathbb{R}^n$ cuyas entradas correspondientes son las mismas.

vectores linealmente independientes: Un conjunto indexado $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ con la propiedad de que la ecuación vectorial $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$ *solamente* tiene la solución trivial, $c_1 = \cdots = c_p = 0$.

vectores singulares derechos (de A): Las columnas de V en la descomposición de valores singulares $A = U\Sigma V^T$.

vectores singulares izquierdos (de A): Las columnas de U en la descomposición en valores singulares de $A = U\Sigma V^T$.

Respuestas a los ejercicios con numeración impar

Capítulo 1

Sección 1.1, página 10

1. La solución es $(x_1, x_2) = (-8, 3)$, o simplemente, $(-8, 3)$.

3. $(2, 1)$

5. Remplazar la Fila 2 por su suma con -4 por la Fila 3, y después sustituya la Fila 1 por su suma con 3 por la Fila 3.

7. El conjunto solución es vacío.

9. $(16, 21, 14, 4)$

11. Inconsistente

13. $(5, 3, -1)$

15. Inconsistente

17. Los cálculos indican que el sistema es inconsistente, de manera que las tres rectas no tienen punto en común.

19. $h \neq 2$

21. Para toda h.

23. *Sugerencia:* Marque un enunciado como verdadero solamente si *siempre* es verdadero. Dar aquí las respuestas sería frustrar la finalidad de las preguntas falso-verdadero, las cuales pretenden ayudarle a leer con cuidado el libro.

25. $k - 2g + h = 0$

27. La reducción por filas de

$$\begin{bmatrix} a & b & f \\ c & d & g \end{bmatrix} \text{ a } \begin{bmatrix} a & b & f \\ 0 & d - b(\frac{c}{a}) & g - f(\frac{c}{a}) \end{bmatrix}$$

indica que $d - b(\frac{c}{a})$ debe ser diferente de cero, ya que f y g son arbitrarias. De otra forma, para algunas elecciones de f y g la segunda fila podría corresponder a una ecuación de la forma $0 = q$, donde q es diferente de cero. Por lo tanto, $ad \neq bc$.

29. Intercambie las Filas 1 y 3; intercambie las Filas 1 y 3.

31. Remplace la Fila 3 por la Fila 3 $+ (-4)$Fila 1; remplace la Fila 3 por la Fila 3 $+ (4)$Fila 1.

33. *Sugerencia:* Revise el problema de práctica 1 y después *escriba* una solución.

Sección 1.2, página 21

1. Forma escalonada reducida: a y b. Forma escalonada: d. No está en forma escalonada: c.

3. $$\begin{bmatrix} 1 & 2 & 0 & -8 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Columnas pivote 1 y 3: $$\begin{bmatrix} 1 & 2 & 4 & 8 \\ 2 & 4 & 6 & 8 \\ 3 & 6 & 9 & 12 \end{bmatrix}$$

5. $\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \end{bmatrix}$, $\begin{bmatrix} \blacksquare & * \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \end{bmatrix}$

7. $\begin{cases} x_1 = -5 - 3x_2 \\ x_2 \text{ es libre.} \\ x_3 = 3 \end{cases}$

9. $\begin{cases} x_1 = 3 + 2x_3 \\ x_2 = 3 + 2x_3 \\ x_3 \text{ es libre.} \end{cases}$

11. $\begin{cases} x_1 = \frac{2}{3}x_2 - \frac{4}{3}x_3 \\ x_2 \text{ es libre.} \\ x_3 \text{ es libre.} \end{cases}$

13. $\begin{cases} x_1 = 5 + 3x_5 \\ x_2 = 1 + 4x_5 \\ x_3 \text{ es libre.} \\ x_4 = 4 - 9x_5 \\ x_5 \text{ es libre.} \end{cases}$

15. ***a***) Consistente, con muchas soluciones.

b) Consistente, con muchas soluciones.

17. Para toda h.

19. ***a***) Inconsistente cuando $h = 2$ y $k \neq 8$.

b) Solución única cuando $h \neq 2$.

c) Muchas soluciones cuando $h = 2$ y $k = 8$.

21. *Sugerencia:* Lea el texto con cuidado y escriba las respuestas. Recuerde, un enunciado es verdadero solo si es cierto en todos los casos.

23. Como hay cuatro pivotes (uno en cada columna de la matriz de coeficientes), la matriz aumentada se debe reducir a la forma

$$\begin{bmatrix} 1 & 0 & 0 & 0 & a \\ 0 & 1 & 0 & 0 & b \\ 0 & 0 & 1 & 0 & c \\ 0 & 0 & 0 & 1 & d \end{bmatrix}$$

y así,

$$\begin{aligned} x_1 \qquad\qquad\quad &= a \\ x_2 \qquad\quad &= b \\ x_3 \quad &= c \\ x_4 &= d \end{aligned}$$

Sin importar los valores de a, b, c y d, la solución existe y es única.

25. Si la matriz de coeficientes tiene una posición pivote en cada fila, entonces existe una posición pivote en la fila inferior, y no hay lugar para un pivote en la columna aumentada. Así, de acuerdo con el teorema 2, el sistema es consistente.

27. Si un sistema lineal es consistente, entonces la solución es única si y solo si *cada columna de la matriz de coeficientes es una columna pivote; de lo contrario, existe una infinidad de soluciones.*

29. Un sistema subdeterminado siempre tiene más variables que ecuaciones. No pueden existir más variables básicas que ecuaciones, así que al menos una variable es libre. A tal variable se le puede asignar una infinidad de valores distintos. Si el sistema es consistente, entonces cada valor diferente de una variable libre producirá una solución distinta.

31. Sí, un sistema de ecuaciones lineales con más ecuaciones que incógnitas puede ser consistente. El siguiente sistema tiene una solución ($x_1 = x_2 = 1$):

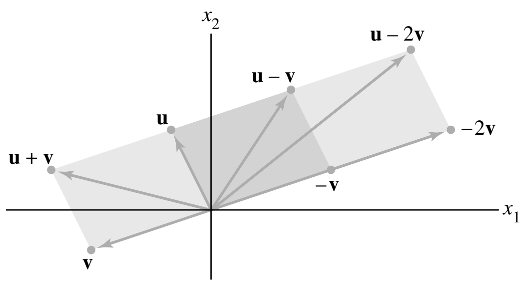
$$\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 0 \\ 3x_1 + 2x_2 &= 5 \end{aligned}$$

33. $p(t) = 1 + 3t + 2t^2$

Sección 1.3, página 32

1. $\begin{bmatrix} -4 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}$

3.



5. $x_1 \begin{bmatrix} 3 \\ -2 \\ 8 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ 0 \\ -9 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$,

$\begin{bmatrix} 3x_1 \\ -2x_1 \\ 8x_1 \end{bmatrix} + \begin{bmatrix} 5x_2 \\ 0 \\ -9x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$,

$\begin{bmatrix} 3x_1 + 5x_2 \\ -2x_1 \\ 8x_1 - 9x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \\ 8 \end{bmatrix}$

$$\begin{aligned} 3x_1 + 5x_2 &= 2 \\ -2x_1 &= -3 \\ 8x_1 - 9x_2 &= 8 \end{aligned}$$

En general, no se muestran los pasos intermedios.

7. $\mathbf{a} = \mathbf{u} - 2\mathbf{v}$, $\mathbf{b} = 2\mathbf{u} - 2\mathbf{v}$, $\mathbf{c} = 2\mathbf{u} - 3.5\mathbf{v}$, $\mathbf{d} = 3\mathbf{u} - 4\mathbf{v}$
Sí, cada vector en $\mathbb{R}^2$ es una combinación lineal de $\mathbf{u}$ y $\mathbf{v}$.

9. $x_1 \begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 6 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 5 \\ -1 \\ -8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

11. No, $\mathbf{b}$ *no* es una combinación lineal de $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.

13. No, $\mathbf{b}$ *no* es una combinación lineal de las columnas de A.

15. $h = 3$.

17. Desde luego, son aceptables los pesos que no son enteros, pero algunas elecciones sencillas son $0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 = \mathbf{0}$, y

$$1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \quad 0 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$$

$$1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix}, \quad 1 \cdot \mathbf{v}_1 - 1 \cdot \mathbf{v}_2 = \begin{bmatrix} 7 \\ 1 \\ 1 \end{bmatrix}$$

19. Gen $\{\mathbf{v}_1, \mathbf{v}_2\}$ es el conjunto de puntos sobre la recta que pasa por $\mathbf{v}_1$ y $\mathbf{0}$, ya que $\mathbf{v}_2$ es un múltiplo de $\mathbf{v}_1$.

21. *Sugerencia:* Demuestre que $\begin{bmatrix} 2 & 2 & h \\ -1 & 1 & k \end{bmatrix}$ es consistente para toda h y k. Explique qué indica este cálculo acerca de Gen $\{\mathbf{u}, \mathbf{v}\}$.

23. *Sugerencia:* Lea con cuidado toda la sección. Ponga especial atención a las definiciones y los enunciados de teoremas, y también a cualquier observación que aparezca antes o después de ellos.

25. ***a)*** No, tres. ***b)*** Sí, un número infinito
c) $\mathbf{a}_1 = 1 \cdot \mathbf{a}_1 + 0 \cdot \mathbf{a}_2 + 0 \cdot \mathbf{a}_3$

27. ***a)*** $5\mathbf{v}_1$ es la salida de 5 días de operación de la mina #1.
b) La salida total es $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$, así que x_1 y x_2 deberían satisfacer $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 = \begin{bmatrix} 240 \\ 2824 \end{bmatrix}$.
c) [**M**] 1.73 días para la mina #1 y 4.70 días para la mina #2.

29. $(17/14, -34/14, 16/14) = (17/14, -17/7, 8/7)$

31. ***a)*** $\begin{bmatrix} 10/3 \\ 2 \end{bmatrix}$
b) Sume 3.5 g en (0, 1), sume 0.5 g en (8, 1), y sume 2 g en (2, 4).

33. *Sugerencia:* Revise el problema de práctica 1 y después *escriba* una solución.

Sección 1.4, página 40

1. El producto no está definido porque el número de columnas (2) en la matriz de 3×2 no concuerda con el número de entradas (3) en el vector.

3. ***a)*** $A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = -2 \cdot \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix}$

$$= \begin{bmatrix} -2 \\ 6 \\ -2 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \\ 18 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \\ 16 \end{bmatrix}$$

b) $A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 6 \end{bmatrix}\begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1\cdot(-2)+2\cdot(3) \\ (-3)\cdot(-2)+1\cdot(3) \\ 1\cdot(-2)+6\cdot(3) \end{bmatrix}$

$= \begin{bmatrix} 4 \\ 9 \\ 16 \end{bmatrix}.$

Presente su trabajo hasta aquí y también para los ejercicios 4 a 6, pero de aquí en adelante efectúe los cálculos mentalmente.

5. $2\cdot\begin{bmatrix} 1 \\ -2 \end{bmatrix} - 1\cdot\begin{bmatrix} 2 \\ -3 \end{bmatrix} + 1\cdot\begin{bmatrix} -3 \\ 1 \end{bmatrix} - 1\cdot\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$

7. $\begin{bmatrix} 4 & -5 & 7 \\ -1 & 3 & -8 \\ 7 & -5 & 0 \\ -4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -8 \\ 0 \\ -7 \end{bmatrix}$

9. $x_1\begin{bmatrix} 5 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 2 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$ y

$\begin{bmatrix} 5 & 1 & -3 \\ 0 & 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 3 & -4 & -2 \\ 1 & 5 & 2 & 4 \\ -3 & -7 & 6 & 12 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -11 \\ 3 \\ 0 \end{bmatrix}$

13. Sí. (Justifique su respuesta).

15. La ecuación $A\mathbf{x} = \mathbf{b}$ no es consistente cuando $3b_1 + b_2$ es diferente de cero. (Muestre su trabajo). El conjunto de vectores $\mathbf{b}$ para los cuales la ecuación *es* consistente es una recta que pasa por el origen [el conjunto de todos los puntos (b_1, b_2) que satisfacen $b_2 = -3b_1$].

17. Solo tres filas contienen una posición pivote. De acuerdo con el teorema 4, la ecuación $A\mathbf{x} = \mathbf{b}$ *no* tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^4$.

19. El trabajo del ejercicio 17 indica que el enunciado *d*) en el teorema 4 es falso. De manera que los cuatro enunciados del teorema 4 son falsos. Por lo tanto, no todos los vectores en $\mathbb{R}^4$ se pueden escribir como una combinación lineal de las columnas de A. Además, las columnas de A *no* generan a $\mathbb{R}^4$.

21. La matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3]$ no tiene un pivote en cada fila, de manera que, de acuerdo con el teorema 4, las columnas de la matriz no generan a $\mathbb{R}^4$. Es decir, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ no genera a $\mathbb{R}^4$.

23. *Sugerencia:* Lea el texto con cuidado y marque cada enunciado como verdadero o falso. Varias partes de los ejercicios 29 y 30 son *implicaciones* de la forma

"Si ⟨enunciado 1⟩, entonces, ⟨enunciado 2⟩"

o, de manera equivalente,

"⟨enunciado 2⟩, si ⟨enunciado 1⟩".

Marque tal implicación como verdadera si el ‹enunciado 2› es válido en todos los casos cuando el ‹enunciado 1› es verdadero.

25. $c_1 = -3, c_2 = -1, c_3 = 2$.

27. La ecuación matricial se puede escribir como $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + c_4\mathbf{v}_4 + c_5\mathbf{v}_5 = \mathbf{v}_6$, donde $c_1 = -3$, $c_2 = 1, c_3 = 2, c_4 = -1, c_5 = 2$, y

$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 8 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -4 \\ 1 \end{bmatrix},$

$\mathbf{v}_4 = \begin{bmatrix} 9 \\ -2 \end{bmatrix}, \quad \mathbf{v}_5 = \begin{bmatrix} 7 \\ -4 \end{bmatrix}, \quad \mathbf{v}_6 = \begin{bmatrix} 11 \\ -11 \end{bmatrix}$

29. *Sugerencia:* Inicie con cualquier matriz B de 3×3 en forma escalonada que tenga tres posiciones pivote.

33. *Sugerencia:* Pregúntese cuántas posiciones pivote tiene A y por qué.

35. Suponga que $\mathbf{y}$ y $\mathbf{z}$ satisfacen $A\mathbf{y} = \mathbf{z}$. Entonces, $5\mathbf{z} = 5A\mathbf{y}$. De acuerdo con el teorema 5(b), $5A\mathbf{y} = A(5\mathbf{y})$. Así, $5\mathbf{z} = A(5\mathbf{y})$, lo que muestra que $5\mathbf{y}$ es una solución de $A\mathbf{x} = 5\mathbf{z}$. Por lo tanto, la ecuación $A\mathbf{x} = 5\mathbf{z}$ es consistente.

37. [**M**] Las columnas no generan a $\mathbb{R}^4$.

39. [**M**] Las columnas generan a $\mathbb{R}^4$.

41. [**M**] Elimine la columna 4 de la matriz del ejercicio 39. También es posible eliminar la columna 3 en vez de la columna 4.

Sección 1.5, página 47

1. El sistema tiene una solución no trivial porque existe una variable libre, x_3.

3. El sistema tiene una solución no trivial porque hay una variable libre, x_3.

5. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

7. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3\begin{bmatrix} -9 \\ 4 \\ 1 \\ 0 \end{bmatrix} + x_4\begin{bmatrix} 8 \\ -5 \\ 0 \\ 1 \end{bmatrix}$

9. $\mathbf{x} = x_2\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

11. *Sugerencia:* El sistema obtenido a partir de la forma escalonada *reducida* es

$$\begin{aligned} x_1 - 4x_2 \qquad\qquad + 5x_6 &= 0 \\ x_3 \qquad - x_6 &= 0 \\ x_5 - 4x_6 &= 0 \\ 0 &= 0 \end{aligned}$$

Las variables básicas son x_1, x_3 y x_5. Las restantes variables son libres.

13. $\mathbf{x} = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix} = \mathbf{p} + x_3\mathbf{q}$. Geométricamente, el conjunto solución es la recta que pasa por $\begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix}$ paralela a $\begin{bmatrix} 4 \\ -7 \\ 1 \end{bmatrix}$.

15. Sean $\mathbf{u} = \begin{bmatrix} -5 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$. La solución de la ecuación homogénea es $\mathbf{x} = x_2\mathbf{u} + x_3\mathbf{v}$, el plano que pasa por el origen generado por **u** y **v**. El conjunto solución del sistema no homogéneo es $\mathbf{x} = \mathbf{p} + x_2\mathbf{u} + x_3\mathbf{v}$, el plano que pasa por **p** paralelo al conjunto solución de la ecuación homogénea.

17. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -4 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$. El conjunto solución es la recta que pasa por $\begin{bmatrix} 8 \\ -4 \\ 0 \end{bmatrix}$, paralela a la recta que es el conjunto solución del sistema homogéneo del ejercicio 5.

19. $\mathbf{x} = \mathbf{a} + t\mathbf{b}$, donde t representa un parámetro, o

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ 3 \end{bmatrix} \text{ o } \begin{cases} x_1 = -2 - 5t \\ x_2 = 3t \end{cases}$$

21. $\mathbf{x} = \mathbf{p} + t(\mathbf{q} - \mathbf{p}) = \begin{bmatrix} 3 \\ -3 \end{bmatrix} + t\begin{bmatrix} 1 \\ 4 \end{bmatrix}$

25. ***a***) $A\mathbf{w} = A(\mathbf{p} + \mathbf{v}_h) = A\mathbf{p} + A\mathbf{v}_h = \mathbf{b} + \mathbf{0} = \mathbf{b}$

b) $A\mathbf{v}_h = A(\mathbf{w} - \mathbf{p}) = A\mathbf{w} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}$

27. (*Argumento geométrico con base en el teorema 6*) Como la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente, entonces, de acuerdo con el teorema 6, su conjunto solución se obtiene trasladando el conjunto solución de $A\mathbf{x} = \mathbf{0}$. Así, el conjunto solución de $A\mathbf{x} = \mathbf{b}$ es un solo vector si y solo si el conjunto solución de $A\mathbf{x} = \mathbf{0}$ es un solo vector, y eso ocurre si y solo si $A\mathbf{x} = \mathbf{0}$ solamente tiene la solución trivial.

(*Demostración con variables libres*) Si $A\mathbf{x} = \mathbf{b}$ tiene una solución, entonces la solución es única si y solo si no hay variables libres en el sistema de ecuaciones correspondiente, es decir, si y solo si toda columna de A es una columna pivote. Esto ocurre si y solo si la ecuación $A\mathbf{x} = 0$ tiene solamente la solución trivial.

29. ***a***) Cuando A es una matriz de 4×4 con tres posiciones pivote, la ecuación $A\mathbf{x} = \mathbf{0}$ tiene una variable libre y, por lo tanto, tiene soluciones no triviales.

b) Con tres posiciones pivote, A no tiene una posición pivote en cada una de sus cuatro filas. Según el teorema 4 de la sección 1.4, la ecuación $A\mathbf{x} = \mathbf{b}$ no tiene una solución para cada posible **b**. La palabra “posible” en el ejercicio significa que los únicos vectores considerados en este caso son los que están en $\mathbb{R}^4$, esto porque A tiene cuatro filas.

31. ***a***) Cuando A es una matriz de 3×2 con dos posiciones pivotes, cada columna es una columna pivote. Así, la ecuación $A\mathbf{x} = \mathbf{0}$ no tiene variables libres y, por consiguiente, no tiene solución no trivial.

b) Con dos posiciones pivote y tres renglones, A no puede tener un pivote en cada renglón. Así, de acuerdo con el teorema 4 de la sección 1.4, la ecuación $A\mathbf{x} = \mathbf{b}$ no puede tener una solución para cada posible **b** (en $\mathbb{R}^3$).

33. Su ejemplo debería tener la propiedad de que la suma de las entradas en cada fila es cero. ¿Por qué?

35. Una respuesta: $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$.

37. Una respuesta es $A = \begin{bmatrix} 1 & -4 \\ 1 & -4 \end{bmatrix}$. Si **b** es cualquier vector que *no* es múltiplo de la primera columna de A, entonces el conjunto solución de $A\mathbf{x} = \mathbf{b}$ es vacío y, por lo tanto, no se puede formar al trasladar el conjunto solución de $A\mathbf{x} = \mathbf{0}$. Esto no contradice el teorema 6 porque este último se aplica cuando la ecuación $A\mathbf{x} = \mathbf{b}$ tiene un conjunto solución que no es vacío.

39. Suponga que $A\mathbf{v} = \mathbf{0}$ y $A\mathbf{w} = \mathbf{0}$. Entonces, como $A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w}$, de acuerdo con el teorema 5(a) de la sección 1.4, $A(\mathbf{v} + \mathbf{w}) = A\mathbf{v} + A\mathbf{w} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Ahora, sean c y d escalares. Utilizando ambas partes del teorema 5, $A(c\mathbf{v} + d\mathbf{w}) = A(c\mathbf{v}) + A(d\mathbf{w}) = cA\mathbf{v} + dA\mathbf{w} = c\mathbf{0} + d\mathbf{0} = \mathbf{0}$.

Sección 1.6, página 54

1. La solución general es $p_B = 0.875p_S$, con p_S libre. Una solución de equilibrio es $p_S = 1000$ y $p_B = 875$. Empleando fracciones, la solución general se podría escribir como $p_B = (7/8)p_S$, y una elección natural de precios podría ser $p_S = 80$ y $p_B = 70$. Solo la *razón* de precios es importante. El equilibrio económico no se altera por el cambio proporcional en los precios.

3. ***a***)

	Distribución de la producción de:				
Salida	C y E	Man.	Serv.	Entrada	Comprados por:
	.1	.1	.2	→	C y E
	.8	.1	.4	→	Man.
	.1	.8	.4	→	Serv.

b) $\begin{bmatrix} .9 & -.1 & -.2 & 0 \\ -.8 & .9 & -.4 & 0 \\ -.1 & -.8 & .6 & 0 \end{bmatrix}$

c) [M] $p_{CyE} \approx 30$, $p_M \approx 71$, $p_S = 100$.

5. ***a***)

	Distribución de la producción de:					
Salida	Agr.	Man.	Serv.	Transp.	Entrada	Comprados por:
	.20	.35	.10	.20	→	Agr.
	.20	.10	.20	.30	→	Man.
	.30	.35	.50	.20	→	Serv.
	.30	.20	.20	.30	→	Transp.

b) Una solución es $p_A = 7.99$, $p_M = 8.36$, $p_S = 14.65$, y $p_T = 10.00$.

c) Distribución de la producción de:

Salida	Agr. ↓	Man. ↓	Serv. ↓	Transp. ↓	Entrada	Comprados por:
	.20	.35	.10	.20	→	Agr.
	.10	.10	.20	.30	→	Man.
	.40	.35	.50	.20	→	Serv.
	.30	.20	.20	.30	→	Transp.

d) Una solución es $p_A = 7.81$, $p_M = 7.67$, $p_S = 15.62$, y $p_T = 10.00$.

La campaña ha beneficiado sobre todo al sector de servicios.

7. $3NaHCO_3 + H_3C_6H_5O_7 \to Na_3C_6H_5O_7 + 3H_2O + 3CO_2$

9. $B_2S_3 + 6H_2O \to 2H_3BO_3 + 3H_2S$

11. [**M**] $16MnS + 13As_2Cr_{10}O_{35} + 374H_2SO_4$
$\to 16HMnO_4 + 26AsH_3 + 130CrS_3O_{12} + 327H_2O$

13. **a.** $\begin{cases} x_1 = x_3 - 40 \\ x_2 = x_3 + 10 \\ x_3 \text{ es libre} \\ x_4 = x_6 + 50 \\ x_5 = x_6 + 60 \\ x_6 \text{ es libre} \end{cases}$ **b.** $\begin{cases} x_2 = 50 \\ x_3 = 40 \\ x_4 = 50 \\ x_5 = 60 \end{cases}$

15. $\begin{cases} x_1 = 60 + x_6 \\ x_2 = -10 + x_6 \\ x_3 = 90 + x_6 \\ x_4 = x_6 \\ x_5 = 80 + x_6 \\ x_6 \text{ es libre} \end{cases}$

Para que el flujo sea no negativo, $x_6 \geq 10$.

Sección 1.7, pág. 60

Justifique sus respuestas a los ejercicios 1 a 22.

1. Linealmente independientes. **3.** Linealmente dependientes.

5. Linealmente independiente. **7.** Linealmente dependiente.

9. *a*) Ninguna h. *b*) Para toda h.

11. $h = -4$. **13.** Para toda h.

15. Linealmente dependientes

17. Linealmente dependientes

19. Linealmente independientes

23. $\begin{bmatrix} \blacksquare & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

25. $\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ y $\begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

27. Las cuatro columnas de la matriz A de 6×4 deben ser columnas pivote. De lo contrario, la ecuación $A\mathbf{x} = \mathbf{0}$ tendría una variable libre; en tal caso, las columnas de A serían linealmente dependientes.

29. A: Cualquier matriz de 3×2 cuya segunda columna es un múltiplo de la primera tendrá la propiedad deseada.

B: Funcionará cualquier matriz de 3×2 con dos columnas diferentes de cero tales que ninguna de ellas sea múltiplo de la otra. En este caso, las columnas forman un conjunto linealmente independiente, y así la ecuación $B\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial.

31. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

33. Verdadero, según el teorema 7.

35. Verdadero, de acuerdo con el teorema 9.

37. Verdadero. Una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ se puede ampliar a una relación de dependencia lineal entre $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$ colocando un peso cero en $\mathbf{v}_4$.

41. [**M**] Utilice las columnas pivote de A,

$$B = \begin{bmatrix} 3 & -4 & 7 \\ -5 & -3 & -11 \\ 4 & 3 & 2 \\ 8 & -7 & 4 \end{bmatrix}$$

Otras elecciones son posibles.

43. [**M**] Cada columna de A que no es columna de B está en el conjunto generado por las columnas de B.

Sección 1.8, página 68

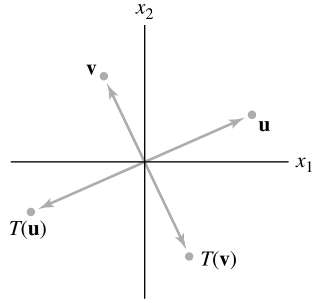
1. $\begin{bmatrix} 2 \\ -6 \end{bmatrix}, \begin{bmatrix} 2a \\ 2b \end{bmatrix}$ **3.** $\mathbf{x} = \begin{bmatrix} 7 \\ 6 \\ 3 \end{bmatrix}$, solución única

5. $\mathbf{x} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$, no es única **7.** $a = 5$, $b = 6$

9. $\mathbf{x} = x_3 \begin{bmatrix} 4 \\ 3 \\ 1 \\ 0 \end{bmatrix}$

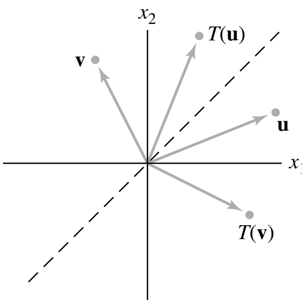
11. Sí, porque el sistema representado por $[A \quad \mathbf{b}]$ es consistente.

13.



Reflexión a través del origen.

15.



Reflexión a través de la recta $x_1 = x_2$.

17. $\begin{bmatrix} 8 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 5 \\ 11 \end{bmatrix}$ **19.** $\begin{bmatrix} 13 \\ 7 \end{bmatrix}, \begin{bmatrix} 2x_1 - x_2 \\ 5x_1 + 6x_2 \end{bmatrix}$

21. *Sugerencia:* Lea el texto con cuidado y escriba sus respuestas. Observe que el ejercicio 21*e*) es una oración de la forma

"⟨enunciado 1⟩ si y solo si ⟨enunciado 2⟩".

Marque dicha oración como verdadera si el ⟨enunciado 1⟩ es válido siempre que el ⟨enunciado 2⟩ también lo sea *y* el ⟨enunciado 2⟩ es verdadero siempre que el ⟨enunciado 1⟩ sea verdadero.

23. ***a*)** Cuando $b = 0$, $f(x) = mx$. En este caso, para toda x, y en $\mathbb{R}$ y todos los escalares c y d,

$$\begin{aligned} f(cx + dy) &= m(cx + dy) = mcx + mdy \\ &= c(mx) + d(my) \\ &= c \cdot f(x) + d \cdot f(y) \end{aligned}$$

Esto demuestra que f es lineal.

***b*)** Cuando $f(x) = mx + b$, con b distinto de cero, $f(0) = m(0) + b = b \neq 0$.

***c*)** En cálculo, f es una "función lineal" porque la gráfica de f es una recta.

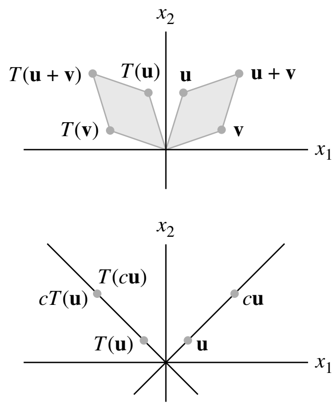
25. *Sugerencia:* Demuestre que la imagen de una recta (es decir, el conjunto de imágenes de todos los puntos sobre una recta) se puede representar por la ecuación paramétrica de una recta.

27. Cualquier punto $\mathbf{x}$ sobre el plano P satisface la ecuación paramétrica $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ para algunos valores de s y t. Por linealidad, la imagen $T(\mathbf{x})$ satisface la ecuación paramétrica

$$T(\mathbf{x}) = sT(\mathbf{u}) + tT(\mathbf{v}) \quad (s, t \text{ en } \mathbb{R}) \qquad (*)$$

El conjunto de imágenes es justamente Gen $\{T(\mathbf{u}), T(\mathbf{v})\}$. Si $T(\mathbf{u})$ y $T(\mathbf{v})$ son linealmente independientes, Gen $\{T(\mathbf{u}), T(\mathbf{v})\}$ es un plano a través de $T(\mathbf{u})$, $T(\mathbf{v})$, y $\mathbf{0}$. Si $T(\mathbf{u})$ y $T(\mathbf{v})$ son linealmente dependientes y no son nulos a la vez, entonces Gen $\{T(\mathbf{u}), T(\mathbf{v})\}$ es una recta que pasa por $\mathbf{0}$. Si $T(\mathbf{u}) = T(\mathbf{v}) = \mathbf{0}$, entonces Gen $\{T(\mathbf{u}), T(\mathbf{v})\}$ es $\{\mathbf{0}\}$.

