Linear Algebra

and its applications

FOURTH EDITION

David C. Lay

FOURTH EDITION

Linear Algebra
and Its Applications

David C. Lay
University of Maryland—College Park

Addison-Wesley
Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montréal Toronto
Delhi Mexico City São Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor-in-Chief: *Deirdre Lynch*
Senior Acquisitions Editor: *William Hoffmann*
Sponsoring Editor: *Caroline Celano*
Senior Content Editor: *Chere Bemelmans*
Editorial Assistant: *Brandon Rawnsley*
Senior Managing Editor: *Karen Wernholm*
Associate Managing Editor: *Tamela Ambush*
Digital Assets Manager: *Marianne Groth*
Supplements Production Coordinator: *Kerri McQueen*
Senior Media Producer: *Carl Cottrell*
QA Manager, Assessment Content: *Marty Wright*
Executive Marketing Manager: *Jeff Weidenaar*
Marketing Assistant: *Kendra Bassi*
Senior Author Support/Technology Specialist: *Joe Vetere*
Rights and Permissions Advisor: *Michael Joyce*
Image Manager: *Rachel Youdelman*
Senior Manufacturing Buyer: *Carol Melville*
Senior Media Buyer: *Ginny Michaud*
Design Manager: *Andrea Nix*
Senior Designer: *Beth Paquin*
Text Design: *Andrea Nix*
Production Coordination: *Tamela Ambush*
Composition: *Dennis Kletzing*
Illustrations: *Scientific Illustrators*

Cover Design: *Nancy Goulet, Studiowink*
Cover Image: *Shoula/Stone/Getty Images*

For permission to use copyrighted material, grateful acknowledgment is made to the copyright holders on page P1, which is hereby made part of this copyright page.

Library of Congress Cataloging-in-Publication Data

Lay, David C.
Linear algebra and its applications / David C. Lay. – 4th ed. update.
p. cm.
Includes index.
ISBN-13: 978-0-321-38517-8
ISBN-10: 0-321-38517-9
1. Algebras, Linear–Textbooks. I. Title.
QA184.2.L39 2012
512′.5–dc22
2010048460

1 2 3 4 5 6 7 8 9 10—DOW—14 13 12 11 10

Addison-Wesley
is an imprint of

PEARSON

www.pearsonhighered.com

ISBN 13: 978-0-321-38517-8
ISBN 10: 0-321-38517-9

To my wife, Lillian, and our children, Christina, Deborah, and Melissa, whose support, encouragement, and faithful prayers made this book possible.

About the Author

David C. Lay holds a B.A. from Aurora University (Illinois), and an M.A. and Ph.D. from the University of California at Los Angeles. Lay has been an educator and research mathematician since 1966, mostly at the University of Maryland, College Park. He has also served as a visiting professor at the University of Amsterdam, the Free University in Amsterdam, and the University of Kaiserslautern, Germany. He has published more than 30 research articles on functional analysis and linear algebra.

As a founding member of the NSF-sponsored Linear Algebra Curriculum Study Group, Lay has been a leader in the current movement to modernize the linear algebra curriculum. Lay is also a co-author of several mathematics texts, including *Introduction to Functional Analysis* with Angus E. Taylor, *Calculus and Its Applications*, with L. J. Goldstein and D. I. Schneider, and *Linear Algebra Gems—Assets for Undergraduate Mathematics,* with D. Carlson, C. R. Johnson, and A. D. Porter.

Professor Lay has received four university awards for teaching excellence, including, in 1996, the title of Distinguished Scholar–Teacher of the University of Maryland. In 1994, he was given one of the Mathematical Association of America's Awards for Distinguished College or University Teaching of Mathematics. He has been elected by the university students to membership in Alpha Lambda Delta National Scholastic Honor Society and Golden Key National Honor Society. In 1989, Aurora University conferred on him the Outstanding Alumnus award. Lay is a member of the American Mathematical Society, the Canadian Mathematical Society, the International Linear Algebra Society, the Mathematical Association of America, Sigma Xi, and the Society for Industrial and Applied Mathematics. Since 1992, he has served several terms on the national board of the Association of Christians in the Mathematical Sciences.

Contents

Chapter 1 Linear Equations in Linear Algebra 1

Chapter 2 Matrix Algebra 91

Chapter 3 Determinants 163

Chapter 4 Vector Spaces 189

Chapter 5 Eigenvalues and Eigenvectors 265

Chapter 6 Orthogonality and Least Squares 329

Chapter 7 Symmetric Matrices and Quadratic Forms 393

Chapter 8 The Geometry of Vector Spaces 435

Chapter 9 Optimization (Online)

Chapter 10 Finite-State Markov Chains (Online)

Appendixes

Preface

The response of students and teachers to the first three editions of *Linear Algebra and Its Applications* has been most gratifying. This *Fourth Edition* provides substantial support both for teaching and for using technology in the course. As before, the text provides a modern elementary introduction to linear algebra and a broad selection of interesting applications. The material is accessible to students with the maturity that should come from successful completion of two semesters of college-level mathematics, usually calculus.

The main goal of the text is to help students master the basic concepts and skills they will use later in their careers. The topics here follow the recommendations of the Linear Algebra Curriculum Study Group, which were based on a careful investigation of the real needs of the students and a consensus among professionals in many disciplines that use linear algebra. Hopefully, this course will be one of the most useful and interesting mathematics classes taken by undergraduates.

WHAT'S NEW IN THIS EDITION

The main goal of this revision was to update the exercises and provide additional content, both in the book and online.

1. More than 25 percent of the exercises are new or updated, especially the computational exercises. The exercise sets remain one of the most important features of this book, and these new exercises follow the same high standard of the exercise sets of the past three editions. They are crafted in a way that retells the substance of each of the sections they follow, developing the students' confidence while challenging them to practice and generalize the new ideas they have just encountered.
2. Twenty-five percent of chapter openers are new. These introductory vignettes provide applications of linear algebra and the motivation for developing the mathematics that follows. The text returns to that application in a section toward the end of the chapter.
3. A New Chapter: Chapter 8, The Geometry of Vector Spaces, provides a fresh topic that my students have really enjoyed studying. Sections 1, 2, and 3 provide the basic geometric tools. Then Section 6 uses these ideas to study Bezier curves and surfaces, which are used in engineering and online computer graphics (in Adobe® Illustrator® and Macromedia® FreeHand®). These four sections can be covered in four or five 50-minute class periods.

   A second course in linear algebra applications typically begins with a substantial review of key ideas from the first course. If part of Chapter 8 is in the first course, the second course could include a brief review of sections 1 to 3 and then a focus on the geometry in sections 4 and 5. That would lead naturally into the online chapters 9 and 10, which have been used with Chapter 8 at a number of schools for the past five years.
4. The *Study Guide*, which has always been an integral part of the book, has been updated to cover the new Chapter 8. As with past editions, the *Study Guide* incorporates

detailed solutions to every third odd-numbered exercise as well as solutions to every odd-numbered writing exercise for which the text only provides a hint.

5. Two new chapters are now available online, and can be used in a second course:

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

   An access code is required and is available to qualified adopters. For more information, visit *www.pearsonhighered.com/irc* or contact your Pearson representative.
6. PowerPoint® slides are now available for the 25 core sections of the text; also included are 75 figures from the text.

DISTINCTIVE FEATURES

Early Introduction of Key Concepts

Many fundamental ideas of linear algebra are introduced within the first seven lectures, in the concrete setting of $\mathbb{R}^n$, and then gradually examined from different points of view. Later generalizations of these concepts appear as natural extensions of familiar ideas, visualized through the geometric intuition developed in Chapter 1. A major achievement of this text is that the level of difficulty is fairly even throughout the course.

A Modern View of Matrix Multiplication

Good notation is crucial, and the text reflects the way scientists and engineers actually use linear algebra in practice. The definitions and proofs focus on the columns of a matrix rather than on the matrix entries. A central theme is to view a matrix–vector product $A\mathbf{x}$ as a linear combination of the columns of A. This modern approach simplifies many arguments, and it ties vector space ideas into the study of linear systems.

Linear Transformations

Linear transformations form a “thread” that is woven into the fabric of the text. Their use enhances the geometric flavor of the text. In Chapter 1, for instance, linear transformations provide a dynamic and graphical view of matrix–vector multiplication.

Eigenvalues and Dynamical Systems

Eigenvalues appear fairly early in the text, in Chapters 5 and 7. Because this material is spread over several weeks, students have more time than usual to absorb and review these critical concepts. Eigenvalues are motivated by and applied to discrete and continuous dynamical systems, which appear in Sections 1.10, 4.8, and 4.9, and in five sections of Chapter 5. Some courses reach Chapter 5 after about five weeks by covering Sections 2.8 and 2.9 instead of Chapter 4. These two optional sections present all the vector space concepts from Chapter 4 needed for Chapter 5.

Orthogonality and Least-Squares Problems

These topics receive a more comprehensive treatment than is commonly found in beginning texts. The Linear Algebra Curriculum Study Group has emphasized the need for a substantial unit on orthogonality and least-squares problems, because orthogonality plays such an important role in computer calculations and numerical linear algebra and because inconsistent linear systems arise so often in practical work.

PEDAGOGICAL FEATURES

Applications

A broad selection of applications illustrates the power of linear algebra to explain fundamental principles and simplify calculations in engineering, computer science, mathematics, physics, biology, economics, and statistics. Some applications appear in separate sections; others are treated in examples and exercises. In addition, each chapter opens with an introductory vignette that sets the stage for some application of linear algebra and provides a motivation for developing the mathematics that follows. Later, the text returns to that application in a section near the end of the chapter.

A Strong Geometric Emphasis

Every major concept in the course is given a geometric interpretation, because many students learn better when they can visualize an idea. There are substantially more drawings here than usual, and some of the figures have never before appeared in a linear algebra text.

Examples

This text devotes a larger proportion of its expository material to examples than do most linear algebra texts. There are more examples than an instructor would ordinarily present in class. But because the examples are written carefully, with lots of detail, students can read them on their own.

Theorems and Proofs

Important results are stated as theorems. Other useful facts are displayed in tinted boxes, for easy reference. Most of the theorems have formal proofs, written with the beginning student in mind. In a few cases, the essential calculations of a proof are exhibited in a carefully chosen example. Some routine verifications are saved for exercises, when they will benefit students.

Practice Problems

A few carefully selected Practice Problems appear just before each exercise set. Complete solutions follow the exercise set. These problems either focus on potential trouble spots in the exercise set or provide a "warm-up" for the exercises, and the solutions often contain helpful hints or warnings about the homework.

Exercises

The abundant supply of exercises ranges from routine computations to conceptual questions that require more thought. A good number of innovative questions pinpoint conceptual difficulties that I have found on student papers over the years. Each exercise set is carefully arranged in the same general order as the text; homework assignments are readily available when only part of a section is discussed. A notable feature of the exercises is their numerical simplicity. Problems "unfold" quickly, so students spend little time on numerical calculations. The exercises concentrate on teaching understanding rather than mechanical calculations. The exercises in the *Fourth Edition* maintain the integrity of the exercises from the third edition, while providing fresh problems for students and instructors.

Exercises marked with the symbol [**M**] are designed to be worked with the aid of a "**M**atrix program" (a computer program, such as **M**ATLAB®, **M**aple™, **M**athematica®, **M**athCad®, or Derive™, or a programmable calculator with matrix capabilities, such as those manufactured by Texas Instruments).

True/False Questions

To encourage students to read all of the text and to think critically, I have developed 300 simple true/false questions that appear in 33 sections of the text, just after the computational problems. They can be answered directly from the text, and they prepare students for the conceptual problems that follow. Students appreciate these questions—after they get used to the importance of reading the text carefully. Based on class testing and discussions with students, I decided not to put the answers in the text. (The *Study Guide* tells the students where to find the answers to the odd-numbered questions.) An additional 150 true/false questions (mostly at the ends of chapters) test understanding of the material. The text does provide simple T/F answers to most of these questions, but it omits the justifications for the answers (which usually require some thought).

Writing Exercises

An ability to write coherent mathematical statements in English is essential for all students of linear algebra, not just those who may go to graduate school in mathematics. The text includes many exercises for which a written justification is part of the answer. Conceptual exercises that require a short proof usually contain hints that help a student get started. For all odd-numbered writing exercises, either a solution is included at the back of the text or a hint is provided and the solution is given in the *Study Guide*, described below.

Computational Topics

The text stresses the impact of the computer on both the development and practice of linear algebra in science and engineering. Frequent Numerical Notes draw attention to issues in computing and distinguish between theoretical concepts, such as matrix inversion, and computer implementations, such as LU factorizations.

WEB SUPPORT

This Web site at *www.pearsonhighered.com/lay* contains support material for the textbook. For students, the Web site contains **review sheets** and **practice exams** (with solutions) that cover the main topics in the text. They come directly from courses I have taught in past years. Each review sheet identifies key definitions, theorems, and skills from a specified portion of the text.

Applications by Chapters

The Web site also contains seven Case Studies, which expand topics introduced at the beginning of each chapter, adding real-world data and opportunities for further exploration. In addition, more than 20 Application Projects either extend topics in the text or introduce new applications, such as cubic splines, airline flight routes, dominance matrices in sports competition, and error-correcting codes. Some mathematical applications are integration techniques, polynomial root location, conic sections, quadric surfaces, and extrema for functions of two variables. Numerical linear algebra topics, such as condition numbers, matrix factorizations, and the QR method for finding eigenvalues, are also included. Woven into each discussion are exercises that may involve large data sets (and thus require technology for their solution).

Getting Started with Technology

If your course includes some work with MATLAB, Maple, Mathematica, or TI calculators, you can read one of the projects on the Web site for an introduction to the technology. In addition, the *Study Guide* provides introductory material for first-time users.

Data Files

Hundreds of files contain data for about 900 numerical exercises in the text, Case Studies, and Application Projects. The data are available at *www.pearsonhighered.com/lay* in a variety of formats—for MATLAB, Maple, Mathematica, and the TI-83+/86/89 graphic calculators. By allowing students to access matrices and vectors for a particular problem with only a few keystrokes, the data files eliminate data entry errors and save time on homework.

MATLAB Projects

These exploratory projects invite students to discover basic mathematical and numerical issues in linear algebra. Written by Rick Smith, they were developed to accompany a computational linear algebra course at the University of Florida, which has used *Linear Algebra and Its Applications* for many years. The projects are referenced by an icon **WEB** at appropriate points in the text. About half of the projects explore fundamental concepts such as the column space, diagonalization, and orthogonal projections; several projects focus on numerical issues such as flops, iterative methods, and the SVD; and a few projects explore applications such as Lagrange interpolation and Markov chains.

SUPPLEMENTS

Study Guide

A printed version of the *Study Guide* is available at low cost. I wrote this *Guide* to be an integral part of the course. An icon **SG** in the text directs students to special subsections of the *Guide* that suggest how to master key concepts of the course. The *Guide* supplies a detailed solution to every third odd-numbered exercise, which allows students to check their work. A complete explanation is provided whenever an odd-numbered writing exercise has only a "Hint" in the answers. Frequent "Warnings" identify common errors and show how to prevent them. MATLAB boxes introduce commands as they are needed. Appendixes in the *Study Guide* provide comparable information about Maple, Mathematica, and TI graphing calculators (ISBN: 0-321-38883-6).

Instructor's Edition

For the convenience of instructors, this special edition includes brief answers to all exercises. A *Note to the Instructor* at the beginning of the text provides a commentary on the design and organization of the text, to help instructors plan their courses. It also describes other support available for instructors. (ISBN: 0-321-38518-7)

Instructor's Technology Manuals

Each manual provides detailed guidance for integrating a specific software package or graphic calculator throughout the course, written by faculty who have already used the technology with this text. The following manuals are available to qualified instructors through the 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), and the TI-83+/86/89 (ISBN: 0-321-38887-9).

ACKNOWLEDGMENTS

I am indeed grateful to many groups of people who have helped me over the years with various aspects of this book.

I want to thank Israel Gohberg and Robert Ellis for more than fifteen years of research collaboration, which greatly shaped my view of linear algebra. And, it has been a privilege to be a member of the Linear Algebra Curriculum Study Group along with David Carlson, Charles Johnson, and Duane Porter. Their creative ideas about teaching linear algebra have influenced this text in significant ways.