29.



31. *Sugerencia:* Como $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente, se puede escribir cierta ecuación y trabajar con ella.

33. Una posibilidad es demostrar que T no mapea el vector cero en el vector cero, algo que *sí* hace toda transformación lineal: $T(0, 0) = (0, -3, 0)$.

35. Tome $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^3$ y sean c y d escalares. Entonces,

$$c\mathbf{u} + d\mathbf{v} = (cu_1 + dv_1, cu_2 + dv_2, cu_3 + dv_3)$$

La transformación T es lineal porque

$$\begin{aligned} T(c\mathbf{u} + d\mathbf{v}) &= (cu_1 + dv_1, 0, cu_3 + dv_3) \\ &= (cu_1, 0, cu_3) + (dv_1, 0, dv_3) \\ &= c(u_1, 0, u_3) + d(v_1, 0, v_3) \\ &= cT(\mathbf{u}) + dT(\mathbf{v}) \end{aligned}$$

37. [**M**] Todos los múltiplos de $(-1, -1, 1, 0)$

39. [**M**] Sí. Una elección para $\mathbf{x}$ es $(1, 2, 0, 0)$.

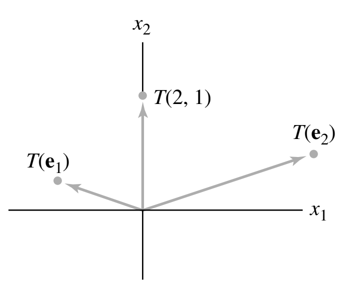
Sección 1.9, página 78

1. $\begin{bmatrix} 3 & -5 \\ 1 & 2 \\ 3 & 0 \\ 1 & 0 \end{bmatrix}$ **3.** $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}$ **5.** $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

7. $\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ **9.** $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

11. La transformación descrita T mapea $\mathbf{e}_1$ en $-\mathbf{e}_1$ y mapea $\mathbf{e}_2$ en $-\mathbf{e}_2$. Una rotación de π radianes también mapea $\mathbf{e}_1$ en $-\mathbf{e}_1$ y mapea $\mathbf{e}_2$ en $-\mathbf{e}_2$. Como una transformación lineal está completamente determinada por su acción sobre las columnas de la matriz identidad, entonces la transformación de rotación tiene el mismo efecto que T sobre cada vector en $\mathbb{R}^2$.

13.



15. $\begin{bmatrix} 2 & -4 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 3 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix}$

19. $\begin{bmatrix} 1 & -5 & 4 \\ 0 & 1 & -6 \end{bmatrix}$

21. $\mathbf{x} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}$

25. No es uno a uno, y no mapea $\mathbb{R}^4$ sobre $\mathbb{R}^4$.

27. No es uno a uno, pero mapea $\mathbb{R}^3$ sobre $\mathbb{R}^2$.

29. $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{bmatrix}$

33. *Sugerencia:* Si $\mathbf{e}_j$ es la j-ésima columna de I_n, entonces $B\mathbf{e}_j$ es la j-ésima columna de B.

35. *Sugerencia:* ¿Es posible que $m > n$? ¿Qué ocurre con $m < n$?

37. [**M**] No. (Explique por qué).

39. [**M**] No. (Explique por qué).

Sección 1.10, página 86

1. ***a***) $x_1 \begin{bmatrix} 110 \\ 4 \\ 20 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 130 \\ 3 \\ 18 \\ 5 \end{bmatrix} = \begin{bmatrix} 295 \\ 9 \\ 48 \\ 8 \end{bmatrix}$, donde x_1 es el número de porciones de Cheerios, y x_2 es el número de porciones de 100% Natural Cereal.

b) $\begin{bmatrix} 110 & 130 \\ 4 & 3 \\ 20 & 18 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 295 \\ 9 \\ 48 \\ 8 \end{bmatrix}$. Mezcle 1.5 porciones de Cheerios con 1 porción de 100% Natural Cereal.

3. ***a***) Debería mezclar 0.99 porciones de Mac and Cheese, 1.54 porciones de brócoli, y 0.79 porciones de pollo para obtener el contenido nutricional deseado.

b) Debería mezclar 1.09 porciones de pasta integral y queso cheddar, 0.88 porciones de brócoli, y 1.03 porciones de pollo para lograr el contenido nutricional deseado. Observe que esta mezcla contiene significativamente menos brócoli, así que a ella le debería gustar más.

5. $R\mathbf{i} = \mathbf{v}$, $\begin{bmatrix} 11 & -5 & 0 & 0 \\ -5 & 10 & -1 & 0 \\ 0 & -1 & 9 & -2 \\ 0 & 0 & -2 & 10 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 50 \\ -40 \\ 30 \\ -30 \end{bmatrix}$

[**M**]: $\mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 3.68 \\ -1.90 \\ 2.57 \\ -2.49 \end{bmatrix}$

7. $R\mathbf{i} = \mathbf{v}$, $\begin{bmatrix} 12 & -7 & 0 & -4 \\ -7 & 15 & -6 & 0 \\ 0 & -6 & 14 & -5 \\ -4 & 0 & -5 & 13 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 40 \\ 30 \\ 20 \\ -10 \end{bmatrix}$

[**M**]: $\mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} 11.43 \\ 10.55 \\ 8.04 \\ 5.84 \end{bmatrix}$

9. $\mathbf{x}_{k+1} = M\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$, donde

$M = \begin{bmatrix} .93 & .05 \\ .07 & .95 \end{bmatrix}$ y $\mathbf{x}_0 = \begin{bmatrix} 800{,}000 \\ 500{,}000 \end{bmatrix}$

La población en el 2012 (para $k = 2$) es $\mathbf{x}_2 = \begin{bmatrix} 741{,}720 \\ 558{,}280 \end{bmatrix}$.

11. ***a***) $M = \begin{bmatrix} .98363 & .00167 \\ .01637 & .99833 \end{bmatrix}$

b) [**M**] $\mathbf{x}_6 = \begin{bmatrix} 30{,}754{,}500 \\ 229{,}449{,}000 \end{bmatrix}$

13. [**M**]

a) Disminuye la población de la ciudad. Después de 7 años, las poblaciones son aproximadamente iguales, pero la población de la ciudad continúa disminuyendo. Después de 20 años, en la ciudad solo hay 417,000 habitantes (valor redondeado a partir de 417,456). Sin embargo, cada año parecen reducirse los cambios en la población.

b) La población de la ciudad se incrementa lentamente y disminuye la población suburbana. Después de 20 años, la población de la ciudad ha crecido de 350,000 a cerca de 370,000.

Capítulo 1 Ejercicios complementarios, página 88

1. ***a***) F ***b***) F ***c***) V ***d***) F ***e***) V ***f***) V
g) F ***h***) F ***i***) V ***j***) F ***k***) V ***l***) F
m) V ***n***) V ***o***) V ***p***) V ***q***) F ***r***) V
s) F ***t***) V ***u***) F ***v***) F ***w***) F ***x***) V
y) V ***z***) F

3. ***a***) Cualquier sistema lineal consistente cuya forma escalonada es

$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$ o $\begin{bmatrix} \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$

o $\begin{bmatrix} 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$

b) Cualquier sistema lineal consistente cuya forma escalonada reducida es I_3.

c) Cualquier sistema lineal inconsistente de tres ecuaciones con tres variables.

5. ***a***) El conjunto solución: **i.** es vacío si $h = 12$ y $k \neq 2$; **ii.** contiene una única solución si $h \neq 12$; **iii.** contiene una infinidad de soluciones si $h = 12$ y $k = 2$.

b) El conjunto solución es vacío si $k + 3h = 0$; de lo contrario, el conjunto solución contiene una única solución.

7. *a*) Sean $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -5 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ -5 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. "Determine si $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ generan $\mathbb{R}^3$".

Solución: No.

b) Sea $A = \begin{bmatrix} 2 & -4 & -2 \\ -5 & 1 & 1 \\ 7 & -5 & -3 \end{bmatrix}$. "Determine si las columnas de A generan $\mathbb{R}^3$".

c) Defina $T(\mathbf{x}) = A\mathbf{x}$. "Determine si T mapea $\mathbb{R}^3$ sobre $\mathbb{R}^3$".

9. $\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \frac{4}{3}\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \frac{7}{3}\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ o $\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 8/3 \\ 4/3 \end{bmatrix} + \begin{bmatrix} 7/3 \\ 14/3 \end{bmatrix}$

10. *Sugerencia:* Construya una "malla" sobre el plano x_1x_2 determinada por $\mathbf{a}_1$ y $\mathbf{a}_2$.

11. Un conjunto solución es una recta cuando el sistema tiene una variable libre. Si la matriz de coeficientes es de 2×3, entonces dos de las columnas deberían ser columnas pivote. Por ejemplo, tome $\begin{bmatrix} 1 & 2 & * \\ 0 & 3 & * \end{bmatrix}$. Ponga lo que sea en la columna 3. La matriz resultante estará en forma escalonada. Haga una operación de remplazo de filas sobre la segunda fila para crear una matriz que *no* esté en forma escalonada, tal como $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 1 & 5 & 2 \end{bmatrix}$.

12. *Sugerencia:* ¿Cuántas variables libres hay en la ecuación $A\mathbf{x} = \mathbf{0}$?

13. $E = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

15. *a*) Si los tres vectores son linealmente independientes, entonces a, c y f deben ser diferentes de cero.
b) Los números $a,\ldots, f$ pueden tener cualesquiera valores.

16. *Sugerencia:* Liste las columnas de derecha a izquierda como $\mathbf{v}_1,\ldots, \mathbf{v}_4$.

17. *Sugerencia:* Utilice el teorema 7.

19. Sea M la recta que pasa por el origen que es paralela a la recta que pasa por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$. Entonces, $\mathbf{v}_2 - \mathbf{v}_1$ y $\mathbf{v}_3 - \mathbf{v}_1$ están en M. De manera que uno de esos dos vectores es múltiplo del otro, por ejemplo, $\mathbf{v}_2 - \mathbf{v}_1 = k(\mathbf{v}_3 - \mathbf{v}_1)$. Esta ecuación produce una relación de dependencia lineal: $(k - 1)\mathbf{v}_1 + \mathbf{v}_2 - k\mathbf{v}_3 = \mathbf{0}$.

Una segunda solución: Una ecuación paramétrica de la recta es $\mathbf{x} = \mathbf{v}_1 + t(\mathbf{v}_2 - \mathbf{v}_1)$. Como $\mathbf{v}_3$ está sobre la recta, existe alguna t_0 tal que $\mathbf{v}_3 = \mathbf{v}_1 + t_0(\mathbf{v}_2 - \mathbf{v}_1) = (1 - t_0)\mathbf{v}_1 + t_0\mathbf{v}_2$. Así que $\mathbf{v}_3$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$, y $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente dependiente.

21. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **23.** $a = 4/5, b = -3/5$

25. *a*) El vector lista los números de apartamentos de 1, 2 y 3 recámaras cuando se construyen x_1 pisos del plan A.
b) $x_1\begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix} + x_2\begin{bmatrix} 4 \\ 4 \\ 8 \end{bmatrix} + x_3\begin{bmatrix} 5 \\ 3 \\ 9 \end{bmatrix}$
c) [**M**] Utilice 2 pisos del plan A y 15 pisos del plan B. O bien, use 6 pisos del plan A, 2 pisos del plan B, y 8 pisos del plan C. Esas son las únicas soluciones factibles. Hay otras soluciones matemáticas, pero requieren un número negativo (o una fracción) de pisos en uno o dos de los planes, lo cual carece de sentido físico.

Capítulo 2

Sección 2.1, página 100

1. $\begin{bmatrix} -4 & 0 & 2 \\ -8 & 10 & -4 \end{bmatrix}, \begin{bmatrix} 3 & -5 & 3 \\ -7 & 6 & -7 \end{bmatrix}$, no está definida, $\begin{bmatrix} 1 & 13 \\ -7 & -6 \end{bmatrix}$

3. $\begin{bmatrix} 1 & 5 \\ -3 & 5 \end{bmatrix}, \begin{bmatrix} 6 & -15 \\ 9 & -6 \end{bmatrix}$

5. **a.** $A\mathbf{b}_1 = \begin{bmatrix} -10 \\ 0 \\ 26 \end{bmatrix},\ A\mathbf{b}_2 = \begin{bmatrix} 11 \\ 8 \\ -19 \end{bmatrix}$,

$$AB = \begin{bmatrix} -10 & 11 \\ 0 & 8 \\ 26 & -19 \end{bmatrix}$$

b. $$AB = \begin{bmatrix} -1\cdot 4 + 3(-2) & -1(-2) + 3\cdot 3 \\ 2\cdot 4 + 4(-2) & 2(-2) + 4\cdot 3 \\ 5\cdot 4 - 3(-2) & 5(-2) - 3\cdot 3 \end{bmatrix} = \begin{bmatrix} -10 & 11 \\ 0 & 8 \\ 26 & -19 \end{bmatrix}$$

7. 3×7 **9.** $k = -2$

11. $AD = \begin{bmatrix} 5 & 6 & 6 \\ 10 & 12 & 10 \\ 15 & 15 & 12 \end{bmatrix}, DA = \begin{bmatrix} 5 & 10 & 15 \\ 6 & 12 & 15 \\ 6 & 10 & 12 \end{bmatrix}$

La multiplicación por la derecha por D multiplica cada *columna* de A por la entrada diagonal correspondiente en D. La multiplicación por la izquierda por D multiplica cada *fila* de A por la entrada diagonal correspondiente en D.

13. *Sugerencia:* Considere que una de las dos matrices es Q.

17. $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$

19. La tercera columna de AB es la suma de las primeras dos columnas de AB. He aquí por qué. Escriba $B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]$. Por definición, la tercera columna de AB es $A\mathbf{b}_3$. Si $\mathbf{b}_3 = \mathbf{b}_1 + \mathbf{b}_2$, entonces $A\mathbf{b}_3 = A(\mathbf{b}_1 + \mathbf{b}_2) = A\mathbf{b}_1 + A\mathbf{b}_2$, debido a una propiedad de la multiplicación matriz-vector.

21. Las columnas de A son linealmente dependientes. ¿Por qué?

23. *Sugerencia:* Suponga que $\mathbf{x}$ satisface $A\mathbf{x} = \mathbf{0}$, y demuestre que $\mathbf{x}$ debe ser $\mathbf{0}$.

25. *Sugerencia:* Utilice los resultados de los ejercicios 23 y 24, y aplique la ley asociativa de la multiplicación al producto CAD.

27. $\mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u} = -3a + 2b - 5c$,

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -3a & -3b & -3c \\ 2a & 2b & 2c \\ -5a & -5b & -5c \end{bmatrix},$$

$$\mathbf{v}\mathbf{u}^T = \begin{bmatrix} -3a & 2a & -5a \\ -3b & 2b & -5b \\ -3c & 2c & -5c \end{bmatrix}$$

29. *Sugerencia:* De acuerdo con el teorema 2*b*), demuestre que la entrada (i, j) de $A(B + C)$ es igual a la entrada (i, j) de $AB + AC$.

31. *Sugerencia:* Utilice la definición del producto I_mA y considere el hecho de que $I_m\mathbf{x} = \mathbf{x}$ para toda $\mathbf{x}$ en $\mathbb{R}^m$.

33. *Sugerencia:* Primero escriba la entrada (i, j) de $(AB)^T$, que es la entrada (j, i) de AB. Después, calcule la entrada (i, j) en B^TA^T, y considere el hecho de que las entradas en la fila i de B^T son $b_{1i},\ldots, b_{ni}$, porque provienen de la columna i de B, y el hecho de que las entradas en la columna j de A^T son $a_{j1},\ldots, a_{jn}$, ya que provienen de la fila j de A.

35. [**M**] Aquí la respuesta depende de la elección del programa de matrices. Para MATLAB, utilice el comando `help` para leer sobre `zeros`, `ones`, `eye` y `diag`. Para la TI-86, estudie las instrucciones `dim`, `fill` e `iden`. La TI-86 no tiene un comando "diagonal".

37. [**M**] Muestre sus resultados e informe sus conclusiones.

39. [**M**] Muestre sus resultados e informe sus conclusiones.

41. Las matrices parecen aproximarse a la matriz

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

Sección 2.2, página 109

1. $\begin{bmatrix} 2 & -3 \\ -5/2 & 4 \end{bmatrix}$ **3.** $-\frac{1}{3}\begin{bmatrix} -3 & -3 \\ 6 & 7 \end{bmatrix}$ o $\begin{bmatrix} 1 & 1 \\ -2 & -7/3 \end{bmatrix}$

5. $x_1 = 7$ y $x_2 = -9$

7. **a** y **b**: $\begin{bmatrix} -9 \\ 4 \end{bmatrix}, \begin{bmatrix} 11 \\ -5 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \end{bmatrix}$ y $\begin{bmatrix} 13 \\ -5 \end{bmatrix}$

11. La demostración se puede modelar después de la demostración del teorema 5.

13. $AB = AC \Rightarrow A^{-1}AB = A^{-1}AC \Rightarrow IB = IC \Rightarrow B = C$. No, en general, B y C pueden ser diferentes cuando A no es invertible. Véase el ejercicio 10 de la sección 2.1.

15. *Sugerencia:* A la matriz $[A, B]$, aplique las matrices elementales empleadas para reducir por filas A a I.

17. $D = C^{-1}B^{-1}A^{-1}$. Demuestre que D funciona.

19. Después de encontrar $X = CB - A$, demuestre que X es una solución.

21. *Sugerencia:* Considere la ecuación $A\mathbf{x} = \mathbf{0}$.

23. *Sugerencia:* Si $A\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial, entonces no existen variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$, y cada columna de A es una columna pivote.

25. *Sugerencia:* Considere el caso $a = b = 0$. Después, considere el vector $\begin{bmatrix} -b \\ a \end{bmatrix}$, y considere el hecho de que $ad - bc = 0$.

27. *Sugerencia:* Para el inciso *a*), intercambie A y B en el cuadro que sigue al ejemplo 6 de la sección 2.1, y luego remplace B por la matriz identidad. Para los incisos *b*) y *c*), inicie escribiendo

$$A = \begin{bmatrix} \text{fila}_1(A) \\ \text{fila}_2(A) \\ \text{fila}_3(A) \end{bmatrix}$$

29. $\frac{1}{3}\begin{bmatrix} -9 & 3 \\ -4 & 1 \end{bmatrix}$ **31.** $\begin{bmatrix} 8 & 3 & 1 \\ 10 & 4 & 1 \\ 7/2 & 3/2 & 1/2 \end{bmatrix}$

33. La forma general de A^{-1} es

$$A^{-1} = B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & & 0 \\ 0 & -1 & 1 & & \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

Sugerencia: Para $j = 1,\ldots, n$, sean $\mathbf{a}_j$, $\mathbf{b}_j$ y $\mathbf{e}_j$ las j-ésimas columnas de A, B e I, respectivamente. Considere que $\mathbf{a}_j - \mathbf{a}_{j+1} = \mathbf{e}_j$ y $\mathbf{b}_j = \mathbf{e}_j - \mathbf{e}_{j+1}$ para $j = 1,\ldots, n - 1$, y $\mathbf{a}_n = \mathbf{b}_n = \mathbf{e}_n$.

35. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}$. Encuentre esto reduciendo por filas a $[A \quad \mathbf{e}_3]$.

37. $C = \begin{bmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \end{bmatrix}$

39. [**M**] Las deflexiones son .62, .66 y .52 pulgadas, respectivamente.

41. [**M**] .95, 6.19, 11.43 y 3.81 newtons, respectivamente.

Sección 2.3, página 115

La abreviación IMT denota el teorema de la matriz invertible (*Invertible Matrix Theorem*, teorema 8).

1. Invertible, de acuerdo con el IMT. Ninguna columna de la matriz es un múltiplo de la otra columna, de manera que son linealmente independientes. Además, la matriz es invertible según el teorema 4 de la sección 2.2, ya que el determinante es diferente de cero.

3. Observe que A^T tiene un pivote en cada columna, de modo que de acuerdo con el IMT, A^T es invertible. Entonces, según el IMT, A también es invertible.

5. No es invertible, de acuerdo con el IMT. Como esta matriz tiene una columna de ceros, sus columnas forman un conjunto linealmente dependiente y, por consiguiente, la matriz no es invertible.

7. Invertible, de acuerdo con el IMT. La matriz se reduce por filas a

$$\begin{bmatrix} \boxed{-1} & -3 & 0 & 1 \\ 0 & \boxed{-4} & 8 & 0 \\ 0 & 0 & \boxed{3} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

y tiene cuatro posiciones pivote.

9. [**M**] La matriz de 4×4 tiene cuatro posiciones pivote, así que es invertible por el IMT.

11. *Sugerencia:* Intente responder las preguntas apoyándose en una cuidadosa lectura del texto.

13. Una matriz cuadrada triangular superior es invertible si y solo si todas las entradas de la diagonal son diferentes de cero. ¿Por qué?

Nota: Las respuestas que se muestran a continuación para los ejercicios 15 a 29 mencionan al IMT. En muchos casos, una parte o la totalidad de una respuesta aceptable también podría basarse en resultados que se emplearon para establecer el IMT.

15. No, porque el enunciado *h*) del IMT entonces sería falso. Una matriz de 4×4 no puede ser invertible cuando sus columnas no generan $\mathbb{R}^4$.

17. Si A tiene dos columnas idénticas, entonces sus columnas son linealmente dependientes. El inciso *e*) del IMT indica que A no puede ser invertible.

19. D es invertible por el enunciado *e*) del IMT. Así, la ecuación $D\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^7$, de acuerdo con el enunciado *g*) del IMT. ¿Podría agregar algo más?

21. La matriz C no puede ser invertible, de acuerdo con el teorema 5 de la sección 2.2 o según el párrafo después del IMT. Así, el enunciado *g*) del IMT es falso y también lo es *h*). Las columnas de C no generan $\mathbb{R}^n$.

23. El enunciado *g*) del IMT es falso para F, de manera que también es falso el enunciado *d*). Es decir, la ecuación $F\mathbf{x} = \mathbf{0}$ tiene una solución no trivial.

25. *Sugerencia:* Primero utilice el IMT.

27. Sea W la inversa de AB. Entonces, $ABW = I$ y $A(BW) = I$. Por desgracia, esta ecuación por sí misma no prueba que A es invertible. ¿Por qué no? Termine la demostración.

29. Como la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es uno a uno, entonces el enunciado *f*) del IMT es verdadero. También es verdadero el enunciado *i*), y la transformación $\mathbf{x} \mapsto A\mathbf{x}$ mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$. También, A es invertible, lo cual, de acuerdo con el teorema 9, implica que la transformación $\mathbf{x} \mapsto A\mathbf{x}$ es invertible.

31. *Sugerencia:* Si la ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$, entonces A tiene un pivote en cada fila (teorema 4 de la sección 1.4). ¿Podrían existir variables libres en una ecuación $A\mathbf{x} = \mathbf{b}$?

33. *Sugerencia:* Primero demuestre que la matriz estándar de T es invertible. Entonces, utilice un teorema o más para demostrar que $T^{-1}(\mathbf{x}) = B\mathbf{x}$, donde $B = \begin{bmatrix} 7 & 9 \\ 4 & 5 \end{bmatrix}$.

35. *Sugerencia:* Para demostrar que T es uno a uno, suponga que $T(\mathbf{u}) = T(\mathbf{v})$ para algunos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. Deduzca que $\mathbf{u} = \mathbf{v}$. Para demostrar que T es sobre, suponga que $\mathbf{y}$ representa un vector arbitrario en $\mathbb{R}^n$ y utilice la inversa S para producir una $\mathbf{x}$ tal que $T(\mathbf{x}) = \mathbf{y}$. Es posible dar una segunda demostración empleando el teorema 9 junto con un teorema de la sección 1.9.

37. *Sugerencia:* Considere las matrices estándar de T y U.

39. Si T mapea $\mathbb{R}^n$ sobre $\mathbb{R}^n$, entonces, de acuerdo con el teorema 12 de la sección 1.9, las columnas de su matriz estándar A generan $\mathbb{R}^n$. Según el IMT, A es invertible. Así, de acuerdo con el teorema 9, T es invertible, y A^{-1} es la matriz estándar de T^{-1}. Como A^{-1} también es invertible, entonces, según el IMT, sus columnas son linealmente independientes y generan $\mathbb{R}^n$. Al aplicar el teorema 12 de la sección 1.9 a la transformación T^{-1} se muestra que T^{-1} es un mapeo uno a uno de $\mathbb{R}^n$ sobre $\mathbb{R}^n$.

41. [**M**]

a) La solución exacta de (3) es $x_1 = 3.94$ y $x_2 = .49$. La solución exacta de (4) es $x_1 = 2.90$ y $x_2 = 2.00$.

b) Cuando la solución de (4) se emplea como una aproximación para la solución de (3), el error al usar el valor de 2.90 para x_1 es aproximadamente del 26%, y el error al utilizar 2.0 para x_2 es de alrededor del 308%.

c) El número de condición de la matriz de coeficientes es 3363. El cambio porcentual en la solución de (3) a (4) es cerca de 7700 veces el cambio porcentual en el lado derecho de la ecuación. Este es el mismo orden de magnitud que el número de condición. Este último da una medida aproximada de qué tan sensible es la solución de $A\mathbf{x} = \mathbf{b}$ ante cambios en $\mathbf{b}$. Información adicional sobre el número de condición se presenta al final del capítulo 6 y en el capítulo 7.

43. [**M**] cond(A) $\approx$ 69,000, que está entre 10^4 y 10^5. Así que se pueden perder 4 o 5 dígitos de exactitud. Varios experimentos con MATLAB deberían comprobar que $\mathbf{x}$ y $\mathbf{x}_1$ concuerdan a 11 o 12 dígitos.

45. [**M**] Algunas versiones de MATLAB emiten una señal de alerta cuando se pide invertir una matriz de Hilbert de orden cercano a 12 o mayor, empleando aritmética de punto flotante. El producto AA^{-1} debería tener varias entradas fuera de la diagonal que están lejos de ser cero. Si no, intente con una matriz más grande.

Sección 2.4, página 121

1. $\begin{bmatrix} A & B \\ EA + C & EB + D \end{bmatrix}$ **3.** $\begin{bmatrix} C & D \\ A & B \end{bmatrix}$

5. $Y = B^{-1}$ (explique por qué), $X = -B^{-1}A$, $Z = C$

7. $X = A^{-1}$ (¿por qué?), $Y = -BA^{-1}$, $Z = 0$ (¿por qué?)

9. $A_{21} = -B_{21}B_{11}^{-1}$, $A_{31} = -B_{31}B_{11}^{-1}$, $C_{22} = B_{22} - B_{21}B_{11}^{-1}B_{12}$

13. *Sugerencia:* Suponga que A es invertible, y sea $A^{-1} = \begin{bmatrix} D & E \\ F & G \end{bmatrix}$. Demuestre que $BD = I$ y $CG = I$.

Esto implica que B y C son invertibles. (¡Explique por qué!). Al contrario, suponga que B y C son invertibles. Para probar que A es invertible, indique cómo debería ser A^{-1} y compruebe que su propuesta funciona.

15. $$G_{k+1} = \begin{bmatrix} X_k & \mathbf{x}_{k+1} \end{bmatrix} \begin{bmatrix} X_k^T \\ \mathbf{x}_{k+1}^T \end{bmatrix} = X_k X_k^T + \mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$$
$$= G_k + \mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$$
Solo se necesita calcular el producto matricial exterior $\mathbf{x}_{k+1}\mathbf{x}_{k+1}^T$ (y después se suma a G_k).

17. La inversa de $\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$ es $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$. De manera similar, $\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix}$ tiene una inversa. De la ecuación (7), se obtiene
$$\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I & -Y \\ 0 & I \end{bmatrix} = \begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix} \qquad (*)$$
Si A es invertible, entonces la matriz del lado derecho de (*) es un producto de matrices invertibles y, por lo tanto, es invertible. De acuerdo con el ejercicio 13, A_{11} y S deben ser invertibles.

19. $W(s) = I_m - C(A - sI_n)^{-1}B$. Este es el complemento de Schur de $A - sI_n$ en la matriz del sistema.

21. ***a)*** $$A^2 = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 1+0 & 0+0 \\ 2-2 & 0+(-1)^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
b) $$M^2 = \begin{bmatrix} A & 0 \\ I & -A \end{bmatrix} \begin{bmatrix} A & 0 \\ I & -A \end{bmatrix} = \begin{bmatrix} A^2+0 & 0+0 \\ A-A & 0+(-A)^2 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

23. Si A_1 y B_1 son $(k+1) \times (k+1)$ y triangulares inferiores, entonces escriba $A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}$ y $B_1 = \begin{bmatrix} b & \mathbf{0}^T \\ \mathbf{w} & B \end{bmatrix}$, donde A y B son de $k \times k$ y triangulares inferiores, $\mathbf{v}$ y $\mathbf{w}$ están en $\mathbb{R}^k$, y a y b son escalares adecuados. Suponga que el producto de matrices triangulares inferiores de $k \times k$ es una matriz triangular inferior, y calcule el producto A_1B_1. ¿Qué se concluye?

25. Utilice el ejercicio 13 para encontrar la inversa de una matriz de la forma $B = \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix}$, donde B_{11} es de $p \times p$, B_{22} es de $q \times q$, y B es invertible. Particione la matriz A, y aplique dos veces su resultado para encontrar que
$$A^{-1} = \begin{bmatrix} -5 & 2 & 0 & 0 & 0 \\ 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 3 & -4 \\ 0 & 0 & 0 & -5/2 & 7/2 \end{bmatrix}$$

27. ***a), b)*** Las instrucciones empleadas en estos ejercicios dependen del programa de matrices.

c) El álgebra requerida viene de la ecuación matricial por bloques
$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$$
donde $\mathbf{x}_1$ y $\mathbf{b}_1$ están en $\mathbb{R}^{20}$, y $\mathbf{x}_2$ y $\mathbf{b}_2$ están en $\mathbb{R}^{30}$. Entonces, $A_{11}\mathbf{x}_1 = \mathbf{b}_1$, lo cual se puede resolver para obtener $\mathbf{x}_1$. La ecuación $A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2 = \mathbf{b}_1$ conduce a $A_{22}\mathbf{x}_2 = \mathbf{b}_1 - A_{21}\mathbf{x}_1$, de la que se puede obtener $\mathbf{x}_2$ mediante reducción por filas de la matriz $[A_{22} \;\; \mathbf{c}]$, donde $\mathbf{c} = \mathbf{b}_2 - A_{21}\mathbf{x}_1$.