I sincerely thank the following reviewers for their careful analyses and constructive suggestions:

Rafal Ablamowicz, *Tennessee Technological University*
Brian E. Blank, *Washington University in St. 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 at Austin*
Douglas B. Meade, *University of South Carolina*
Tim Olson, *University of Florida*
Albert L. Vitter III, *Tulane University*

For this Fourth Edition, I thank my brother, Steven Lay, at Lee University, for his generous help and encouragement, and for his newly revised Chapters 8 and 9. I also thank Thomas Polaski, of Winthrop University, for his newly revised Chapter 10. For good advice and help with chapter introductory examples, I thank Raymond Rosentrater, of Westmont College. Another gifted professor, Judith McDonald, of Washington State University, developed many new exercises for the text. Her help and enthusiasm for the book was refreshing and inspiring.

I thank the technology experts who labored on the various supplements for the Fourth Edition, preparing the data, writing notes for the instructors, writing technology notes for the students in the *Study Guide*, and sharing their projects with us: Jeremy Case (MATLAB), Taylor University; Douglas Meade (Maple), University of South Carolina; Michael Miller (TI Calculator), Western Baptist College; and Marie Vanisko (Mathematica), Carroll College.

I thank Professor John Risley and graduate students David Aulicino, Sean Burke, and Hersh Goldberg for their technical expertise in helping develop online homework support for the text. I am grateful for the class testing of this online homework support by the following: Agnes Boskovitz, Malcolm Brooks, Elizabeth Ormerod, Alexander Isaev, and John Urbas at the Australian National University; John Scott and Leben Wee at Montgomery College, Maryland; and Xingru Zhang at SUNY University of Buffalo.

I appreciate the mathematical assistance provided by Blaise DeSesa, Jean Horn, Roger Lipsett, Paul Lorczak, Thomas Polaski, Sarah Streett, and Marie Vanisko, who checked the accuracy of calculations in the text.

Finally, I sincerely thank the staff at Addison-Wesley for all their help with the development and production of the Fourth Edition: Caroline Celano, sponsoring editor, Chere Bemelmans, senior content editor; Tamela Ambush, associate managing editor; Carl Cottrell, senior media producer; Jeff Weidenaar, executive marketing manager; Kendra Bassi, marketing assistant; and Andrea Nix, text design. Saved for last are the three good friends who have guided the development of the book nearly from the beginning—giving wise counsel and encouragement—Greg Tobin, publisher, Laurie Rosatone, former editor, and William Hoffman, current editor. Thank you all so much.

David C. Lay

A Note to Students

This course is potentially the most interesting and worthwhile undergraduate mathematics course you will complete. In fact, some students have written or spoken to me after graduation and said that they still use this text occasionally as a reference in their careers at major corporations and engineering graduate schools. The following remarks offer some practical advice and information to help you master the material and enjoy the course.

In linear algebra, the *concepts* are as important as the *computations*. The simple numerical exercises that begin each exercise set only help you check your understanding of basic procedures. Later in your career, computers will do the calculations, but you will have to choose the calculations, know how to interpret the results, and then explain the results to other people. For this reason, many exercises in the text ask you to explain or justify your calculations. A written explanation is often required as part of the answer. For odd-numbered exercises, you will find either the desired explanation or at least a good hint. You must avoid the temptation to look at such answers before you have tried to write out the solution yourself. Otherwise, you are likely to think you understand something when in fact you do not.

To master the concepts of linear algebra, you will have to read and reread the text carefully. New terms are in boldface type, sometimes enclosed in a definition box. A glossary of terms is included at the end of the text. Important facts are stated as theorems or are enclosed in tinted boxes, for easy reference. I encourage you to read the first five pages of the Preface to learn more about the structure of this text. This will give you a framework for understanding how the course may proceed.

In a practical sense, linear algebra is a language. You must learn this language the same way you would a foreign language—with daily work. Material presented in one section is not easily understood unless you have thoroughly studied the text and worked the exercises for the preceding sections. Keeping up with the course will save you lots of time and distress!

Numerical Notes

I hope you read the Numerical Notes in the text, even if you are not using a computer or graphic calculator with the text. In real life, most applications of linear algebra involve numerical computations that are subject to some numerical error, even though that error may be extremely small. The Numerical Notes will warn you of potential difficulties in using linear algebra later in your career, and if you study the notes now, you are more likely to remember them later.

If you enjoy reading the Numerical Notes, you may want to take a course later in numerical linear algebra. Because of the high demand for increased computing power, computer scientists and mathematicians work in numerical linear algebra to develop faster and more reliable algorithms for computations, and electrical engineers design faster and smaller computers to run the algorithms. This is an exciting field, and your first course in linear algebra will help you prepare for it.

Study Guide

To help you succeed in this course, I suggest that you purchase the *Study Guide* (*www.mypearsonstore.com*; 0-321-38883-6). Not only will it help you learn linear algebra, it also will show you how to study mathematics. At strategic points in your textbook, an icon SG will direct you to special subsections in the *Study Guide* entitled "Mastering Linear Algebra Concepts." There you will find suggestions for constructing effective review sheets of key concepts. The act of preparing the sheets is one of the secrets to success in the course, because you will construct *links between ideas*. These links are the "glue" that enables you to build a solid foundation for learning and *remembering* the main concepts in the course.

The *Study Guide* contains a detailed solution to every third odd-numbered exercise, plus solutions to all odd-numbered writing exercises for which only a hint is given in the Answers section of this book. The *Guide* is separate from the text because you must learn to write solutions by yourself, without much help. (I know from years of experience that easy access to solutions in the back of the text slows the mathematical development of most students.) The *Guide* also provides warnings of common errors and helpful hints that call attention to key exercises and potential exam questions.

If you have access to technology—MATLAB, Maple, Mathematica, or a TI graphing calculator—you can save many hours of homework time. The *Study Guide* is your "lab manual" that explains how to use each of these matrix utilities. It introduces new commands when they are needed. You can download from the website *www.pearsonhighered.com/lay* the data for more than 850 exercises in the text. (With a few keystrokes, you can display any numerical homework problem on your screen.) Special matrix commands will perform the computations for you!

What you do in your first few weeks of studying this course will set your pattern for the term and determine how well you finish the course. Please read "How to Study Linear Algebra" in the *Study Guide* as soon as possible. My students have found the strategies there very helpful, and I hope you will, too.

1

Linear Equations in Linear Algebra

INTRODUCTORY EXAMPLE

Linear Models in Economics and Engineering

It was late summer in 1949. Harvard Professor Wassily Leontief was carefully feeding the last of his punched cards into the university's Mark II computer. The cards contained economic information about the U.S. economy and represented a summary of more than 250,000 pieces of information produced by the U.S. Bureau of Labor Statistics after two years of intensive work. Leontief had divided the U.S. economy into 500 "sectors," such as the coal industry, the automotive industry, communications, and so on. For each sector, he had written a linear equation that described how the sector distributed its output to the other sectors of the economy. Because the Mark II, one of the largest computers of its day, could not handle the resulting system of 500 equations in 500 unknowns, Leontief had distilled the problem into a system of 42 equations in 42 unknowns.

Programming the Mark II computer for Leontief's 42 equations had required several months of effort, and he was anxious to see how long the computer would take to solve the problem. The Mark II hummed and blinked for 56 hours before finally producing a solution. We will discuss the nature of this solution in Sections 1.6 and 2.6.

Leontief, who was awarded the 1973 Nobel Prize in Economic Science, opened the door to a new era in mathematical modeling in economics. His efforts at Harvard in 1949 marked one of the first significant uses of computers to analyze what was then a large-scale mathematical model. Since that time, researchers in many other fields have employed computers to analyze mathematical models. Because of the massive amounts of data involved, the models are usually *linear*; that is, they are described by *systems of linear equations*.

The importance of linear algebra for applications has risen in direct proportion to the increase in computing power, with each new generation of hardware and software triggering a demand for even greater capabilities. Computer science is thus intricately linked with linear algebra through the explosive growth of parallel processing and large-scale computations.

Scientists and engineers now work on problems far more complex than even dreamed possible a few decades ago. Today, linear algebra has more potential value for students in many scientific and business fields than any other undergraduate mathematics subject! The material in this text provides the foundation for further work in many interesting areas. Here are a few possibilities; others will be described later.

- *Oil exploration.* When a ship searches for offshore oil deposits, its computers solve thousands of separate systems of linear equations *every day*. The seismic data for the equations are obtained from underwater shock waves created by explosions from air guns. The waves bounce off subsurface rocks and are measured by geophones attached to mile-long cables behind the ship.

- *Linear programming*. Many important management decisions today are made on the basis of linear programming models that utilize hundreds of variables. The airline industry, for instance, employs linear programs that schedule flight crews, monitor the locations of aircraft, or plan the varied schedules of support services such as maintenance and terminal operations.
- *Electrical networks*. Engineers use simulation software to design electrical circuits and microchips involving millions of transistors. Such software relies on linear algebra techniques and systems of linear equations.

WEB

Systems of linear equations lie at the heart of linear algebra, and this chapter uses them to introduce some of the central concepts of linear algebra in a simple and concrete setting. Sections 1.1 and 1.2 present a systematic method for solving systems of linear equations. This algorithm will be used for computations throughout the text. Sections 1.3 and 1.4 show how a system of linear equations is equivalent to a *vector equation* and to a *matrix equation*. This equivalence will reduce problems involving linear combinations of vectors to questions about systems of linear equations. The fundamental concepts of spanning, linear independence, and linear transformations, studied in the second half of the chapter, will play an essential role throughout the text as we explore the beauty and power of linear algebra.

1.1 SYSTEMS OF LINEAR EQUATIONS

A **linear equation** in the variables $x_1, \dots, x_n$ is an equation that can be written in the form

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

where b and the **coefficients** $a_1, \dots, a_n$ are real or complex numbers, usually known in advance. The subscript n may be any positive integer. In textbook examples and exercises, n is normally between 2 and 5. In real-life problems, n might be 50 or 5000, or even larger.

The equations

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

are both linear because they can be rearranged algebraically as in equation (1):

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

The equations

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

are not linear because of the presence of x_1x_2 in the first equation and $\sqrt{x_1}$ in the second.

A **system of linear equations** (or a **linear system**) is a collection of one or more linear equations involving the same variables—say, $x_1, \dots, x_n$. An example is

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

A **solution** of the system is a list $(s_1, s_2, \dots, s_n)$ of numbers that makes each equation a true statement when the values $s_1, \dots, s_n$ are substituted for $x_1, \dots, x_n$, respectively. For instance, $(5, 6.5, 3)$ is a solution of system (2) because, when these values are substituted in (2) for x_1, x_2, x_3, respectively, the equations simplify to $8 = 8$ and $-7 = -7$.

The set of all possible solutions is called the **solution set** of the linear system. Two linear systems are called **equivalent** if they have the same solution set. That is, each solution of the first system is a solution of the second system, and each solution of the second system is a solution of the first.

Finding the solution set of a system of two linear equations in two variables is easy because it amounts to finding the intersection of two lines. A typical problem is

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

The graphs of these equations are lines, which we denote by ℓ_1 and ℓ_2. A pair of numbers (x_1, x_2) satisfies *both* equations in the system if and only if the point (x_1, x_2) lies on both ℓ_1 and ℓ_2. In the system above, the solution is the single point $(3, 2)$, as you can easily verify. See Fig. 1.

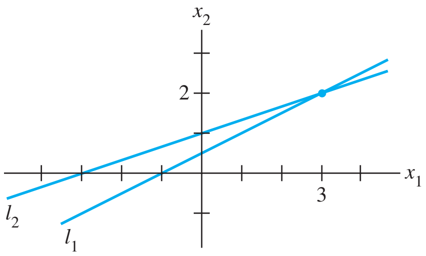
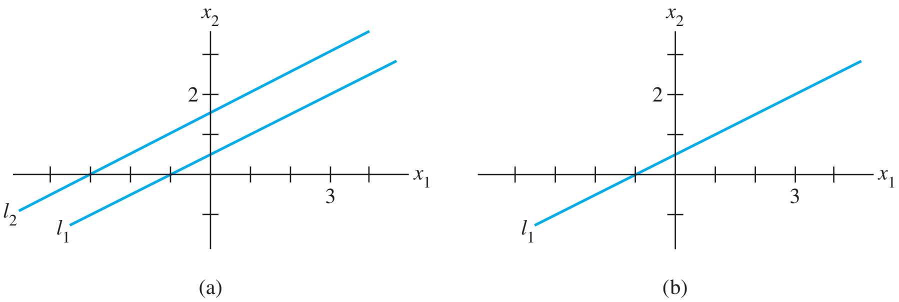


FIGURE 1 Exactly one solution.

Of course, two lines need not intersect in a single point—they could be parallel, or they could coincide and hence "intersect" at every point on the line. Figure 2 shows the graphs that correspond to the following systems:

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



(a) (b)

FIGURE 2 (a) No solution. (b) Infinitely many solutions.

Figures 1 and 2 illustrate the following general fact about linear systems, to be verified in Section 1.2.

A system of linear equations has

1. no solution, or
2. exactly one solution, or
3. infinitely many solutions.

A system of linear equations is said to be **consistent** if it has either one solution or infinitely many solutions; a system is **inconsistent** if it has no solution.

Matrix Notation

The essential information of a linear system can be recorded compactly in a rectangular array called a **matrix**. Given the system

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

with the coefficients of each variable aligned in columns, the matrix

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

is called the **coefficient matrix** (or **matrix of coefficients**) of the system (3), and

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

is called the **augmented matrix** of the system. (The second row here contains a zero because the second equation could be written as $0 \cdot x_1 + 2x_2 - 8x_3 = 8$.) An augmented matrix of a system consists of the coefficient matrix with an added column containing the constants from the right sides of the equations.

The **size** of a matrix tells how many rows and columns it has. The augmented matrix (4) above has 3 rows and 4 columns and is called a 3×4 (read "3 by 4") matrix. If m and n are positive integers, an $\boldsymbol{m \times n}$ **matrix** is a rectangular array of numbers with m rows and n columns. (The number of rows always comes first.) Matrix notation will simplify the calculations in the examples that follow.

Solving a Linear System

This section and the next describe an algorithm, or a systematic procedure, for solving linear systems. The basic strategy is *to replace one system with an equivalent system (i.e., one with the same solution set) that is easier to solve*.

Roughly speaking, use the x_1 term in the first equation of a system to eliminate the x_1 terms in the other equations. Then use the x_2 term in the second equation to eliminate the x_2 terms in the other equations, and so on, until you finally obtain a very simple equivalent system of equations.

Three basic operations are used to simplify a linear system: Replace one equation by the sum of itself and a multiple of another equation, interchange two equations, and multiply all the terms in an equation by a nonzero constant. After the first example, you will see why these three operations do not change the solution set of the system.

EXAMPLE 1 Solve system (3).

SOLUTION The elimination procedure is shown here with and without matrix notation, and the results are placed side by side for comparison:

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

Keep x_1 in the first equation and eliminate it from the other equations. To do so, add 4 times equation 1 to equation 3. After some practice, this type of calculation is usually performed mentally:

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

The result of this calculation is written in place of the original third equation:

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

Now, multiply equation 2 by $1/2$ in order to obtain 1 as the coefficient for x_2. (This calculation will simplify the arithmetic in the next step.)

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

Use the x_2 in equation 2 to eliminate the $-3x_2$ in equation 3. The "mental" computation is

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

The new system has a *triangular* form:[1]

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

Eventually, you want to eliminate the $-2x_2$ term from equation 1, but it is more efficient to use the x_3 in equation 3 first, to eliminate the $-4x_3$ and $+x_3$ terms in equations 2 and 1. The two "mental" calculations are

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

---

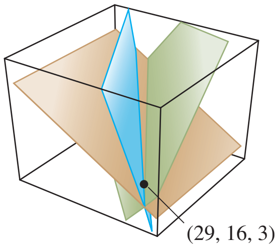
[1] The intuitive term *triangular* will be replaced by a precise term in the next section.

It is convenient to combine the results of these two operations:

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

Now, having cleaned out the column above the x_3 in equation 3, move back to the x_2 in equation 2 and use it to eliminate the $-2x_2$ above it. Because of the previous work with x_3, there is now no arithmetic involving x_3 terms. Add 2 times equation 2 to equation 1 and obtain the system:

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



Each of the original equations determines a plane in three-dimensional space. The point (29, 16, 3) lies in all three planes.