Sección 2.5, página 129

1. $L\mathbf{y} = \mathbf{b} \Rightarrow \mathbf{y} = \begin{bmatrix} -7 \\ -2 \\ 6 \end{bmatrix}$, $U\mathbf{x} = \mathbf{y} \Rightarrow \mathbf{x} = \begin{bmatrix} 3 \\ 4 \\ -6 \end{bmatrix}$

3. $\mathbf{y} = \begin{bmatrix} 6 \\ 12 \\ 0 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -5 \\ -4 \\ 0 \end{bmatrix}$ **5.** $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 38 \\ 16 \\ 8.5 \\ -4 \end{bmatrix}$

7. $LU = \begin{bmatrix} 1 & 0 \\ -3/2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ 0 & 7/2 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 7 & 2 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 4 & 5 & 1 & 0 \\ -2 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -5 & -3 \\ 0 & -2 & 3 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 5 & 2 \\ 0 & 3 & 2 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

17. $U^{-1} = \begin{bmatrix} 1/2 & -3/4 & -2 \\ 0 & -1/4 & -1 \\ 0 & 0 & -1/2 \end{bmatrix}$,
$L^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -2 & -1 & 1 \end{bmatrix}$,
$A^{-1} = \begin{bmatrix} 3 & 5/4 & -2 \\ 3/2 & 3/4 & -1 \\ 1 & 1/2 & -1/2 \end{bmatrix}$

19. *Sugerencia:* Piense en aplicar reducción por filas a $[A \;\; I]$.

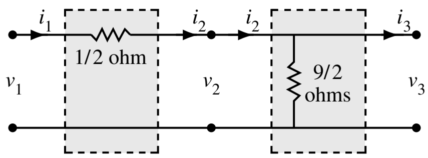
21. *Sugerencia:* Represente las operaciones de fila mediante una secuencia de matrices elementales.

23. ***a)*** Denote las filas de D como transpuestas de vectores columna. Entonces, con la multiplicación matricial particionada se obtiene
$$A = CD = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_4 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_4^T \end{bmatrix} = \mathbf{c}_1\mathbf{v}_1^T + \cdots + \mathbf{c}_4\mathbf{v}_4^T$$

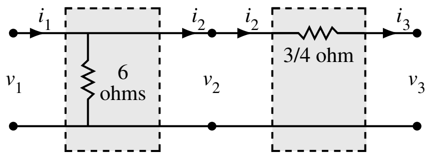
b) A tiene 40,000 entradas. Como C tiene 1600 entradas y D tiene 400 entradas, entonces juntas solo ocupan el 5% de la memoria necesaria para almacenar A.

25. Explique por qué U, D y V^T son invertibles. Después, utilice un teorema sobre la inversa de un producto de matrices invertibles.

27. *a*)



b)



29. *a*) $\begin{bmatrix} 1+R_3/R_2 & -R_1-R_3-(R_1R_3)/R_2 \\ -1/R_2 & 1+R_1/R_2 \end{bmatrix}$

b) $A = \begin{bmatrix} 3 & -12 \\ -1/3 & 5/3 \end{bmatrix}$
$= \begin{bmatrix} 1 & -6 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1/3 & 1 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

Sean $R_1 = 2$ ohms, $R_2 = 3$ ohms y $R_3 = 6$ ohms

31. [**M**]

a) $L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.25 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -.25 & -.0667 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -.2667 & -.2857 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -.2679 & -.0833 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -.2917 & -.2921 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -.2697 & -.0861 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -.2948 & -.2931 & 1 \end{bmatrix}$

b) $U = \begin{bmatrix} 4 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.75 & -.25 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3.7333 & -1.0667 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3.4286 & -.2857 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.7083 & -1.0833 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3.3919 & -.2921 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3.7052 & -1.0861 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3.3868 \end{bmatrix}$

c) $A^{-1} = \begin{bmatrix} .2953 & .0866 & .0945 & .0509 & .0318 & .0227 & .0100 & .0082 \\ .0866 & .2953 & .0509 & .0945 & .0227 & .0318 & .0082 & .0100 \\ .0945 & .0509 & .3271 & .1093 & .1045 & .0591 & .0318 & .0227 \\ .0509 & .0945 & .1093 & .3271 & .0591 & .1045 & .0227 & .0318 \\ .0318 & .0227 & .1045 & .0591 & .3271 & .1093 & .0945 & .0509 \\ .0227 & .0318 & .0591 & .1045 & .1093 & .3271 & .0509 & .0945 \\ .0100 & .0082 & .0318 & .0227 & .0945 & .0509 & .2953 & .0866 \\ .0082 & .0100 & .0227 & .0318 & .0509 & .0945 & .0866 & .2953 \end{bmatrix}$

Obtenga A^{-1} directamente y después calcule $A^{-1} - U^{-1}L^{-1}$ para comparar ambos métodos de inversión matricial.

Sección 2.6, página 136

1. $C = \begin{bmatrix} .10 & .60 & .60 \\ .30 & .20 & 0 \\ .30 & .10 & .10 \end{bmatrix}, \left\{\begin{matrix}\text{demanda} \\ \text{intermedia}\end{matrix}\right\} = \begin{bmatrix} 60 \\ 20 \\ 10 \end{bmatrix}$

3. $\mathbf{x} = \begin{bmatrix} 44.44 \\ 16.67 \\ 16.67 \end{bmatrix}$ **5.** $\mathbf{x} = \begin{bmatrix} 110 \\ 120 \end{bmatrix}$

7. *a*) $\begin{bmatrix} 1.6 \\ 1.2 \end{bmatrix}$ *b*) $\begin{bmatrix} 111.6 \\ 121.2 \end{bmatrix}$

c) *Sugerencia:* Encuentre una fórmula que implique $(I - C)^{-1}$.

9. $\mathbf{x} = \begin{bmatrix} 82.8 \\ 131.0 \\ 110.3 \end{bmatrix}$

11. *Sugerencia:* Utilice propiedades de transpuestas para obtener $\mathbf{p}^T = \mathbf{p}^T C + \mathbf{v}^T$ de manera que $\mathbf{p}^T\mathbf{x} = (\mathbf{p}^T C + \mathbf{v}^T)\mathbf{x} = \mathbf{p}^T C\mathbf{x} + \mathbf{v}^T\mathbf{x}$. Ahora calcule $\mathbf{p}^T\mathbf{x}$ con la ecuación de producción.

13. [**M**] $\mathbf{x} =$ (99576, 97703, 51231, 131570, 49488, 329554, 13835). Las entradas en $\mathbf{x}$ sugieren más precisión en la respuesta que está garantizada por las entradas en $\mathbf{d}$, que parece ser exacta quizá solamente al millar más cercano. Así que una respuesta más realista para $\mathbf{x}$ podría ser $\mathbf{x} = 1000 \times (100, 98, 51, 132, 49, 330, 14)$.

15. [**M**] $\mathbf{x}^{(12)}$ es el primer vector cuyas entradas son exactas al millar más cercano. El cálculo de $\mathbf{x}^{(12)}$ requiere cerca de 1260 flops, mientras que la reducción por filas de $[(I - C)\mathbf{d}]$ necesita solamente alrededor de 550 flops. Si C es mayor de 20×20, entonces se necesitan menos flops para calcular $\mathbf{x}^{(12)}$ mediante iteración que para determinar el vector de equilibrio $\mathbf{x}$ por reducción de filas. Conforme aumenta el tamaño de C, se incrementa la ventaja del método iterativo. También, cuando C se vuelve más dispersa para modelos grandes de economía, entonces se requieren menos iteraciones para una exactitud razonable.

Sección 2.7, página 144

1. $\begin{bmatrix} 1 & .25 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **3.** $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1/2 & -\sqrt{3}/2 & 3+4\sqrt{3} \\ \sqrt{3}/2 & 1/2 & 4-3\sqrt{3} \\ 0 & 0 & 1 \end{bmatrix}$

Véase el problema de práctica.

9. $A(BD)$ requiere 800 multiplicaciones. $(AB)D$ necesita 408 multiplicaciones. El primer método utiliza casi el doble de multiplicaciones. Si D tuviera 10,000 columnas, las cuentas serían 80,000 y 40,008, respectivamente.

11. Considere el hecho de que

$$\sec\varphi - \tan\varphi \operatorname{sen}\varphi = \frac{1}{\cos\varphi} - \frac{\operatorname{sen}^2\varphi}{\cos\varphi} = \cos\varphi$$

13. $\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix} = \begin{bmatrix} I & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}\begin{bmatrix} A & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}$. Primero aplique la transformación lineal A, y después traslade mediante $\mathbf{p}$.

15. $(12, -6, -3)$ **17.** $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & -\sqrt{3}/2 & 0 \\ 0 & \sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

19. El triángulo con vértices en (7, 2, 0), (7.5, 5, 0) y (5, 5, 0)

21. [**M**] $\begin{bmatrix} 2.2586 & -1.0395 & -.3473 \\ -1.3495 & 2.3441 & .0696 \\ .0910 & -.3046 & 1.2777 \end{bmatrix}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix}$

Sección 2.8, página 151

1. El conjunto es cerrado bajo sumas, pero no bajo multiplicación por un escalar negativo. (Bosqueje un ejemplo).

3. El conjunto no es cerrado bajo sumas o múltiplos escalares.

5. Sí. El sistema correspondiente a $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{w}]$ es consistente.

7. ***a***) Los tres vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

b) Una infinidad de vectores.

c) Sí, porque $A\mathbf{x} = \mathbf{p}$ tiene una solución.

9. No, porque $A\mathbf{p} \neq \mathbf{0}$.

11. $p = 4$ y $q = 3$. Nul A es un subespacio de $\mathbb{R}^4$ porque las soluciones de $A\mathbf{x} = \mathbf{0}$ deben tener cuatro entradas, para ajustar con las columnas de A. Col A es un subespacio de $\mathbb{R}^3$ porque cada vector columna tiene tres entradas.

13. Para Nul A, elija $(1, -2, 1, 0)$ o $(-1, 4, 0, 1)$, por ejemplo. Para Col A, seleccione cualquier columna de A.

15. Sea A la matriz cuyas columnas son los vectores dados. Entonces, A es invertible porque su determinante es diferente de cero, y así sus columnas forman una base para $\mathbb{R}^2$, de acuerdo con el IMT (o con el ejemplo 5). (Se podrían dar otras razones para la invertibilidad de A).

17. Sea A la matriz cuyas columnas son los vectores dados. La reducción por filas muestra tres pivotes, así que A es invertible. Según el IMT, las columnas de A forman una base para $\mathbb{R}^3$.

19. Sea A la matriz de 3×2 cuyas columnas son los vectores dados. Las columnas de A posiblemente no puedan generar $\mathbb{R}^3$ porque A no puede tener un pivote en cada fila. Así, las columnas no son una base para $\mathbb{R}^3$. (Estas son una base para un plano en $\mathbb{R}^3$).

21. *Sugerencia:* Lea la sección con cuidado, y escriba sus respuestas. Esta sección tiene términos y conceptos clave que debe aprender antes de seguir adelante.

23. Base para Col A: $\begin{bmatrix} 4 \\ 6 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \\ 4 \end{bmatrix}$

Base para Nul A: $\begin{bmatrix} 4 \\ -5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 6 \\ 0 \\ 1 \end{bmatrix}$

25. Base para Col A: $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 5 \\ -5 \end{bmatrix}$

Base para Nul A: $\begin{bmatrix} 2 \\ -2.5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ .5 \\ 0 \\ -4 \\ 1 \end{bmatrix}$

27. Construya una matriz A de 3×3 distinta de cero, y construya $\mathbf{b}$ como una combinación lineal conveniente de las columnas de A.

29. *Sugerencia:* Se necesita una matriz distinta de cero cuyas columnas sean linealmente dependientes.

31. Si Col $F \neq \mathbb{R}^5$, entonces las columnas de F no generan $\mathbb{R}^5$. Puesto que F es cuadrada, el IMT indica que F no es invertible y la ecuación $F\mathbf{x} = \mathbf{0}$ tiene una solución no trivial. Es decir, Nul F contiene un vector diferente de cero. Otra manera de describir esto es escribir Nul $F \neq \{\mathbf{0}\}$.

33. Si Nul $C = \{\mathbf{0}\}$, entonces la ecuación $C\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial. Como C es cuadrada, el IMT indica que C es invertible y la ecuación $C\mathbf{x} = \mathbf{b}$ tiene una solución para cada $\mathbf{b}$ en $\mathbb{R}^6$. Además, cada solución es única de acuerdo con el teorema 5 de la sección 2.2.

35. Si Nul B contiene vectores diferentes de cero, entonces la ecuación $B\mathbf{x} = \mathbf{0}$ tiene soluciones no triviales. Como B es cuadrada, el IMT indica que B no es invertible y las columnas de B no generan $\mathbb{R}^5$. Así, Col B es un subespacio de $\mathbb{R}^5$, pero Col $B \neq \mathbb{R}^5$.

37. [**M**] Muestre la forma escalonada reducida de A, y seleccione las columnas pivote de A como una base para Col A. Para Nul A, escriba la solución de $A\mathbf{x} = \mathbf{0}$ en forma vectorial paramétrica.

Base para Col A: $\begin{bmatrix} 3 \\ -7 \\ -5 \\ 3 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ 7 \\ -7 \end{bmatrix}$

Base para Nul A: $\begin{bmatrix} -2.5 \\ -1.5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.5 \\ 2.5 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3.5 \\ -1.5 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Sección 2.9, página 157

1. $\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ **5.** $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ **7.** $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1.5 \\ .5 \end{bmatrix}$

9. Base para Col A: $\begin{bmatrix} 1 \\ 3 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ 5 \\ 9 \\ 14 \end{bmatrix}$; dim Col $A = 3$

Base para Nul A: $\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}$; dim Nul $A = 1$

11. Base para Col A: $\begin{bmatrix} 2 \\ 3 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} -5 \\ -8 \\ 9 \\ -7 \end{bmatrix}$; dim Col $A = 2$

Base para Nul A: $\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$; dim Nul $A = 3$

13. Los vectores $\mathbf{v}_1$, $\mathbf{v}_3$ y $\mathbf{v}_4$ forman una base para el subespacio dado, H. Así, dim $H = 3$.

15. Col $A = \mathbb{R}^4$, porque A tiene un pivote en cada fila, y así las columnas de A generan $\mathbb{R}^4$. Nul A *no puede* ser igual a $\mathbb{R}^2$ porque Nul A es un subespacio de $\mathbb{R}^6$. Sin embargo, es cierto que Nul A es bidimensional. La razón es que la ecuación $A\mathbf{x} = \mathbf{0}$ tiene dos variables libres, porque A tiene seis columnas, y solo cuatro de ellas son columnas pivote.

19. El hecho de que el espacio solución de $A\mathbf{x} = \mathbf{0}$ tenga una base de tres vectores significa que dim Nul $A = 3$. Como una matriz A de 5×7 tiene siete columnas, entonces el teorema del rango indica que rango $A = 7 -$ dim Nul $A = 4$.

21. Una matriz de 9×8 tiene ocho columnas. De acuerdo con el teorema del rango, dim Nul $A = 8 -$ rango A. Como el rango es 7, entonces dim Nul $A = 1$. Es decir, la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$ es 1.

23. Cree una matriz A de 3×5 con dos columnas pivote. Las restantes tres columnas corresponderán a variables libres en la ecuación $A\mathbf{x} = \mathbf{0}$. De manera que sí es posible la construcción deseada.

25. Por definición, las p columnas de A generan Col A. Si dim Col $A = p$, entonces el conjunto generador de p columnas automáticamente es una base para Col A, de acuerdo con el teorema de la base. En particular, las columnas son linealmente independientes.

27. ***a)*** *Sugerencia:* Las columnas de B generan W, y cada vector $\mathbf{a}_j$ está en W. El vector $\mathbf{c}_j$ está en $\mathbb{R}^p$ ya que B tiene p columnas.

b) *Sugerencia:* ¿Cuál es el tamaño de C?

c) *Sugerencia*: ¿Cómo se relacionan B y C con A?

29. [**M**] Los cálculos deberían mostrar que la matriz $[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{x}]$ corresponden a un sistema consistente. El vector de $\mathcal{B}$-coordenadas de $\mathbf{x}$ es $(2, -1)$.

Capítulo 2 Ejercicios complementarios, página 160

1. ***a)*** V ***b)*** F ***c)*** V ***d)*** F ***e)*** F ***f)*** F
g) V ***h)*** V ***i)*** V ***j)*** F ***k)*** V ***l)*** F
m) F ***n)*** V ***o)*** F ***p)*** V

3. I

5. $A^2 = 2A - I$. Multiplique por A: $A^3 = 2A^2 - A$.
Sustituya $A^2 = 2A - I$:
$A^3 = 2(2A - I) - A = 3A - 2I$.
Multiplique por A otra vez: $A^4 = A(3A - 2I) = 3A^2 - 2A$.
Otra vez sustituya la identidad $A^2 = 2A - I$:
$A^4 = 3(2A - I) - 2A = 4A - 3I$.

7. $\begin{bmatrix} 10 & -1 \\ 9 & 10 \\ -5 & -3 \end{bmatrix}$ **9.** $\begin{bmatrix} -3 & 13 \\ -8 & 27 \end{bmatrix}$

11. ***a)*** $p(x_i) = c_0 + c_1 x_i + \cdots + c_{n-1} x_i^{n-1}$

$$= \text{fila}_i(V) \cdot \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \end{bmatrix} = \text{fila}_i(V\mathbf{c}) = y_i$$

b) Suponga que $x_1, \ldots, x_n$ son distintos y suponga que $V\mathbf{c} = \mathbf{0}$ para algún vector $\mathbf{c}$. Entonces, las entradas en $\mathbf{c}$ son los coeficientes de un polinomio cuyo valor es cero en los distintos puntos $x_1, \ldots, x_n$. Sin embargo, un polinomio diferente de cero de grado $n - 1$ no puede tener n ceros, así que el polinomio debe ser idénticamente cero. Es decir, todas las entradas en $\mathbf{c}$ deben ser cero. Esto muestra que las columnas de V son linealmente independientes.

c) *Sugerencia:* Cuando $x_1, \ldots, x_n$ son diferentes, existe un vector $\mathbf{c}$ tal que $V\mathbf{c} = \mathbf{y}$. ¿Por qué?

13. ***a)*** $P^2 = (\mathbf{u}\mathbf{u}^T)(\mathbf{u}\mathbf{u}^T) = \mathbf{u}(\mathbf{u}^T\mathbf{u})\mathbf{u}^T = \mathbf{u}(1)\mathbf{u}^T = P$

b) $P^T = (\mathbf{u}\mathbf{u}^T)^T = \mathbf{u}^{TT}\mathbf{u}^T = \mathbf{u}\mathbf{u}^T = P$

c) $Q^2 = (I - 2P)(I - 2P)$
$= I - I(2P) - 2PI + 2P(2P)$
$= I - 4P + 4P^2 = I$, debido al inciso *a*).

15. La multiplicación por la izquierda por una matriz elemental produce una operación de fila elemental:

$$B \sim E_1 B \sim E_2 E_1 B \sim E_3 E_2 E_1 B = C$$

Así, B es equivalente por filas a C. Como las operaciones de fila son reversibles, C es equivalente por filas a B. (Alternativamente, muestre a C que cambia a B mediante operaciones de fila empleando las inversas de las E_i).

17. Como B es de 4×6 (con más columnas que filas), sus seis columnas son linealmente dependientes y existe una $\mathbf{x}$ diferente de cero tal que $B\mathbf{x} = \mathbf{0}$. Así, $AB\mathbf{x} = A\mathbf{0} = \mathbf{0}$, lo que indica que la matriz AB no es invertible, de acuerdo con el IMT.

19. [**M**] Con cuatro cifras decimales, conforme k se incrementa,

$$A^k \to \begin{bmatrix} .2857 & .2857 & .2857 \\ .4286 & .4286 & .4286 \\ .2857 & .2857 & .2857 \end{bmatrix} \quad \text{y}$$

$$B^k \to \begin{bmatrix} .2022 & .2022 & .2022 \\ .3708 & .3708 & .3708 \\ .4270 & .4270 & .4270 \end{bmatrix}$$

o, en formato racional,

$$A^k \to \begin{bmatrix} 2/7 & 2/7 & 2/7 \\ 3/7 & 3/7 & 3/7 \\ 2/7 & 2/7 & 2/7 \end{bmatrix} \quad \text{y}$$

$$B^k \to \begin{bmatrix} 18/89 & 18/89 & 18/89 \\ 33/89 & 33/89 & 33/89 \\ 38/89 & 38/89 & 38/89 \end{bmatrix}$$

Capítulo 3

Sección 3.1, página 167

1. 1 **3.** -5 **5.** -23 **7.** 4

9. 10. Inicie con la fila 3.

11. -12. Inicie con la columna 1 o la fila 4.

13. 6. Inicie con la fila 2 o la columna 2.

15. 1 **17.** -5

19. $ad - bc$, $cb - da$. Intercambiando dos filas cambia el signo del determinante.

21. -2, $(18 + 12k) - (20 + 12k) = -2$. Un remplazo de fila no cambia el valor de un determinante.

23. -5, $k(4) - k(2) + k(-7) = -5k$. Escalar una fila por una constante k multiplica el determinante por k.

25. 1 **27.** k **29.** -1

31. 1. La matriz es triangular superior o inferior, únicamente con 1 sobre la diagonal. El determinante es 1, el producto de las entradas diagonales.

33. $\det EA = \det \begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - ad = (-1)(ad - bc)$
$= (\det E)(\det A)$

35. $\det EA = \det \begin{bmatrix} a + kc & b + kd \\ c & d \end{bmatrix}$
$= (a + kc)d - (b + kd)c$
$= ad + kcd - bc - kdc = (+1)(ad - bc)$
$= (\det E)(\det A)$

37. $5A = \begin{bmatrix} 15 & 5 \\ 20 & 10 \end{bmatrix}$; no

41. El área del paralelogramo y el determinante de $[\mathbf{u}\ \ \mathbf{v}]$ son iguales a 6. Si $\mathbf{v} = \begin{bmatrix} x \\ 2 \end{bmatrix}$ para cualquier x, el área continúa siendo 6. En cada caso no cambia la base del paralelogramo, y la altura permanece en 2 porque la segunda coordenada de $\mathbf{v}$ siempre es 2.

43. [**M**] En general, $\det(A + B)$ no es igual a $\det A + \det B$.

45. [**M**] Puede revisar sus conjeturas cuando llegue a la sección 3.2.

Sección 3.2, página 175

1. El intercambio de dos filas invierte el signo del determinante.

3. Una operación de remplazo de filas no modifica el determinante.

5. 3 **7.** 0 **9.** 3 **11.** 120

13. 6 **15.** 35 **17.** -7 **19.** 14

21. Invertible **23.** No es invertible

25. Linealmente independiente.

29. -32

31. *Sugerencia:* Demuestre que $(\det A)(\det A^{-1}) = 1$.

33. *Sugerencia:* Utilice el teorema 6.

35. *Sugerencia:* Aplique el teorema 6 y algún otro.

37. $\det AB = \det \begin{bmatrix} 6 & 0 \\ 17 & 4 \end{bmatrix} = 24$;
$(\det A)(\det B) = 3 \cdot 8 = 24$

39. ***a***) 212 ***b***) 500 ***c***) -3 ***d***) $\frac{1}{4}$ ***e***) 64

41. $\det A = (a + e)d - (b + f)c = ad + ed - bc - fc$
$= (ad - bc) + (ed - fc) = \det B + \det C$

43. *Sugerencia:* Calcule $\det A$ con desarrollo por cofactores a lo largo de la columna 3.

45. [**M**] Exprese su conjetura sobre A^TA y AA^T.

Sección 3.3, página 184

1. $\begin{bmatrix} 5/6 \\ -1/6 \end{bmatrix}$ **3.** $\begin{bmatrix} 4 \\ 5/2 \end{bmatrix}$ **5.** $\begin{bmatrix} 3/2 \\ 4 \\ -7/2 \end{bmatrix}$

7. $s \neq \pm\sqrt{3}$; $x_1 = \dfrac{5s + 4}{6(s^2 - 3)}$, $x_2 = \dfrac{-4s - 15}{4(s^2 - 3)}$

9. $s \neq 0, -1$; $x_1 = \dfrac{1}{3(s + 1)}$, $x_2 = \dfrac{4s + 3}{6s(s + 1)}$

11. $\operatorname{adj} A = \begin{bmatrix} 0 & 1 & 0 \\ -3 & -1 & -3 \\ 3 & 2 & 6 \end{bmatrix}$, $A^{-1} = \dfrac{1}{3}\begin{bmatrix} 0 & 1 & 0 \\ -3 & -1 & -3 \\ 3 & 2 & 6 \end{bmatrix}$

13. $\operatorname{adj} A = \begin{bmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{bmatrix}$, $A^{-1} = \dfrac{1}{6}\begin{bmatrix} -1 & -1 & 5 \\ 1 & -5 & 1 \\ 1 & 7 & -5 \end{bmatrix}$

15. $\operatorname{adj} A = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ -1 & -9 & 3 \end{bmatrix}$, $A^{-1} = \dfrac{1}{6}\begin{bmatrix} 2 & 0 & 0 \\ 2 & 6 & 0 \\ -1 & -9 & 3 \end{bmatrix}$

17. Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$. La matriz adjunta es la transpuesta de la matriz de cofactores:

$$\operatorname{adj} A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Siguiendo al teorema 8, divida entre det A; esto produce la fórmula de la sección 2.2.

19. 8 **21.** 14 **23.** 22

25. Una matriz A de 3×3 no es invertible si y solo si sus columnas son linealmente dependientes (de acuerdo con el teorema de la matriz invertible). Esto ocurre si y solo si una de las columnas está en el plano generado por las otras dos columnas, lo cual es equivalente a la condición de que el paralelepípedo determinado por esas columnas tiene volumen cero, lo cual, a la vez, equivale a la condición $\det A = 0$.

27. 24 **29.** $\frac{1}{2}|\det[\,\mathbf{v}_1 \quad \mathbf{v}_2\,]|$

31. ***a)*** Véase el ejemplo 5. ***b)*** $4\pi abc/3$

33. [**M**] En MATLAB, las entradas en $B - \text{inv}(A)$ son aproximadamente 10^{-15} o más pequeñas.

35. [**M**] La versión estudiantil 4.0 de MATLAB utiliza 57,771 flops para inv(A), y 14,269,045 flops para la fórmula inversa. El comando `inv(A)` solamente requiere cerca del 0.4% de las operaciones para la fórmula inversa.

Capítulo 3 Ejercicios complementarios, página 185

1. ***a)*** V ***b)*** V ***c)*** F ***d)*** F ***e)*** F ***f)*** F
g) V ***h)*** V ***i)*** F ***j)*** F ***k)*** V ***l)*** F
m) F ***n)*** V ***o)*** F ***p)*** V

La solución del ejercicio 3 se basa en el hecho de que si una matriz contiene dos filas (o dos columnas) que son múltiplos entre sí, entonces, de acuerdo con el teorema 4, el determinante de la matriz es cero, ya que la matriz no puede ser invertible.

3. Realice dos operaciones de remplazo de filas, y después factorice un múltiplo común en las filas 2 y 3.

$$\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix} = \begin{vmatrix} 1 & a & b+c \\ 0 & b-a & a-b \\ 0 & c-a & a-c \end{vmatrix}$$

$$= (b-a)(c-a)\begin{vmatrix} 1 & a & b+c \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{vmatrix}$$

$$= 0$$

5. -12

7. Cuando el determinante se desarrolle por cofactores de la primera fila, la ecuación tiene la forma $ax + by + c = 0$, donde al menos a o b no valen cero. Esta es la ecuación de una recta. Es claro que (x_1, y_1) y (x_2, y_2) están sobre la recta, porque cuando las coordenadas de uno de los puntos se sustituyen por x y y, dos filas de la matriz son iguales y, de esta forma, el determinante es cero.

9. $T \sim \begin{bmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{bmatrix}$. Así, de acuerdo con el teorema 3,

$$\det T = (b-a)(c-a)\det\begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 1 & c+a \end{bmatrix}$$

$$= (b-a)(c-a)\det\begin{bmatrix} 1 & a & a^2 \\ 0 & 1 & b+a \\ 0 & 0 & c-b \end{bmatrix}$$

$$= (b-a)(c-a)(c-b)$$

11. Área = 12. Si se resta un vértice de los cuatro vértices, y si los nuevos vértices son $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$, entonces la figura trasladada (y por lo tanto, la figura original) será un paralelogramo si y solo si uno de los vectores $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ es la suma de los otros dos vectores.

13. Según la fórmula inversa, $(\text{adj } A) \cdot \dfrac{1}{\det A} A = A^{-1}A = I$.

Según el teorema de la matriz invertida, adj A es invertible y $(\text{adj } A)^{-1} = \dfrac{1}{\det A} A$.

15. ***a)*** $X = CA^{-1}$, $Y = D - CA^{-1}B$. Ahora utilice el ejercicio 14*c*).

b) A partir del inciso *a*), y de la propiedad multiplicativa de los determinantes,

$$\begin{aligned} \det\begin{bmatrix} A & B \\ C & D \end{bmatrix} &= \det[\,A(D - CA^{-1}B)\,] \\ &= \det[\,AD - ACA^{-1}B\,] \\ &= \det[\,AD - CAA^{-1}B\,] \\ &= \det[\,AD - CB\,] \end{aligned}$$

donde la igualdad $AC = CA$ se aplicó en el tercer paso.

17. Primero considere el caso $n = 2$, y demuestre que el resultado es válido mediante cálculo directo de los determinantes de B y C. Ahora, suponga que la fórmula es válida para todas las matrices $(k-1) \times (k-1)$, y sean A, B y C matrices de $k \times k$. Utilice un desarrollo por factores a lo largo la primera columna y la hipótesis inductiva para obtener det B. En C aplique operaciones de remplazo de filas para generar ceros debajo del primer pivote y así producir una matriz triangular. Encuentre el determinante de esta matriz y súmelo a det B para obtener el resultado.

19. [**M**] Calcule:

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{vmatrix} = 1, \quad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{vmatrix} = 1,$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 \\ 1 & 2 & 3 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{vmatrix} = 1$$

Suposición:

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & & 2 \\ 1 & 2 & 3 & & 3 \\ \vdots & & & \ddots & \vdots \\ 1 & 2 & 3 & \dots & n \end{vmatrix} = 1$$

Para confirmar la suposición, utilice operaciones de remplazo de filas para generar ceros bajo el primer pivote, después bajo el segundo pivote, y así sucesivamente. La matriz resultante es

$$\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & & 1 \\ 0 & 0 & 1 & & 1 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}$$

que es una matriz triangular superior con determinante 1.

Capítulo 4

Sección 4.1, página 195

1. ***a*)** $\mathbf{u} + \mathbf{v}$ está en V porque sus entradas son no negativas.

***b*)** *Ejemplo:* Si $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ y $c = -1$, entonces $\mathbf{u}$ está en V, pero $c\mathbf{u}$ no está en V.

3. *Ejemplo*: Si $\mathbf{u} = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$ y $c = 4$, entonces $\mathbf{u}$ está en H, pero $c\mathbf{u}$ no está en H.

5. Sí, de acuerdo con el teorema 1, porque el conjunto es Gen $\{t^2\}$.

7. No, el conjunto no es cerrado bajo multiplicación por escalares que no sean enteros.

9. $H = \text{Gen}\,\{\mathbf{v}\}$, donde $\mathbf{v} = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}$. Según el teorema 1, H es un subespacio de $\mathbb{R}^3$.

11. $W = \text{Gen}\,\{\mathbf{u}, \mathbf{v}\}$, donde $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$. Según el teorema 1, W es un subespacio de $\mathbb{R}^3$.

13. ***a*)** Solo existen tres vectores en $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, y $\mathbf{w}$ no es uno de ellos.

***b*)** Hay una infinidad de vectores en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

***c*)** $\mathbf{w}$ está en Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ porque $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$.

15. W no es un espacio vectorial porque el vector cero no está en W.

17. $S = \left\{ \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 3 \\ 0 \end{bmatrix} \right\}$

19. *Sugerencia:* Utilice el teorema 1.

21. Sí. Evidentemente, se satisfacen las condiciones para un subespacio: la matriz cero está en H, la suma de dos matrices triangulares superiores es triangular superior, y cualquier múltiplo escalar de una matriz triangular superior es otra vez triangular superior.

25. 4 **27.** ***a*)** 8 ***b*)** 3 ***c*)** 5 ***d*)** 4

29.
$$\begin{aligned} \mathbf{u} + (-1)\mathbf{u} &= 1\mathbf{u} + (-1)\mathbf{u} && \text{Axioma 10} \\ &= [1 + (-1)]\mathbf{u} && \text{Axioma 8} \\ &= 0\mathbf{u} = \mathbf{0} && \text{Ejercicio 27} \end{aligned}$$

Del ejercicio 26, se deduce que $(-1)\mathbf{u} = -\mathbf{u}$.

31. Cualquier subespacio H que contiene a $\mathbf{u}$ y $\mathbf{v}$ también debe contener todos los múltiplos escalares de $\mathbf{u}$ y $\mathbf{v}$; por lo tanto, debe contener todas las sumas de los múltiplos escalares de $\mathbf{u}$ y de $\mathbf{v}$. Por consiguiente, H debe contener a Gen $\{\mathbf{u}, \mathbf{v}\}$.

33. *Sugerencia:* Para una parte de la solución, considere $\mathbf{w}_1$ y $\mathbf{w}_2$ en $H + K$, y escriba $\mathbf{w}_1$ y $\mathbf{w}_2$ en la forma $\mathbf{w}_1 = \mathbf{u}_1 + \mathbf{v}_1$ y $\mathbf{w}_2 = \mathbf{u}_2 + \mathbf{v}_2$, donde $\mathbf{u}_1$ y $\mathbf{u}_2$ están en H, y $\mathbf{v}_1$ y $\mathbf{v}_2$ están en K.

35. [**M**] La forma escalonada reducida de $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{w}]$ muestra que $\mathbf{w} = \mathbf{v}_1 - 2\mathbf{v}_2 + \mathbf{v}_3$. Por lo tanto, $\mathbf{w}$ está en el subespacio generado por $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$.

37. [**M**] Las funciones son $\cos(4t)$ y $\cos(6t)$. Véase el ejercicio 34 de la sección 4.5.

Sección 4.2, página 205

1. $\begin{bmatrix} 3 & -5 & -3 \\ 6 & -2 & 0 \\ -8 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$,

de manera que $\mathbf{w}$ está en Nul A.

3. $\begin{bmatrix} 2 \\ -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ 0 \\ 1 \end{bmatrix}$ **5.** $\begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 5 \\ 1 \\ 0 \end{bmatrix}$

7. W no es un subespacio de $\mathbb{R}^3$ ya que el vector cero $(0, 0, 0)$ no está en W.

9. W es un subespacio de $\mathbb{R}^4$ porque W es el conjunto de soluciones del sistema

$$\begin{aligned} p - 3q - 4s \quad &= 0 \\ 2p \quad - \quad s - 5r &= 0 \end{aligned}$$

11. W no es un subespacio porque $\mathbf{0}$ no está en W. *Justificación:* Si un elemento típico $(s - 2t, 3 + 3s, 3s + t, 2s)$ fuera cero, entonces $3 + 3s = 0$ y $2s = 0$, lo que es imposible.

13. $W = \text{Col}\, A$ para $A = \begin{bmatrix} 1 & -6 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$, así que W es un espacio vectorial, de acuerdo con el teorema 3.

15. $\begin{bmatrix} 0 & 2 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & 0 \\ 2 & -1 & -1 \end{bmatrix}$

17. *a*) 2 *b*) 4 **19.** *a*) 5 *b*) 2

21. El vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ está en Nul A y el vector $\begin{bmatrix} 6 \\ -3 \\ -9 \\ 9 \end{bmatrix}$ está en Col A.

Son posibles otras respuestas.

23. **w** está tanto en Nul A como en Col A.

27. Sean $\mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$ y $A = \begin{bmatrix} 1 & -3 & -3 \\ -2 & 4 & 2 \\ -1 & 5 & 7 \end{bmatrix}$. Así, **x** está en Nul A, y como este es un subespacio de $\mathbb{R}^3$, entonces $10\mathbf{x}$ se encuentra en Nul A.

29. *a*) $A\mathbf{0} = \mathbf{0}$, por lo que el vector cero está en Col A.

b) Por una propiedad de la multiplicación matricial, $A\mathbf{x} + A\mathbf{w} = A(\mathbf{x} + \mathbf{w})$, lo que indica que $A\mathbf{x} + A\mathbf{w}$ es una combinación lineal de las columnas de A y, por lo tanto, está en Col A.

c) $c(A\mathbf{x}) = A(c\mathbf{x})$, lo que muestra que $c(A\mathbf{x})$ se encuentra en Col A para todos los escalares c.

31. *a*) Para polinomios arbitrarios **p**, **q** en $\mathbb{P}_2$ y cualquier escalar c,

$$T(\mathbf{p}+\mathbf{q}) = \begin{bmatrix} (\mathbf{p}+\mathbf{q})(0) \\ (\mathbf{p}+\mathbf{q})(1) \end{bmatrix} = \begin{bmatrix} \mathbf{p}(0)+\mathbf{q}(0) \\ \mathbf{p}(1)+\mathbf{q}(1) \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix} + \begin{bmatrix} \mathbf{q}(0) \\ \mathbf{q}(1) \end{bmatrix} = T(\mathbf{p}) + T(\mathbf{q})$$
$$T(c\mathbf{p}) = \begin{bmatrix} c\mathbf{p}(0) \\ c\mathbf{p}(1) \end{bmatrix} = c\begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix} = cT(\mathbf{p})$$

Así, T es una transformación lineal de $\mathbb{P}_2$ en $\mathbb{P}_2$.

b) Cualquier polinomio cuadrático que se anula en 0 y 1 debe ser múltiplo de $\mathbf{p}(t) = t(t-1)$. El rango de T es $\mathbb{R}^2$.

33. *a*) Para A, B en $M_{2\times 2}$ y cualquier escalar c,

$$\begin{aligned} T(A+B) &= (A+B) + (A+B)^T \\ &= A + B + A^T + B^T \quad \text{Propiedad de la transpuesta} \\ &= (A + A^T) + (B + B^T) = T(A) + T(B) \\ T(cA) &= (cA) + (cA)^T = cA + cA^T \\ &= c(A + A^T) = cT(A) \end{aligned}$$

Así, T es una transformación lineal de $M_{2\times 2}$ en $M_{2\times 2}$.

b) Si B es cualquier elemento en $M_{2\times 2}$ con la propiedad $B^T = B$, y si $A = B$, entonces,

$$T(A) = \tfrac{1}{2}B + \left(\tfrac{1}{2}B\right)^T = \tfrac{1}{2}B + \tfrac{1}{2}B = B$$

c) El inciso *b*) demostró que el rango de T contiene todas las B tales que $B^T = B$. Así, es suficiente demostrar que cualquier B en el rango de T tiene esta propiedad. Si $B = T(A)$, entonces, de acuerdo con las propiedades de las transpuestas,

$$B^T = (A + A^T)^T = A^T + A^{TT} = A^T + A = B$$

d) El núcleo de T es $\left\{\begin{bmatrix} 0 & b \\ -b & 0 \end{bmatrix} : b \text{ real}\right\}$.

35. *Sugerencia:* Compruebe las tres condiciones para un subespacio. Elementos típicos de $T(U)$ tienen la forma $T(\mathbf{u}_1)$ y $T(\mathbf{u}_2)$, donde $\mathbf{u}_1$, $\mathbf{u}_2$ están en U.

37. [**M**] **w** está en Col A, pero no en Nul A. (Explique por qué).

39. [**M**] La forma escalonada reducida de A es

$$\begin{bmatrix} 1 & 0 & 1/3 & 0 & 10/3 \\ 0 & 1 & 1/3 & 0 & -26/3 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Sección 4.3, página 213

1. La matriz de 3×3 $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ tiene tres posiciones pivote. De acuerdo con el teorema de la matriz invertible, A es invertible y sus columnas forman una base para $\mathbb{R}^3$. (Véase el ejemplo 3).

3. Este conjunto no forma una base para $\mathbb{R}^3$. El conjunto es linealmente dependiente y no genera $\mathbb{R}^3$.

5. Este conjunto no forma una base para $\mathbb{R}^3$. El conjunto es linealmente dependiente porque el vector cero está en el conjunto. Sin embargo,

$$\begin{bmatrix} 3 & -3 & 0 & 0 \\ -3 & 7 & 0 & -3 \\ 0 & 0 & 0 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 4 & 0 & -3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$$

La matriz tiene un pivote en cada fila y, por lo tanto, sus columnas generan $\mathbb{R}^3$.

7. Este conjunto no forma una base para $\mathbb{R}^3$. El conjunto es linealmente independiente porque un vector no es un múltiplo del otro. Sin embargo, los vectores no generan $\mathbb{R}^3$. La matriz $\begin{bmatrix} -2 & 6 \\ 3 & -1 \\ 0 & 5 \end{bmatrix}$ puede tener a lo sumo dos pivotes ya que solo tiene dos columnas. Por lo tanto, no habrá un pivote en cada fila.

9. $\begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$ **11.** $\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$

13. Base para Nul A: $\begin{bmatrix} -6 \\ -5/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -3/2 \\ 0 \\ 1 \end{bmatrix}$

Base para Col A: $\begin{bmatrix} -2 \\ 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \\ 8 \end{bmatrix}$

15. $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_5\}$ **17.** [M] $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_5\}$

19. Las tres respuestas más sencillas son $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_1, \mathbf{v}_3\}$ y $\{\mathbf{v}_2, \mathbf{v}_3\}$. Son posibles otras respuestas.

23. *Sugerencia:* Utilice el teorema de matriz invertible.

25. No. (¿Por qué el conjunto no es una base para H?).

27. $\{\cos \omega t, \operatorname{sen} \omega t\}$

29. Sea A la matriz de $n \times k$ $[\mathbf{v}_1 \cdots \mathbf{v}_k]$. Como A tiene menos columnas que filas, entonces no puede existir una posición pivote en cada fila de A. De acuerdo con el teorema 4 de la sección 1.4, las columnas de A no generan $\mathbb{R}^n$ y, por lo tanto, no son una base para $\mathbb{R}^n$.

31. *Sugerencia:* Si $\{\mathbf{v}_1,\ldots, \mathbf{v}_p\}$ es linealmente dependiente, entonces existen $c_1,\ldots, c_p$, no todos cero, tales que $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$. Utilice esta ecuación.

33. Ningún polinomio es un múltiplo del otro polinomio, de manera que $\{\mathbf{p}_1, \mathbf{p}_2\}$ es un conjunto linealmente independiente en $\mathbb{P}_3$.

35. Sea $\{\mathbf{v}_1, \mathbf{v}_3\}$ cualquier conjunto linealmente independiente en el espacio vectorial V, y sean $\mathbf{v}_2$ y $\mathbf{v}_4$ una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_3$. Entonces, $\{\mathbf{v}_1, \mathbf{v}_3\}$ es una base para Gen $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

37. [**M**] Hay que ser hábil para encontrar valores especiales de t que den varios ceros en (5), y después crear un sistema de ecuaciones que se pueda resolver fácilmente a mano. O bien, se podrían usar valores de t tales como $t = 0, .1, .2,\ldots$ para crear un sistema de ecuaciones que pueda resolverse con un programa de matrices.

Sección 4.4, página 222

1. $\begin{bmatrix} 3 \\ -7 \end{bmatrix}$ **3.** $\begin{bmatrix} -7 \\ 4 \\ 3 \end{bmatrix}$ **5.** $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ **7.** $\begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 2 \\ -3 & -5 \end{bmatrix}$ **11.** $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ **13.** $\begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix}$

17. $\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5\mathbf{v}_1 - 2\mathbf{v}_2 = 10\mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3 = -\mathbf{v}_2 - \mathbf{v}_3$

(un número infinito de respuestas)

19. *Sugerencia:* Por hipótesis, el vector cero tiene una única representación como una combinación lineal de elementos de S.

21. $\begin{bmatrix} 9 & 2 \\ 4 & 1 \end{bmatrix}$

23. *Sugerencia:* Suponga que $[\mathbf{u}]_{\mathcal{B}} = [\mathbf{w}]_{\mathcal{B}}$ para algunas $\mathbf{u}$ y $\mathbf{w}$ en V, y denote las entradas en $[\mathbf{u}]_{\mathcal{B}}$ como $c_1,\ldots, c_n$. Utilice la definición de $[\mathbf{u}]_{\mathcal{B}}$.

25. Un posible método: Primero, demuestre que si $\mathbf{u}_1,\ldots, \mathbf{u}_p$ son linealmente *dependientes*, entonces $[\mathbf{u}_1]_{\mathcal{B}},\ldots, [\mathbf{u}_p]_{\mathcal{B}}$ son linealmente dependientes. Segundo, demuestre que si $[\mathbf{u}_1]_{\mathcal{B}},\ldots, [\mathbf{u}_p]_{\mathcal{B}}$ son linealmente dependientes, entonces $\mathbf{u}_1,\ldots, \mathbf{u}_p$ son linealmente *dependientes*. Utilice las dos ecuaciones que se presentan en el ejercicio.

27. Linealmente independiente. (Justifique las respuestas a los ejercicios 27 a 34).

29. Linealmente dependiente.

31. *a*) Los vectores de coordenadas $\begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \\ -7 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ no generan $\mathbb{R}^3$. Debido al isomorfismo entre $\mathbb{R}^3$ y $\mathbb{P}_2$, los polinomios correspondientes no generan $\mathbb{P}_2$.

b) Los vectores de coordenadas $\begin{bmatrix} 0 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -8 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix}$ generan $\mathbb{R}^3$. Debido al isomorfismo entre $\mathbb{R}^3$ y $\mathbb{P}_2$, los polinomios correspondientes generan $\mathbb{P}_2$.

33. [**M**] Los vectores de coordenadas $\begin{bmatrix} 3 \\ 7 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 16 \\ -6 \\ 2 \end{bmatrix}$ son un subconjunto linealmente dependiente de $\mathbb{R}^4$. Debido al isomorfismo entre $\mathbb{R}^4$ y $\mathbb{P}_3$, los polinomios correspondientes forman un subconjunto linealmente dependiente de $\mathbb{P}_3$ y, por lo tanto, no forman una base para $\mathbb{P}_3$.

35. [**M**] $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -5/3 \\ 8/3 \end{bmatrix}$ **37.** [**M**] $\begin{bmatrix} 1.3 \\ 0 \\ 0.8 \end{bmatrix}$

Sección 4.5, página 229

1. $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$; dim es 2

3. $\begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -3 \\ 0 \end{bmatrix}$; dim es 3

5. $\begin{bmatrix} 1 \\ 2 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 2 \\ 6 \end{bmatrix}$; dim es 3

7. No hay base; dim es 0 **9.** 2 **11.** 3 **13.** 2, 3

15. 2, 3 **17.** 0, 3

21. *Sugerencia:* Solo se necesita demostrar que los primeros cuatro polinomios de Hermite son linealmente independientes. ¿Por qué?

23. $[\mathbf{p}]_{\mathcal{B}} = (3, 6, 2, 1)$

25. *Sugerencia:* Suponga que S genera V, y utilice el teorema del conjunto generador. Esto conduce a una contradicción, lo que demuestra que es falsa la hipótesis de generación.

27. *Sugerencia:* Considere el hecho de que cada $\mathbb{P}_n$ es un subespacio de $\mathbb{P}$.

29. Justifique cada respuesta:

***a*)** Verdadero ***b*)** Verdadero ***c*)** Verdadero

31. *Sugerencia:* Como H es un subespacio diferente de cero de un espacio de dimensión finita, entonces H también es de dimensión finita y tiene una base, por ejemplo, $\mathbf{v}_1,\ldots,\mathbf{v}_p$. Primero demuestre que $\{T(\mathbf{v}_1),\ldots,T(\mathbf{v}_p)\}$ genera $T(H)$.

33. **[M]** ***a*)** Una base es $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_2, \mathbf{e}_3\}$. De hecho, cualesquiera dos de los vectores $\mathbf{e}_2,\ldots,\mathbf{e}_5$ extenderá $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a una base de $\mathbb{R}^5$.

Sección 4.6, página 236

1. rango $A = 2$; dim Nul $A = 2$;

Base para Col A: $\begin{bmatrix} 1 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \\ -6 \end{bmatrix}$

Base para Fil A: $(1, 0, -1, 5)$, $(0, -2, 5, -6)$

Base para Nul A: $\begin{bmatrix} 1 \\ 5/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ -3 \\ 0 \\ 1 \end{bmatrix}$

3. rango $A = 3$; dim Nul $A = 3$;

Base para Col A: $\begin{bmatrix} 2 \\ -2 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 6 \\ -3 \\ 9 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 3 \\ 3 \end{bmatrix}$

Base para Nul A: $(2, 6, -6, 6, 3, 6)$, $(0, 3, 0, 3, 3, 0)$, $(0, 0, 0, 0, 3, 0)$

Base para Fil A: $\begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

5. 4, 3, 3

7. Sí; no. Como Col A es un subespacio de dimensión 4 de $\mathbb{R}^4$, entonces coincide con $\mathbb{R}^4$. El espacio nulo no puede ser $\mathbb{R}^3$, porque los vectores en Nul A tienen 7 entradas. Nul A es un subespacio tridimensional de $\mathbb{R}^7$, de acuerdo con el teorema del rango.

9. 3, no. Observe que las columnas de una matriz de 4×6 están en $\mathbb{R}^4$, más que en $\mathbb{R}^3$. Col A es un subespacio tridimensional de $\mathbb{R}^4$.

11. 2

13. 5, 5. En ambos casos, el número de pivotes no puede exceder el número de columnas o el número de filas.

15. 4

19. Sí. Intente escribir una explicación.

21. No. Explique por qué.

23. Sí. Solo se necesitan seis ecuaciones lineales homogéneas.

25. No. Explique por qué.

27. Fil A y Nul A están en $\mathbb{R}^n$; Col A y Nul A^T están en $\mathbb{R}^m$. Solo hay cuatro subespacios distintos porque Fil A^T = Col A y Col A^T = Fil A.

29. Recuerde que dim Col $A = m$ precisamente cuando Col $A = \mathbb{R}^m$ o, de manera equivalente, cuando la ecuación $A\mathbf{x} = \mathbf{b}$ es consistente para toda $\mathbf{b}$. De acuerdo con el ejercicio 28*b*), dim Col $A = m$ precisamente cuando dim Nul $A^T = 0$ o, de manera equivalente, cuando la ecuación $A^T\mathbf{x} = \mathbf{0}$ solo tiene la solución trivial.

31. $\mathbf{u}\mathbf{v}^T = \begin{bmatrix} 2a & 2b & 2c \\ -3a & -3b & -3c \\ 5a & 5b & 5c \end{bmatrix}$. Todas las columnas son múltiplos de $\mathbf{u}$, de manera que Col $\mathbf{u}\mathbf{v}^T$ es unidimensional, a menos que $a = b = c = 0$.

33. *Sugerencia:* Sea $A = [\mathbf{u} \quad \mathbf{u}_2 \quad \mathbf{u}_3]$. Si $\mathbf{u} \neq \mathbf{0}$, entonces $\mathbf{u}$ es una base para Col A. ¿Por qué?

35. **[M]** *Sugerencia:* Véase el ejercicio 28 y las observaciones antes del ejemplo 4.

37. **[M]** Las matrices C y R para el ejercicio 35 funcionan aquí, y $A = CR$.

Sección 4.7, página 242

1. a. $\begin{bmatrix} 6 & 9 \\ -2 & -4 \end{bmatrix}$ **b.** $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$

3. (ii)

5. a. $\begin{bmatrix} 4 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{bmatrix}$ **b.** $\begin{bmatrix} 8 \\ 2 \\ 2 \end{bmatrix}$

7. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} -3 & 1 \\ -5 & 2 \end{bmatrix}$, $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} -2 & 1 \\ -5 & 3 \end{bmatrix}$

9. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 2 & 3 \\ 0 & -1 \end{bmatrix}$, $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \frac{1}{2}\begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix}$

13. $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} 1 & 3 & 0 \\ -2 & -5 & 2 \\ 1 & 4 & 3 \end{bmatrix}$, $[-1 + 2t]_{\mathcal{B}} = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$

15. ***a*)** $\mathcal{B}$ es una base para V.

***b*)** El mapeo de coordenadas es una transformación lineal.

***c*)** El producto de una matriz y un vector.

***d*)** El vector de coordenadas de $\mathbf{v}$ respecto de $\mathcal{B}$.

17. ***a*)** **[M]**

$$P^{-1} = \frac{1}{32}\begin{bmatrix} 32 & 0 & 16 & 0 & 12 & 0 & 10 \\ & 32 & 0 & 24 & 0 & 20 & 0 \\ & & 16 & 0 & 16 & 0 & 15 \\ & & & 8 & 0 & 10 & 0 \\ & & & & 4 & 0 & 6 \\ & & & & & 2 & 0 \\ & & & & & & 1 \end{bmatrix}$$

b) $\cos^2 t = (1/2)[1 + \cos 2t]$
$\cos^3 t = (1/4)[3\cos t + \cos 3t]$
$\cos^4 t = (1/8)[3 + 4\cos 2t + \cos 4t]$
$\cos^5 t = (1/16)[10\cos t + 5\cos 3t + \cos 5t]$
$\cos^6 t = (1/32)[10 + 15\cos 2t + 6\cos 4t + \cos 6t]$

19. [**M**] *Sugerencia:* Sea C la base $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Entonces las columnas de P son $[\mathbf{u}_1]_{\mathcal{C}}$, $[\mathbf{u}_2]_{\mathcal{C}}$ y $[\mathbf{u}_3]_{\mathcal{C}}$. Utilice la definición de vectores de $\mathcal{C}$-coordenadas y álgebra matricial para calcular $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$. A continuación se presentan las respuestas numéricas:

a) $\mathbf{u}_1 = \begin{bmatrix} -6 \\ -5 \\ 21 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -6 \\ -9 \\ 32 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} -5 \\ 0 \\ 3 \end{bmatrix}$

b) $\mathbf{w}_1 = \begin{bmatrix} 28 \\ -9 \\ -3 \end{bmatrix}$, $\mathbf{w}_2 = \begin{bmatrix} 38 \\ -13 \\ 2 \end{bmatrix}$, $\mathbf{w}_3 = \begin{bmatrix} 21 \\ -7 \\ 3 \end{bmatrix}$

Sección 4.8, página 251

1. Si $y_k = 2^k$, entonces $y_{k+1} = 2^{k+1}$ y $y_{k+2} = 2^{k+2}$.
Al sustituir estas fórmulas en el lado izquierdo de la ecuación se obtiene

$$\begin{aligned} y_{k+2} + 2y_{k+1} - 8y_k &= 2^{k+2} + 2 \cdot 2^{k+1} - 8 \cdot 2^k \\ &= 2^k(2^2 + 2 \cdot 2 - 8) \\ &= 2^k(0) = 0 \quad \text{para toda } k \end{aligned}$$

Puesto que la ecuación en diferencias es válida para toda k, entonces 2^k es una solución. Un cálculo similar funciona para $y_k = (-4)^k$.

3. Las señales 2^k y $(-4)^k$ son linealmente independientes porque ninguna es múltiplo de la otra. Por ejemplo, no existe un escalar c tal que $2^k = c(-4)^k$ *para toda* k. De acuerdo con el teorema 17, el conjunto solución H de la ecuación en diferencias del ejercicio 1 es bidimensional. De acuerdo con el teorema de la base de la sección 4.5, las dos señales linealmente independientes 2^k y $(-4)^k$ forman una base para H.

5. Si $y_k = (-2)^k$, entonces

$$\begin{aligned} y_{k+2} + 4y_{k+1} + 4y_k &= (-2)^{k+2} + 4(-2)^{k+1} + 4(-2)^k \\ &= (-2)^k[(-2)^2 + 4(-2) + 4] \\ &= (-2)^k(0) = 0 \quad \text{para toda } k \end{aligned}$$

De manera similar, si $y_k = k(-2)^k$, entonces

$$\begin{aligned} & y_{k+2} + 4y_{k+1} + 4y_k \\ &\quad = (k+2)(-2)^{k+2} + 4(k+1)(-2)^{k+1} + 4k(-2)^k \\ &\quad = (-2)^k[(k+2)(-2)^2 + 4(k+1)(-2) + 4k] \\ &\quad = (-2)^k[4k + 8 - 8k - 8 + 4k] \\ &\quad = (-2)^k(0) \quad \text{para toda } k \end{aligned}$$

Así, $(-2)^k$ y $k(-2)^k$ están en el espacio solución H de la ecuación en diferencias. Además, no existe escalar c tal que $k(-2)^k = c(-2)^k$ *para toda* k, porque c se debe seleccionar independientemente de k. De igual manera, no existe un escalar c tal que $(-2)^k = ck(-2)^k$ para toda k. Así, ambas señales son linealmente independientes. Como $\dim H = 2$, entonces las señales forman una base para H, de acuerdo con el teorema de la base.

7. Sí **9.** Sí

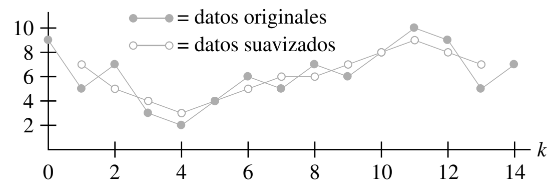
11. No, dos señales no pueden generar el espacio solución tridimensional.

13. $\left(\frac{1}{3}\right)^k$, $\left(\frac{2}{3}\right)^k$ **15.** $\left(\frac{1}{2}\right)^k$, $\left(-\frac{2}{3}\right)^k$

17. $Y_k = c_1(.8)^k + c_2(.5)^k + 10 \to 10$ conforme $k \to \infty$

19. $y_k = c_1(-2 + \sqrt{3})^k + c_2(-2 - \sqrt{3})^k$

21. 7, 5, 4, 3, 4, 5, 6, 6, 7, 8, 9, 8, 7; véase la figura:



23. *a*) $y_{k+1} - 1.01y_k = -450$, $y_0 = 10{,}000$

25. $k^2 + c_1 \cdot (-4)^k + c_2$ **27.** $-2 + k + c_1 \cdot 2^k + c_2 \cdot (-2)^k$

29. $\mathbf{x}_{k+1} = A\mathbf{x}_k$, donde

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -6 & 8 & -3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} y_k \\ y_{k+1} \\ y_{k+2} \\ y_{k+3} \end{bmatrix}$$

31. La ecuación es válida para toda k, así que esta vale para $k - 1$ en lugar de k, lo que transforma la ecuación en

$y_{k+2} + 5y_{k+1} + 6y_k = 0 \quad$ para toda k

La ecuación es de orden 2.

33. Para toda k, la matriz de Casorati $C(k)$ no es invertible. En este caso, la matriz de Casorati no da información acerca de la independencia o dependencia lineal del conjunto de señales. De hecho, ninguna señal es un múltiplo de la otra, así que son linealmente independientes.

35. *Sugerencia:* Compruebe las dos propiedades que definen una transformación lineal. Para $\{y_k\}$ y $\{z_k\}$ en $\mathbb{S}$, estudie $T(\{y_k\} + \{z_k\})$. Observe que si r es cualquier escalar, entonces el k-ésimo término de $r\{y_k\}$ es ry_k; así, $T(r\{y_k\})$ es la secuencia $\{w_k\}$ dada por

$$w_k = ry_{k+2} + a(ry_{k+1}) + b(ry_k)$$

37. *Sugerencia:* Encuentre $TD(y_0, y_1, y_2, \ldots)$ y $DT(y_0, y_1, y_2, \ldots)$.