The work is essentially done. It shows that the only solution of the original system is (29, 16, 3). However, since there are so many calculations involved, it is a good practice to check the work. To verify that (29, 16, 3) *is* a solution, substitute these values into the left side of the original system, and compute:

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

The results agree with the right side of the original system, so (29, 16, 3) is a solution of the system. ■

Example 1 illustrates how operations on equations in a linear system correspond to operations on the appropriate rows of the augmented matrix. The three basic operations listed earlier correspond to the following operations on the augmented matrix.

ELEMENTARY ROW OPERATIONS

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.[2]
2. (Interchange) Interchange two rows.
3. (Scaling) Multiply all entries in a row by a nonzero constant.

Row operations can be applied to any matrix, not merely to one that arises as the augmented matrix of a linear system. Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other.

It is important to note that row operations are *reversible*. If two rows are interchanged, they can be returned to their original positions by another interchange. If a row is scaled by a nonzero constant c, then multiplying the new row by $1/c$ produces the original row. Finally, consider a replacement operation involving two rows—say, rows 1 and 2—and suppose that c times row 1 is added to row 2 to produce a new row 2. To "reverse" this operation, add $-c$ times row 1 to (new) row 2 and obtain the original row 2. See Exercises 29–32 at the end of this section.

---

[2] A common paraphrase of row replacement is "Add to one row a multiple of another row."

At the moment, we are interested in row operations on the augmented matrix of a system of linear equations. Suppose a system is changed to a new one via row operations. By considering each type of row operation, you can see that any solution of the original system remains a solution of the new system. Conversely, since the original system can be produced via row operations on the new system, each solution of the new system is also a solution of the original system. This discussion justifies the following statement.

> If the augmented matrices of two linear systems are row equivalent, then the two systems have the same solution set.

Though Example 1 is lengthy, you will find that after some practice, the calculations go quickly. Row operations in the text and exercises will usually be extremely easy to perform, allowing you to focus on the underlying concepts. Still, you must learn to perform row operations accurately because they will be used throughout the text.

The rest of this section shows how to use row operations to determine the size of a solution set, without completely solving the linear system.

Existence and Uniqueness Questions

Section 1.2 will show why a solution set for a linear system contains either no solutions, one solution, or infinitely many solutions. Answers to the following two questions will determine the nature of the solution set for a linear system.

To determine which possibility is true for a particular system, we ask two questions.

TWO FUNDAMENTAL QUESTIONS ABOUT A LINEAR SYSTEM

1. Is the system consistent; that is, does at least one solution *exist*?
2. If a solution exists, is it the *only* one; that is, is the solution *unique*?

These two questions will appear throughout the text, in many different guises. This section and the next will show how to answer these questions via row operations on the augmented matrix.

EXAMPLE 2 Determine if the following system is consistent:

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

SOLUTION This is the system from Example 1. Suppose that we have performed the row operations necessary to obtain the triangular form

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

At this point, we know x_3. Were we to substitute the value of x_3 into equation 2, we could compute x_2 and hence could determine x_1 from equation 1. So a solution exists; the system is consistent. (In fact, x_2 is uniquely determined by equation 2 since x_3 has only one possible value, and x_1 is therefore uniquely determined by equation 1. So the solution is unique.) ■

EXAMPLE 3 Determine if the following system is consistent:

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

SOLUTION The augmented matrix is

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

To obtain an x_1 in the first equation, interchange rows 1 and 2:

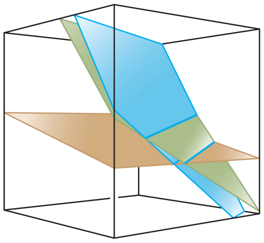
$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 5 & -8 & 7 & 1 \end{bmatrix}$$

To eliminate the $5x_1$ term in the third equation, add $-5/2$ times row 1 to row 3:

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

Next, use the x_2 term in the second equation to eliminate the $-(1/2)x_2$ term from the third equation. Add $1/2$ times row 2 to row 3:

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



This system is inconsistent because there is no point that lies in all three planes.

The augmented matrix is now in triangular form. To interpret it correctly, go back to equation notation:

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

The equation $0 = 5/2$ is a short form of $0x_1 + 0x_2 + 0x_3 = 5/2$. This system in triangular form obviously has a built-in contradiction. There are no values of x_1, x_2, x_3 that satisfy (8) because the equation $0 = 5/2$ is never true. Since (8) and (5) have the same solution set, the original system is inconsistent (i.e., has no solution). ■

Pay close attention to the augmented matrix in (7). Its last row is typical of an inconsistent system in triangular form.

NUMERICAL NOTE

In real-world problems, systems of linear equations are solved by a computer. For a square coefficient matrix, computer programs nearly always use the elimination algorithm given here and in Section 1.2, modified slightly for improved accuracy.

The vast majority of linear algebra problems in business and industry are solved with programs that use *floating point arithmetic*. Numbers are represented as decimals $\pm .d_1 \cdots d_p \times 10^r$, where r is an integer and the number p of digits to the right of the decimal point is usually between 8 and 16. Arithmetic with such numbers typically is inexact, because the result must be rounded (or truncated) to the number of digits stored. “Roundoff error” is also introduced when a number such as $1/3$ is entered into the computer, since its decimal representation must be approximated by a finite number of digits. Fortunately, inaccuracies in floating point arithmetic seldom cause problems. The numerical notes in this book will occasionally warn of issues that you may need to consider later in your career.

PRACTICE PROBLEMS

Throughout the text, practice problems should be attempted before working the exercises. Solutions appear after each exercise set.

1. State in words the next elementary row operation that should be performed on the system in order to solve it. [More than one answer is possible in (a).]

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

2. The augmented matrix of a linear system has been transformed by row operations into the form below. Determine if the system is consistent.

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

3. Is $(3, 4, -2)$ a solution of the following system?

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

4. For what values of h and k is the following system consistent?

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

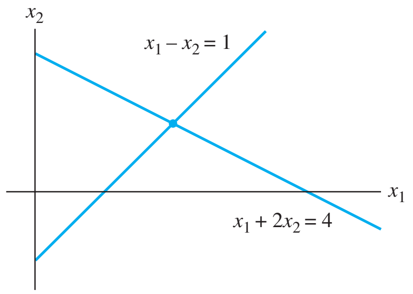
1.1 EXERCISES

Solve each system in Exercises 1–4 by using elementary row operations on the equations or on the augmented matrix. Follow the systematic elimination procedure described in this section.

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. Find the point (x_1, x_2) that lies on the line $x_1 + 2x_2 = 4$ and on the line $x_1 - x_2 = 1$. See the figure.



4. Find the point of intersection of the lines $x_1 + 2x_2 = -13$ and $3x_1 - 2x_2 = 1$

Consider each matrix in Exercises 5 and 6 as the augmented matrix of a linear system. State in words the next two elementary row operations that should be performed in the process of solving the system.

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

In Exercises 7–10, the augmented matrix of a linear system has been reduced by row operations to the form shown. In each case, continue the appropriate row operations and describe the solution set of the original system.

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

Solve the systems in Exercises 11–14.

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

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

Determine if the systems in Exercises 15 and 16 are consistent. Do not completely solve the systems.

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

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

17. Do the three lines $2x_1 + 3x_2 = -1$, $6x_1 + 5x_2 = 0$, and $2x_1 - 5x_2 = 7$ have a common point of intersection? Explain.

18. Do the three planes $2x_1 + 4x_2 + 4x_3 = 4$, $x_2 - 2x_3 = -2$, and $2x_1 + 3x_2 = 0$ have at least one common point of intersection? Explain.

In Exercises 19–22, determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system.

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

In Exercises 23 and 24, key statements from this section are either quoted directly, restated slightly (but still true), or altered in some way that makes them false in some cases. Mark each statement True or False, and *justify* your answer. (If true, give the approximate location where a similar statement appears, or refer to a definition or theorem. If false, give the location of a statement that has been quoted or used incorrectly, or cite an example that shows the statement is not true in all cases.) Similar true/false questions will appear in many sections of the text.

23. a. Every elementary row operation is reversible.

b. A 5 × 6 matrix has six rows.

c. The solution set of a linear system involving variables $x_1, \dots, x_n$ is a list of numbers $(s_1, \dots, s_n)$ that makes each equation in the system a true statement when the values $s_1, \dots, s_n$ are substituted for $x_1, \dots, x_n$, respectively.

d. Two fundamental questions about a linear system involve existence and uniqueness.

24. a. Two matrices are row equivalent if they have the same number of rows.

b. Elementary row operations on an augmented matrix never change the solution set of the associated linear system.

c. Two equivalent linear systems can have different solution sets.

d. A consistent system of linear equations has one or more solutions.

25. Find an equation involving g, h, and k that makes this augmented matrix correspond to a consistent system:

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

26. Suppose the system below is consistent for all possible values of f and g. What can you say about the coefficients c and d? Justify your answer.

$$\begin{aligned} 2x_1 + 4x_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

27. Suppose a, b, c, and d are constants such that a is not zero and the system below is consistent for all possible values of f and g. What can you say about the numbers a, b, c, and d? Justify your answer.

$$\begin{aligned} ax_1 + bx_2 &= f \\ cx_1 + dx_2 &= g \end{aligned}$$

28. Construct three different augmented matrices for linear systems whose solution set is $x_1 = 3$, $x_2 = -2$, $x_3 = -1$.

In Exercises 29–32, find the elementary row operation that transforms the first matrix into the second, and then find the reverse row operation that transforms the second matrix into the first.

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

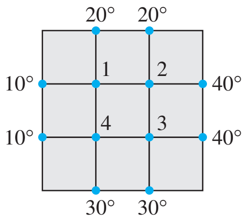
30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & 1 & -2 \end{bmatrix}$

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

An important concern in the study of heat transfer is to determine the steady-state temperature distribution of a thin plate when the temperature around the boundary is known. Assume the plate shown in the figure represents a cross section of a metal beam, with negligible heat flow in the direction perpendicular to the plate. Let $T_1, \dots, T_4$ denote the temperatures at the four interior nodes of the mesh in the figure. The temperature at a node is approximately equal to the average of the four nearest nodes—to the left, above, to the right, and below.[3] For instance,

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



33. Write a system of four equations whose solution gives estimates for the temperatures $T_1, \dots, T_4$.

34. Solve the system of equations from Exercise 33. [*Hint*: To speed up the calculations, interchange rows 1 and 4 before starting "replace" operations.]

---

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

SOLUTIONS TO PRACTICE PROBLEMS

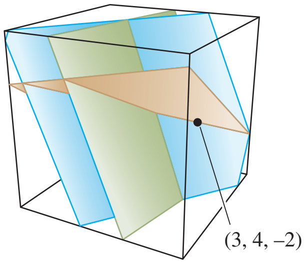
1. a. For "hand computation," the best choice is to interchange equations 3 and 4. Another possibility is to multiply equation 3 by 1/5. Or, replace equation 4 by its sum with −1/5 times row 3. (In any case, do not use the x_2 in equation 2 to eliminate the $4x_2$ in equation 1. Wait until a triangular form has been reached and the x_3 terms and x_4 terms have been eliminated from the first two equations.)

b. The system is in triangular form. Further simplification begins with the x_4 in the fourth equation. Use the x_4 to eliminate all x_4 terms above it. The appropriate step now is to add 2 times equation 4 to equation 1. (After that, move to equation 3, multiply it by $1/2$, and then use the equation to eliminate the x_3 terms above it.)

2. The system corresponding to the augmented matrix is

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

The third equation makes $x_3 = 0$, which is certainly an allowable value for x_3. After eliminating the x_3 terms in equations 1 and 2, you could go on to solve for unique values for x_2 and x_1. Hence a solution exists, and it is unique. Contrast this situation with that in Example 3.



Since $(3, 4, -2)$ satisfies the first two equations, it is on the line of the intersection of the first two planes. Since $(3, 4, -2)$ does not satisfy all three equations, it does not lie on all three planes.

3. It is easy to check if a specific list of numbers is a solution. Set $x_1 = 3$, $x_2 = 4$, and $x_3 = -2$, and find that

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

Although the first two equations are satisfied, the third is not, so $(3, 4, -2)$ is not a solution of the system. Notice the use of parentheses when making the substitutions. They are strongly recommended as a guard against arithmetic errors.

4. When the second equation is replaced by its sum with 3 times the first equation, the system becomes

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

If $k + 3h$ is nonzero, the system has no solution. The system is consistent for any values of h and k that make $k + 3h = 0$.

1.2 ROW REDUCTION AND ECHELON FORMS

This section refines the method of Section 1.1 into a row reduction algorithm that will enable us to analyze any system of linear equations.[1] By using only the first part of the algorithm, we will be able to answer the fundamental existence and uniqueness questions posed in Section 1.1.

The algorithm applies to any matrix, whether or not the matrix is viewed as an augmented matrix for a linear system. So the first part of this section concerns an arbitrary rectangular matrix and begins by introducing two important classes of matrices that include the "triangular" matrices of Section 1.1. In the definitions that follow, a *nonzero* row or column in a matrix means a row or column that contains at least one nonzero entry; a **leading entry** of a row refers to the leftmost nonzero entry (in a nonzero row).

---

[1] The algorithm here is a variant of what is commonly called *Gaussian elimination*. A similar elimination method for linear systems was used by Chinese mathematicians in about 250 B.C. The process was unknown in Western culture until the nineteenth century, when a famous German mathematician, Carl Friedrich Gauss, discovered it. A German engineer, Wilhelm Jordan, popularized the algorithm in an 1888 text on geodesy.

DEFINITION

A rectangular matrix is in **echelon form** (or **row echelon form**) if it has the following three properties:

1. All nonzero rows are above any rows of all zeros.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional conditions, then it is in **reduced echelon form** (or **reduced row echelon form**):

4. The leading entry in each nonzero row is 1.
5. Each leading 1 is the only nonzero entry in its column.

An **echelon matrix** (respectively, **reduced echelon matrix**) is one that is in echelon form (respectively, reduced echelon form). Property 2 says that the leading entries form an *echelon* ("steplike") pattern that moves down and to the right through the matrix. Property 3 is a simple consequence of property 2, but we include it for emphasis.

The "triangular" matrices of Section 1.1, such as

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

are in echelon form. In fact, the second matrix is in reduced echelon form. Here are additional examples.

EXAMPLE 1 The following matrices are in echelon form. The leading entries ($\blacksquare$) may have any nonzero value; the starred entries ($*$) may have any value (including zero).

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

The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0's below *and above* each leading 1.

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

■

Any nonzero matrix may be **row reduced** (that is, transformed by elementary row operations) into more than one matrix in echelon form, using different sequences of row operations. However, the reduced echelon form one obtains from a matrix is unique. The following theorem is proved in Appendix A at the end of the text.

THEOREM 1

Uniqueness of the Reduced Echelon Form

Each matrix is row equivalent to one and only one reduced echelon matrix.

If a matrix A is row equivalent to an echelon matrix U, we call U **an echelon form** (or row echelon form) **of** $\boldsymbol{A}$; if U is in reduced echelon form, we call U **the reduced echelon form of** $\boldsymbol{A}$. [Most matrix programs and calculators with matrix capabilities use the abbreviation RREF for reduced (row) echelon form. Some use REF for (row) echelon form.]

Pivot Positions

When row operations on a matrix produce an echelon form, further row operations to obtain the reduced echelon form do not change the positions of the leading entries. Since the reduced echelon form is unique, *the leading entries are always in the same positions in any echelon form obtained from a given matrix*. These leading entries correspond to leading 1's in the reduced echelon form.

DEFINITION

A **pivot position** in a matrix A is a location in A that corresponds to a leading 1 in the reduced echelon form of A. A **pivot column** is a column of A that contains a pivot position.

In Example 1, the squares (■) identify the pivot positions. Many fundamental concepts in the first four chapters will be connected in one way or another with pivot positions in a matrix.

EXAMPLE 2 Row reduce the matrix A below to echelon form, and locate the pivot columns of A.

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

SOLUTION Use the same basic strategy as in Section 1.1. The top of the leftmost nonzero column is the first pivot position. A nonzero entry, or *pivot*, must be placed in this position. A good choice is to interchange rows 1 and 4 (because the mental computations in the next step will not involve fractions).

$$\begin{bmatrix} \overset{\text{Pivot}}{1} & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ \underset{\text{Pivot column}}{0} & -3 & -6 & 4 & 9 \end{bmatrix}$$

Create zeros below the pivot, 1, by adding multiples of the first row to the rows below, and obtain matrix (1) below. The pivot position in the second row must be as far left as possible—namely, in the second column. Choose the 2 in this position as the next pivot.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & \overset{\text{Pivot}}{2} & 4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & \underset{\text{Next pivot column}}{-3} & -6 & 4 & 9 \end{bmatrix} \tag{1}$$

Add $-5/2$ times row 2 to row 3, and add $3/2$ times row 2 to row 4.

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

The matrix in (2) is different from any encountered in Section 1.1. There is no way to create a leading entry in column 3! (We can't use row 1 or 2 because doing so would destroy the echelon arrangement of the leading entries already produced.) However, if we interchange rows 3 and 4, we can produce a leading entry in column 4.

Pivot

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

Pivot columns

The matrix is in echelon form and thus reveals that columns 1, 2, and 4 of A are pivot columns.

Pivot positions

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

Pivot columns ■

A **pivot**, as illustrated in Example 2, is a nonzero number in a pivot position that is used as needed to create zeros via row operations. The pivots in Example 2 were 1, 2, and -5. Notice that these numbers are not the same as the actual elements of A in the highlighted pivot positions shown in (3).

With Example 2 as a guide, we are ready to describe an efficient procedure for transforming a matrix into an echelon or reduced echelon matrix. Careful study and mastery of this procedure now will pay rich dividends later in the course.

The Row Reduction Algorithm

The algorithm that follows consists of four steps, and it produces a matrix in echelon form. A fifth step produces a matrix in reduced echelon form. We illustrate the algorithm by an example.

EXAMPLE 3 Apply elementary row operations to transform the following matrix first into echelon form and then into reduced echelon form:

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

SOLUTION

STEP 1

Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.

$$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ \underset{\substack{\uparrow \\ \text{Pivot column}}}{3} & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

STEP 2

Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.

Interchange rows 1 and 3. (We could have interchanged rows 1 and 2 instead.)

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

STEP 3

Use row replacement operations to create zeros in all positions below the pivot.

As a preliminary step, we could divide the top row by the pivot, 3. But with two 3's in column 1, it is just as easy to add -1 times row 1 to row 2.

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

STEP 4

Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1–3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

With row 1 covered, step 1 shows that column 2 is the next pivot column; for step 2, select as a pivot the "top" entry in that column.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & \overset{\text{Pivot}}{2} & -4 & 4 & 2 & -6 \\ 0 & \underset{\substack{\uparrow \\ \text{New pivot column}}}{3} & -6 & 6 & 4 & -5 \end{bmatrix}$$

For step 3, we could insert an optional step of dividing the "top" row of the submatrix by the pivot, 2. Instead, we add $-3/2$ times the "top" row to the row below. This produces

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

When we cover the row containing the second pivot position for step 4, we are left with a new submatrix having only one row:

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

Steps 1–3 require no work for this submatrix, and we have reached an echelon form of the full matrix. If we want the reduced echelon form, we perform one more step.

STEP 5

Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

The rightmost pivot is in row 3. Create zeros above it, adding suitable multiples of row 3 to rows 2 and 1.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 2 & -4 & 4 & 0 & -14 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{Row } 1 + (-6) \cdot \text{row } 3 \\ \leftarrow \text{Row } 2 + (-2) \cdot \text{row } 3 \\ \end{matrix}$$

The next pivot is in row 2. Scale this row, dividing by the pivot.

$$\begin{bmatrix} 3 & -9 & 12 & -9 & 0 & -9 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \\ \leftarrow \text{Row scaled by } \frac{1}{2} \\ \end{matrix}$$

Create a zero in column 2 by adding 9 times row 2 to row 1.

$$\begin{bmatrix} 3 & 0 & -6 & 9 & 0 & -72 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{Row } 1 + (9) \cdot \text{row } 2 \\ \\ \end{matrix}$$

Finally, scale row 1, dividing by the pivot, 3.

$$\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \begin{matrix} \leftarrow \text{Row scaled by } \frac{1}{3} \\ \\ \end{matrix}$$

This is the reduced echelon form of the original matrix. ■

The combination of steps 1–4 is called the **forward phase** of the row reduction algorithm. Step 5, which produces the unique reduced echelon form, is called the **backward phase**.

NUMERICAL NOTE

In step 2 above, a computer program usually selects as a pivot the entry in a column having the largest absolute value. This strategy, called **partial pivoting**, is used because it reduces roundoff errors in the calculations.

Solutions of Linear Systems

The row reduction algorithm leads directly to an explicit description of the solution set of a linear system when the algorithm is applied to the augmented matrix of the system.

Suppose, for example, that the augmented matrix of a linear system has been changed into the equivalent *reduced* echelon form

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

There are three variables because the augmented matrix has four columns. The associated system of equations is

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

The variables x_1 and x_2 corresponding to pivot columns in the matrix are called **basic variables**.[2] The other variable, x_3, is called a **free variable**.

Whenever a system is consistent, as in (4), the solution set can be described explicitly by solving the *reduced* system of equations for the basic variables in terms of the free variables. This operation is possible because the reduced echelon form places each basic variable in one and only one equation. In (4), solve the first equation for x_1 and the second for x_2. (Ignore the third equation; it offers no restriction on the variables.)

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

The statement "x_3 is free" means that you are free to choose any value for x_3. Once that is done, the formulas in (5) determine the values for x_1 and x_2. For instance, when $x_3 = 0$, the solution is $(1, 4, 0)$; when $x_3 = 1$, the solution is $(6, 3, 1)$. *Each different choice of x_3 determines a (different) solution of the system, and every solution of the system is determined by a choice of x_3.*

EXAMPLE 4 Find the general solution of the linear system whose augmented matrix has been reduced to

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

SOLUTION The matrix is in echelon form, but we want the reduced echelon form before solving for the basic variables. The row reduction is completed next. The symbol $\sim$ before a matrix indicates that the matrix is row equivalent to the preceding matrix.

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

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

---

[2] Some texts use the term *leading variables* because they correspond to the columns containing leading entries.

There are five variables because the augmented matrix has six columns. The associated system now is

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

The pivot columns of the matrix are 1, 3, and 5, so the basic variables are x_1, x_3, and x_5. The remaining variables, x_2 and x_4, must be free. Solve for the basic variables to obtain the general solution:

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

Note that the value of x_5 is already fixed by the third equation in system (6). ■

Parametric Descriptions of Solution Sets

The descriptions in (5) and (7) are *parametric descriptions* of solution sets in which the free variables act as parameters. *Solving a system* amounts to finding a parametric description of the solution set or determining that the solution set is empty.

Whenever a system is consistent and has free variables, the solution set has many parametric descriptions. For instance, in system (4), we may add 5 times equation 2 to equation 1 and obtain the equivalent system

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

We could treat x_2 as a parameter and solve for x_1 and x_3 in terms of x_2, and we would have an accurate description of the solution set. However, to be consistent, we make the (arbitrary) convention of always using the free variables as the parameters for describing a solution set. (The answer section at the end of the text also reflects this convention.)

Whenever a system is inconsistent, the solution set is empty, even when the system has free variables. In this case, the solution set has *no* parametric representation.

Back-Substitution

Consider the following system, whose augmented matrix is in echelon form but is *not* in reduced echelon form:

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

A computer program would solve this system by back-substitution, rather than by computing the reduced echelon form. That is, the program would solve equation 3 for x_4 in terms of x_5 and substitute the expression for x_4 into equation 2, solve equation 2 for x_2, and then substitute the expressions for x_2 and x_4 into equation 1 and solve for x_1.

Our matrix format for the backward phase of row reduction, which produces the reduced echelon form, has the same number of arithmetic operations as back-substitution. But the discipline of the matrix format substantially reduces the likelihood of errors during hand computations. The best strategy is to use only the *reduced* echelon form to solve a system! The *Study Guide* that accompanies this text offers several helpful suggestions for performing row operations accurately and rapidly.

NUMERICAL NOTE

In general, the forward phase of row reduction takes much longer than the backward phase. An algorithm for solving a system is usually measured in flops (or floating point operations). A **flop** is one arithmetic operation ($+, -, *, /$) on two real floating point numbers.[3] For an $n \times (n+1)$ matrix, the reduction to echelon form can take $2n^3/3 + n^2/2 - 7n/6$ flops (which is approximately $2n^3/3$ flops when n is moderately large—say, $n \geq 30$). In contrast, further reduction to reduced echelon form needs at most n^2 flops.

Existence and Uniqueness Questions

Although a nonreduced echelon form is a poor tool for solving a system, this form is just the right device for answering two fundamental questions posed in Section 1.1.

EXAMPLE 5 Determine the existence and uniqueness of the solutions to the system

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

SOLUTION The augmented matrix of this system was row reduced in Example 3 to

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

The basic variables are x_1, x_2, and x_5; the free variables are x_3 and x_4. There is no equation such as $0 = 1$ that would indicate an inconsistent system, so we could use back-substitution to find a solution. But the *existence* of a solution is already clear in (8). Also, the solution is *not unique* because there are free variables. Each different choice of x_3 and x_4 determines a different solution. Thus the system has infinitely many solutions. ■

When a system is in echelon form and contains no equation of the form $0 = b$, with b nonzero, every nonzero equation contains a basic variable with a nonzero coefficient. Either the basic variables are completely determined (with no free variables) or at least one of the basic variables may be expressed in terms of one or more free variables. In the former case, there is a unique solution; in the latter case, there are infinitely many solutions (one for each choice of values for the free variables).

These remarks justify the following theorem.

---

[3] Traditionally, a *flop* was only a multiplication or division, because addition and subtraction took much less time and could be ignored. The definition of *flop* given here is preferred now, as a result of advances in computer architecture. See Golub and Van Loan, *Matrix Computations*, 2nd ed. (Baltimore: The Johns Hopkins Press, 1989), pp. 19–20.

THEOREM 2

Existence and Uniqueness Theorem

A linear system is consistent if and only if the rightmost column of the augmented matrix is *not* a pivot column—that is, if and only if an echelon form of the augmented matrix has *no* row of the form

$$[\,0 \quad \cdots \quad 0 \quad b\,] \qquad \text{with } b \text{ nonzero}$$

If a linear system is consistent, then the solution set contains either (i) a unique solution, when there are no free variables, or (ii) infinitely many solutions, when there is at least one free variable.

The following procedure outlines how to find and describe all solutions of a linear system.

USING ROW REDUCTION TO SOLVE A LINEAR SYSTEM

1. Write the augmented matrix of the system.
2. Use the row reduction algorithm to obtain an equivalent augmented matrix in echelon form. Decide whether the system is consistent. If there is no solution, stop; otherwise, go to the next step.
3. Continue row reduction to obtain the reduced echelon form.
4. Write the system of equations corresponding to the matrix obtained in step 3.
5. Rewrite each nonzero equation from step 4 so that its one basic variable is expressed in terms of any free variables appearing in the equation.

PRACTICE PROBLEMS

1. Find the general solution of the linear system whose augmented matrix is

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

2. Find the general solution of the system

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

1.2 EXERCISES

In Exercises 1 and 2, determine which matrices are in reduced echelon form and which others are only in echelon form.

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. $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

Row reduce the matrices in Exercises 3 and 4 to reduced echelon form. Circle the pivot positions in the final matrix and in the original matrix, and list the pivot columns.

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. Describe the possible echelon forms of a nonzero 2×2 matrix. Use the symbols $\blacksquare$, $*$, and 0, as in the first part of Example 1.

6. Repeat Exercise 5 for a nonzero 3×2 matrix.

Find the general solutions of the systems whose augmented matrices are given in Exercises 7–14.

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

Exercises 15 and 16 use the notation of Example 1 for matrices in echelon form. Suppose each matrix represents the augmented matrix for a system of linear equations. In each case, determine if the system is consistent. If the system is consistent, determine if the solution is unique.

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

In Exercises 17 and 18, determine the value(s) of h such that the matrix is the augmented matrix of a consistent linear system.

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

In Exercises 19 and 20, choose h and k such that the system has (a) no solution, (b) a unique solution, and (c) many solutions. Give separate answers for each part.

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

In Exercises 21 and 22, mark each statement True or False. Justify each answer.[4]

21. a. In some cases, a matrix may be row reduced to more than one matrix in reduced echelon form, using different sequences of row operations.

b. The row reduction algorithm applies only to augmented matrices for a linear system.

c. A basic variable in a linear system is a variable that corresponds to a pivot column in the coefficient matrix.

d. Finding a parametric description of the solution set of a linear system is the same as *solving* the system.

e. If one row in an echelon form of an augmented matrix is $[\,0\;\;0\;\;0\;\;5\;\;0\,]$, then the associated linear system is inconsistent.

22. a. The reduced echelon form of a matrix is unique.

b. If every column of an augmented matrix contains a pivot, then the corresponding system is consistent.

c. The pivot positions in a matrix depend on whether row interchanges are used in the row reduction process.

d. A general solution of a system is an explicit description of all solutions of the system.

e. Whenever a system has free variables, the solution set contains many solutions.

23. Suppose the coefficient matrix of a linear system of four equations in four variables has a pivot in each column. Explain why the system has a unique solution.

24. Suppose a system of linear equations has a 3×5 *augmented* matrix whose fifth column is not a pivot column. Is the system consistent? Why (or why not)?

---

[4] True/false questions of this type will appear in many sections. Methods for justifying your answers were described before Exercises 23 and 24 in Section 1.1.

25. Suppose the coefficient matrix of a system of linear equations has a pivot position in every row. Explain why the system is consistent.

26. Suppose a 3×5 *coefficient* matrix for a system has three pivot columns. Is the system consistent? Why or why not?

27. Restate the last sentence in Theorem 2 using the concept of pivot columns: "If a linear system is consistent, then the solution is unique if and only if ____________."

28. What would you have to know about the pivot columns in an augmented matrix in order to know that the linear system is consistent and has a unique solution?

29. A system of linear equations with fewer equations than unknowns is sometimes called an *underdetermined system*. Can such a system have a unique solution? Explain.

30. Give an example of an inconsistent underdetermined system of two equations in three unknowns.

31. A system of linear equations with more equations than unknowns is sometimes called an *overdetermined system*. Can such a system be consistent? Illustrate your answer with a specific system of three equations in two unknowns.

32. Suppose an $n \times (n+1)$ matrix is row reduced to reduced echelon form. Approximately what fraction of the total number of operations (flops) is involved in the backward phase of the reduction when $n = 20$? when $n = 200$?

Suppose experimental data are represented by a set of points in the plane. An **interpolating polynomial** for the data is a polynomial whose graph passes through every point. In scientific work, such a polynomial can be used, for example, to estimate values between the known data points. Another use is to create curves for graphical images on a computer screen. One method for finding an interpolating polynomial is to solve a system of linear equations.

WEB

33. Find the interpolating polynomial $p(t) = a_0 + a_1 t + a_2 t^2$ for the data $(1, 6)$, $(2, 15)$, $(3, 28)$. That is, find a_0, a_1, and a_2 such that

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

34. [**M**] In a wind tunnel experiment, the force on a projectile due to air resistance was measured at different velocities:

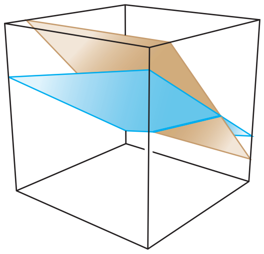
Velocity (100 ft/sec)	0	2	4	6	8	10
Force (100 lb)	0	2.90	14.8	39.6	74.3	119

Find an interpolating polynomial for these data and estimate the force on the projectile when the projectile is traveling at 750 ft/sec. Use $p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 t^5$. What happens if you try to use a polynomial of degree less than 5? (Try a cubic polynomial, for instance.)[5]

SOLUTIONS TO PRACTICE PROBLEMS

1. The reduced echelon form of the augmented matrix and the corresponding system are

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



The general solution of the system of equations is the line of intersection of the two planes.

The basic variables are x_1 and x_2, and the general solution is

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

Note: It is essential that the general solution describe each variable, with any parameters clearly identified. The following statement does *not* describe the solution:

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

This description implies that x_2 and x_3 are *both* free, which certainly is not the case.

---

[5] Exercises marked with the symbol [**M**] are designed to be worked with the aid of a "**M**atrix program" (a computer program, such as **M**ATLAB®, **M**aple™, **M**athematica®, **M**athCad®, or Derive™, or a programmable calculator with matrix capabilities, such as those manufactured by Texas Instruments or Hewlett-Packard).

2. Row reduce the system's augmented matrix:

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

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

This echelon matrix shows that the system is *inconsistent*, because its rightmost column is a pivot column; the third row corresponds to the equation $0 = 5$. There is no need to perform any more row operations. Note that the presence of the free variables in this problem is irrelevant because the system is inconsistent.