Sección 4.9, página 260

1. *a*) De:

$$\begin{array}{cl} \begin{array}{cc} \text{N} & \text{M} \end{array} & \text{A:} \\ \begin{bmatrix} .7 & .6 \\ .3 & .4 \end{bmatrix} & \begin{array}{l} \text{Noticias} \\ \text{Música} \end{array} \end{array}$$

b) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *c*) 33%

3. *a*) De:

$$\begin{array}{cl} \begin{array}{cc} \text{H} & \text{I} \end{array} & \text{A:} \\ \begin{bmatrix} .95 & .45 \\ .05 & .55 \end{bmatrix} & \begin{array}{l} \text{Sano} \\ \text{Enfermo} \end{array} \end{array}$$

b) 15%, 12.5%

c) .925; utilice $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

5. $\begin{bmatrix} 5/14 \\ 9/14 \end{bmatrix}$ **7.** $\begin{bmatrix} 1/4 \\ 1/2 \\ 1/4 \end{bmatrix}$

9. Sí, porque todas las entradas de P^2 son positivas.

11. *a*) $\begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$ *b*) 2/3

13. *a*) $\begin{bmatrix} .9 \\ .1 \end{bmatrix}$ *b*) .10, no

15. [**M**] Aproximadamente el 13.9% de la población de Estados Unidos.

17. *a*) Las entradas en una columna de P suman 1. Una columna en la matriz $P - I$ tiene las mismas entradas que en P excepto que una de las entradas ha disminuido en 1. Por lo tanto, la suma de cada columna es 0.

b) De acuerdo con el inciso *a*), la fila inferior de $P - I$ es el negativo de la suma de las otras filas.

c) Según el inciso *b*) y el teorema del conjunto generador, la fila inferior de $P - I$ se puede eliminar y las restantes $(n - 1)$ filas continuarán generando el espacio fila. De forma alternativa, utilice el inciso *a*) y considere el hecho de que las operaciones fila no cambian el espacio fila. Sea A la matriz obtenida de $P - I$ sumando la fila inferior a todas las demás filas. Según el inciso *a*), el espacio fila se genera por las primeras $(n - 1)$ filas de A.

d) De acuerdo con el teorema del rango y el inciso *c*), la dimensión del espacio columna de $P - I$ es menor que n, y así el espacio nulo es no trivial. En vez del teorema del rango, se puede recurrir al teorema de la matriz invertida, ya que $P - I$ es una matriz cuadrada.

19. *a*) El producto $S\mathbf{x}$ es igual a la suma de las entradas en $\mathbf{x}$. Para un vector de probabilidad, esta suma debe ser 1.

b) $P = [\mathbf{p}_1 \quad \mathbf{p}_2 \cdots \mathbf{p}_n]$, donde las $\mathbf{p}_i$ son vectores de probabilidad. Mediante la multiplicación matricial y el inciso *a*),

$$SP = \begin{bmatrix} S\mathbf{p}_1 & S\mathbf{p}_2 & \cdots & S\mathbf{p}_n \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix} = S$$

c) Por el inciso *b*), $S(P\mathbf{x}) = (SP)\mathbf{x} = S\mathbf{x} = 1$. También, las entradas en $P\mathbf{x}$ son no negativas (porque P y $\mathbf{x}$ tienen entradas no negativas). Así, por el inciso *a*), $P\mathbf{x}$ es un vector de probabilidad.

Capítulo 4 Ejercicios complementarios, página 262

1. *a*) V *b*) V *c*) F *d*) F *e*) V *f*) V *g*) F *h*) F *i*) V *j*) F *k*) F *l*) F *m*) V *n*) F *o*) V *p*) V *q*) F *r*) V *s*) V *t*) F

3. El conjunto de todas las (b_1, b_2, b_3) que satisfacen $b_1 + 2b_2 + b_3 = 0$.

5. El vector $\mathbf{p}_1$ no es cero y $\mathbf{p}_2$ no es un múltiplo de $\mathbf{p}_1$, de manera que conserve ambos vectores. Como $\mathbf{p}_3 = 2\mathbf{p}_1 + 2\mathbf{p}_2$, elimine $\mathbf{p}_3$. Ya que $\mathbf{p}_4$ tiene un término t^2, no puede ser una combinación lineal de $\mathbf{p}_1$ y $\mathbf{p}_2$, entonces conserve $\mathbf{p}_4$. Finalmente, $\mathbf{p}_5 = \mathbf{p}_1 + \mathbf{p}_4$, entonces elimine $\mathbf{p}_5$. La base resultante es $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_4\}$.

7. Se debería saber que el conjunto solución del sistema homogéneo se genera con dos soluciones. En este caso, el espacio nulo de la matriz de coeficientes A de 18×20, a lo sumo, es bidimensional. De acuerdo con el teorema del rango, dim Col $A \geq 20 - 2 = 18$, lo que significa que Col $A = \mathbb{R}^{18}$, porque A tiene 18 filas, y cada ecuación $A\mathbf{x} = \mathbf{b}$ es consistente.

9. Sea A la matriz estándar de $m \times n$ de la transformación T.

a) Si T es uno a uno, entonces las columnas de A son linealmente independientes (teorema 12 de la sección 1.9), por lo que dim Nul $A = 0$. De acuerdo con el teorema del rango, dim Col A = rango $A = n$. Como el rango de T es Col A, la dimensión del rango de T es n.

b) Si T es sobre, entonces las columnas de A generan $\mathbb{R}^m$ (teorema 12 de la sección 1.9), de manera que dim Col $A = m$. Según el teorema del rango, dim Nul $A = n -$ dim Col $A = n - m$. Como el núcleo de T es Nul A, la dimensión del núcleo de T es $n - m$.

11. Si S es un conjunto generador finito para V, entonces un subconjunto de S (por ejemplo, S') es una base para V. Como S' debe generar a V, S' no puede ser un subconjunto propio de S debido a la minimalidad de S. Por lo tanto, $S' = S$, lo que demuestra que S es una base para V.

12. *a*) *Sugerencia:* Cualquier $\mathbf{y}$ en Col AB tiene la forma $\mathbf{y} = AB\mathbf{x}$ para alguna $\mathbf{x}$.

13. De acuerdo con el ejercicio 9, rango $PA \leq$ rango A, y rango A = rango $P^{-1}PA \leq$ rango PA. Por lo tanto, rango PA = rango A.

15. La ecuación $AB = 0$ indica que cada columna de B está en Nul A. Como Nul A es un subespacio, entonces todas las combinaciones lineales de las columnas de B están en Nul A, de manera que Col B es un subespacio de Nul A. Según el teorema 11 de la sección 4.5, dim Col $B \leq$ dim Nul A. Al aplicar el teorema del rango, se encuentra que

$$N = \text{rango } A + \dim \text{Nul } A \geq \text{rango } A + \text{rango } B$$

17. *a*) Suponga que A_1 consiste en las r columnas pivote en A. Las columnas de A_1 son linealmente independientes. Así, A_1 es una submatriz de $m \times r$ con rango r.

b) De acuerdo con el teorema del rango aplicado a A_1, la dimensión de Fil A es r, de manera que A_1 tiene r filas linealmente independientes. Utilícelas para formar A_2. Entonces, A_2 es de $r \times r$ con filas linealmente independientes. Según el teorema de la matriz invertida, A_2 es invertible.

19. $$\begin{bmatrix} B & AB & A^2B \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -.9 & .81 \\ 1 & .5 & .25 \end{bmatrix} \sim \begin{bmatrix} 1 & -.9 & .81 \\ 0 & 1 & 0 \\ 0 & 0 & -.56 \end{bmatrix}$$

Esta matriz tiene rango 3, así que el par (A, B) es controlable.

21. [**M**] rango $[B \quad AB \quad A^2B \quad A^3B] = 3$. El par (A, B) no es controlable.

Capítulo 5

Sección 5.1, página 271

1. Sí **3.** Sí, $\lambda = -2$ **5.** Sí, $\lambda = -5$

7. Sí, $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

9. $\lambda = 1$: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$; $\lambda = 3$: $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

11. $\lambda = -1$: $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$; $\lambda = 7$: $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$

13. $\lambda = 1$: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$; $\lambda = 2$: $\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$; $\lambda = 3$: $\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

15. $\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ **17.** 0, 3, -2

19. 0. Justifique su respuesta.

23. *Sugerencia:* Aplique el teorema 2.

25. *Sugerencia:* Utilice la ecuación $A\mathbf{x} = \lambda\mathbf{x}$ para encontrar una ecuación que implique a A^{-1}.

27. *Sugerencia:* Para cualquier λ, $(A - \lambda I)^T = A^T - \lambda I$. De acuerdo con un teorema (¿cuál?), $A^T - \lambda I$ es invertible si y solo si $A - \lambda I$ es invertible.

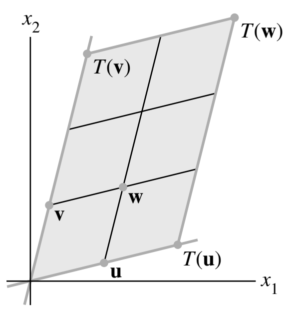
29. Sea $\mathbf{v}$ el vector en $\mathbb{R}^n$ con todas sus entradas igual a 1. Entonces, $A\mathbf{v} = s\mathbf{v}$.

31. *Sugerencia:* Si A es la matriz estándar de T, busque un vector $\mathbf{v}$ diferente de cero (un punto en el plano) tal que $A\mathbf{v} = \mathbf{v}$.

33. ***a)*** $\mathbf{x}_{k+1} = c_1\lambda^{k+1}\mathbf{u} + c_2\mu^{k+1}\mathbf{v}$

b)
$$\begin{aligned} A\mathbf{x}_k &= A(c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v}) \\ &= c_1\lambda^k A\mathbf{u} + c_2\mu^k A\mathbf{v} && \text{Linealidad} \\ &= c_1\lambda^k\lambda\mathbf{u} + c_2\mu^k\mu\mathbf{v} && \mathbf{u} \text{ y } \mathbf{v} \text{ son vectores propios.} \\ &= \mathbf{x}_{k+1} \end{aligned}$$

35.



37. [**M**] $\lambda = 5$: $\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$; $\lambda = 10$: $\begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$; $\lambda = 15$: $\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}$

39. [**M**] $\lambda = -4$: $\begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$; $\lambda = -8$: $\begin{bmatrix} 6 \\ 3 \\ 3 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$

$\lambda = 12$: $\begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

Sección 5.2, página 279

1. $\lambda^2 - 4\lambda - 45$; 9, -5 **3.** $\lambda^2 - 3\lambda - 40$; -5, 8

5. $\lambda^2 - 16\lambda + 48$; 4, 12

7. $\lambda^2 - 9\lambda + 32$; ningún valor propio real

9. $-\lambda^3 + 10\lambda^2 - 33\lambda + 36$ **11.** $-\lambda^3 + 8\lambda^2 - 19\lambda + 12$

13. $-\lambda^3 + 18\lambda^2 - 95\lambda + 150$ **15.** 2, 3, 5, 5

17. 3, 3, 1, 1, 0

19. *Sugerencia:* La ecuación dada es válida para toda λ.

23. *Sugerencia:* Encuentre una matriz invertible P tal que $RQ = P^{-1}AP$.

25. ***a)*** $\{\mathbf{v}_1, \mathbf{v}_2\}$, donde $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ es un vector propio para $\lambda = .3$

b) $\mathbf{x}_0 = \mathbf{v}_1 - \frac{1}{14}\mathbf{v}_2$

c) $\mathbf{x}_1 = \mathbf{v}_1 - \frac{1}{14}(.3)\mathbf{v}_2$, $\mathbf{x}_2 = \mathbf{v}_1 - \frac{1}{14}(.3)^2\mathbf{v}_2$ y $\mathbf{x}_k = \mathbf{v}_1 - \frac{1}{14}(.3)^k\mathbf{v}_2$. As $k \to \infty$, $(.3)^k \to 0$ y $\mathbf{x}_k \to \mathbf{v}_1$.

27. ***a)*** $A\mathbf{v}_1 = \mathbf{v}_1$, $A\mathbf{v}_2 = .5\mathbf{v}_2$, $A\mathbf{v}_3 = .2\mathbf{v}_3$ (Esto también muestra que los valores propios de A son 1, .5 y .2).

b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es linealmente independiente porque los vectores propios corresponden a distintos valores propios (teorema 2). Como hay 3 vectores en el conjunto, entonces el conjunto es una base para $\mathbb{R}^3$. Así, existen constantes (únicas) tales que

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

Entonces,

$$\mathbf{w}\mathbf{x}_0 = c_1\mathbf{w}^T\mathbf{v}_1 + c_2\mathbf{w}^T\mathbf{v}_2 + c_3\mathbf{w}^T\mathbf{v}_3 \qquad (*)$$

Como $\mathbf{x}_0$ y $\mathbf{v}_1$ son vectores de probabilidad y puesto que las entradas en $\mathbf{v}_2$ y en $\mathbf{v}_3$ suman 0, (*) indica que $1 = c_1$.

c) De acuerdo con el inciso *b*),

$$\mathbf{x}_0 = \mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

Con base en el inciso *a*),

$$\begin{aligned} \mathbf{x}_k &= A^k\mathbf{x}_0 = A^k\mathbf{v}_1 + c_2A^k\mathbf{v}_2 + c_3A^k\mathbf{v}_3 \\ &= \mathbf{v}_1 + c_2(.5)^k\mathbf{v}_2 + c_3(.2)^k\mathbf{v}_3 \\ &\to \mathbf{v}_1 \text{ como } k \to \infty \end{aligned}$$

29. [**M**] Informe sus resultados y conclusiones.

Sección 5.3, página 286

1. $\begin{bmatrix} 226 & -525 \\ 90 & -209 \end{bmatrix}$ **3.** $\begin{bmatrix} a^k & 0 \\ 2(a^k - b^k) & b^k \end{bmatrix}$

5. $\lambda = 2$: $\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$; $\lambda = 3$: $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

Cuando una respuesta implica una diagonalización, $A = PDP^{-1}$, los factores P y D no son únicos, de manera que su respuesta tal vez difiera de la que aquí se presenta.

7. $P = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

9. No es diagonalizable

11. $P = \begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

13. $P = \begin{bmatrix} -1 & 2 & 1 \\ -1 & -1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

15. $P = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

17. No es diagonalizable

19. $P = \begin{bmatrix} 1 & 3 & -1 & -1 \\ 0 & 2 & -1 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

23. Sí. (Explique por qué).

25. No, A debe ser diagonalizable. (Explique por qué).

27. *Sugerencia:* Escriba $A = PDP^{-1}$. Como A es invertible, 0 no es un valor propio de A, de manera que D tiene entradas diferentes de cero sobre su diagonal.

29. Una respuesta es $P_1 = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$, cuyas columnas son vectores propios correspondientes a los valores propios en D_1.

31. *Sugerencia:* Construya una adecuada matriz triangular de 2×2.

[**M**] $P = \begin{bmatrix} 2 & 0 & -1 & 1 \\ 7 & -1 & 0 & -4 \\ 7 & 0 & 2 & 0 \\ 0 & 1 & 0 & 3 \end{bmatrix}$

$D = \begin{bmatrix} -12 & 0 & 0 & 0 \\ 0 & -12 & 0 & 0 \\ 0 & 0 & 13 & 0 \\ 0 & 0 & 0 & 13 \end{bmatrix}$

35. [**M**] $P = \begin{bmatrix} 2 & 1 & -3 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 \end{bmatrix}$,

$D = \begin{bmatrix} 7 & 0 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & -14 & 0 \\ 0 & 0 & 0 & 0 & -14 \end{bmatrix}$

Sección 5.4, página 293

1. $\begin{bmatrix} 3 & -1 & 0 \\ -5 & 6 & 4 \end{bmatrix}$

3. ***a)*** $T(\mathbf{e}_1) = \mathbf{b}_3$, $T(\mathbf{e}_2) = -\mathbf{b}_1 - 2\mathbf{b}_2$, $T(\mathbf{e}_3) = 2\mathbf{b}_1 + 3\mathbf{b}_3$

b) $[\,T(\mathbf{e}_1)\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, [\,T(\mathbf{e}_2)\,]_{\mathcal{B}} = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix}$,

$[\,T(\mathbf{e}_3)\,]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$

c) $\begin{bmatrix} 0 & -1 & 2 \\ 0 & -2 & 0 \\ 1 & 0 & 3 \end{bmatrix}$

5. ***a)*** $9 - 3t + t^2 + t^3$

b) Para cualesquiera $\mathbf{p}$, $\mathbf{q}$ en $\mathbb{P}_2$ y cualquier escalar c,

$$\begin{aligned} T[\mathbf{p}(t) + \mathbf{q}(t)] &= (t+3)[\mathbf{p}(t) + \mathbf{q}(t)] \\ &= (t+3)\mathbf{p}(t) + (t+3)\mathbf{q}(t) \\ &= T[\mathbf{p}(t)] + T[\mathbf{q}(t)] \\ T[c \cdot \mathbf{p}(t)] &= (t+3)[c \cdot \mathbf{p}(t)] = c \cdot (t+3)\mathbf{p}(t) \\ &= c \cdot T[\mathbf{p}(t)] \end{aligned}$$

c) $\begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 3 & 0 & 0 \\ 5 & -2 & 0 \\ 0 & 4 & 1 \end{bmatrix}$

9. ***a)*** $\begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$

b) *Sugerencia:* Calcule $T(\mathbf{p} + \mathbf{q})$ y $T(c \cdot \mathbf{p})$ para $\mathbf{p}$, $\mathbf{q}$ arbitrarios en $\mathbb{P}_2$ y cualquier escalar c.

c) $\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

11. $\begin{bmatrix} -2 & -2 \\ 0 & -1 \end{bmatrix}$ **13.** $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

15. $\mathbf{b}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

17. *a*) $A\mathbf{b}_1 = 3\mathbf{b}_1$, por lo que $\mathbf{b}_1$ es un vector propio de A. Sin embargo, A solo tiene un valor propio, $\lambda = 3$, y el espacio propio es unidimensional, por lo que A no es diagonalizable.

b) $\begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$

19. Por definición, Si A es semejante a B, existe una matriz invertible P tal que $P^{-1}AP = B$. (Véase la sección 5.2). Entonces, B es invertible porque es el producto de matrices invertibles. Utilice la ecuación $P^{-1}AP = B$ para demostrar que A^{-1} es semejante a B^{-1}.

21. *Sugerencia:* Revise el problema de práctica 2.

23. *Sugerencia:* Calcule $B(P^{-1}\mathbf{x})$.

25. *Sugerencia:* Escriba $A = PBP^{-1} = (PB)P^{-1}$, y utilice la propiedad de la traza.

27. Para cada j, $I(\mathbf{b}_j) = \mathbf{b}_j$. Como el vector de coordenadas estándar de cualquier vector en $\mathbb{R}^n$ es justamente el vector mismo, entonces $[I(\mathbf{b}_j)]_{\mathcal{E}} = \mathbf{b}_j$. Por lo tanto, la matriz para I respecto de $\mathcal{B}$ y a la base estándar $\mathcal{E}$ es simplemente $[\mathbf{b}_1 \ \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_n]$. Esta matriz es precisamente la matriz de *cambio de coordenadas* $P_{\mathcal{B}}$ definida en la sección 4.4.

29. La matriz $\mathcal{B}$ para la transformación identidad es I_n, porque el vector de $\mathcal{B}$-coordenadas del j-ésimo vector base $\mathbf{b}_j$ es la j-ésima columna de I_n.

31. [**M**] $\begin{bmatrix} -7 & -2 & -6 \\ 0 & -4 & -6 \\ 0 & 0 & -1 \end{bmatrix}$

Sección 5.5, página 300

1. $\lambda = 2 + i$, $\begin{bmatrix} -1+i \\ 1 \end{bmatrix}$; $\lambda = 2 - i$, $\begin{bmatrix} -1-i \\ 1 \end{bmatrix}$

3. $\lambda = 3 + 2i$, $\begin{bmatrix} -1-i \\ 4 \end{bmatrix}$; $\lambda = 3 - 2i$, $\begin{bmatrix} -1+i \\ 4 \end{bmatrix}$

5. $\lambda = 4 + i$, $\begin{bmatrix} 1-i \\ 2 \end{bmatrix}$; $\lambda = 4 - i$, $\begin{bmatrix} 1+i \\ 2 \end{bmatrix}$

7. $\lambda = \sqrt{3} \pm i$, $\varphi = \pi/6$ radián, $r = 2$

9. $\lambda = \pm 2i$, $\varphi = -\pi/2$ radianes, $r = 2$

11. $\lambda = -\sqrt{3} \pm i$, $\varphi = -5\pi/6$ radián, $r = 2$

En los ejercicios 13 a 20, son posibles otras respuestas. Cualquier P que haga $P^{-1}AP$ igual a C o a C^T es una respuesta satisfactoria. Primero, encuentre P; después, calcule $P^{-1}AP$.

13. $P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, C = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}$

15. $P = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}, C = \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$

17. $P = \begin{bmatrix} 1 & 2 \\ 0 & -5 \end{bmatrix}, C = \begin{bmatrix} -3 & 4 \\ -4 & -3 \end{bmatrix}$

19. $P = \begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}, C = \begin{bmatrix} .96 & -.28 \\ .28 & .96 \end{bmatrix}$

21. $\mathbf{y} = \begin{bmatrix} 2 \\ -1+2i \end{bmatrix} = \dfrac{-1+2i}{5}\begin{bmatrix} -2-4i \\ 5 \end{bmatrix}$

23. *a*) Propiedades de conjugados y el hecho de que $\overline{\mathbf{x}}^T = \overline{\mathbf{x}^T}$.

b) $\overline{A\mathbf{x}} = A\overline{\mathbf{x}}$ y A es real; *c*) porque $\mathbf{x}^T A\overline{\mathbf{x}}$ es un escalar y entonces puede considerarse como una matriz de 1×1; *d*) propiedades de transpuestas; *e*) $A^T = A$, definición de q.

25. *Sugerencia:* Primero escriba $\mathbf{x} = \text{Re}\,\mathbf{x} + i\,(\text{Im}\,\mathbf{x})$.

27. [**M**] $P = \begin{bmatrix} -1 & 1 & -1 & -1 \\ 0 & -1 & 0 & 2 \\ 1 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$,

$C = \begin{bmatrix} -2 & 5 & 0 & 0 \\ -5 & -2 & 0 & 0 \\ 0 & 0 & -4 & 10 \\ 0 & 0 & -10 & -4 \end{bmatrix}$

Son posibles otras elecciones, pero C debe ser igual a $P^{-1}AP$.

Sección 5.6, página 309

1. *a*) *Sugerencia:* Encuentre c_1, c_2 tales que $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Utilice esta representación y considere el hecho de que $\mathbf{v}_1$ y $\mathbf{v}_2$ son vectores propios de A para calcular $\mathbf{x}_1 = \begin{bmatrix} 49/3 \\ 41/3 \end{bmatrix}$.

b) En general, $\mathbf{x}_k = 5(3)^k\mathbf{v}_1 - 4(\frac{1}{3})^k\mathbf{v}_2$ para $k \geq 0$.

3. Cuando $p = .2$, los valores propios de A son .9 y .7, y

$\mathbf{x}_k = c_1(.9)^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2(.7)^k\begin{bmatrix} 2 \\ 1 \end{bmatrix} \to \mathbf{0}$ conforme $k \to \infty$

La elevada tasa de depredación reduce el abastecimiento de comida de los búhos y, con el tiempo, perecen ambas poblaciones: presas y depredadores.

5. Si $p = .325$, los valores propios son 1.05 y .55. Como $1.05 > 1$, ambas poblaciones crecerán en un 5% anual. Un vector propio para 1.05 es (6, 13), de manera que, con el tiempo, habrá aproximadamente 6 búhos manchados por cada 13 (mil) ardillas voladoras.

7. *a*) El origen es un punto silla porque A tiene un valor propio más grande que 1 y uno más pequeño que 1 (en valor absoluto).

b) La dirección de más fuerte atracción está dada por el vector propio correspondiente al valor propio $1/3$, a saber, $\mathbf{v}_2$. Todos los vectores que son múltiplos de $\mathbf{v}_2$ son atraídos hacia el origen. La dirección de mayor repulsión está dada por el vector propio $\mathbf{v}_1$. Todos los múltiplos de $\mathbf{v}_1$ se repelen.

9. Punto silla; valores propios: 2, .5; dirección de mayor repulsión: la recta a través de (0, 0) y $(-1, 1)$; dirección de mayor atracción: la recta por (0, 0) y (1, 4).

11. Atractor; valores propios: .9, .8; mayor atracción: recta que pasa por (0, 0) y (5, 4).

13. Repulsor; valores propios: 1.2, 1.1; mayor repulsión: recta que pasa por (0, 0) y (3, 4).

15. $\mathbf{x}_k = \mathbf{v}_1 + .1(.5)^k \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} + .3(.2)^k \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \to \mathbf{v}_1$

conforme $k \to \infty$

17. ***a)*** $A = \begin{bmatrix} 0 & 1.6 \\ .3 & .8 \end{bmatrix}$

b) La población está creciendo porque el mayor valor propio de A es 1.2, que es más grande que 1 en magnitud. La tasa de crecimiento final es 1.2, que es del 20% anual. El vector propio (4, 3) para $\lambda_1 = 1.2$ indica que habrá 4 juveniles por cada 3 adultas.

c) [**M**] La razón entre juveniles y adultas parece estabilizarse después de 5 o 6 años.

Sección 5.7, página 317

1. $\mathbf{x}(t) = \frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{4t} - \frac{3}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{2t}$

3. $-\frac{5}{2}\begin{bmatrix} -3 \\ 1 \end{bmatrix} e^{t} + \frac{9}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix} e^{-t}$. El origen es un punto silla. La dirección de mayor atracción es la recta que pasa por $(-1, 1)$ y el origen. La dirección de mayor repulsión es la recta que pasa por $(-3, 1)$ y por el origen.

5. $-\frac{1}{2}\begin{bmatrix} 1 \\ 3 \end{bmatrix} e^{4t} + \frac{7}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{6t}$. El origen es un repulsor. La dirección de mayor repulsión es la recta que pasa por (1, 1) y por el origen.

7. Sean $P = \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}$ y $D = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix}$. Entonces, $A = PDP^{-1}$. Al sustituir $\mathbf{x} = P\mathbf{y}$ en $\mathbf{x}' = A\mathbf{x}$, se tiene

$$\frac{d}{dt}(P\mathbf{y}) = A(P\mathbf{y})$$
$$P\mathbf{y}' = PDP^{-1}(P\mathbf{y}) = PD\mathbf{y}$$

Multiplicando por la izquierda por P^{-1} resulta

$$\mathbf{y}' = D\mathbf{y} \quad \text{o} \quad \begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$$

9. (solución compleja):

$c_1 \begin{bmatrix} 1-i \\ 1 \end{bmatrix} e^{(-2+i)t} + c_2 \begin{bmatrix} 1+i \\ 1 \end{bmatrix} e^{(-2-i)t}$

(solución real):

$c_1 \begin{bmatrix} \cos t + \operatorname{sen} t \\ \cos t \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} \operatorname{sen} t - \cos t \\ \operatorname{sen} t \end{bmatrix} e^{-2t}$

Las trayectorias caen en espiral hacia el origen.

11. (compleja): $c_1 \begin{bmatrix} -3+3i \\ 2 \end{bmatrix} e^{3it} + c_2 \begin{bmatrix} -3-3i \\ 2 \end{bmatrix} e^{-3it}$

(real):

$c_1 \begin{bmatrix} -3\cos 3t - 3\operatorname{sen} 3t \\ 2\cos 3t \end{bmatrix} + c_2 \begin{bmatrix} -3\operatorname{sen} 3t + 3\cos 3t \\ 2\operatorname{sen} 3t \end{bmatrix}$

Las trayectorias son elipses alrededor del origen.

13. (compleja): $c_1 \begin{bmatrix} 1+i \\ 2 \end{bmatrix} e^{(1+3i)t} + c_2 \begin{bmatrix} 1-i \\ 2 \end{bmatrix} e^{(1-3i)t}$

(real): $c_1 \begin{bmatrix} \cos 3t - \operatorname{sen} 3t \\ 2\cos 3t \end{bmatrix} e^{t} + c_2 \begin{bmatrix} \operatorname{sen} 3t + \cos 3t \\ 2\operatorname{sen} 3t \end{bmatrix} e^{t}$

Las trayectorias son espirales que se alejan del origen.

15. [**M**] $\mathbf{x}(t) = c_1 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} e^{-2t} + c_2 \begin{bmatrix} -6 \\ 1 \\ 5 \end{bmatrix} e^{-t} + c_3 \begin{bmatrix} -4 \\ 1 \\ 4 \end{bmatrix} e^{t}$

El origen es un punto silla. Una solución con $c_3 = 0$ es atraída hacia el origen. Una solución con $c_1 = c_2 = 0$ es repelida.

17. [**M**] (compleja):

$c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23-34i \\ -9+14i \\ 3 \end{bmatrix} e^{(5+2i)t} +$

$c_3 \begin{bmatrix} 23+34i \\ -9-14i \\ 3 \end{bmatrix} e^{(5-2i)t}$

(real): $c_1 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} e^{t} + c_2 \begin{bmatrix} 23\cos 2t + 34\operatorname{sen} 2t \\ -9\cos 2t - 14\operatorname{sen} 2t \\ 3\cos 2t \end{bmatrix} e^{5t} +$

$c_3 \begin{bmatrix} 23\operatorname{sen} 2t - 34\cos 2t \\ -9\operatorname{sen} 2t + 14\cos 2t \\ 3\operatorname{sen} 2t \end{bmatrix} e^{5t}$

El origen es un repulsor. Las trayectorias son espirales hacia afuera, alejándose del origen.