1.3 VECTOR EQUATIONS

Important properties of linear systems can be described with the concept and notation of vectors. This section connects equations involving vectors to ordinary systems of equations. The term *vector* appears in a variety of mathematical and physical contexts, which we will discuss in Chapter 4, "Vector Spaces." Until then, *vector* will mean an *ordered list of numbers*. This simple idea enables us to get to interesting and important applications as quickly as possible.

Vectors in $\mathbb{R}^2$

A matrix with only one column is called a **column vector**, or simply a **vector**. Examples of vectors with two entries are

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

where w_1 and w_2 are any real numbers. The set of all vectors with two entries is denoted by $\mathbb{R}^2$ (read "r-two"). The $\mathbb{R}$ stands for the real numbers that appear as entries in the vectors, and the exponent 2 indicates that each vector contains two entries.[1]

Two vectors in $\mathbb{R}^2$ are **equal** if and only if their corresponding entries are equal. Thus $\begin{bmatrix} 4 \\ 7 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ are *not* equal, because vectors in $\mathbb{R}^2$ are *ordered pairs* of real numbers.

Given two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, their **sum** is the vector $\mathbf{u} + \mathbf{v}$ obtained by adding corresponding entries of $\mathbf{u}$ and $\mathbf{v}$. For example,

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

Given a vector $\mathbf{u}$ and a real number c, the **scalar multiple** of $\mathbf{u}$ by c is the vector $c\mathbf{u}$ obtained by multiplying each entry in $\mathbf{u}$ by c. For instance,

$$\text{if} \quad \mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{and} \quad c = 5, \qquad \text{then} \quad c\mathbf{u} = 5\begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$$

---

[1] Most of the text concerns vectors and matrices that have only real entries. However, all definitions and theorems in Chapters 1–5, and in most of the rest of the text, remain valid if the entries are complex numbers. Complex vectors and matrices arise naturally, for example, in electrical engineering and physics.

The number c in $c\mathbf{u}$ is called a **scalar**; it is written in lightface type to distinguish it from the boldface vector $\mathbf{u}$.

The operations of scalar multiplication and vector addition can be combined, as in the following example.

EXAMPLE 1 Given $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$, find $4\mathbf{u}$, $(-3)\mathbf{v}$, and $4\mathbf{u} + (-3)\mathbf{v}$.

SOLUTION

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

and

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

Sometimes, for convenience (and also to save space), this text may write a column vector such as $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ in the form $(3, -1)$. In this case, the parentheses and the comma distinguish the vector $(3, -1)$ from the 1×2 row matrix $\begin{bmatrix} 3 & -1 \end{bmatrix}$, written with brackets and no comma. Thus

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

because the matrices have different shapes, even though they have the same entries.

Geometric Descriptions of $\mathbb{R}^2$

Consider a rectangular coordinate system in the plane. Because each point in the plane is determined by an ordered pair of numbers, *we can identify a geometric point* (a, b) *with the column vector* $\begin{bmatrix} a \\ b \end{bmatrix}$. So we may regard $\mathbb{R}^2$ as the set of all points in the plane. See Fig. 1.

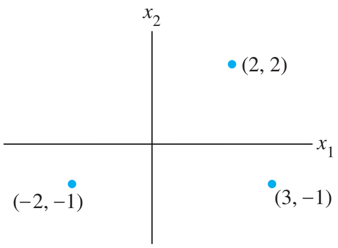


FIGURE 1 Vectors as points.

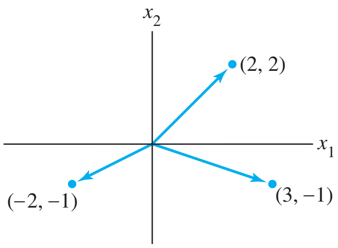


FIGURE 2 Vectors with arrows.

The geometric visualization of a vector such as $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$ is often aided by including an arrow (directed line segment) from the origin $(0, 0)$ to the point $(3, -1)$, as in Fig. 2. In this case, the individual points along the arrow itself have no special significance.[2]

The sum of two vectors has a useful geometric representation. The following rule can be verified by analytic geometry.

---

[2] In physics, arrows can represent forces and usually are free to move about in space. This interpretation of vectors will be discussed in Section 4.1.

Parallelogram Rule for Addition

If **u** and **v** in $\mathbb{R}^2$ are represented as points in the plane, then $\mathbf{u} + \mathbf{v}$ corresponds to the fourth vertex of the parallelogram whose other vertices are **u**, **0**, and **v**. See Fig. 3.

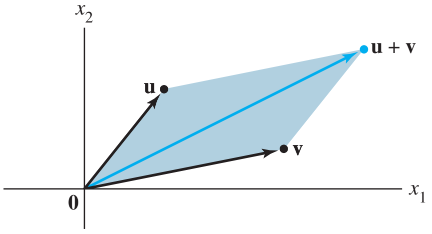


FIGURE 3 The parallelogram rule.

EXAMPLE 2 The vectors $\mathbf{u} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \end{bmatrix}$, and $\mathbf{u} + \mathbf{v} = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ are displayed in Fig. 4. ■

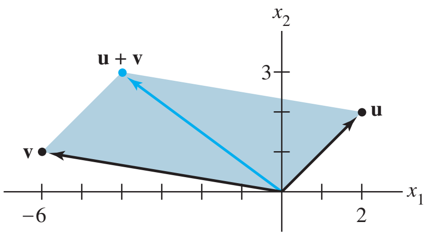
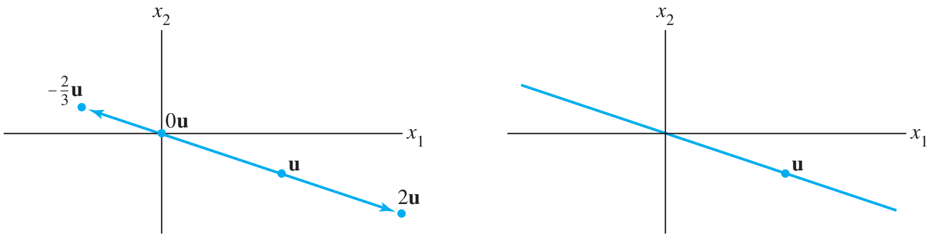


FIGURE 4

The next example illustrates the fact that the set of all scalar multiples of one fixed nonzero vector is a line through the origin, (0, 0).

EXAMPLE 3 Let $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Display the vectors **u**, 2**u**, and $-\frac{2}{3}\mathbf{u}$ on a graph.

SOLUTION See Fig. 5, where **u**, $2\mathbf{u} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$, and $-\frac{2}{3}\mathbf{u} = \begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$ are displayed. The arrow for 2**u** is twice as long as the arrow for **u**, and the arrows point in the same direction. The arrow for $-\frac{2}{3}\mathbf{u}$ is two-thirds the length of the arrow for **u**, and the arrows point in opposite directions. In general, the length of the arrow for $c\mathbf{u}$ is $|c|$ times the



Typical multiples of **u**

The set of all multiples of **u**

FIGURE 5

length of the arrow for **u**. [Recall that the length of the line segment from $(0, 0)$ to (a, b) is $\sqrt{a^2 + b^2}$. We shall discuss this further in Chapter 6.] ■

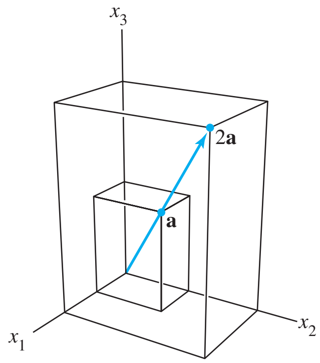


FIGURE 6
Scalar multiples .

Vectors in $\mathbb{R}^3$

Vectors in $\mathbb{R}^3$ are 3×1 column matrices with three entries. They are represented geometrically by points in a three-dimensional coordinate space, with arrows from the origin sometimes included for visual clarity. The vectors $\mathbf{a} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ and $2\mathbf{a}$ are displayed in Fig. 6.

Vectors in $\mathbb{R}^n$

If n is a positive integer, $\mathbb{R}^n$ (read "r-n") denotes the collection of all lists (or *ordered n-tuples*) of n real numbers, usually written as $n \times 1$ column matrices, such as

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

The vector whose entries are all zero is called the **zero vector** and is denoted by **0**. (The number of entries in **0** will be clear from the context.)

Equality of vectors in $\mathbb{R}^n$ and the operations of scalar multiplication and vector addition in $\mathbb{R}^n$ are defined entry by entry just as in $\mathbb{R}^2$. These operations on vectors have the following properties, which can be verified directly from the corresponding properties for real numbers. See Practice Problem 1 and Exercises 33 and 34 at the end of this section.

Algebraic Properties of $\mathbb{R}^n$

For all **u**, **v**, **w** in $\mathbb{R}^n$ and all scalars c and 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) $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$,
where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$
(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) $1\mathbf{u} = \mathbf{u}$

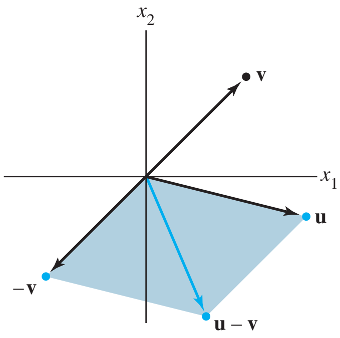


FIGURE 7
Vector subtraction.

For simplicity of notation, a vector such as $\mathbf{u} + (-1)\mathbf{v}$ is often written as $\mathbf{u} - \mathbf{v}$. Figure 7 shows $\mathbf{u} - \mathbf{v}$ as the sum of **u** and $-\mathbf{v}$.

Linear Combinations

Given vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1, c_2, \dots, c_p$, the vector **y** defined by

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

is called a **linear combination** of $\mathbf{v}_1, \dots, \mathbf{v}_p$ with **weights** $c_1, \dots, c_p$. Property (ii) above permits us to omit parentheses when forming such a linear combination. The weights in a linear combination can be any real numbers, including zero. For example, some linear combinations of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are

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

EXAMPLE 4 Figure 8 identifies selected linear combinations of $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. (Note that sets of parallel grid lines are drawn through integer multiples of $\mathbf{v}_1$ and $\mathbf{v}_2$.) Estimate the linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$ that generate the vectors $\mathbf{u}$ and $\mathbf{w}$.

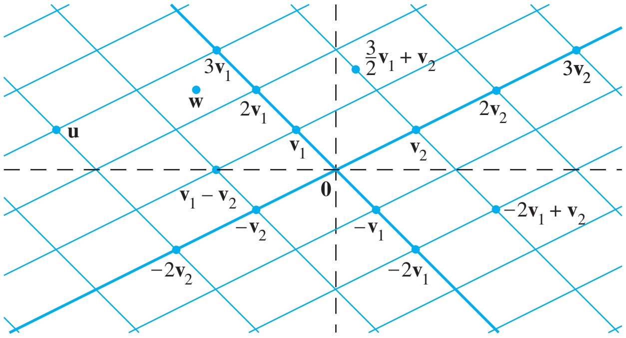


FIGURE 8 Linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$.

SOLUTION The parallelogram rule shows that $\mathbf{u}$ is the sum of $3\mathbf{v}_1$ and $-2\mathbf{v}_2$; that is,

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

This expression for $\mathbf{u}$ can be interpreted as instructions for traveling from the origin to $\mathbf{u}$ along two straight paths. First, travel 3 units in the $\mathbf{v}_1$ direction to $3\mathbf{v}_1$, and then travel -2 units in the $\mathbf{v}_2$ direction (parallel to the line through $\mathbf{v}_2$ and $\mathbf{0}$). Next, although the vector $\mathbf{w}$ is not on a grid line, $\mathbf{w}$ appears to be about halfway between two pairs of grid lines, at the vertex of a parallelogram determined by $(5/2)\mathbf{v}_1$ and $(-1/2)\mathbf{v}_2$. (See Fig. 9.) Thus a reasonable estimate for $\mathbf{w}$ is

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

■

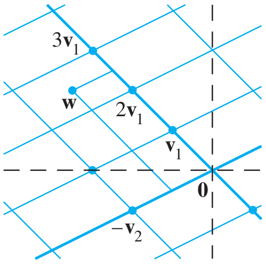


FIGURE 9

The next example connects a problem about linear combinations to the fundamental existence question studied in Sections 1.1 and 1.2.

EXAMPLE 5 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Determine whether $\mathbf{b}$ can be generated (or written) as a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$. That is, determine whether weights x_1 and x_2 exist such that

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b} \tag{1}$$

If vector equation (1) has a solution, find it.

SOLUTION Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

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

which is the same as

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

and

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

The vectors on the left and right sides of (2) are equal if and only if their corresponding entries are both equal. That is, x_1 and x_2 make the vector equation (1) true if and only if x_1 and x_2 satisfy the system

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

To solve this system, row reduce the augmented matrix of the system as follows:[3]

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

The solution of (3) is $x_1 = 3$ and $x_2 = 2$. Hence **b** is a linear combination of $\mathbf{a}_1$ and $\mathbf{a}_2$, with weights $x_1 = 3$ and $x_2 = 2$. That is,

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

Observe in Example 5 that the original vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and **b** are the columns of the augmented matrix that we row reduced:

$$\begin{array}{c} \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{array} \end{array}$$

For brevity, write this matrix in a way that identifies its columns—namely,

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{b} \end{bmatrix} \tag{4}$$

It is clear how to write this augmented matrix immediately from vector equation (1), without going through the intermediate steps of Example 5. Take the vectors in the order in which they appear in (1) and put them into the columns of a matrix as in (4).

The discussion above is easily modified to establish the following fundamental fact.

A vector equation

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

has the same solution set as the linear system whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{5}$$

In particular, **b** can be generated by a linear combination of $\mathbf{a}_1, \dots, \mathbf{a}_n$ if and only if there exists a solution to the linear system corresponding to the matrix (5).

---

[3] The symbol $\sim$ between matrices denotes row equivalence (Section 1.2).

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of vectors.

DEFINITION

If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \dots, \mathbf{v}_p$ is denoted by Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ and is called the **subset of** $\mathbb{R}^n$ **spanned** (or **generated**) **by** $\mathbf{v}_1, \dots, \mathbf{v}_p$. That is, Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the collection of all vectors that can be written in the form

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

with $c_1, \dots, c_p$ scalars.

Asking whether a vector $\mathbf{b}$ is in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ amounts to asking whether the vector equation

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

has a solution, or, equivalently, asking whether the linear system with augmented matrix $[\,\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_p \quad \mathbf{b}\,]$ has a solution.

Note that Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ contains every scalar multiple of $\mathbf{v}_1$ (for example), since $c\mathbf{v}_1 = c\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. In particular, the zero vector must be in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$.

A Geometric Description of Span $\{\mathbf{v}\}$ and Span $\{\mathbf{u}, \mathbf{v}\}$

Let $\mathbf{v}$ be a nonzero vector in $\mathbb{R}^3$. Then Span $\{\mathbf{v}\}$ is the set of all scalar multiples of $\mathbf{v}$, which is the set of points on the line in $\mathbb{R}^3$ through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 10.

If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in $\mathbb{R}^3$, with $\mathbf{v}$ not a multiple of $\mathbf{u}$, then Span $\{\mathbf{u}, \mathbf{v}\}$ is the plane in $\mathbb{R}^3$ that contains $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{0}$. In particular, Span $\{\mathbf{u}, \mathbf{v}\}$ contains the line in $\mathbb{R}^3$ through $\mathbf{u}$ and $\mathbf{0}$ and the line through $\mathbf{v}$ and $\mathbf{0}$. See Fig. 11.

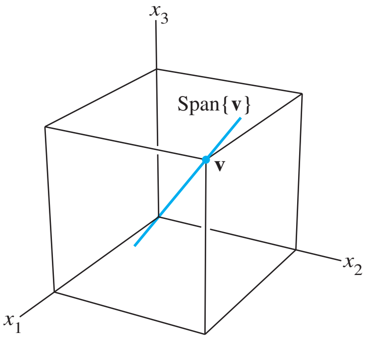


FIGURE 10 Span $\{\mathbf{v}\}$ as a line through the origin.

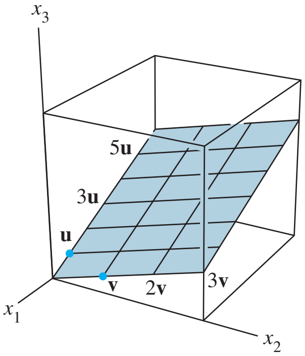


FIGURE 11 Span $\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

EXAMPLE 6 Let $\mathbf{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Then Span $\{\mathbf{a}_1, \mathbf{a}_2\}$ is a plane through the origin in $\mathbb{R}^3$. Is $\mathbf{b}$ in that plane?

SOLUTION Does the equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ have a solution? To answer this, row reduce the augmented matrix $[\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}\,]$:

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

The third equation is $0 = -2$, which shows that the system has no solution. The vector equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 = \mathbf{b}$ has no solution, and so $\mathbf{b}$ is *not* in Span $\{\mathbf{a}_1, \mathbf{a}_2\}$. ■

Linear Combinations in Applications

The final example shows how scalar multiples and linear combinations can arise when a quantity such as "cost" is broken down into several categories. The basic principle for the example concerns the cost of producing several units of an item when the cost per unit is known:

$$\begin{Bmatrix} \text{number} \\ \text{of units} \end{Bmatrix} \cdot \begin{Bmatrix} \text{cost} \\ \text{per unit} \end{Bmatrix} = \begin{Bmatrix} \text{total} \\ \text{cost} \end{Bmatrix}$$

EXAMPLE 7 A company manufactures two products. For \$1.00 worth of product B, the company spends \$.45 on materials, \$.25 on labor, and \$.15 on overhead. For \$1.00 worth of product C, the company spends \$.40 on materials, \$.30 on labor, and \$.15 on overhead. Let

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

Then $\mathbf{b}$ and $\mathbf{c}$ represent the "costs per dollar of income" for the two products.

a. What economic interpretation can be given to the vector $100\mathbf{b}$?

b. Suppose the company wishes to manufacture x_1 dollars worth of product B and x_2 dollars worth of product C. Give a vector that describes the various costs the company will have (for materials, labor, and overhead).

SOLUTION

a. Compute

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

The vector $100\mathbf{b}$ lists the various costs for producing \$100 worth of product B—namely, \$45 for materials, \$25 for labor, and \$15 for overhead.

b. The costs of manufacturing x_1 dollars worth of B are given by the vector $x_1\mathbf{b}$, and the costs of manufacturing x_2 dollars worth of C are given by $x_2\mathbf{c}$. Hence the total costs for both products are given by the vector $x_1\mathbf{b} + x_2\mathbf{c}$. ■

PRACTICE PROBLEMS

1. Prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$.

2. For what value(s) of h will $\mathbf{y}$ be in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if

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

1.3 EXERCISES

In Exercises 1 and 2, compute $\mathbf{u}+\mathbf{v}$ and $\mathbf{u}-2\mathbf{v}$.

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

In Exercises 3 and 4, display the following vectors using arrows on an xy-graph: $\mathbf{u}$, $\mathbf{v}$, $-\mathbf{v}$, $-2\mathbf{v}$, $\mathbf{u}+\mathbf{v}$, $\mathbf{u}-\mathbf{v}$, and $\mathbf{u}-2\mathbf{v}$. Notice that $\mathbf{u}-\mathbf{v}$ is the vertex of a parallelogram whose other vertices are $\mathbf{u}$, $\mathbf{0}$, and $-\mathbf{v}$.

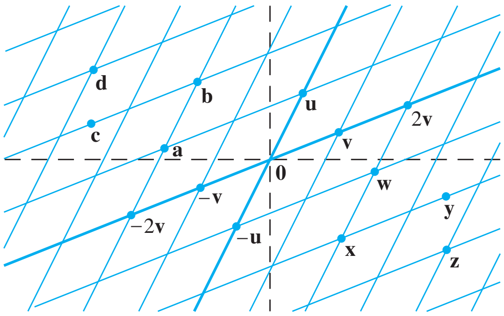
3. $\mathbf{u}$ and $\mathbf{v}$ as in Exercise 1 **4.** $\mathbf{u}$ and $\mathbf{v}$ as in Exercise 2

In Exercises 5 and 6, write a system of equations that is equivalent to the given vector equation.

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

Use the accompanying figure to write each vector listed in Exercises 7 and 8 as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. Is every vector in $\mathbb{R}^2$ a linear combination of $\mathbf{u}$ and $\mathbf{v}$?



7. Vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\mathbf{d}$

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

In Exercises 9 and 10, write a vector equation that is equivalent to the given system of equations.

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

In Exercises 11 and 12, determine if $\mathbf{b}$ is a linear combination of $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$.

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

In Exercises 13 and 14, determine if $\mathbf{b}$ is a linear combination of the vectors formed from the columns of the matrix 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. Let $\mathbf{a}_1=\begin{bmatrix}1\\3\\-1\end{bmatrix}, \mathbf{a}_2=\begin{bmatrix}-5\\-8\\2\end{bmatrix}$, and $\mathbf{b}=\begin{bmatrix}3\\-5\\h\end{bmatrix}$. For what value(s) of h is $\mathbf{b}$ in the plane spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$?

16. Let $\mathbf{v}_1=\begin{bmatrix}1\\0\\-2\end{bmatrix}, \mathbf{v}_2=\begin{bmatrix}-2\\1\\7\end{bmatrix}$, and $\mathbf{y}=\begin{bmatrix}h\\-3\\-5\end{bmatrix}$. For what value(s) of h is $\mathbf{y}$ in the plane generated by $\mathbf{v}_1$ and $\mathbf{v}_2$?

In Exercises 17 and 18, list five vectors in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$. For each vector, show the weights on $\mathbf{v}_1$ and $\mathbf{v}_2$ used to generate the vector and list the three entries of the vector. Do not make a sketch.

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. Give a geometric description of Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the vectors $\mathbf{v}_1=\begin{bmatrix}8\\2\\-6\end{bmatrix}$ and $\mathbf{v}_2=\begin{bmatrix}12\\3\\-9\end{bmatrix}$.

20. Give a geometric description of Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the vectors in Exercise 18.

21. Let $\mathbf{u}=\begin{bmatrix}2\\-1\end{bmatrix}$ and $\mathbf{v}=\begin{bmatrix}2\\1\end{bmatrix}$. Show that $\begin{bmatrix}h\\k\end{bmatrix}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$ for all h and k.

22. Construct a 3×3 matrix A, with nonzero entries, and a vector $\mathbf{b}$ in $\mathbb{R}^3$ such that $\mathbf{b}$ is *not* in the set spanned by the columns of A.

In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. Another notation for the vector $\begin{bmatrix}-4\\3\end{bmatrix}$ is $[\,-4 \quad 3\,]$.

b. The points in the plane corresponding to $\begin{bmatrix}-2\\5\end{bmatrix}$ and $\begin{bmatrix}-5\\2\end{bmatrix}$ lie on a line through the origin.

c. An example of a linear combination of vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is the vector $\frac{1}{2}\mathbf{v}_1$.

d. The solution set of the linear system whose augmented matrix is $[\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}\,]$ is the same as the solution set of the equation $x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b}$.

e. The set Span $\{\mathbf{u}, \mathbf{v}\}$ is always visualized as a plane through the origin.

24. a. When $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors, Span $\{\mathbf{u}, \mathbf{v}\}$ contains only the line through $\mathbf{u}$ and the origin, and the line through $\mathbf{v}$ and the origin.

b. Any list of five real numbers is a vector in $\mathbb{R}^5$.

c. Asking whether the linear system corresponding to an augmented matrix $[\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}\,]$ has a solution amounts to asking whether $\mathbf{b}$ is in Span $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

d. The vector $\mathbf{v}$ results when a vector $\mathbf{u} - \mathbf{v}$ is added to the vector $\mathbf{v}$.

e. The weights $c_1, \dots, c_p$ in a linear combination $c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$ cannot all be zero.

25. Let $A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 3 & -2 \\ -2 & 6 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -4 \end{bmatrix}$. Denote the columns of A by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, and let $W = \text{Span}\,\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$.

a. Is $\mathbf{b}$ in $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$? How many vectors are in $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$?

b. Is $\mathbf{b}$ in W? How many vectors are in W?

c. Show that $\mathbf{a}_1$ is in W. [*Hint:* Row operations are unnecessary.]

26. Let $A = \begin{bmatrix} 2 & 0 & 6 \\ -1 & 8 & 5 \\ 1 & -2 & 1 \end{bmatrix}$, let $\mathbf{b} = \begin{bmatrix} 10 \\ 3 \\ 7 \end{bmatrix}$, and let W be the set of all linear combinations of the columns of A.

a. Is $\mathbf{b}$ in W?

b. Show that the second column of A is in W.

27. A mining company has two mines. One day's operation at mine #1 produces ore that contains 30 metric tons of copper and 600 kilograms of silver, while one day's operation at mine #2 produces ore that contains 40 metric tons of copper and 380 kilograms of silver. Let $\mathbf{v}_1 = \begin{bmatrix} 30 \\ 600 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 40 \\ 380 \end{bmatrix}$. Then $\mathbf{v}_1$ and $\mathbf{v}_2$ represent the "output per day" of mine #1 and mine #2, respectively.

a. What physical interpretation can be given to the vector $5\mathbf{v}_1$?

b. Suppose the company operates mine #1 for x_1 days and mine #2 for x_2 days. Write a vector equation whose solution gives the number of days each mine should operate in order to produce 240 tons of copper and 2824 kilograms of silver. Do not solve the equation.

c. **[M]** Solve the equation in (b).

28. A steam plant burns two types of coal: anthracite (A) and bituminous (B). For each ton of A burned, the plant produces 27.6 million Btu of heat, 3100 grams (g) of sulfur dioxide, and 250 g of particulate matter (solid-particle pollutants). For each ton of B burned, the plant produces 30.2 million Btu, 6400 g of sulfur dioxide, and 360 g of particulate matter.

a. How much heat does the steam plant produce when it burns x_1 tons of A and x_2 tons of B?

b. Suppose the output of the steam plant is described by a vector that lists the amounts of heat, sulfur dioxide, and particulate matter. Express this output as a linear combination of two vectors, assuming that the plant burns x_1 tons of A and x_2 tons of B.

c. **[M]** Over a certain time period, the steam plant produced 162 million Btu of heat, 23,610 g of sulfur dioxide, and 1623 g of particulate matter. Determine how many tons of each type of coal the steam plant must have burned. Include a vector equation as part of your solution.

29. Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be points in $\mathbb{R}^3$ and suppose that for $j = 1, \dots, k$ an object with mass m_j is located at point $\mathbf{v}_j$. Physicists call such objects *point masses*. The total mass of the system of point masses is

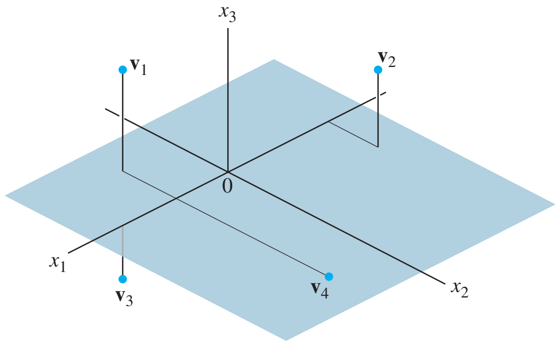
$$m = m_1 + \cdots + m_k$$

The *center of gravity* (or *center of mass*) of the system is

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

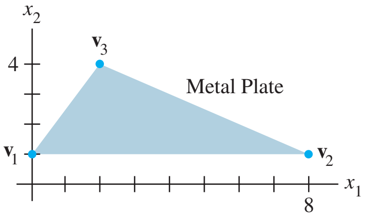
Compute the center of gravity of the system consisting of the following point masses (see the figure):

Point	**Mass**
$\mathbf{v}_1 = (2, -2, 4)$	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. Let $\mathbf{v}$ be the center of mass of a system of point masses located at $\mathbf{v}_1, \dots, \mathbf{v}_k$ as in Exercise 29. Is $\mathbf{v}$ in Span $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$? Explain.

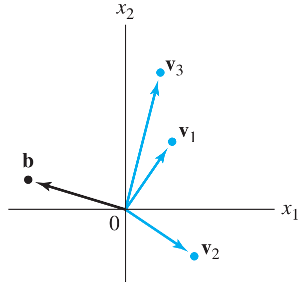
31. A thin triangular plate of uniform density and thickness has vertices at $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$, and $\mathbf{v}_3 = (2, 4)$, as in the figure below, and the mass of the plate is 3 g.



a. Find the (x, y)-coordinates of the center of mass of the plate. This "balance point" of the plate coincides with the center of mass of a system consisting of three 1-gram point masses located at the vertices of the plate.

b. Determine how to distribute an additional mass of 6 g at the three vertices of the plate to move the balance point of the plate to $(2, 2)$. [*Hint:* Let w_1, w_2, and w_3 denote the masses added at the three vertices, so that $w_1 + w_2 + w_3 = 6$.]

32. Consider the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{b}$ in $\mathbb{R}^2$, shown in the figure. Does the equation $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}$ have a solution? Is the solution unique? Use the figure to explain your answers.



33. Use the vectors $\mathbf{u} = (u_1, \dots, u_n)$, $\mathbf{v} = (v_1, \dots, v_n)$, and $\mathbf{w} = (w_1, \dots, w_n)$ to verify the following algebraic properties of $\mathbb{R}^n$.

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

b. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ for each scalar c

34. Use the vector $\mathbf{u} = (u_1, \dots, u_n)$ to verify the following algebraic properties of $\mathbb{R}^n$.

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

b. $c(d\mathbf{u}) = (cd)\mathbf{u}$ for all scalars c and d

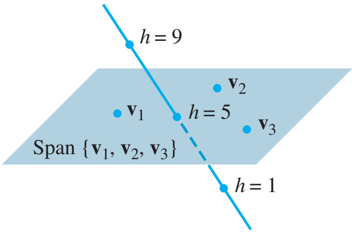
SOLUTIONS TO PRACTICE PROBLEMS

1. Take arbitrary vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ in $\mathbb{R}^n$, and compute

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1 + v_1, \dots, u_n + v_n) && \text{Definition of vector addition} \\ &= (v_1 + u_1, \dots, v_n + u_n) && \text{Commutativity of addition in } \mathbb{R} \\ &= \mathbf{v} + \mathbf{u} && \text{Definition of vector addition} \end{aligned}$$

2. The vector $\mathbf{y}$ belongs to Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if there exist scalars x_1, x_2, x_3 such that

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



The points $\begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$ lie on a line that intersects the plane when $h = 5$.

This vector equation is equivalent to a system of three linear equations in three unknowns. If you row reduce the augmented matrix for this system, you find that

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

The system is consistent if and only if there is no pivot in the fourth column. That is, $h - 5$ must be 0. So $\mathbf{y}$ is in Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if $h = 5$.

Remember: The presence of a free variable in a system does not guarantee that the system is consistent.

1.4 THE MATRIX EQUATION $A\mathbf{x} = \mathbf{b}$

A fundamental idea in linear algebra is to view a linear combination of vectors as the product of a matrix and a vector. The following definition permits us to rephrase some of the concepts of Section 1.3 in new ways.

DEFINITION

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if $\mathbf{x}$ is in $\mathbb{R}^n$, then the **product of A and $\mathbf{x}$**, denoted by $A\mathbf{x}$, is **the linear combination of the columns of A using the corresponding entries in $\mathbf{x}$ as weights**; that is,

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Note that $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

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

EXAMPLE 2 For $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in $\mathbb{R}^m$, write the linear combination $3\mathbf{v}_1 - 5\mathbf{v}_2 + 7\mathbf{v}_3$ as a matrix times a vector.

SOLUTION Place $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ into the columns of a matrix A and place the weights $3, -5$, and 7 into a vector $\mathbf{x}$. That is,

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

Section 1.3 showed how to write a system of linear equations as a vector equation involving a linear combination of vectors. For example, the system

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

is equivalent to

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

As in Example 2, the linear combination on the left side is a matrix times a vector, so that (2) becomes

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

Equation (3) has the form $A\mathbf{x} = \mathbf{b}$. Such an equation is called a **matrix equation**, to distinguish it from a vector equation such as is shown in (2).

Notice how the matrix in (3) is just the matrix of coefficients of the system (1). Similar calculations show that any system of linear equations, or any vector equation such as (2), can be written as an equivalent matrix equation in the form $A\mathbf{x} = \mathbf{b}$. This simple observation will be used repeatedly throughout the text.

Here is the formal result.