19. [**M**] $A = \begin{bmatrix} -2 & 3/4 \\ 1 & -1 \end{bmatrix}$,

$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \frac{5}{2}\begin{bmatrix} 1 \\ 2 \end{bmatrix} e^{-.5t} - \frac{1}{2}\begin{bmatrix} -3 \\ 2 \end{bmatrix} e^{-2.5t}$

21. [**M**] $A = \begin{bmatrix} -1 & -8 \\ 5 & -5 \end{bmatrix}$,

$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} -20\operatorname{sen} 6t \\ 15\cos 6t - 5\operatorname{sen} 6t \end{bmatrix} e^{-3t}$

Sección 5.8, página 324

1. Vector propio: $\mathbf{x}_4 = \begin{bmatrix} 1 \\ .3326 \end{bmatrix}$ o $A\mathbf{x}_4 = \begin{bmatrix} 4.9978 \\ 1.6652 \end{bmatrix}$;

$\lambda \approx 4.9978$

3. Vector propio: $\mathbf{x}_4 = \begin{bmatrix} .5188 \\ 1 \end{bmatrix}$ o $A\mathbf{x}_4 = \begin{bmatrix} .4594 \\ .9075 \end{bmatrix}$;

$\lambda \approx .9075$

5. $\mathbf{x} = \begin{bmatrix} -.7999 \\ 1 \end{bmatrix}$, $A\mathbf{x} = \begin{bmatrix} 4.0015 \\ -5.0020 \end{bmatrix}$;

λ estimada $= -5.0020$

7. [**M**]

$\mathbf{x}_k$: $\begin{bmatrix} .75 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .9565 \end{bmatrix}, \begin{bmatrix} .9932 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ .9990 \end{bmatrix}, \begin{bmatrix} .9998 \\ 1 \end{bmatrix}$

μ_k: 11.5, 12.78, 12.96, 12.9948, 12.9990

9. [**M**] $\mu_5 = 8.4233$, $\mu_6 = 8.4246$; valor real: 8.42443 (con 5 lugares de precisión)

11. μ_k: 5.8000, 5.9655, 5.9942, 5.9990 $(k = 1, 2, 3, 4)$; $R(\mathbf{x}_k)$: 5.9655, 5.9990, 5.99997, 5.9999993

13. Sí, pero las secuencias pueden tener lenta convergencia.

15. *Sugerencia:* Escriba $A\mathbf{x} - \alpha\mathbf{x} = (A - \alpha I)\mathbf{x}$, y considere el hecho de que $(A - \alpha I)$ es invertible cuando α *no* es un valor propio de A.

17. [**M**] $\nu_0 = 3.3384$, $\nu_1 = 3.32119$ (con 4 lugares de precisión con redondeo), $\nu_2 = 3.3212209$. Valor real: 3.3212201 (con 7 lugares de precisión)

19. [**M**] ***a***) $\mu_6 = 30.2887 = \mu_7$ con cuatro cifras decimales. Con seis lugares, el mayor valor propio es (.957629, .688937, 1, .943782).

b) El método de potencias inverso (con $\alpha = 0$) produce $\mu_1^{-1} = .010141$, $\mu_2^{-1} = .010150$. Con siete lugares, el menor valor propio es 0.0101500, con vector propio $(-.603972, 1, -.251135, .148953)$. La razón de la convergencia rápida es que el siguiente valor propio más pequeño (segundo en tamaño) está cerca de .85.

21. ***a***) Si todos los valores propios de A son menores que 1 en magnitud, y si $\mathbf{x} \neq \mathbf{0}$, entonces $A^k\mathbf{x}$ es aproximadamente un vector propio para k grande.

b) Si el valor propio estrictamente dominante es 1, y si $\mathbf{x}$ tiene una componente en la dirección del vector propio correspondiente, entonces $\{A^k\mathbf{x}\}$ convergerá a un múltiplo de ese vector propio.

c) Si todos los valores propios de A son mayores que 1 en magnitud, y si $\mathbf{x}$ no es un vector propio, entonces la distancia de $A^k\mathbf{x}$ al vector propio más cercano *se incrementará* conforme $k \to \infty$.

Capítulo 5 Ejercicios complementarios, página 326

1.

a) V	***b***) F	***c***) V	***d***) F	***e***) V	***f***) V
g) F	***h***) V	***i***) F	***j***) V	***k***) F	***l***) F
m) F	***n***) V	***o***) F	***p***) V	***q***) F	***r***) V
s) F	***t***) V	***u***) V	***v***) V	***w***) F	***x***) V

3. ***a***) Suponga que $A\mathbf{x} = \lambda\mathbf{x}$, con $\mathbf{x} \neq \mathbf{0}$. Entonces, $(5I - A)\mathbf{x} = 5\mathbf{x} - A\mathbf{x} = 5\mathbf{x} - \lambda\mathbf{x} = (5 - \lambda)\mathbf{x}$. El valor propio es $5 - \lambda$.

b)
$$\begin{aligned}(5I - 3A + A^2)\mathbf{x} &= 5\mathbf{x} - 3A\mathbf{x} + A(A\mathbf{x}) \\ &= 5\mathbf{x} - 3\lambda\mathbf{x} + \lambda^2\mathbf{x} \\ &= (5 - 3\lambda + \lambda^2)\mathbf{x}.\end{aligned}$$
El valor propio es $5 - 3\lambda + \lambda^2$.

5. Suponga que $A\mathbf{x} = \lambda\mathbf{x}$, con $\mathbf{x} \neq \mathbf{0}$. Entonces,
$$\begin{aligned}p(A)\mathbf{x} &= (c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n)\mathbf{x} \\ &= c_0\mathbf{x} + c_1 A\mathbf{x} + c_2 A^2\mathbf{x} + \cdots + c_n A^n\mathbf{x} \\ &= c_0\mathbf{x} + c_1\lambda\mathbf{x} + c_2\lambda^2\mathbf{x} + \cdots + c_n\lambda^n\mathbf{x} = p(\lambda)\mathbf{x}\end{aligned}$$
Así, $p(\lambda)$ es un valor propio de la matriz $p(A)$.

7. Si $A = PDP^{-1}$, entonces $p(A) = Pp(D)P^{-1}$, como se muestra en el ejercicio 6. Si la entrada (j, j) en D es λ, entonces la entrada (j, j) en D^k es λ^k, de manera que la entrada (j, j) en $p(D)$ es $p(\lambda)$. Si p es el polinomio característico de A, entonces $p(\lambda) = 0$ para cada entrada diagonal de D, porque las entradas en D son los valores propios de A. Por lo tanto, $p(D)$ es la matriz cero. Así, $p(A) = P \cdot 0 \cdot P^{-1} = 0$.

9. Si $I - A$ no fuera invertible, entonces la ecuación $(I - A)\mathbf{x} = \mathbf{0}$ tendría una solución no trivial. Por lo tanto, $\mathbf{x} - A\mathbf{x} = \mathbf{0}$ y $A\mathbf{x} = 1\cdot\mathbf{x}$, lo que muestra que A tendría 1 como un valor propio. Esto no puede suceder si todos los valores propios son menores que 1 en magnitud. Así que $I - A$ debe ser invertible.

11. ***a***) Tome $\mathbf{x}$ en H. Entonces, $\mathbf{x} = c\mathbf{u}$ para algún escalar c. Así, $A\mathbf{x} = A(c\mathbf{u}) = c(A\mathbf{u}) = c(\lambda\mathbf{u}) = (c\lambda)\mathbf{u}$, lo que muestra que $A\mathbf{x}$ está en H.

b) Sea $\mathbf{x}$ un vector diferente de cero en K. Como K es unidimensional, entonces K debe ser el conjunto de todos los múltiplos escalares de $\mathbf{x}$. Si K es invariante bajo A, entonces $A\mathbf{x}$ está en K y, por lo tanto, $A\mathbf{x}$ es un múltiplo de $\mathbf{x}$. Por consiguiente, $\mathbf{x}$ es un vector propio de A.

13. 1, 3, 7

15. Sustituya a por $a - \lambda$ en la fórmula del determinante del ejercicio complementario 16 del capítulo 3:
$$\det(A - \lambda I) = (a - b - \lambda)^{n-1}[a - \lambda + (n - 1)b]$$
Este determinante es cero solo si $a - b - \lambda = 0$ o $a - \lambda + (n - 1)b = 0$. Así, λ es un valor propio de A si y solo si $\lambda = a - b$ o $\lambda = a + (n - 1)b$. De la fórmula anterior para $\det(A - \lambda I)$, la multiplicidad algebraica es $n - 1$ para $a - b$ y 1 para $a + (n - 1)b$.

17. $\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = \lambda^2 - (\text{tr } A)\lambda + \det A$. Utilice la fórmula cuadrática para resolver la ecuación característica
$$\lambda = \frac{\text{tr } A \pm \sqrt{(\text{tr } A)^2 - 4 \det A}}{2}$$
Ambos valores propios son reales si y solo si el discriminante es no negativo, es decir, $(\text{tr } A)^2 - 4 \det A \geq 0$. Esta desigualdad se simplifica a $(\text{tr } A)^2 \geq 4 \det A$ y $\left(\dfrac{\text{tr } A}{2}\right)^2 \geq \det A$.

19. $C_p = \begin{bmatrix} 0 & 1 \\ -6 & 5 \end{bmatrix}$; $\det(C_p - \lambda I) = 6 - 5\lambda + \lambda^2 = p(\lambda)$

21. Si p es un polinomio de orden 2, entonces un cálculo tal como el del ejercicio 19 muestra que el polinomio característico de C_p es $p(\lambda) = (-1)^2 p(\lambda)$, de manera que el resultado es verdadero para $n = 2$. Suponga que el resultado es válido para $n = k$ para alguna $k \geq 2$, y considere un polinomio p de grado $k + 1$. Entonces, al desarrollar $\det(C_p - \lambda I)$ por cofactores en la primera columna, el determinante de $C_p - \lambda I$ es igual a
$$(-\lambda)\det\begin{bmatrix} -\lambda & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & & & 1 \\ -a_1 & -a_2 & \cdots & -a_k - \lambda \end{bmatrix} + (-1)^{k+1}a_0$$

La matriz de $k \times k$ que se muestra es $C_q - \lambda I$, donde $q(t) = a_1 + a_2 t + \cdots + a_k t^{k-1} + t^k$. Por la suposición de inducción, el determinante de $C_q - \lambda I$ es $(-1)^k q(\lambda)$. Así,

$$\begin{aligned}\det(C_p - \lambda I) &= (-1)^{k+1} a_0 + (-\lambda)(-1)^k q(\lambda) \\ &= (-1)^{k+1}[a_0 + \lambda(a_1 + \cdots + a_k \lambda^{k-1} + \lambda^k)] \\ &= (-1)^{k+1} p(\lambda)\end{aligned}$$

Entonces la fórmula es válida para $n = k + 1$ cuando es verdadera para $n = k$. Por el principio de inducción, la fórmula para $\det(C_p - \lambda I)$ es correcta para toda $n \geq 2$.

23. Del ejercicio 22, las columnas de la matriz de Vandermonde V son vectores propios de C_p, que corresponden a los valores propios λ_1, λ_2, λ_3 (las raíces del polinomio p). Como esos valores propios son distintos, los vectores propios forman un conjunto linealmente independiente, de acuerdo con el teorema 2 de la sección 5.1. Así, V tiene columnas linealmente independientes y, por lo tanto, es invertible, según el teorema de la matriz invertible. Por último, como las columnas de V son vectores propios de C_p, el teorema de diagonalización (teorema 5 de la sección 5.3) indica que $V^{-1}C_pV$ es diagonal.

25. [**M**] Si su programa de matrices calcula valores propios y vectores propios mediante métodos iterativos en vez de utilizar cálculos simbólicos, podría tener algunas dificultades. Encontrará que $AP - PD$ tiene entradas extremadamente pequeñas y que PDP^{-1} es muy cercana a la matriz A. (Esto sucedía hasta hace pocos años, pero la situación podría cambiar conforme mejoren los programas de matrices). Si P se construye con los vectores propios del programa, entonces compruebe el número de condición de P. Esto puede indicar que en realidad no se tienen tres vectores propios linealmente independientes.

Capítulo 6

Sección 6.1, página 336

1. 5, 8, $\frac{8}{5}$ **3.** $\begin{bmatrix} 3/35 \\ -1/35 \\ -1/7 \end{bmatrix}$ **5.** $\begin{bmatrix} 8/13 \\ 12/13 \end{bmatrix}$

7. $\sqrt{35}$ **9.** $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}$ **11.** $\begin{bmatrix} 7/\sqrt{69} \\ 2/\sqrt{69} \\ 4/\sqrt{69} \end{bmatrix}$

13. $5\sqrt{5}$ **15.** No son ortogonales **17.** Ortogonales

21. *Sugerencia:* Utilice los teoremas 2 y 3 de la sección 2.1.

23. $\mathbf{u}\cdot\mathbf{v} = 0$, $\|\mathbf{u}\|^2 = 30$, $\|\mathbf{v}\|^2 = 101$, $\|\mathbf{u} + \mathbf{v}\|^2 = (-5)^2 + (-9)^2 + 5^2 = 131 = 30 + 101$

25. El conjunto de todos los múltiplos de $\begin{bmatrix} -b \\ a \end{bmatrix}$ (cuando $\mathbf{v} \neq \mathbf{0}$)

27. *Sugerencia:* Aplique la definición de ortogonalidad.

29. *Sugerencia:* Considere un vector típico $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ en W.

31. *Sugerencia:* Si $\mathbf{x}$ está en $W^\perp$, entonces $\mathbf{x}$ es ortogonal a todo vector en W.

33. [**M**] Establezca su suposición y compruébela algebraicamente.

Sección 6.2, página 344

1. No son ortogonales **3.** No son ortogonales **5.** Ortogonales

7. Demuestre que $\mathbf{u}_1\cdot\mathbf{u}_2 = 0$, mencione el teorema 4, y observe que dos vectores linealmente independientes en $\mathbb{R}^2$ forman una base. Después obtenga

$$\mathbf{x} = \tfrac{39}{13}\begin{bmatrix} 2 \\ -3 \end{bmatrix} + \tfrac{26}{52}\begin{bmatrix} 6 \\ 4 \end{bmatrix} = 3\begin{bmatrix} 2 \\ -3 \end{bmatrix} + \tfrac{1}{2}\begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

9. Demuestre que $\mathbf{u}_1\cdot\mathbf{u}_2 = 0$, $\mathbf{u}_1\cdot\mathbf{u}_3 = 0$, y $\mathbf{u}_2\cdot\mathbf{u}_3 = 0$. Mencione el teorema 4, y observe que tres vectores linealmente independientes en $\mathbb{R}^3$ forman una base. Después obtenga

$$\mathbf{x} = \tfrac{5}{2}\mathbf{u}_1 - \tfrac{27}{18}\mathbf{u}_2 + \tfrac{18}{9}\mathbf{u}_3 = \tfrac{5}{2}\mathbf{u}_1 - \tfrac{3}{2}\mathbf{u}_2 + 2\mathbf{u}_3$$

11. $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ **13.** $\mathbf{y} = \begin{bmatrix} -4/5 \\ 7/5 \end{bmatrix} + \begin{bmatrix} 14/5 \\ 8/5 \end{bmatrix}$

15. $\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} .6 \\ -.8 \end{bmatrix}$, la distancia es 1

17. $\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$

19. Ortonormales **21.** Ortonormales

25. *Sugerencia:* $\|U\mathbf{x}\|^2 = (U\mathbf{x})^T(U\mathbf{x})$. Además, los incisos *a*) y *c*) se deducen de *b*).

27. *Sugerencia:* Se necesitan dos teoremas, uno de los cuales solamente se aplica a matrices *cuadradas*.

29. *Sugerencia:* Si se tiene un candidato para una inversa, es posible revisar si el candidato funciona.

31. Suponga que $\hat{\mathbf{y}} = \dfrac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. Remplace $\mathbf{u}$ por $c\mathbf{u}$ con $c \neq 0$; entonces,

$$\frac{\mathbf{y}\cdot(c\mathbf{u})}{(c\mathbf{u})\cdot(c\mathbf{u})}(c\mathbf{u}) = \frac{c(\mathbf{y}\cdot\mathbf{u})}{c^2\mathbf{u}\cdot\mathbf{u}}(c)\mathbf{u} = \hat{\mathbf{y}}$$

33. Sea $L = \text{Gen}\{\mathbf{u}\}$, donde $\mathbf{u}$ es diferente de cero, y $T(\mathbf{x}) = \text{proy}_L\,\mathbf{x}$. Por definición,

$$T(\mathbf{x}) = \frac{\mathbf{x}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = (\mathbf{x}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u}$$

Para $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$ y cualesquiera escalares c y d, las propiedades del producto interior (teorema 1) indican que

$$\begin{aligned}T(c\mathbf{x} + d\mathbf{y}) &= [(c\mathbf{x} + d\mathbf{y})\cdot\mathbf{u}](\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= [c(\mathbf{x}\cdot\mathbf{u}) + d(\mathbf{y}\cdot\mathbf{u})](\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= c(\mathbf{x}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} + d(\mathbf{y}\cdot\mathbf{u})(\mathbf{u}\cdot\mathbf{u})^{-1}\mathbf{u} \\ &= cT(\mathbf{x}) + dT(\mathbf{y})\end{aligned}$$

Por lo tanto, T es lineal.

Sección 6.3, página 352

1. $\mathbf{x} = -\frac{8}{9}\mathbf{u}_1 - \frac{2}{9}\mathbf{u}_2 + \frac{2}{3}\mathbf{u}_3 + 2\mathbf{u}_4$; $\mathbf{x} = \begin{bmatrix} 0 \\ -2 \\ 4 \\ -2 \end{bmatrix} + \begin{bmatrix} 10 \\ -6 \\ -2 \\ 2 \end{bmatrix}$

3. $\begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}$ **5.** $\begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix} = \mathbf{y}$

7. $\mathbf{y} = \begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix} + \begin{bmatrix} -7/3 \\ 7/3 \\ 7/3 \end{bmatrix}$ **9.** $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ 3 \\ -1 \end{bmatrix}$

11. $\begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}$ **13.** $\begin{bmatrix} -1 \\ -3 \\ -2 \\ 3 \end{bmatrix}$ **15.** $\sqrt{40}$

17. *a*) $U^TU = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $UU^T = \begin{bmatrix} 8/9 & -2/9 & 2/9 \\ -2/9 & 5/9 & 4/9 \\ 2/9 & 4/9 & 5/9 \end{bmatrix}$

b) $\text{proy}_W\ \mathbf{y} = 6\mathbf{u}_1 + 3\mathbf{u}_2 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$, $(UU^T)\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$

19. Cualquier múltiplo de $\begin{bmatrix} 0 \\ 2/5 \\ 1/5 \end{bmatrix}$, tal como $\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$.

23. *Sugerencia:* Utilice el teorema 3 y el teorema de descomposición ortogonal. Para la unicidad, suponga que $A\mathbf{p} = \mathbf{b}$ y $A\mathbf{p}_1 = \mathbf{b}$, y considere las ecuaciones $\mathbf{p} = \mathbf{p}_1 + (\mathbf{p} - \mathbf{p}_1)$ y $\mathbf{p} = \mathbf{p} + \mathbf{0}$.

Sección 6.4, página 358

1. $\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}$ **3.** $\begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3/2 \\ 3/2 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \\ -1 \end{bmatrix}$ **7.** $\begin{bmatrix} 2/\sqrt{30} \\ -5/\sqrt{30} \\ 1/\sqrt{30} \end{bmatrix}, \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$

9. $\begin{bmatrix} 3 \\ 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 1 \\ 3 \end{bmatrix}$ **11.** $\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 3 \\ -3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 2 \\ -2 \end{bmatrix}$

13. $R = \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix}$

15. $Q = \begin{bmatrix} 1/\sqrt{5} & 1/2 & 1/2 \\ -1/\sqrt{5} & 0 & 0 \\ -1/\sqrt{5} & 1/2 & 1/2 \\ 1/\sqrt{5} & -1/2 & 1/2 \\ 1/\sqrt{5} & 1/2 & -1/2 \end{bmatrix}$,

$R = \begin{bmatrix} \sqrt{5} & -\sqrt{5} & 4\sqrt{5} \\ 0 & 6 & -2 \\ 0 & 0 & 4 \end{bmatrix}$

19. Suponga que $\mathbf{x}$ satisface $R\mathbf{x} = \mathbf{0}$; entonces, $QR\mathbf{x} = Q\mathbf{0} = \mathbf{0}$, y $A\mathbf{x} = \mathbf{0}$. Como las columnas de A son linealmente independientes, $\mathbf{x}$ debe ser cero. A la vez, este hecho indica que las columnas de R son linealmente independientes. Como R es cuadrada, entonces es invertible, de acuerdo con el teorema de la matriz invertible.

21. Denote las columnas de Q como $\mathbf{q}_1, \ldots, \mathbf{q}_n$. Observe que $n \le m$, porque A es de $m \times n$ y tiene columnas linealmente independientes. Considere el hecho de que las columnas de Q pueden ampliarse a una base ortonormal para $\mathbb{R}^m$, por ejemplo, $\{\mathbf{q}_1, \ldots, \mathbf{q}_m\}$. (La *Guía de estudio* describe un método). Sea $Q_0 = [\,\mathbf{q}_{n+1} \ \cdots \ \mathbf{q}_m\,]$ y $Q_1 = [\,Q \quad Q_0\,]$. Entonces, al utilizar la multiplicación matricial particionada, $Q_1\begin{bmatrix} R \\ 0 \end{bmatrix} = QR = A$.

23. *Sugerencia:* Particione R como una matriz por bloques de 2×2.

25. [**M**] Las entradas diagonales de R son 20, 6, 10.3923 y 7.0711, con cuatro cifras decimales.

Sección 6.5, página 366

1. *a*) $\begin{bmatrix} 6 & -11 \\ -11 & 22 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 11 \end{bmatrix}$ *b*) $\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

3. *a*) $\begin{bmatrix} 6 & 6 \\ 6 & 42 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$ *b*) $\hat{\mathbf{x}} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix}$

5. $\hat{\mathbf{x}} = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ **7.** $2\sqrt{5}$

9. *a*) $\hat{\mathbf{b}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ *b*) $\hat{\mathbf{x}} = \begin{bmatrix} 2/7 \\ 1/7 \end{bmatrix}$

11. *a*) $\hat{\mathbf{b}} = \begin{bmatrix} 3 \\ 1 \\ 4 \\ -1 \end{bmatrix}$ *b*) $\hat{\mathbf{x}} = \begin{bmatrix} 2/3 \\ 0 \\ 1/3 \end{bmatrix}$

13. $A\mathbf{u} = \begin{bmatrix} 11 \\ -11 \\ 11 \end{bmatrix}$, $A\mathbf{v} = \begin{bmatrix} 7 \\ -12 \\ 7 \end{bmatrix}$,

$\mathbf{b} - A\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ -6 \end{bmatrix}$, $\mathbf{b} - A\mathbf{v} = \begin{bmatrix} 4 \\ 3 \\ -2 \end{bmatrix}$. No, $\mathbf{u}$ posiblemente no podría ser solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. ¿Por qué?

15. $\hat{\mathbf{x}} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$

19. ***a*)** Si $A\mathbf{x} = \mathbf{0}$, entonces $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$. Esto muestra que Nul A está contenido en Nul A^TA.

***b*)** Si $A^TA\mathbf{x} = \mathbf{0}$, entonces $\mathbf{x}^TA^TA\,\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. Así, $(A\mathbf{x})^T(A\mathbf{x}) = 0$ (lo que significa que $\|A\mathbf{x}\|^2 = 0$), y por lo tanto, $A\mathbf{x} = \mathbf{0}$. Esto muestra que Nul A^TA está contenido en Nul A.

21. *Sugerencia:* Para el inciso *a*), utilice un importante teorema del capítulo 2.

23. De acuerdo con el teorema 14, $\hat{\mathbf{b}} = A\hat{\mathbf{x}} = A(A^TA)^{-1}A^T\mathbf{b}$. La matriz $A(A^TA)^{-1}A^T$ se presenta con frecuencia en estadística, donde en ocasiones se le conoce como *matriz sombrero*.

25. Las ecuaciones normales son $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$, cuya solución es el conjunto de (x, y) tal que $x + y = 3$. Las soluciones corresponden a puntos sobre la recta intermedia entre las rectas $x + y = 2$ y $x + y = 4$.

Sección 6.6, página 374

1. $y = .9 + .4x$ **3.** $y = 1.1 + 1.3x$

5. Si dos puntos de datos tienen diferentes coordenadas x, entonces las dos columnas de la matriz de diseño X no pueden ser múltiplos entre sí y, por lo tanto, son linealmente independientes. De acuerdo con el teorema 14 de la sección 6.5, las ecuaciones normales tienen solución única.

7. ***a*)** $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde $\mathbf{y} = \begin{bmatrix} 1.8 \\ 2.7 \\ 3.4 \\ 3.8 \\ 3.9 \end{bmatrix}$, $X = \begin{bmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \\ 5 & 25 \end{bmatrix}$,

$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$, $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \end{bmatrix}$

***b*)** [**M**] $y = 1.76x - .20x^2$

9. $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, donde $\mathbf{y} = \begin{bmatrix} 7.9 \\ 5.4 \\ -.9 \end{bmatrix}$, $X = \begin{bmatrix} \cos 1 & \operatorname{sen} 1 \\ \cos 2 & \operatorname{sen} 2 \\ \cos 3 & \operatorname{sen} 3 \end{bmatrix}$,

$\boldsymbol{\beta} = \begin{bmatrix} A \\ B \end{bmatrix}$, $\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix}$

11. [**M**] $\beta = 1.45$ y $e = .811$; la órbita es una elipse. La ecuación $r = \beta/(1 - e \cdot \cos\vartheta)$ produce $r = 1.33$ cuando $\vartheta = 4.6$.

13. [**M**] ***a*)** $y = -.8558 + 4.7025t + 5.5554t^2 - .0274t^2$

***b*)** La función velocidad es $v(t) = 4.7025 + 11.1108t - .0822t^2$, y $v(4.5) = 53.0$ ft/seg.

15. *Sugerencia:* Escriba X y $\mathbf{y}$ como en la ecuación (1), y calcule X^TX y $X^T\mathbf{y}$.

17. ***a*)** La media de los datos en x es $\bar{x} = 5.5$. Los datos en forma de desviación media son $(-3.5, 1)$, $(-.5, 2)$, $(1.5, 3)$ y $(2.5, 3)$. Las columnas de X son ortogonales porque las entradas en la segunda columna suman 0.

***b*)** $\begin{bmatrix} 4 & 0 \\ 0 & 21 \end{bmatrix}\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 7.5 \end{bmatrix}$,

$y = \frac{9}{4} + \frac{5}{14}x^* = \frac{9}{4} + \frac{5}{14}(x - 5.5)$

19. *Sugerencia:* La ecuación tiene una agradable interpretación geométrica.

Sección 6.7, página 382

1. ***a*)** $3, \sqrt{105}, 225$ ***b*)** Todos los múltiplos de $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

3. 28 **5.** $5\sqrt{2}, 3\sqrt{3}$ **7.** $\frac{56}{25} + \frac{14}{25}t$

9. ***a*)** Polinomio constante, $p(t) = 5$.

***b*)** $t^2 - 5$ es ortogonal a p_0 y p_1; valores: $(4, -4, -4, 4)$; respuesta: $q(t) = \frac{1}{4}(t^2 - 5)$

11. $\frac{17}{5}t$

13. Compruebe cada uno de los cuatro axiomas. Por ejemplo:

$$\begin{aligned} 1. \quad \langle \mathbf{u}, \mathbf{v} \rangle &= (A\mathbf{u})\cdot(A\mathbf{v}) && \text{Definición} \\ &= (A\mathbf{v})\cdot(A\mathbf{u}) && \text{Propiedad del producto} \\ &= \langle \mathbf{v}, \mathbf{u} \rangle && \text{Definición} \end{aligned}$$

15.
$$\begin{aligned} \langle \mathbf{u}, c\mathbf{v} \rangle &= \langle c\mathbf{v}, \mathbf{u} \rangle && \text{Axioma 1} \\ &= c\langle \mathbf{v}, \mathbf{u} \rangle && \text{Axioma 3} \\ &= c\langle \mathbf{u}, \mathbf{v} \rangle && \text{Axioma 1} \end{aligned}$$

17. *Sugerencia:* Calcule 4 veces el lado derecho.

19. $\langle \mathbf{u}, \mathbf{v} \rangle = \sqrt{a}\sqrt{b} + \sqrt{b}\sqrt{a} = 2\sqrt{ab}$, $\|\mathbf{u}\|^2 = (\sqrt{a})^2 + (\sqrt{b})^2 = a + b$. Como a y b son no negativos, $\|\mathbf{u}\| = \sqrt{a + b}$. De forma similar, $\|\mathbf{v}\| = \sqrt{b + a}$. De acuerdo con Cauchy-Schwarz, $2\sqrt{ab} \le \sqrt{a + b}\sqrt{b + a} = a + b$. Por lo que $\sqrt{ab} \le \dfrac{a + b}{2}$.

21. 0 **23.** $2/\sqrt{5}$ **25.** $1, t, 3t^2 - 1$

27. [**M**] Los nuevos polinomios ortogonales son múltiplos de $-17t + 5t^3$ y $72 - 155t^2 + 35t^4$. Escale esos polinomios para que sus valores en $-2, -1, 0, 1$ y 2 sean enteros pequeños.

Sección 6.8, página 389

1. $y = 2 + \frac{3}{2}t$

3. $p(t) = 4p_0 - .1p_1 - .5p_2 + .2p_3$
$= 4 - .1t - .5(t^2 - 2) + .2\left(\frac{5}{6}t^3 - \frac{17}{6}t\right)$
(Este polinomio ajusta los datos de manera exacta).

5. Utilice la identidad
$\operatorname{sen} mt \operatorname{sen} nt = \frac{1}{2}[\cos(mt - nt) - \cos(mt + nt)]$

7. Aplique la identidad $\cos^2 kt = \dfrac{1 + \cos 2kt}{2}$.

9. $\pi + 2\operatorname{sen} t + \operatorname{sen} 2t + \frac{2}{3}\operatorname{sen} 3t$ [*Sugerencia:* Ahorre tiempo empleando los resultados del ejemplo 4].

11. $\frac{1}{2} - \frac{1}{2}\cos 2t$ ¿Por qué?

13. *Sugerencia:* Tome las funciones f y g en $C[0, 2\pi]$, y fije un entero $m \geq 0$. Escriba el coeficiente de Fourier de $f + g$ que implique a cos mt, y también escriba el coeficiente de Fourier que implique a sen mt ($m > 0$).

15. [**M**] La curva cúbica es la gráfica de $g(t) = -.2685 + 3.6095t + 5.8576t^2 - .0477t^3$. La velocidad en $t = 4.5$ segundos es $g'(4.5) = 53.4$ ft/seg. Esto es aproximadamente .7% más rápido que la estimación obtenida en el ejercicio 13 de la sección 6.6.