THEOREM 3

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1, \dots, \mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

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

has the same solution set as the vector equation

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

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \tag{6}$$

Theorem 3 provides a powerful tool for gaining insight into problems in linear algebra, because a system of linear equations may now be viewed in three different but equivalent ways: as a matrix equation, as a vector equation, or as a system of linear equations. Whenever you construct a mathematical model of a problem in real life, you are free to choose whichever viewpoint is most natural. Then you may switch from one formulation of a problem to another whenever it is convenient. In any case, the matrix equation (4), the vector equation (5), and the system of equations are all solved in the same way—by row reducing the augmented matrix (6). Other methods of solution will be discussed later.

Existence of Solutions

The definition of $A\mathbf{x}$ leads directly to the following useful fact.

The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A.

Section 1.3 considered the existence question, "Is $\mathbf{b}$ in Span $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$?" Equivalently, "Is $A\mathbf{x} = \mathbf{b}$ consistent?" A harder existence problem is to determine whether the equation $A\mathbf{x} = \mathbf{b}$ is consistent *for all* possible $\mathbf{b}$.

EXAMPLE 3 Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. Is the equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible b_1, b_2, b_3?

SOLUTION Row reduce the augmented matrix for $A\mathbf{x} = \mathbf{b}$:

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

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

The third entry in column 4 equals $b_1 - \frac{1}{2}b_2 + b_3$. The equation $A\mathbf{x} = \mathbf{b}$ is *not* consistent for every $\mathbf{b}$ because some choices of $\mathbf{b}$ can make $b_1 - \frac{1}{2}b_2 + b_3$ nonzero. ■

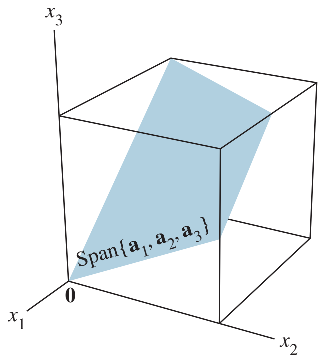


FIGURE 1

The columns of $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]$ span a plane through $\mathbf{0}$.

The reduced matrix in Example 3 provides a description of all $\mathbf{b}$ for which the equation $A\mathbf{x} = \mathbf{b}$ *is* consistent: The entries in $\mathbf{b}$ must satisfy

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

This is the equation of a plane through the origin in $\mathbb{R}^3$. The plane is the set of all linear combinations of the three columns of A. See Fig. 1.

The equation $A\mathbf{x} = \mathbf{b}$ in Example 3 fails to be consistent for all $\mathbf{b}$ because the echelon form of A has a row of zeros. If A had a pivot in all three rows, we would not care about the calculations in the augmented column because in this case an echelon form of the augmented matrix could not have a row such as $[\,0 \quad 0 \quad 0 \quad 1\,]$.

In the next theorem, the sentence "The columns of A span $\mathbb{R}^m$" means that *every* $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A. In general, a set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^m$ **spans** (or **generates**) $\mathbb{R}^m$ if every vector in $\mathbb{R}^m$ is a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_p$—that is, if Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\} = \mathbb{R}^m$.

THEOREM 4

Let A be an $m \times n$ matrix. Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.
b. Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.
c. The columns of A span $\mathbb{R}^m$.
d. A has a pivot position in every row.

Theorem 4 is one of the most useful theorems in this chapter. Statements (a), (b), and (c) are equivalent because of the definition of $A\mathbf{x}$ and what it means for a set of vectors to span $\mathbb{R}^m$. The discussion after Example 3 suggests why (a) and (d) are equivalent; a proof is given at the end of the section. The exercises will provide examples of how Theorem 4 is used.

Warning: Theorem 4 is about a *coefficient matrix*, not an augmented matrix. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ may or may not be consistent.

Computation of $A\mathbf{x}$

The calculations in Example 1 were based on the definition of the product of a matrix A and a vector $\mathbf{x}$. The following simple example will lead to a more efficient method for calculating the entries in $A\mathbf{x}$ when working problems by hand.

EXAMPLE 4 Compute $A\mathbf{x}$, where $A = \begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$.

SOLUTION From the definition,

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

The first entry in the product $A\mathbf{x}$ is a sum of products (sometimes called a *dot product*), using the first row of A and the entries in $\mathbf{x}$. That is,

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

This matrix shows how to compute the first entry in $A\mathbf{x}$ directly, without writing down all the calculations shown in (7). Similarly, the second entry in $A\mathbf{x}$ can be calculated at once by multiplying the entries in the second row of A by the corresponding entries in $\mathbf{x}$ and then summing the resulting products:

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

Likewise, the third entry in $A\mathbf{x}$ can be calculated from the third row of A and the entries in $\mathbf{x}$. ■

Row–Vector Rule for Computing $A\mathbf{x}$

If the product $A\mathbf{x}$ is defined, then the ith entry in $A\mathbf{x}$ is the sum of the products of corresponding entries from row i of A and from the vector $\mathbf{x}$.

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

By definition, the matrix in Example 5(c) with 1's on the diagonal and 0's elsewhere is called an **identity matrix** and is denoted by I. The calculation in part (c) shows that $I\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^3$. There is an analogous $n \times n$ identity matrix, sometimes written as I_n. As in part (c), $I_n\mathbf{x} = \mathbf{x}$ for every $\mathbf{x}$ in $\mathbb{R}^n$.

Properties of the Matrix–Vector Product $A\mathbf{x}$

The facts in the next theorem are important and will be used throughout the text. The proof relies on the definition of $A\mathbf{x}$ and the algebraic properties of $\mathbb{R}^n$.

THEOREM 5

If A is an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:

a. $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$;

b. $A(c\mathbf{u}) = c(A\mathbf{u})$.

PROOF For simplicity, take $n = 3$, $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]$, and $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^3$. (The proof of the general case is similar.) For $i = 1, 2, 3$, let u_i and v_i be the ith entries in $\mathbf{u}$ and $\mathbf{v}$, respectively. To prove statement (a), compute $A(\mathbf{u}+\mathbf{v})$ as a linear combination of the columns of A using the entries in $\mathbf{u}+\mathbf{v}$ as weights.

$$\begin{aligned}
A(\mathbf{u}+\mathbf{v}) &= [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ u_3+v_3 \end{bmatrix} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{Entries in } \mathbf{u}+\mathbf{v} \\
&= \overset{\downarrow}{(u_1+v_1)}\underset{\uparrow}{\mathbf{a}_1} + \overset{\downarrow}{(u_2+v_2)}\underset{\uparrow}{\mathbf{a}_2} + \overset{\downarrow}{(u_3+v_3)}\underset{\uparrow}{\mathbf{a}_3} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{Columns of } A \\
&= (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + u_3\mathbf{a}_3) + (v_1\mathbf{a}_1 + v_2\mathbf{a}_2 + v_3\mathbf{a}_3) \\
&= A\mathbf{u} + A\mathbf{v}
\end{aligned}$$

To prove statement (b), compute $A(c\mathbf{u})$ as a linear combination of the columns of A using the entries in $c\mathbf{u}$ as weights.

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

■

NUMERICAL NOTE

To optimize a computer algorithm to compute $A\mathbf{x}$, the sequence of calculations should involve data stored in contiguous memory locations. The most widely used professional algorithms for matrix computations are written in Fortran, a language that stores a matrix as a set of columns. Such algorithms compute $A\mathbf{x}$ as a linear combination of the columns of A. In contrast, if a program is written in the popular language C, which stores matrices by rows, $A\mathbf{x}$ should be computed via the alternative rule that uses the rows of A.

PROOF OF THEOREM 4 As was pointed out after Theorem 4, statements (a), (b), and (c) are logically equivalent. So, it suffices to show (for an arbitrary matrix A) that (a) and (d) are either both true or both false. That will tie all four statements together.

Let U be an echelon form of A. Given $\mathbf{b}$ in $\mathbb{R}^m$, we can row reduce the augmented matrix $[\,A \quad \mathbf{b}\,]$ to an augmented matrix $[\,U \quad \mathbf{d}\,]$ for some $\mathbf{d}$ in $\mathbb{R}^m$:

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

If statement (d) is true, then each row of U contains a pivot position and there can be no pivot in the augmented column. So $A\mathbf{x} = \mathbf{b}$ has a solution for any $\mathbf{b}$, and (a) is true. If (d) is false, the last row of U is all zeros. Let $\mathbf{d}$ be any vector with a 1 in its last entry. Then $[\, U \quad \mathbf{d}\,]$ represents an *inconsistent* system. Since row operations are reversible, $[\, U \quad \mathbf{d}\,]$ can be transformed into the form $[\, A \quad \mathbf{b}\,]$. The new system $A\mathbf{x} = \mathbf{b}$ is also inconsistent, and (a) is false. ■

PRACTICE PROBLEMS

1. Let $A = \begin{bmatrix} 1 & 5 & -2 & 0 \\ -3 & 1 & 9 & -5 \\ 4 & -8 & -1 & 7 \end{bmatrix}$, $\mathbf{p} = \begin{bmatrix} 3 \\ -2 \\ 0 \\ -4 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} -7 \\ 9 \\ 0 \end{bmatrix}$. It can be shown that $\mathbf{p}$ is a solution of $A\mathbf{x} = \mathbf{b}$. Use this fact to exhibit $\mathbf{b}$ as a specific linear combination of the columns of A.

2. Let $A = \begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$. Verify Theorem 5(a) in this case by computing $A(\mathbf{u} + \mathbf{v})$ and $A\mathbf{u} + A\mathbf{v}$.

1.4 EXERCISES

Compute the products in Exercises 1–4 using (a) the definition, as in Example 1, and (b) the row–vector rule for computing $A\mathbf{x}$. If a product is undefined, explain why.

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

In Exercises 5–8, use the definition of $A\mathbf{x}$ to write the matrix equation as a vector equation, or vice versa.

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

In Exercises 9 and 10, write the system first as a vector equation and then as a matrix equation.

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

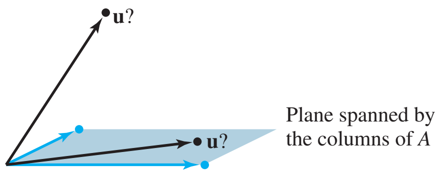
10. $\begin{aligned} 4x_1 - x_2 &= 8 \\ 5x_1 + 3x_2 &= 2 \\ 3x_1 - x_2 &= 1 \end{aligned}$

Given A and $\mathbf{b}$ in Exercises 11 and 12, write the augmented matrix for the linear system that corresponds to the matrix equation $A\mathbf{x} = \mathbf{b}$. Then solve the system and write the solution as a 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. Let $\mathbf{u} = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$ and $A = \begin{bmatrix} 3 & -5 \\ -2 & 6 \\ 1 & 1 \end{bmatrix}$. Is $\mathbf{u}$ in the plane in $\mathbb{R}^3$ spanned by the columns of A? (See the figure.) Why or why not?



Where is $\mathbf{u}$?

14. Let $\mathbf{u} = \begin{bmatrix} 4 \\ -1 \\ 4 \end{bmatrix}$ and $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & 1 & -1 \\ 1 & 2 & 0 \end{bmatrix}$. Is $\mathbf{u}$ in the subset of $\mathbb{R}^3$ spanned by the columns of A? Why or why not?

15. Let $A = \begin{bmatrix} 3 & -1 \\ -9 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. Show that the equation $A\mathbf{x} = \mathbf{b}$ does not have a solution for all possible $\mathbf{b}$, and describe the set of all $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ *does* have a solution.

16. Repeat the requests from Exercise 15 with

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

Exercises 17–20 refer to the matrices A and B below. Make appropriate calculations that justify your answers and mention an appropriate theorem.

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

17. How many rows of A contain a pivot position? Does the equation $A\mathbf{x} = \mathbf{b}$ have a solution for each $\mathbf{b}$ in $\mathbb{R}^4$?

18. Can every vector in $\mathbb{R}^4$ be written as a linear combination of the columns of the matrix B above? Do the columns of B span $\mathbb{R}^3$?

19. Can each vector in $\mathbb{R}^4$ be written as a linear combination of the columns of the matrix A above? Do the columns of A span $\mathbb{R}^4$?

20. Do the columns of B span $\mathbb{R}^4$? Does the equation $B\mathbf{x} = \mathbf{y}$ have a solution for each $\mathbf{y}$ in $\mathbb{R}^4$?

21. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$. Does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ span $\mathbb{R}^4$? Why or why not?

22. Let $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -3 \\ 9 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ -2 \\ -6 \end{bmatrix}$. Does $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ span $\mathbb{R}^3$? Why or why not?

In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. The equation $A\mathbf{x} = \mathbf{b}$ is referred to as a *vector equation.*

b. A vector $\mathbf{b}$ is a linear combination of the columns of a matrix A if and only if the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution.

c. The equation $A\mathbf{x} = \mathbf{b}$ is consistent if the augmented matrix $[\,A \quad \mathbf{b}\,]$ has a pivot position in every row.

d. The first entry in the product $A\mathbf{x}$ is a sum of products.

e. If the columns of an $m \times n$ matrix A span $\mathbb{R}^m$, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for each $\mathbf{b}$ in $\mathbb{R}^m$.

f. If A is an $m \times n$ matrix and if the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent for some $\mathbf{b}$ in $\mathbb{R}^m$, then A cannot have a pivot position in every row.

24. a. Every matrix equation $A\mathbf{x} = \mathbf{b}$ corresponds to a vector equation with the same solution set.

b. If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then $\mathbf{b}$ is in the set spanned by the columns of A.

c. Any linear combination of vectors can always be written in the form $A\mathbf{x}$ for a suitable matrix A and vector $\mathbf{x}$.

d. If the coefficient matrix A has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent.

e. The solution set of a linear system whose augmented matrix is $[\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3 \quad \mathbf{b}\,]$ is the same as the solution set of $A\mathbf{x} = \mathbf{b}$, if $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3\,]$.

f. If A is an $m \times n$ matrix whose columns do not span $\mathbb{R}^m$, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$.

25. Note that $\begin{bmatrix} 4 & -3 & 1 \\ 5 & -2 & 5 \\ -6 & 2 & -3 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix}$. Use this fact (and no row operations) to find scalars c_1, c_2, c_3 such that $\begin{bmatrix} -7 \\ -3 \\ 10 \end{bmatrix} = c_1\begin{bmatrix} 4 \\ 5 \\ -6 \end{bmatrix} + c_2\begin{bmatrix} -3 \\ -2 \\ 2 \end{bmatrix} + c_3\begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}$.

26. Let $\mathbf{u} = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$. It can be shown that $2\mathbf{u} - 3\mathbf{v} - \mathbf{w} = \mathbf{0}$. Use this fact (and no row operations) to find x_1 and x_2 that satisfy the equation $\begin{bmatrix} 7 & 3 \\ 2 & 1 \\ 5 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$.

27. Rewrite the (numerical) matrix equation below in symbolic form as a vector equation, using symbols $\mathbf{v}_1, \mathbf{v}_2, \ldots$ for the vectors and $c_1, c_2, \ldots$ for scalars. Define what each symbol represents, using the data given in the matrix equation.

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

28. Let $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{v}$ represent vectors in $\mathbb{R}^5$, and let x_1, x_2, and x_3 denote scalars. Write the following vector equation as a matrix equation. Identify any symbols you choose to use.

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

29. Construct a 3×3 matrix, not in echelon form, whose columns span $\mathbb{R}^3$. Show that the matrix you construct has the desired property.

30. Construct a 3×3 matrix, not in echelon form, whose columns do *not* span $\mathbb{R}^3$. Show that the matrix you construct has the desired property.

31. Let A be a 3×2 matrix. Explain why the equation $A\mathbf{x} = \mathbf{b}$ cannot be consistent for all $\mathbf{b}$ in $\mathbb{R}^3$. Generalize your argument to the case of an arbitrary A with more rows than columns.

32. Could a set of three vectors in $\mathbb{R}^4$ span all of $\mathbb{R}^4$? Explain. What about n vectors in $\mathbb{R}^m$ when n is less than m?

33. Suppose A is a 4×3 matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^4$ with the property that $A\mathbf{x} = \mathbf{b}$ has a unique solution. What can you say about the reduced echelon form of A? Justify your answer.