Capítulo 6 Ejercicios complementarios, página 390

1. *a*) F *b*) V *c*) V *d*) F *e*) F *f*) V *g*) V *h*) V *i*) F *j*) V *k*) V *l*) F *m*) V *n*) F *o*) F *p*) V *q*) V *r*) F *s*) F

2. *Sugerencia:* Si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortonormal y $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, entonces los vectores $c_1\mathbf{v}_1$ y $\mathbf{c}_2\mathbf{v}_2$ son ortogonales, y

$$\|\mathbf{x}\|^2 = \|c_1\mathbf{v}_1 + c_2\mathbf{v}_2\|^2 = \|c_1\mathbf{v}_1\|^2 + \|c_2\mathbf{v}_2\|^2$$
$$= (|c_1|\|\mathbf{v}_1\|)^2 + (|c_2|\|\mathbf{v}_2\|)^2 = |c_1|^2 + |c_2|^2$$

(Explique por qué). Así que la igualdad establecida es válida para $p = 2$. Suponga que la igualdad es verdadera para $p = k$, con $k \geq 2$, sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k+1}\}$ un conjunto ortonormal, y considere que $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} = \mathbf{u}_k + c_{k+1}\mathbf{v}_{k+1}$, donde $\mathbf{u}_k = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$.

3. A partir de $\mathbf{x}$ y un conjunto ortonormal $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ en $\mathbb{R}^n$, sea $\hat{\mathbf{x}}$ la proyección ortogonal de $\mathbf{x}$ sobre el subespacio generado por $\mathbf{v}_1, \ldots, \mathbf{v}_p$. De acuerdo con el teorema 10 de la sección 6.3,

$$\hat{\mathbf{x}} = (\mathbf{x}\cdot\mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{x}\cdot\mathbf{v}_p)\mathbf{v}_p$$

Por el ejercicio 2, $\|\hat{\mathbf{x}}\|^2 = |\mathbf{x}\cdot\mathbf{v}_1|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_p|^2$. La desigualdad de Bessel se deduce del hecho de que $\|\hat{\mathbf{x}}\|^2 \leq \|\mathbf{x}\|^2$, observado antes del enunciado de la desigualdad de Cauchy-Schwarz, en la sección 6.7.

5. Suponga que $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ para toda $\mathbf{x}$, $\mathbf{y}$ en $\mathbb{R}^n$, y sea $\mathbf{e}_1, \ldots, \mathbf{e}_n$ la base estándar para $\mathbb{R}^n$. Para $j = 1, \ldots, n$, $U\mathbf{e}_j$ es la j-ésima columna de U. Puesto que $\|U\mathbf{e}_j\|^2 = (U\mathbf{e}_j)\cdot(U\mathbf{e}_j) = \mathbf{e}_j\cdot\mathbf{e}_j = 1$, las columnas de U son vectores unitarios; puesto que $(U\mathbf{e}_j)\cdot(U\mathbf{e}_k) = \mathbf{e}_j\cdot\mathbf{e}_k = 0$ para $j \neq k$, las columnas son ortogonales en pares.

7. *Sugerencia:* Calcule Q^TQ, considerando el hecho de que $(\mathbf{u}\mathbf{u}^T)^T = \mathbf{u}^{TT}\mathbf{u}^T = \mathbf{u}\mathbf{u}^T$.

9. Sea $W = \text{Gen}\{\mathbf{u}, \mathbf{v}\}$. Dada $\mathbf{z}$ en $\mathbb{R}^n$, sea $\hat{\mathbf{z}} = \text{proy}_W\mathbf{z}$. Entonces $\hat{\mathbf{z}}$ está en Col A, donde $A = [\mathbf{u} \quad \mathbf{v}]$, por ejemplo, $\hat{\mathbf{z}} = A\hat{\mathbf{x}}$ para alguna $\hat{\mathbf{x}}$ en $\mathbb{R}^2$. Así, $\hat{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{z}$. Las ecuaciones normales se pueden resolver para producir $\hat{\mathbf{x}}$, y después $\hat{\mathbf{z}}$ se determina calculando $A\hat{\mathbf{x}}$.

11. *Sugerencia:* Sean $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$ y

$A = \begin{bmatrix} \mathbf{v}^T \\ \mathbf{v}^T \\ \mathbf{v}^T \end{bmatrix} = \begin{bmatrix} 1 & -2 & 5 \\ 1 & -2 & 5 \\ 1 & -2 & 5 \end{bmatrix}$. El conjunto dado de ecuaciones es $A\mathbf{x} = \mathbf{b}$, y el conjunto de todas las soluciones de mínimos cuadrados coincide con el conjunto de soluciones de $A^TA\mathbf{x} = A^T\mathbf{b}$ (teorema 13 de la sección 6.5). Estudie esta ecuación, y considere el hecho de que $(\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{v}(\mathbf{v}^T\mathbf{x}) = (\mathbf{v}^T\mathbf{x})\mathbf{v}$, porque $\mathbf{v}^T\mathbf{x}$ es un escalar.

13. *a*) El cálculo fila-columna de $A\mathbf{u}$ muestra que cada fila de A es ortogonal a cada $\mathbf{u}$ en Nul A. Así, cada fila de A está en $(\text{Nul } A)^\perp$. Como $(\text{Nul } A)^\perp$ es un subespacio, debe contener a todas las combinaciones lineales de las filas de A; por lo que $(\text{Nul } A)^\perp$ contiene a Fil A.

b) Si rango $A = r$, entonces dim Nul $A = n - r$, de acuerdo con el teorema del rango. Por el ejercicio 24*c*) de la sección 6.3,

dim Nul A + dim$(\text{Nul } A)^\perp = n$

Así, dim $(\text{Nul } A)^\perp$ debe ser r. Pero Fil A es un subespacio r-dimensional de $(\text{Nul } A)^\perp$, de acuerdo con el teorema del rango y el inciso *a*). Por lo tanto, Fil A debe coincidir con $(\text{Nul } A)^\perp$.

c) Sustituya A por A^T en el inciso *b*) y concluya que Fil A^T coincide con $(\text{Nul } A^T)^\perp$. Como Fil A^T = Col A, esto prueba *c*).

15. Si $A = URU^T$ con U ortogonal, entonces A es similar a R (porque U es invertible y $U^T = U^{-1}$) y así A tiene los mismos valores propios que R (de acuerdo con el teorema 4 de la sección 5.2), a saber, los n números reales sobre la diagonal de R.

17. [**M**] $\dfrac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = .4618$,

$$\text{cond}(A) \times \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 3363 \times (1.548 \times 10^{-4}) = .5206.$$

Observe que $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ casi es igual a cond(A) por $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$.

19. [**M**] $\dfrac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = 7.178 \times 10^{-8}$, $\dfrac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 2.832 \times 10^{-4}$.

Observe que el cambio relativo en $\mathbf{x}$ es *mucho* más pequeño que el cambio relativo en $\mathbf{b}$. En efecto, puesto que

$$\text{cond}(A) \times \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} = 23{,}683 \times (2.832 \times 10^{-4}) = 6.707$$

el límite teórico del cambio relativo en $\mathbf{x}$ es 6.707 (con cuatro cifras significativas). Este ejercicio muestra que aun cuando el número de condición es grande, el error relativo en una solución no necesita ser tan grande como se podría esperar.

Capítulo 7

Sección 7.1, página 399

1. Simétrica **3.** No es simétrica **5.** No es simétrica

7. Ortogonal, $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$ **9.** No ortogonal

11. Ortogonal, $\begin{bmatrix} 2/3 & 0 & \sqrt{5}/3 \\ 2/3 & 1/\sqrt{5} & -4/\sqrt{45} \\ 1/3 & -2/\sqrt{5} & -2/\sqrt{45} \end{bmatrix}$.

13. $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, D = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$

15. $P = \begin{bmatrix} -4/\sqrt{17} & 1/\sqrt{17} \\ 1/\sqrt{17} & 4/\sqrt{17} \end{bmatrix}, D = \begin{bmatrix} 17 & 0 \\ 0 & 0 \end{bmatrix}$

17. $P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix},$
$D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

19. $P = \begin{bmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2/\sqrt{45} & -1/3 \\ 0 & 5/\sqrt{45} & 2/3 \end{bmatrix},$
$D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

21. $P = \begin{bmatrix} .5 & -.5 & -1/\sqrt{2} & 0 \\ .5 & .5 & 0 & -1/\sqrt{2} \\ .5 & -.5 & 1/\sqrt{2} & 0 \\ .5 & .5 & 0 & 1/\sqrt{2} \end{bmatrix},$
$D = \begin{bmatrix} 9 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

23. $P = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix},$
$D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

27. $(B^TAB)^T = B^TA^TB^{TT}$ Producto de transpuestas en orden inverso
$= B^TAB$ Porque A es simétrica

El resultado sobre B^TB es un caso especial cuando $A = I$. $(BB^T)^T = B^{TT}B^T = BB^T$, por lo que BB^T es simétrica.

29. *Sugerencia:* Utilice una diagonalización ortogonal de A, o recurra al teorema 2.

31. El teorema de diagonalización de la sección 5.3 dice que las columnas de P son vectores propios (linealmente independientes) correspondientes a los valores propios de A listados en la diagonal de D. Así, P tiene exactamente k columnas de vectores propios correspondientes a λ. Esas k columnas forman una base para el espacio propio.

33. $A = 8\mathbf{u}_1\mathbf{u}_1^T + 6\mathbf{u}_2\mathbf{u}_2^T + 3\mathbf{u}_3\mathbf{u}_3^T$

$$= 8\begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 6\begin{bmatrix} 1/6 & 1/6 & -2/6 \\ 1/6 & 1/6 & -2/6 \\ -2/6 & -2/6 & 4/6 \end{bmatrix} + 3\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

35. *Sugerencia:* $(\mathbf{u}\mathbf{u}^T)\mathbf{x} = \mathbf{u}(\mathbf{u}^T\mathbf{x}) = (\mathbf{u}^T\mathbf{x})\mathbf{u}$, porque $\mathbf{u}^T\mathbf{x}$ es un escalar.

Sección 7.2, página 406

1. *a*) $5x_1^2 + \frac{2}{3}x_1x_2 + x_2^2$ *b*) 185 *c*) 16

3. *a*) $\begin{bmatrix} 10 & -3 \\ -3 & -3 \end{bmatrix}$ *b*) $\begin{bmatrix} 5 & 3/2 \\ 3/2 & 0 \end{bmatrix}$

5. *a*) $\begin{bmatrix} 8 & -3 & 2 \\ -3 & 7 & -1 \\ 2 & -1 & -3 \end{bmatrix}$ *b*) $\begin{bmatrix} 0 & 2 & 3 \\ 2 & 0 & -4 \\ 3 & -4 & 0 \end{bmatrix}$

7. $\mathbf{x} = P\mathbf{y}$, donde $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $\mathbf{y}^T D\mathbf{y} = 6y_1^2 - 4y_2^2$.

En los ejercicios 9 a 14, son posibles otras respuestas (cambio de variables y nueva forma cuadrática).

9. Positiva definida; los valores propios son 2 y 7.

Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}$

Nueva forma cuadrática: $7y_1^2 + 2y_2^2$

11. Indefinida; los valores propios son 7 y -3

Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

Nueva forma cuadrática: $7y_1^2 - 3y_2^2$

13. Positiva semidefinida; los valores propios son 0 y 10

Cambio de variable: $\mathbf{x} = P\mathbf{y}$, con $P = \frac{1}{\sqrt{10}}\begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}$

Nueva forma cuadrática: $10y_1^2$

15. [**M**] Negativa semidefinida; los valores propios son $0, -6, -8, -12$
Cambio de variable: $\mathbf{x} = P\mathbf{y}$;

$$P = \begin{bmatrix} 3/\sqrt{12} & 0 & -1/2 & 0 \\ 1/\sqrt{12} & -2/\sqrt{6} & 1/2 & 0 \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & -1/\sqrt{2} \\ 1/\sqrt{12} & 1/\sqrt{6} & 1/2 & 1/\sqrt{2} \end{bmatrix}$$

Nueva forma cuadrática: $-6y_2^2 - 8y_3^2 - 12y_4^2$

17. [**M**] Indefinida; los valores propios son 8.5 y -6.5
Cambio de variable: $\mathbf{x} = P\mathbf{y}$;

$$P = \frac{1}{\sqrt{50}}\begin{bmatrix} 3 & -4 & 3 & 4 \\ 5 & 0 & -5 & 0 \\ 4 & 3 & 4 & -3 \\ 0 & 5 & 0 & 5 \end{bmatrix}$$

Nueva forma cuadrática: $8.5y_1^2 + 8.5y_2^2 - 6.5y_3^2 - 6.5y_4^2$

19. 8

23. Escriba el polinomio característico de dos formas:

$$\det(A - \lambda I) = \det\begin{bmatrix} a-\lambda & b \\ b & d-\lambda \end{bmatrix}$$
$$= \lambda^2 - (a+d)\lambda + ad - b^2$$

y

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2$$

Iguale los coeficientes para obtener $\lambda_1 + \lambda_2 = a + d$ y $\lambda_1\lambda_2 = ad - b^2 = \det A$.

25. El ejercicio 27 de la sección 7.1 demostró que B^TB es simétrica. Además, $\mathbf{x}^TB^TB\mathbf{x} = (B\mathbf{x})^TB\mathbf{x} = \|B\mathbf{x}\|^2 \geq 0$, así que la forma cuadrática es positiva semidefinida, y se dice que la matriz B^TB es positiva semidefinida. *Sugerencia:* Para demostrar que B^TB es positiva definida cuando B es cuadrada e invertible, suponga que $\mathbf{x}^TB^TB\mathbf{x} = 0$ y deduzca que $\mathbf{x} = \mathbf{0}$.

27. *Sugerencia:* Demuestre que $A + B$ es simétrica y que la forma cuadrática $\mathbf{x}^T(A+B)\mathbf{x}$ es positiva definida.

Sección 7.3, página 413

1. $\mathbf{x} = P\mathbf{y}$, donde $P = \begin{bmatrix} 1/3 & 2/3 & -2/3 \\ 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$

3. *a*) 9 *b*) $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$ *c*) 6

5. *a*) 7 *b*) $\pm\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ *c*) 3

7. $\pm\begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}$ **9.** $5 + \sqrt{5}$ **11.** 3

13. *Sugerencia:* Si $m = M$, tome $\alpha = 0$ en la fórmula para $\mathbf{x}$. Es decir, sea $\mathbf{x} = \mathbf{u}_n$, y compruebe que $\mathbf{x}^TA\mathbf{x} = m$. Si $m < M$ y si t es un número entre m y M, entonces $0 \leq t - m \leq M - m$ y $0 \leq (t-m)/(M-m) \leq 1$. Sea $\alpha = (t-m)/(M-m)$. Resuelva la expresión para α para ver que $t = (1-\alpha)m + \alpha M$. Conforme α va de 0 a 1, t va de m a M. Construya $\mathbf{x}$ como en el enunciado del ejercicio, y compruebe sus propiedades.

15. [M] *a*) 7.5 *b*) $\begin{bmatrix} .5 \\ .5 \\ .5 \\ .5 \end{bmatrix}$ *c*) $-.5$

17. [M] *a*) -4 *b*) $\begin{bmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{bmatrix}$ *c*) -10

Sección 7.4, página 423

1. 3, 1 **3.** 3, 2

Las respuestas en los ejercicios 5 a 13 no son las únicas posibilidades.

5. $\begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$

9. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}\begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}$
$\times\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$

11. $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}\begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\times\begin{bmatrix} 3/\sqrt{10} & -1/\sqrt{10} \\ 1/\sqrt{10} & 3/\sqrt{10} \end{bmatrix}$

13. $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$

$$= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$
$$\times\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{18} & 1/\sqrt{18} & -4/\sqrt{18} \\ -2/3 & 2/3 & 1/3 \end{bmatrix}$$

15. *a*) rango $A = 2$

b) Base para Col A: $\begin{bmatrix} .40 \\ .37 \\ -.84 \end{bmatrix}, \begin{bmatrix} -.78 \\ -.33 \\ -.52 \end{bmatrix}$

Base para Nul A: $\begin{bmatrix} .58 \\ -.58 \\ .58 \end{bmatrix}$

(Recuerde que V^T aparece en la DVS).

17. Sea $A = U\Sigma V^T = U\Sigma V^{-1}$. Puesto que A es cuadrada e invertible, rango $A = n$, y todas las entradas en la diagonal de Σ deben ser distintas de cero. Así, $A^{-1} = (U\Sigma V^{-1})^{-1} = V\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$.

19. *Sugerencia:* Como U y V son ortogonales,

$$A^TA = (U\Sigma V^T)^TU\Sigma V^T = V\Sigma^TU^TU\Sigma V^T$$
$$= V(\Sigma^T\Sigma)V^{-1}$$

Así, V diagonaliza A^TA. ¿Qué indica esto acerca de V?

21. Sea $A = U\Sigma V^T$. La matriz PU es ortogonal, porque P y U son ortogonales. (Véase el ejercicio 29 de la sección 6.2). Así, la ecuación $PA = (PU)\Sigma V^T$ tiene la forma requerida para una descomposición en valores singulares. Según el ejercicio 19, las entradas diagonales en Σ son los valores singulares de PA.

23. *Sugerencia:* Utilice la expansión columna-fila de $(U\Sigma)V^T$.

25. *Sugerencia:* Considere la DVS para la matriz estándar de T, por ejemplo, $A = U\Sigma V^T = U\Sigma V^{-1}$. Sean $\mathcal{B} = \{\mathbf{v}_1,..., \mathbf{v}_n\}$ y $\mathcal{C} = \{\mathbf{u}_1,\ldots, \mathbf{u}_m\}$ bases construidas de las columnas de V y U, respectivamente. Calcule la matriz para T respecto de $\mathcal{B}$ y $\mathcal{C}$, como en la sección 5.4. Para hacer esto, se debe probar que $V^{-1}\mathbf{v}_j = \mathbf{e}_j$, la j-ésima columna de I_n.

27. [**M**] $\begin{bmatrix} -.57 & -.65 & -.42 & .27 \\ .63 & -.24 & -.68 & -.29 \\ .07 & -.63 & .53 & -.56 \\ -.51 & .34 & -.29 & -.73 \end{bmatrix}$

$$\times \begin{bmatrix} 16.46 & 0 & 0 & 0 & 0 \\ 0 & 12.16 & 0 & 0 & 0 \\ 0 & 0 & 4.87 & 0 & 0 \\ 0 & 0 & 0 & 4.31 & 0 \end{bmatrix}$$

$$\times \begin{bmatrix} -.10 & .61 & -.21 & -.52 & .55 \\ -.39 & .29 & .84 & -.14 & -.19 \\ -.74 & -.27 & -.07 & .38 & .49 \\ .41 & -.50 & .45 & -.23 & .58 \\ -.36 & -.48 & -.19 & -.72 & -.29 \end{bmatrix}$$

29. [**M**] 25.9343, 16.7554, 11.2917, 1.0785, .00037793; $\sigma_1/\sigma_5 = 68{,}622$

Sección 7.5, página 430

1. $M = \begin{bmatrix} 12 \\ 10 \end{bmatrix}$; $B = \begin{bmatrix} 7 & 10 & -6 & -9 & -10 & 8 \\ 2 & -4 & -1 & 5 & 3 & -5 \end{bmatrix}$;

$S = \begin{bmatrix} 86 & -27 \\ -27 & 16 \end{bmatrix}$

3. $\begin{bmatrix} .95 \\ -.32 \end{bmatrix}$ para $\lambda = 95.2$, $\begin{bmatrix} .32 \\ .95 \end{bmatrix}$ para $\lambda = 6.8$

5. [**M**] (.130, 874, 468), 75.9% de la varianza

7. $y_1 = .95x_1 - .32x_2$; y_1 explica el 93.3% de la varianza.

9. $c_1 = 1/3$, $c_2 = 2/3$, $c_3 = 2/3$; la varianza de y es 9.

11. ***a)*** Si $\mathbf{w}$ es el vector en $\mathbb{R}^N$ con un 1 en cada posición, entonces

$$\begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_N \end{bmatrix} \mathbf{w} = \mathbf{X}_1 + \cdots + \mathbf{X}_N = \mathbf{0}$$

porque las $\mathbf{X}_k$ están en la forma de desviación media. Entonces,

$$\begin{aligned} &\begin{bmatrix} \mathbf{Y}_1 & \cdots & \mathbf{Y}_N \end{bmatrix} \mathbf{w} \\ &= \begin{bmatrix} P^T\mathbf{X}_1 & \cdots & P^T\mathbf{X}_N \end{bmatrix} \mathbf{w} \qquad \text{Por definición} \\ &= P^T\begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_N \end{bmatrix} \mathbf{w} = P^T\mathbf{0} = \mathbf{0} \end{aligned}$$

Es decir, $\mathbf{Y}_1 + \cdots + \mathbf{Y}_N = \mathbf{0}$, por lo que las $\mathbf{Y}_k$ están en forma de desviación media.

b) *Sugerencia:* Como las $\mathbf{X}_j$ están en forma de desviación media, la matriz de covarianza de las $\mathbf{X}_j$ es

$$1/(N-1)\begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_N \end{bmatrix}\begin{bmatrix} \mathbf{X}_1 & \cdots & \mathbf{X}_N \end{bmatrix}^T$$

Calcule la matriz de covarianza de las $\mathbf{Y}_j$, con base en el inciso *a*).

13. Si $B = \begin{bmatrix} \hat{\mathbf{X}}_1 & \cdots & \hat{\mathbf{X}}_N \end{bmatrix}$, entonces

$$\begin{aligned} S &= \frac{1}{N-1}BB^T = \frac{1}{N-1}\begin{bmatrix} \hat{\mathbf{X}}_1 & \cdots & \hat{\mathbf{X}}_n \end{bmatrix}\begin{bmatrix} \hat{\mathbf{X}}_1^T \\ \vdots \\ \hat{\mathbf{X}}_N^T \end{bmatrix} \\ &= \frac{1}{N-1}\sum_1^N \hat{\mathbf{X}}_k\hat{\mathbf{X}}_k^T = \frac{1}{N-1}\sum_1^N (\mathbf{X}_k - \mathbf{M})(\mathbf{X}_k - \mathbf{M})^T \end{aligned}$$

Capítulo 7 Ejercicios complementarios, página 432

1. ***a)*** V ***b)*** F ***c)*** V ***d)*** F ***e)*** F ***f)*** F ***g)*** F ***h)*** V ***i)*** F ***j)*** F ***k)*** F ***l)*** F ***m)*** V ***n)*** F ***o)*** V ***p)*** V ***q)*** F

3. Si rango $A = r$, entonces dim Nul $A = n - r$, de acuerdo con el teorema del rango. Así, 0 es un valor propio con multiplicidad $n - r$. Por lo tanto, de los n términos en la descomposición espectral de A, exactamente $n - r$ son nulos. Los restantes r términos (correspondientes a los valores propios diferentes de cero) son matrices con rango 1, como se mencionó en el análisis de la descomposición espectral.

5. Si $A\mathbf{v} = \lambda\mathbf{v}$ para alguna λ diferente de cero, entonces $\mathbf{v} = \lambda^{-1}A\mathbf{v} = A(\lambda^{-1}\mathbf{v})$, lo que muestra que $\mathbf{v}$ es una combinación lineal de las columnas de A.

7. *Sugerencia:* Si $A = R^TR$, donde R es invertible, entonces A es positiva definida, por el ejercicio 25 de la sección 7.2. Al contrario, suponga que A es positiva definida. Entonces por el ejercicio 26 de la sección 7.2, $A = B^TB$ para alguna matriz positiva definida B. Explique por qué B admite una factorización QR, y úsela para crear la factorización de Cholesky de A.

9. Si A es de $m \times n$ y $\mathbf{x}$ está en $\mathbb{R}^n$, entonces $\mathbf{x}^TA^TA\mathbf{x} = (A\mathbf{x})^T(A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$. Así, A^TA es positiva semidefinida. Por el ejercicio 22 de la sección 6.5, rango A^TA = rango A.

11. *Sugerencia:* Escriba una DVS de A en la forma $A = U\Sigma V^T = PQ$, donde $P = U\Sigma U^T$ y $Q = UV^T$. Demuestre que P es simétrica y con los mismos valores propios que Σ. Explique por qué Q es una matriz ortogonal.

13. ***a)*** Si $\mathbf{b} = A\mathbf{x}$, entonces $\mathbf{x}^+ = A^+\mathbf{b} = A^+A\mathbf{x}$. De acuerdo con el ejercicio 12*a*), $\mathbf{x}^+$ es la proyección ortogonal de $\mathbf{x}$ sobre Fil A.

b) A partir de *a*) y del ejercicio 12*c*),

$A\mathbf{x}^+ = A(A^+A\mathbf{x}) = (AA^+A)\mathbf{x} = A\mathbf{x} = \mathbf{b}$

c) Como $\mathbf{x}^+$ es la proyección ortogonal sobre Fil A, el teorema de Pitágoras indica que $\|\mathbf{u}\|^2 = \|\mathbf{x}^+\|^2 + \|\mathbf{u} - \mathbf{x}^+\|^2$. El inciso *c*) se deduce inmediatamente.

15. [**M**] $A^+ = \dfrac{1}{40} \cdot \begin{bmatrix} -2 & -14 & 13 & 13 \\ -2 & -14 & 13 & 13 \\ -2 & 6 & -7 & -7 \\ 2 & -6 & 7 & 7 \\ 4 & -12 & -6 & -6 \end{bmatrix}$, $\hat{\mathbf{x}} = \begin{bmatrix} .7 \\ .7 \\ -.8 \\ .8 \\ .6 \end{bmatrix}$

La forma escalonada reducida de $\begin{bmatrix} A \\ \mathbf{x}^T \end{bmatrix}$ es la misma que la forma escalonada reducida de A, excepto por una fila adicional de ceros. Así que al sumar múltiplos escalares de las filas de A a $\mathbf{x}^T$ se puede obtener el vector cero, lo que prueba que $\mathbf{x}^T$ está en Fil A.

Base para Nul A: $\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

Capítulo 8

Sección 8.1, página 442

1. Algunas posibles respuestas: $\mathbf{y} = 2\mathbf{v}_1 - 1.5\mathbf{v}_2 + .5\mathbf{v}_3$, $\mathbf{y} = 2\mathbf{v}_1 - 2\mathbf{v}_3 + \mathbf{v}_4$, $\mathbf{y} = 2\mathbf{v}_1 + 3\mathbf{v}_2 - 7\mathbf{v}_3 + 3\mathbf{v}_4$

3. $\mathbf{y} = -3\mathbf{v}_1 + 2\mathbf{v}_2 + 2\mathbf{v}_3$. Los pesos suman 1, por lo que es una suma afín.

5. *a*) $\mathbf{p}_1 = 3\mathbf{b}_1 - \mathbf{b}_2 - \mathbf{b}_3 \in \text{aff } S$ porque los coeficientes suman 1.

b) $\mathbf{p}_2 = 2\mathbf{b}_1 + 0\mathbf{b}_2 + \mathbf{b}_3 \notin \text{aff } S$ porque los coeficientes no suman 1.

c) $\mathbf{p}_3 = -\mathbf{b}_1 + 2\mathbf{b}_2 + 0\mathbf{b}_3 \in \text{aff } S$ ya que los coeficientes suman 1.

7. *a*) $\mathbf{p}_1 \in \text{Gen } S$, pero $\mathbf{p}_1 \notin \text{aff } S$

b) $\mathbf{p}_2 \in \text{Gen } S$, y $\mathbf{p}_2 \in \text{aff } S$

c) $\mathbf{p}_3 \notin \text{Gen } S$, por lo que $\mathbf{p}_3 \notin \text{aff } S$.

9. $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Otras respuestas son posibles.

13. Gen $\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$ es un plano si y solo si $\{\mathbf{v}_2 - \mathbf{v}_1, \mathbf{v}_3 - \mathbf{v}_1\}$ es linealmente independiente. Suponga que c_2 y c_3 satisfacen $c_2(\mathbf{v}_2 - \mathbf{v}_1) + c_3(\mathbf{v}_3 - \mathbf{v}_1) = \mathbf{0}$. Demuestre que esto implica $c_2 = c_3 = 0$.

15. Sea $S = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$. Para probar que S es afín, es suficiente demostrar que S es un plano afín, de acuerdo con el teorema 3. Sea $W = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$. Entonces, W es un subespacio de $\mathbb{R}^n$, de acuerdo con el teorema 2 de la sección 4.2 (o el teorema 12 de la sección 2.8). Como $S = W + \mathbf{p}$, donde $\mathbf{p}$ satisface $A\mathbf{p} = \mathbf{b}$, de acuerdo con el teorema 6 de la sección 1.5, S es un traslado de W, y por lo tanto, S es un plano afín.

17. Un conjunto conveniente consta de cualesquiera tres vectores que no sean colineales y tengan 5 en su tercer entrada. Si 5 es su tercera entrada, estos se encuentran en el plano $z = 5$. Si los vectores no son colineales, entonces su envoltura afín no puede ser una recta, así que debe ser un plano.

19. Si $\mathbf{p}, \mathbf{q} \in f(S)$, entonces existen $\mathbf{r}, \mathbf{s} \in S$ tales que $f(\mathbf{r}) = \mathbf{p}$ y $f(\mathbf{s}) = \mathbf{q}$. Dado cualquier $t \in \mathbb{R}$, debe demostrarse que $\mathbf{z} = (1 - t)\mathbf{p} + t\mathbf{q}$ está en $f(S)$. Ahora utilice las definiciones de $\mathbf{p}$ y $\mathbf{q}$, y considere el hecho de que f es lineal.

21. Como B es afín, el teorema 1 implica que B contiene todas las combinaciones afines de puntos de B. Por lo tanto, B contiene todas las combinaciones afines de puntos de A. Es decir, aff $A \subset B$.

23. Como $A \subset (A \cup B)$, del ejercicio 22 se deduce que aff $A \subset$ aff $(A \cup B)$. De forma similar, aff $B \subset$ aff $(A \cup B)$, de manera que [aff $A \cup$ aff B] $\subset$ aff $(A \cup B)$.

25. Para demostrar que $D \subset E \cap F$, demuestre que $D \subset E$ y $D \subset F$.

Sección 8.2, página 452

1. Afínmente dependiente y $2\mathbf{v}_1 + \mathbf{v}_2 - 3\mathbf{v}_3 = \mathbf{0}$.

3. El conjunto tiene independencia afín. Si los puntos son $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ y $\mathbf{v}_4$, entonces $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base para $\mathbb{R}^3$, y $\mathbf{v}_4 = 16\mathbf{v}_1 + 5\mathbf{v}_2 - 3\mathbf{v}_3$, pero los pesos en la combinación lineal no suman 1.

5. $-4\mathbf{v}_1 + 5\mathbf{v}_2 - 4\mathbf{v}_3 + 3\mathbf{v}_4 = \mathbf{0}$

7. Las coordenadas baricéntricas son $(-2, 4, -1)$.

11. Cuando un conjunto de cinco puntos se traslada restándole, por ejemplo, el primer punto, entonces el nuevo conjunto de cuatro puntos debe ser linealmente dependiente, de acuerdo con el teorema 8 de la sección 1.7, porque los cuatro puntos están en $\mathbb{R}^3$. Según el teorema 5, el conjunto original de cinco puntos es afínmente dependiente.

13. Si $\{\mathbf{v}_1, \mathbf{v}_2\}$ es afínmente dependiente, entonces existen c_1 y c_2, sin que ambos sean cero, tales que $c_1 + c_2 = 0$ y $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = 0$. Demuestre que esto implica que $\mathbf{v}_1 = \mathbf{v}_2$. Para lo inverso, suponga que $\mathbf{v}_1 = \mathbf{v}_2$ y seleccione c_1 y c_2 específicos que muestren su dependencia afín.

15. *a*) Los vectores $\mathbf{v}_2 - \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v}_3 - \mathbf{v}_1 = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ no son múltiplos y entonces son linealmente independientes. De acuerdo con el teorema 5, S es afínmente independiente.

b) $\mathbf{p}_1 \leftrightarrow \left(-\frac{6}{8}, \frac{9}{8}, \frac{5}{8}\right)$, $\mathbf{p}_2 \leftrightarrow \left(0, \frac{1}{2}, \frac{1}{2}\right)$, $\mathbf{p}_3 \leftrightarrow \left(\frac{14}{8}, -\frac{5}{8}, -\frac{1}{8}\right)$, $\mathbf{p}_4 \leftrightarrow \left(\frac{6}{8}, -\frac{5}{8}, \frac{7}{8}\right)$, $\mathbf{p}_5 \leftrightarrow \left(\frac{1}{4}, \frac{1}{8}, \frac{5}{8}\right)$

c) $\mathbf{p}_6$ es $(-, -, +)$, $\mathbf{p}_7$ es $(0, +, -)$ y $\mathbf{p}_8$ es $(+, +, -)$

17. Suponga que $S = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ es un conjunto afínmente independiente. Entonces, la ecuación (7) tiene una solución porque $\mathbf{p}$ está en aff S. Por lo tanto, la ecuación (8) tiene una solución. Según el teorema 5, las formas homogéneas de los puntos en S son linealmente independientes. Por lo que (8) tiene solución única. Entonces (7) también tiene solución única, porque (8) contiene ambas ecuaciones que aparecen en (7).

El siguiente argumento es similar a la demostración del teorema 7 de la sección 4.4. Si $S = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ es un conjunto afínmente independiente, entonces existen escalares $c_1, \ldots, c_k$ que satisfacen (7), por definición de aff S. Suponga que $\mathbf{x}$ también tiene la representación

$$\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_k\mathbf{b}_k \;\text{ y }\; d_1 + \cdots + d_k = 1 \tag{7a}$$

para escalares $d_1, \ldots, d_k$. Entonces, la resta produce la ecuación

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_k - d_k)\mathbf{b}_k \tag{7b}$$

Los pesos en (7b) suman cero porque las c y las d suman 1 por separado. Esto es imposible, a menos que cada peso en (8) sea igual a 0, ya que S es un conjunto afínmente independiente. Esto demuestra que $c_i = d_i$ para $i = 1, \ldots, k$.

19. Si $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ es un conjunto afínmente dependiente, entonces existen escalares c_1, c_2 y c_3, no todos cero, tales que $c_1\mathbf{p}_1 + c_2\mathbf{p}_2 + c_3\mathbf{p}_3 = \mathbf{0}$ y $c_1 + c_2 + c_3 = 0$. Ahora utilice la linealidad de f.

21. Sea $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ y $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$. Entonces,

$$\det[\,\tilde{\mathbf{a}} \quad \tilde{\mathbf{b}} \quad \tilde{\mathbf{c}}\,] = \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 1 & 1 & 1 \end{bmatrix} =$$

$\det \begin{bmatrix} a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \\ c_1 & c_2 & 1 \end{bmatrix}$, por la propiedad del determinante de la transpuesta (teorema 5 de la sección 3.2). Por el ejercicio 30 de la sección 3.3, este determinante es igual a 2 veces el área del triángulo con vértices en **a**, **b** y **c**.

23. Si $[\,\tilde{\mathbf{a}} \quad \tilde{\mathbf{b}} \quad \tilde{\mathbf{c}}\,]\begin{bmatrix} r \\ s \\ t \end{bmatrix} = \tilde{\mathbf{p}}$, entonces la regla de Cramer da $r = \det[\,\tilde{\mathbf{p}} \quad \tilde{\mathbf{b}} \quad \tilde{\mathbf{c}}\,]/\det[\,\tilde{\mathbf{a}} \quad \tilde{\mathbf{b}} \quad \tilde{\mathbf{c}}\,]$. Por el ejercicio 21, el numerador de este cociente es dos veces el área del $\Delta\mathbf{pbc}$, y el denominador es el doble del área del $\Delta\mathbf{abc}$. Esto demuestra la fórmula para r. Las otras fórmulas se pueden probar empleando la regla de Cramer para s y t.

Sección 8.3, página 459

3. Ninguno está en conv S.

5. $\mathbf{p}_1 = -\frac{1}{6}\mathbf{v}_1 + \frac{1}{3}\mathbf{v}_2 + \frac{2}{3}\mathbf{v}_3 + \frac{1}{6}\mathbf{v}_4$, por lo que $\mathbf{p}_1 \notin$ conv S.
$\mathbf{p}_2 = \frac{1}{3}\mathbf{v}_1 + \frac{1}{3}\mathbf{v}_2 + \frac{1}{6}\mathbf{v}_3 + \frac{1}{6}\mathbf{v}_4$, por lo que $\mathbf{p}_2 \in$ conv S.

7. *a*) Las coordenadas baricéntricas de $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ y $\mathbf{p}_4$ son, respectivamente, $\left(\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right)$, $\left(0, \frac{1}{2}, \frac{1}{2}\right)$, $\left(\frac{1}{2}, -\frac{1}{4}, \frac{3}{4}\right)$ y $\left(\frac{1}{2}, \frac{3}{4}, -\frac{1}{4}\right)$.
b) $\mathbf{p}_3$ y $\mathbf{p}_4$ están fuera de conv T. $\mathbf{p}_1$ está dentro de conv T. $\mathbf{p}_2$ está en la arista $\overline{\mathbf{v}_2\mathbf{v}_3}$ de conv T.

9. $\mathbf{p}_1$ y $\mathbf{p}_3$ están fuera del tetraedro conv S. $\mathbf{p}_2$ está sobre la cara que contiene los vértices $\mathbf{v}_2$, $\mathbf{v}_3$ y $\mathbf{v}_4$. $\mathbf{p}_4$ está dentro de conv S. $\mathbf{p}_5$ está sobre la arista entre $\mathbf{v}_1$ y $\mathbf{v}_3$.

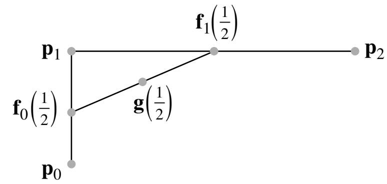
13. Si $\mathbf{p}, \mathbf{q} \in f(S)$, entonces existen $\mathbf{r}, \mathbf{s} \in S$ tales que $f(\mathbf{r}) = \mathbf{p}$ y $f(\mathbf{s}) = \mathbf{q}$. El objetivo es mostrar que el segmento de recta $\mathbf{y} = (1 - t)\mathbf{p} + t\mathbf{q}$, para $0 \le t \le 1$, está en $f(S)$. Utilice la linealidad de f y la convexidad de S para demostrar que $\mathbf{y} = f(\mathbf{w})$ para alguna $\mathbf{w}$ en S. Esto probará que $\mathbf{y}$ está en $f(S)$ y que $f(S)$ es convexo.

15. $\mathbf{p} = \frac{1}{6}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_4$ y $\mathbf{p} = \frac{1}{2}\mathbf{v}_1 + \frac{1}{6}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_3$.

17. Suponga que $A \subset B$, donde B es convexo. Entonces, como B es convexo, el teorema 7 implica que B contiene todas las combinaciones convexas de puntos de B. Por lo tanto, B contiene todas las combinaciones convexas de puntos de A. Es decir, conv $A \subset B$.

19. *a*) Con base en el ejercicio 18, demuestre que conv A y conv B son subconjuntos de conv $(A \cup B)$. Esto implicará que su unión también es un subconjunto de conv $(A \cup B)$.
b) Una posibilidad es dejar que A esté formado por dos esquinas adyacentes de un cuadrado, y B por las otras dos esquinas. Entonces, ¿qué son (conv A) $\cup$ (conv B), y conv $(A \cup B)$?

21.



23.
$$\begin{aligned} \mathbf{g}(t) &= (1-t)\mathbf{f}_0(t) + t\mathbf{f}_1(t) \\ &= (1-t)[(1-t)\mathbf{p}_0 + t\mathbf{p}_1] + t[(1-t)\mathbf{p}_1 + t\mathbf{p}_2] \\ &= (1-t)^2\mathbf{p}_0 + 2t(1-t)\mathbf{p}_1 + t^2\mathbf{p}_2. \end{aligned}$$

La suma de los pesos en la combinación lineal para **g** es $(1-t)^2 + 2t(1-t) + t^2$, que es igual a $(1 - 2t + t^2) + (2t - 2t^2) + t^2 = 1$. Cada uno de los pesos está entre 0 y 1 cuando $0 \le t \le 1$, así que $\mathbf{g}(t)$ está en conv $\{\mathbf{p}_0, \mathbf{p}_1, \mathbf{p}_2\}$.

Sección 8.4, página 467

1. $f(x_1, x_2) = 3x_1 + 4x_2$ y $d = 13$.

3. *a*) Abierto *b*) Cerrado *c*) Ni abierto ni cerrado
d) Cerrado *e*) Cerrado

5. *a*) No es compacto, convexo
b) Compacto, convexo
c) No es compacto, convexo
d) No es compacto, no es convexo
e) No es compacto, convexo

7. *a*) $\mathbf{n} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$ o un múltiplo
b) $f(\mathbf{x}) = 2x_2 + 3x_3$, $d = 11$

9. *a*) $\mathbf{n} = \begin{bmatrix} 3 \\ -1 \\ 2 \\ 1 \end{bmatrix}$ o un múltiplo
b) $f(\mathbf{x}) = 3x_1 - x_2 + 2x_3 + x_4$, $d = 5$

11. $\mathbf{v}_2$ está en el mismo lado que $\mathbf{0}$, $\mathbf{v}_1$ está en el otro lado, y $\mathbf{v}_3$ está en H.

13. Una posibilidad es $\mathbf{p} = \begin{bmatrix} 32 \\ -14 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 10 \\ -7 \\ 1 \\ 0 \end{bmatrix}$,
$\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 1 \end{bmatrix}$.

15. $f(x_1, x_2, x_3) = x_1 - 3x_2 + 4x_3 - 2x_4$ y $d = 5$

17. $f(x_1, x_2, x_3) = x_1 - 2x_2 + x_3$ y $d = 0$

19. $f(x_1, x_2, x_3) = -5x_1 + 3x_2 + x_3$ y $d = 0$

23. $f(x_1, x_2) = 3x_1 - 2x_2$ con d que satisface $9 < d < 10$ es una posibilidad.

25. $f(x, y) = 4x + 1$. Una elección natural para d es 12.75, que es igual a $f(3, .75)$. El punto $(3, .75)$ es 3/4 de la distancia entre el centro $B(\mathbf{0}, 3)$ y el centro de $B(\mathbf{p}, 1)$.

27. El ejercicio 2*a*) de la sección 8.3 da una posibilidad. O bien, sea $S = \{(x, y) : x^2y^2 = 1 \text{ y } y > 0\}$. Entonces, conv S es el semiplano superior (abierto).

29. Sean $\mathbf{x}, \mathbf{y} \in B(\mathbf{p}, \delta)$ y suponga que $\mathbf{z} = (1 - t)\mathbf{x} + t\mathbf{y}$, donde $0 \le t \le 1$. Después, demuestre que

$$\begin{aligned}\|\mathbf{z} - \mathbf{p}\| &= \|[(1 - t)\mathbf{x} + t\mathbf{y}] - \mathbf{p}\| \\ &= \|(1 - t)(\mathbf{x} - \mathbf{p}) + t(\mathbf{y} - \mathbf{p})\| < \delta.\end{aligned}$$

Sección 8.5, página 479

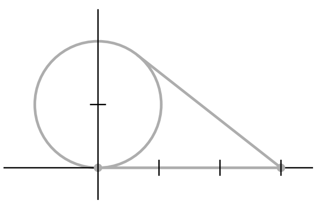
1. *a*) $m = 1$ en el punto $\mathbf{p}_1$ *b*) $m = 5$ en el punto $\mathbf{p}_2$
c) $m = 5$ en el punto $\mathbf{p}_3$

3. *a*) $m = -3$ en el punto $\mathbf{p}_3$
b) $m = 1$ en el conjunto conv $\{\mathbf{p}_1, \mathbf{p}_3\}$
c) $m = -3$ en el conjunto conv $\{\mathbf{p}_1, \mathbf{p}_2\}$

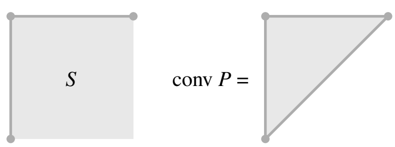
5. $\left\{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}5\\0\end{bmatrix}, \begin{bmatrix}4\\3\end{bmatrix}, \begin{bmatrix}0\\5\end{bmatrix}\right\}$

7. $\left\{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}7\\0\end{bmatrix}, \begin{bmatrix}6\\4\end{bmatrix}, \begin{bmatrix}0\\6\end{bmatrix}\right\}$

9. El origen es un punto extremo, pero no es un vértice. Explique por qué.



11. Una posibilidad es hacer que S sea un cuadrado que incluya parte de la frontera, pero no toda. Por ejemplo, que incluya solamente dos aristas adyacentes. La envoltura convexa del perfil P es una región triangular.



13. *a*) $f_0(C^5) = 32$, $f_1(C^5) = 80$, $f_2(C^5) = 80$, $f_3(C^5) = 40$, $f_4(C^5) = 10$, y $32 - 80 + 80 - 40 + 10 = 2$.

b)

	f_0	f_1	f_2	f_3	f_4
S^1	2				
S^2	4	4			
S^3	8	12	6		
S^4	16	32	24	8	
S^5	32	80	80	40	10

15. *a*) $f_0(P^n) = f_0(Q) + 1$
b) $f_k(P^n) = f_k(Q) + f_{k-1}(Q)$
c) $f_{n-1}(P^n) = f_{n-2}(Q) + 1$

19. Sean S convexo y $\mathbf{x} \in cS + dS$, donde $c > 0$ y $d > 0$. Entonces, existen $\mathbf{s}_1$ y $\mathbf{s}_2$ en S tales que $\mathbf{x} = c\mathbf{s}_1 + d\mathbf{s}_2$. Pero entonces,

$$\mathbf{x} = c\mathbf{s}_1 + d\mathbf{s}_2 = (c + d)\left(\frac{c}{c + d}\mathbf{s}_1 + \frac{d}{c + d}\mathbf{s}_2\right).$$

Ahora demuestre que la expresión en el lado derecho es un miembro de $(c + d)S$.

Para el inverso, tome un punto típico en $(c + d)S$ y demuestre que está en $cS + dS$.

21. *Sugerencia:* Suponga que A y B son convexos. Sean $\mathbf{x}, \mathbf{y} \in A + B$. Entonces, existen $\mathbf{a}, \mathbf{c} \in A$ y $\mathbf{b}, \mathbf{d} \in B$ tales que $\mathbf{x} = \mathbf{a} + \mathbf{b}$ y $\mathbf{y} = \mathbf{c} + \mathbf{d}$. Para cualquier t tal que $0 \le t \le 1$, demuestre que

$$\mathbf{w} = (1 - t)\mathbf{x} + t\mathbf{y} = (1 - t)(\mathbf{a} + \mathbf{b}) + t(\mathbf{c} + \mathbf{d})$$

representa un punto en $A + B$.

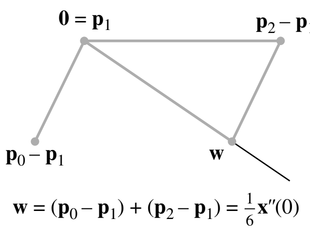
Sección 8.6, página 490

1. Los puntos de control para $\mathbf{x}(t) + \mathbf{b}$ deberían ser $\mathbf{p}_0 + \mathbf{b}$, $\mathbf{p}_1 + \mathbf{b}$ y $\mathbf{p}_3 + \mathbf{b}$. Escriba la curva de Bézier que pasa a través de esos puntos, y pruebe algebraicamente que esta curva es $\mathbf{x}(t) + \mathbf{b}$.

3. *a*) $\mathbf{x}'(t) = (-3 + 6t - 3t_2)\mathbf{p}_0 + (3 - 12t + 9t^2)\mathbf{p}_1 + (6t - 9t^2)\mathbf{p}_2 + 3t^2\mathbf{p}_3$ por lo que $\mathbf{x}'(0) = -3\mathbf{p}_0 + 3\mathbf{p}_1 = 3(\mathbf{p}_1 - \mathbf{p}_0)$ y $\mathbf{x}'(1) = -3\mathbf{p}_2 + 3\mathbf{p}_3 = 3(\mathbf{p}_3 - \mathbf{p}_2)$. Esto demuestra que el vector tangente $\mathbf{x}'(0)$ apunta en la dirección de $\mathbf{p}_0$ a $\mathbf{p}_1$ y es tres veces la longitud de $\mathbf{p}_1 - \mathbf{p}_0$. De igual manera, $\mathbf{x}'(1)$ apunta en la dirección de $\mathbf{p}_2$ a $\mathbf{p}_3$ y es tres veces la longitud de $\mathbf{p}_3 - \mathbf{p}_2$. En particular, $\mathbf{x}'(1) = \mathbf{0}$ si y solo si $\mathbf{p}_3 = \mathbf{p}_2$.

b) $\mathbf{x}''(t) = (6 - 6t)\mathbf{p}_0 + (-12 + 18t)\mathbf{p}_1 + (6 - 18t)\mathbf{p}_2 + 6t\mathbf{p}_3$ así que
$\mathbf{x}''(0) = 6\mathbf{p}_0 - 12\mathbf{p}_1 + 6\mathbf{p}_2 = 6(\mathbf{p}_0 - \mathbf{p}_1) + 6(\mathbf{p}_2 - \mathbf{p}_1)$
y
$\mathbf{x}''(1) = 6\mathbf{p}_1 - 12\mathbf{p}_2 + 6\mathbf{p}_3 = 6(\mathbf{p}_1 - \mathbf{p}_2) + 6(\mathbf{p}_3 - \mathbf{p}_2)$.
Para un esquema de $\mathbf{x}''(0)$, construya un sistema de coordenadas con el origen en $\mathbf{p}_1$, temporalmente, y etiquete $\mathbf{p}_0$ como $\mathbf{p}_0 - \mathbf{p}_1$, y $\mathbf{p}_2$ como $\mathbf{p}_2 - \mathbf{p}_1$. Por último, construya

una recta de este nuevo origen a través de la suma de $\mathbf{p}_0 - \mathbf{p}_1$ y $\mathbf{p}_2 - \mathbf{p}_1$, extendiéndose un poco. Esta recta apunta en la dirección de $\mathbf{x}''(0)$.



5. ***a***) Del ejercicio 3*a*) o de la ecuación (9) en el libro,

$$\mathbf{x}'(1) = 3(\mathbf{p}_3 - \mathbf{p}_2)$$

Utilice la fórmula para $\mathbf{x}'(0)$, con los puntos de control de $\mathbf{y}(t)$, y obtenga

$$\mathbf{y}'(0) = 3\mathbf{p}_3 + 3\mathbf{p}_4 = 3(\mathbf{p}_4 - \mathbf{p}_3)$$

Para continuidad C^1, $3(\mathbf{p}_3 - \mathbf{p}_2) = 3(\mathbf{p}_4 - \mathbf{p}_3)$, de manera que $\mathbf{p}_3 = (\mathbf{p}_4 + \mathbf{p}_2)/2$, y $\mathbf{p}_3$ es el punto medio del segmento de recta de $\mathbf{p}_2$ a $\mathbf{p}_4$.

b) Si $\mathbf{x}'(1) = \mathbf{y}'(0) = \mathbf{0}$, entonces $\mathbf{p}_2 = \mathbf{p}_3$ y $\mathbf{p}_3 = \mathbf{p}_4$. Así, el "segmento de recta" de $\mathbf{p}_2$ a $\mathbf{p}_4$ es justamente el punto $\mathbf{p}_3$. [*Nota:* En este caso, la curva combinada sigue siendo continua C^1, por definición. Sin embargo, algunas elecciones de los otros puntos de "control", $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_5$ y $\mathbf{p}_6$, pueden producir una curva con una esquina visible en $\mathbf{p}_3$; en tal caso la curva no es continua G^1 en $\mathbf{p}_3$].

7. *Sugerencia:* Utilice $\mathbf{x}''(t)$ del ejercicio 3 y adapte esto para la segunda curva para ver que

$$\mathbf{y}''(t) = 6(1-t)\mathbf{p}_3 + 6(-2+3t)\mathbf{p}_4 + 6(1-3t)\mathbf{p}_5 + 6t\mathbf{p}_6$$

Entonces, sea $\mathbf{x}''(1) = \mathbf{y}''(0)$. Como la curva es continua C^1 en $\mathbf{p}_3$, el ejercicio 5*a*) indica que el punto $\mathbf{p}_3$ es el punto medio del segmento de $\mathbf{p}_2$ a $\mathbf{p}_4$. Esto implica que $\mathbf{p}_4 - \mathbf{p}_3 = \mathbf{p}_3 - \mathbf{p}_2$. Utilice esta sustitución para probar que $\mathbf{p}_4$ y $\mathbf{p}_5$ están unívocamente determinados por $\mathbf{p}_1$, $\mathbf{p}_2$ y $\mathbf{p}_3$. Solamente $\mathbf{p}_6$ se puede seleccionar arbitrariamente.

9. Escriba un vector de los pesos polinomiales para $\mathbf{x}(t)$, expanda los polinomios de pesos, y factorice el vector como $M_B\mathbf{u}(t)$:

$$\begin{bmatrix} 1-4t+6t^2-4t^3+t^4 \\ 4t-12t^2+12t^3-4t^4 \\ 6t^2-12t^3+6t^4 \\ 4t^3-4t^4 \\ t^4 \end{bmatrix} = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ 0 & 4 & -12 & 12 & -4 \\ 0 & 0 & 6 & -12 & 6 \\ 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ t \\ t^2 \\ t^3 \\ t^4 \end{bmatrix},$$

$$M_B = \begin{bmatrix} 1 & -4 & 6 & -4 & 1 \\ 0 & 4 & -12 & 12 & -4 \\ 0 & 0 & 6 & -12 & 6 \\ 0 & 0 & 0 & 4 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

13. ***a***) *Sugerencia:* Aplique el hecho de que $\mathbf{q}_0 = \mathbf{p}_0$.

b) Multiplique la primera y la última partes de la ecuación (13) por $\frac{8}{3}$ y despeje $8\mathbf{q}_2$.

c) Utilice la ecuación (8) para sustituir $8\mathbf{q}_3$ y después aplique el inciso *a*).

15. ***a***) De la ecuación (11), $\mathbf{y}'(1) = .5\mathbf{x}'(0.5) = \mathbf{z}'(0)$.

b) Observe que $\mathbf{y}'(1) = 3(\mathbf{q}_3 - \mathbf{q}_2)$. Esto se deduce de la ecuación (9), con $\mathbf{y}(t)$ y sus puntos de control en lugar de $\mathbf{x}(t)$ y los puntos de control de esta. De forma similar, para $\mathbf{z}(t)$ y sus puntos de control, $\mathbf{z}'(0) = 3(\mathbf{r}_1 - \mathbf{r}_0)$. Por el inciso *a*), $3(\mathbf{q}_3 - \mathbf{q}_2) = 3(\mathbf{r}_1 - \mathbf{r}_0)$. Remplace $\mathbf{r}_0$ por $\mathbf{q}_3$, y obtenga $\mathbf{q}_3 - \mathbf{q}_2 = \mathbf{r}_1 - \mathbf{q}_3$, así que $\mathbf{q}_3 = (\mathbf{q}_2 + \mathbf{r}_1)/2$.

c) Sean $\mathbf{q}_0 = \mathbf{p}_0$ y $\mathbf{r}_3 = \mathbf{p}_3$. Calcule $\mathbf{q}_1 = (\mathbf{p}_0 + \mathbf{p}_1)/2$ y $\mathbf{r}_2 = (\mathbf{p}_2 + \mathbf{p}_3)/2$. Calcule $\mathbf{m} = (\mathbf{p}_1 + \mathbf{p}_2)/2$. Calcule $\mathbf{q}_2 = (\mathbf{q}_1 + \mathbf{m})/2$ y $\mathbf{r}_1 = (\mathbf{m} + \mathbf{r}_2)/2$. Calcule $\mathbf{q}_3 = (\mathbf{q}_2 + \mathbf{r}_1)/2$ y sea $\mathbf{r}_0 = \mathbf{q}_3$.

17. ***a***) $\mathbf{r}_0 = \mathbf{p}_0$, $\mathbf{r}_1 = \dfrac{\mathbf{p}_0 + 2\mathbf{p}_1}{3}$, $\mathbf{r}_2 = \dfrac{2\mathbf{p}_1 + \mathbf{p}_2}{3}$, $\mathbf{r}_3 = \mathbf{p}_2$

b) *Sugerencia:* Escriba la fórmula estándar (7) en esta sección, con $\mathbf{r}_i$ en vez de $\mathbf{p}_1$ para $i = 0, \ldots, 3$, y después remplace $\mathbf{r}_0$ y $\mathbf{r}_3$ por $\mathbf{p}_0$ y $\mathbf{p}_2$, respectivamente:

$$\begin{aligned} \mathbf{x}(t) = {} & (1 - 3t + 3t^2 - t^3)\mathbf{p}_0 \\ & + (3t - 6t^2 + 3t^3)\mathbf{r}_1 \\ & + (3t^2 - 3t^3)\mathbf{r}_2 + t^3\mathbf{p}_2 \end{aligned} \qquad (iii)$$

Aplique las fórmulas para $\mathbf{r}_1$ y $\mathbf{r}_2$ del inciso *a*) para examinar el segundo y tercer términos en esta expresión para $\mathbf{x}(t)$.

Índice analítico

Créditos de fotografía

Página 50 Torres de alta tensión eléctrica y cables: Haak78/Shutterstock; tren del lago Llanberis en la estación Llanberis de Snowdonia, Gwynedd, North Wales: DWImages Wales/Alamy; Máquina para pulir acero: Mircea Bezergheanu/Shutterstock.

Página 54 Bienes: Yuri Arcurs/Shutterstock; Servicios: Michael Jung/Shutterstock.

Página 84 Vista aérea del centro de Kuala Lumpur: SF Photo/Shutterstock; Vista aérea del suburbio residencial en las afueras de Phoenix, Arizona: Iofoto/Shutterstock.

Página 122 Sonda espacial Galileo: NASA.

Página 141 Modelado molecular en realidad virtual: Departamento de Ciencias de la Computación, University of North Carolina en Chapel Hill. Foto de Bo Strain.

Página 163 Richard Feynman: AP images.

Página 189 Transbordador espacial Columbia; Centro Espacial Kennedy/NASA.

Página 221 Computadora Laptop: Tatniz/Shutterstock; Smartphone: Kraska/Shutterstock.

Página 254 Chicago, vista frontal: Archana Bhartia/Shutterstock; Inmobilaria Noah Strycker/Shutterstock.

Página 265 Búho manchado del Pacífico del Norte: John y Karen Hollingsworth/US Fish and Wildlife Service.

Página 329 Datum de Norteamérica: Dmitry Kalinovsky/Shutterstock.

Página 374 Cometa Halley: Art Directors & TRIP/Alamy.

Página 393 Satélite Landsat: NASA.

Página 394 Bandas espectrales y bandas principales: MDA Information Systems, Inc. of Gaithersburg, Maryland.

Página 412 Puente: Maslov Dmitry/Shutterstock; Construcción: Dmitry Kalinovsky/Shutterstock; Familia: Dean Mitchell/Shutterstock.

Referencias a las aplicaciones

WEB indica material en el sitio Web.

Álgebra lineal numérica

Biología y ecología

Ciencias físicas

Computadoras y ciencia de la computación

Estadística

Álgebra lineal y sus aplicaciones ofrece una introducción elemental moderna al álgebra lineal y una amplia selección de aplicaciones a la ingeniería, ciencias de la computación, matemáticas, física, biología, economía y estadística.

El texto proporciona un importante apoyo tanto para la enseñanza como para el uso de la tecnología en el curso mediante ejemplos, ejercicios y proyectos que proponen el uso de programas como MATLAB® y Maple™, así como de calculadoras programables con capacidad matricial. La obra ofrece nuevos contenidos tanto en el libro como en línea:

- Más del 25 por ciento de los ejercicios son nuevos o se han actualizado, en especial los ejercicios computacionales.
- El 25 por ciento de los ejemplos introductorios de capítulo son nuevos.
- El nuevo capítulo *La geometría de los espacios vectoriales* ofrece un tema atractivo que los alumnos disfrutan.

Algunas de las características distintivas de esta edición son la introducción temprana a los conceptos clave; una visión moderna de la multiplicación de matrices; el fuerte énfasis geométrico; el tratamiento más completo de la ortogonalidad y los mínimos cuadrados; un gran número de ejercicios que se centran en el razonamiento en lugar de los cálculos mecánicos, y 300 preguntas de verdadero/falso.

Este libro incluye un código de acceso a MathXL®. Para que los estudiantes puedan utilizar el código se requiere que el profesor abra un curso.

Los capítulos en línea y el material complementario están disponibles en el sitio Web: www.pearsonenespañol.com/lay

ISBN 978-607-32-1398-1

9 786073 213981 90000