34. Let A be a 3×4 matrix, let $\mathbf{v}_1$ and $\mathbf{v}_2$ be vectors in $\mathbb{R}^3$, and let $\mathbf{w} = \mathbf{v}_1 + \mathbf{v}_2$. Suppose $\mathbf{v}_1 = A\mathbf{u}_1$ and $\mathbf{v}_2 = A\mathbf{u}_2$ for some vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ in $\mathbb{R}^4$. What fact allows you to conclude that the system $A\mathbf{x} = \mathbf{w}$ is consistent? (*Note:* $\mathbf{u}_1$ and $\mathbf{u}_2$ denote vectors, not scalar entries in vectors.)

35. Let A be a 5×3 matrix, let $\mathbf{y}$ be a vector in $\mathbb{R}^3$, and let $\mathbf{z}$ be a vector in $\mathbb{R}^5$. Suppose $A\mathbf{y} = \mathbf{z}$. What fact allows you to conclude that the system $A\mathbf{x} = 5\mathbf{z}$ is consistent?

36. Suppose A is a 4×4 matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^4$ with the property that $A\mathbf{x} = \mathbf{b}$ has a unique solution. Explain why the columns of A must span $\mathbb{R}^4$.

[**M**] In Exercises 37–40, determine if the columns of the matrix span $\mathbb{R}^4$.

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**] Find a column of the matrix in Exercise 39 that can be deleted and yet have the remaining matrix columns still span $\mathbb{R}^4$.

42. [**M**] Find a column of the matrix in Exercise 40 that can be deleted and yet have the remaining matrix columns still span $\mathbb{R}^4$. Can you delete more than one column?

SG Mastering Linear Algebra Concepts: Span 1–18

WEB

SOLUTIONS TO PRACTICE PROBLEMS

1. The matrix equation

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

is equivalent to the vector equation

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

which expresses $\mathbf{b}$ as a linear combination of the columns of A.

2.
$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 4 \\ -1 \end{bmatrix} + \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

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

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

1.5 SOLUTION SETS OF LINEAR SYSTEMS

Solution sets of linear systems are important objects of study in linear algebra. They will appear later in several different contexts. This section uses vector notation to give explicit and geometric descriptions of such solution sets.

Homogeneous Linear Systems

A system of linear equations is said to be **homogeneous** if it can be written in the form $A\mathbf{x} = \mathbf{0}$, where A is an $m \times n$ matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$. Such a system $A\mathbf{x} = \mathbf{0}$ *always* has at least one solution, namely, $\mathbf{x} = \mathbf{0}$ (the zero vector in $\mathbb{R}^n$). This zero solution is usually called the **trivial solution**. For a given equation $A\mathbf{x} = \mathbf{0}$, the important question is whether there exists a **nontrivial solution**, that is, a nonzero vector $\mathbf{x}$ that satisfies $A\mathbf{x} = \mathbf{0}$. The Existence and Uniqueness Theorem in Section 1.2 (Theorem 2) leads immediately to the following fact.

The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one free variable.

EXAMPLE 1 Determine if the following homogeneous system has a nontrivial solution. Then describe the solution set.

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

SOLUTION Let A be the matrix of coefficients of the system and row reduce the augmented matrix $[\, A \quad \mathbf{0}\,]$ to echelon form:

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

Since x_3 is a free variable, $A\mathbf{x} = \mathbf{0}$ has nontrivial solutions (one for each choice of x_3). To describe the solution set, continue the row reduction of $[\, A \quad \mathbf{0}\,]$ to *reduced* echelon form:

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

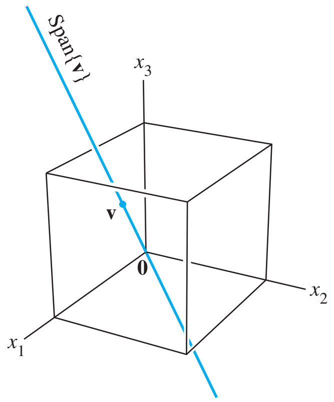


FIGURE 1

Solve for the basic variables x_1 and x_2 and obtain $x_1 = \frac{4}{3}x_3$, $x_2 = 0$, with x_3 free. As a vector, the general solution of $A\mathbf{x} = \mathbf{0}$ has the form

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

Here x_3 is factored out of the expression for the general solution vector. This shows that every solution of $A\mathbf{x} = \mathbf{0}$ in this case is a scalar multiple of $\mathbf{v}$. The trivial solution is obtained by choosing $x_3 = 0$. Geometrically, the solution set is a line through $\mathbf{0}$ in $\mathbb{R}^3$. See Fig. 1. ■

Notice that a nontrivial solution $\mathbf{x}$ can have some zero entries so long as not all of its entries are zero.

EXAMPLE 2 A single linear equation can be treated as a very simple system of equations. Describe all solutions of the homogeneous "system"

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

SOLUTION There is no need for matrix notation. Solve for the basic variable x_1 in terms of the free variables. The general solution is $x_1 = .3x_2 + .2x_3$, with x_2 and x_3 free. As a vector, the general solution is

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

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

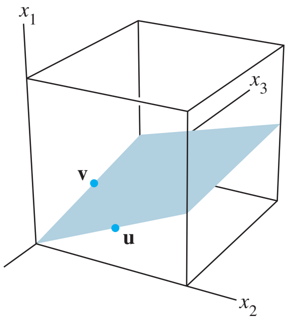


FIGURE 2

This calculation shows that every solution of (1) is a linear combination of the vectors $\mathbf{u}$ and $\mathbf{v}$, shown in (2). That is, the solution set is Span $\{\mathbf{u}, \mathbf{v}\}$. Since neither $\mathbf{u}$ nor $\mathbf{v}$ is a scalar multiple of the other, the solution set is a plane through the origin. See Fig. 2. ■

Examples 1 and 2, along with the exercises, illustrate the fact that the solution set of a homogeneous equation $A\mathbf{x} = \mathbf{0}$ can always be expressed explicitly as Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ for suitable vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$. If the only solution is the zero vector, then the solution set is Span $\{\mathbf{0}\}$. If the equation $A\mathbf{x} = \mathbf{0}$ has only one free variable, the solution set is a line through the origin, as in Fig. 1. A plane through the origin, as in Fig. 2, provides a good mental image for the solution set of $A\mathbf{x} = \mathbf{0}$ when there are two or more free variables. Note, however, that a similar figure can be used to visualize Span $\{\mathbf{u}, \mathbf{v}\}$ even when $\mathbf{u}$ and $\mathbf{v}$ do not arise as solutions of $A\mathbf{x} = \mathbf{0}$. See Fig. 11 in Section 1.3.

Parametric Vector Form

The original equation (1) for the plane in Example 2 is an *implicit* description of the plane. Solving this equation amounts to finding an *explicit* description of the plane as the set spanned by $\mathbf{u}$ and $\mathbf{v}$. Equation (2) is called a **parametric vector equation** of the plane. Sometimes such an equation is written as

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

to emphasize that the parameters vary over all real numbers. In Example 1, the equation $\mathbf{x} = x_3\mathbf{v}$ (with x_3 free), or $\mathbf{x} = t\mathbf{v}$ (with t in $\mathbb{R}$), is a parametric vector equation of a line. Whenever a solution set is described explicitly with vectors as in Examples 1 and 2, we say that the solution is in **parametric vector form**.

Solutions of Nonhomogeneous Systems

When a nonhomogeneous linear system has many solutions, the general solution can be written in parametric vector form as one vector plus an arbitrary linear combination of vectors that satisfy the corresponding homogeneous system.

EXAMPLE 3 Describe all solutions of $A\mathbf{x} = \mathbf{b}$, where

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

SOLUTION Here A is the matrix of coefficients from Example 1. Row operations on $[\, A \quad \mathbf{b} \,]$ produce

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

Thus $x_1 = -1 + \frac{4}{3}x_3$, $x_2 = 2$, and x_3 is free. As a vector, the general solution of $A\mathbf{x} = \mathbf{b}$ has the form

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

The equation $\mathbf{x} = \mathbf{p} + x_3\mathbf{v}$, or, writing t as a general parameter,

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

describes the solution set of $A\mathbf{x} = \mathbf{b}$ in parametric vector form. Recall from Example 1 that the solution set of $A\mathbf{x} = \mathbf{0}$ has the parametric vector equation

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

[with the same $\mathbf{v}$ that appears in (3)]. Thus the solutions of $A\mathbf{x} = \mathbf{b}$ are obtained by adding the vector $\mathbf{p}$ to the solutions of $A\mathbf{x} = \mathbf{0}$. The vector $\mathbf{p}$ itself is just one particular solution of $A\mathbf{x} = \mathbf{b}$ [corresponding to $t = 0$ in (3)]. ■

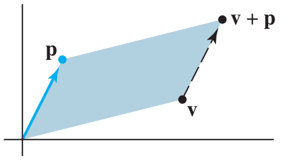


FIGURE 3
Adding $\mathbf{p}$ to $\mathbf{v}$ translates $\mathbf{v}$ to $\mathbf{v} + \mathbf{p}$.

To describe the solution set of $A\mathbf{x} = \mathbf{b}$ geometrically, we can think of vector addition as a *translation*. Given $\mathbf{v}$ and $\mathbf{p}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$, the effect of adding $\mathbf{p}$ to $\mathbf{v}$ is to *move* $\mathbf{v}$ in a direction parallel to the line through $\mathbf{p}$ and $\mathbf{0}$. We say that $\mathbf{v}$ is **translated by p** to $\mathbf{v} + \mathbf{p}$. See Fig. 3. If each point on a line L in $\mathbb{R}^2$ or $\mathbb{R}^3$ is translated by a vector $\mathbf{p}$, the result is a line parallel to L. See Fig. 4.

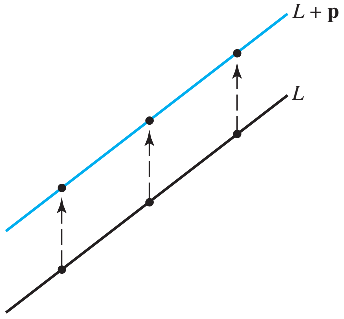


FIGURE 4
Translated line.

Suppose L is the line through $\mathbf{0}$ and $\mathbf{v}$, described by equation (4). Adding $\mathbf{p}$ to each point on L produces the translated line described by equation (3). Note that $\mathbf{p}$ is on the line in equation (3). We call (3) **the equation of the line through p parallel to v**. Thus *the solution set of* $A\mathbf{x} = \mathbf{b}$ *is a line through* $\mathbf{p}$ *parallel to the solution set of* $A\mathbf{x} = \mathbf{0}$. Figure 5 illustrates this case.

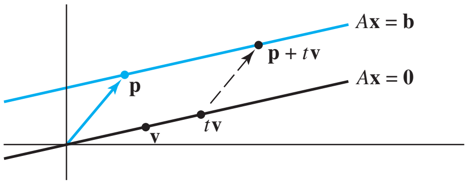


FIGURE 5 Parallel solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$.

The relation between the solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$ shown in Fig. 5 generalizes to any *consistent* equation $A\mathbf{x} = \mathbf{b}$, although the solution set will be larger than a line when there are several free variables. The following theorem gives the precise statement. See Exercise 25 for a proof.

THEOREM 6

Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given $\mathbf{b}$, and let $\mathbf{p}$ be a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$.

Theorem 6 says that if $A\mathbf{x} = \mathbf{b}$ has a solution, then the solution set is obtained by translating the solution set of $A\mathbf{x} = \mathbf{0}$, using any particular solution $\mathbf{p}$ of $A\mathbf{x} = \mathbf{b}$ for the translation. Figure 6 illustrates the case in which there are two free variables. Even when $n > 3$, our mental image of the solution set of a consistent system $A\mathbf{x} = \mathbf{b}$ (with $\mathbf{b} \neq \mathbf{0}$) is either a single nonzero point or a line or plane not passing through the origin.

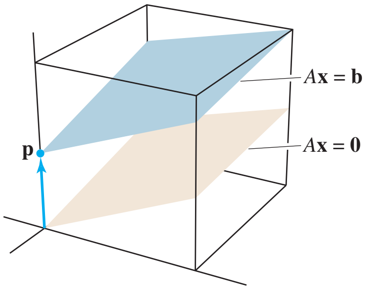


FIGURE 6 Parallel solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$.

Warning: Theorem 6 and Fig. 6 apply only to an equation $A\mathbf{x} = \mathbf{b}$ that has at least one nonzero solution $\mathbf{p}$. When $A\mathbf{x} = \mathbf{b}$ has no solution, the solution set is empty.

The following algorithm outlines the calculations shown in Examples 1, 2, and 3.

WRITING A SOLUTION SET (OF A CONSISTENT SYSTEM) IN PARAMETRIC VECTOR FORM

1. Row reduce the augmented matrix to reduced echelon form.
2. Express each basic variable in terms of any free variables appearing in an equation.
3. Write a typical solution $\mathbf{x}$ as a vector whose entries depend on the free variables, if any.
4. Decompose $\mathbf{x}$ into a linear combination of vectors (with numeric entries) using the free variables as parameters.

PRACTICE PROBLEMS

1. Each of the following equations determines a plane in $\mathbb{R}^3$. Do the two planes intersect? If so, describe their intersection.

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

2. Write the general solution of $10x_1 - 3x_2 - 2x_3 = 7$ in parametric vector form, and relate the solution set to the one found in Example 2.

1.5 EXERCISES

In Exercises 1–4, determine if the system has a nontrivial solution. Try to use as few row operations as possible.

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

In Exercises 5 and 6, follow the method of Examples 1 and 2 to write the solution set of the given homogeneous system in parametric vector form.

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

In Exercises 7–12, describe all solutions of $A\mathbf{x} = \mathbf{0}$ in parametric vector form, where A is row equivalent to the given matrix.

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. Suppose the solution set of a certain system of linear equations can be described as $x_1 = 5 + 4x_3$, $x_2 = -2 - 7x_3$, with x_3 free. Use vectors to describe this set as a line in $\mathbb{R}^3$.

14. Suppose the solution set of a certain system of linear equations can be described as $x_1 = 5x_4$, $x_2 = 3 - 2x_4$, $x_3 = 2 + 5x_4$, with x_4 free. Use vectors to describe this set as a "line" in $\mathbb{R}^4$.

15. Describe and compare the solution sets of $x_1 + 5x_2 - 3x_3 = 0$ and $x_1 + 5x_2 - 3x_3 = -2$.

16. Describe and compare the solution sets of $x_1 - 2x_2 + 3x_3 = 0$ and $x_1 - 2x_2 + 3x_3 = 4$.

17. Follow the method of Example 3 to describe the solutions of the following system in parametric vector form. Also, give a geometric description of the solution set and compare it to that in Exercise 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. As in Exercise 17, describe the solutions of the following system in parametric vector form, and provide a geometric comparison with the solution set in Exercise 6.

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

In Exercises 19 and 20, find the parametric equation of the line through **a** parallel to **b**.

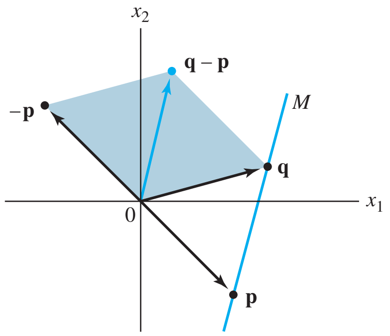
19. $\mathbf{a} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

20. $\mathbf{a} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -7 \\ 6 \end{bmatrix}$

In Exercises 21 and 22, find a parametric equation of the line M through **p** and **q**. [*Hint*: M is parallel to the vector $\mathbf{q} - \mathbf{p}$. See the figure below.]

21. $\mathbf{p} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

22. $\mathbf{p} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$



In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. A homogeneous equation is always consistent.
    b. The equation $A\mathbf{x} = \mathbf{0}$ gives an explicit description of its solution set.
    c. The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution if and only if the equation has at least one free variable.
    d. The equation $\mathbf{x} = \mathbf{p} + t\mathbf{v}$ describes a line through **v** parallel to **p**.
    e. The solution set of $A\mathbf{x} = \mathbf{b}$ is the set of all vectors of the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is any solution of the equation $A\mathbf{x} = \mathbf{0}$.

24. a. A homogeneous system of equations can be inconsistent.
    b. If **x** is a nontrivial solution of $A\mathbf{x} = \mathbf{0}$, then every entry in **x** is nonzero.
    c. The effect of adding **p** to a vector is to move the vector in a direction parallel to **p**.
    d. The equation $A\mathbf{x} = \mathbf{b}$ is homogeneous if the zero vector is a solution.

e. If $A\mathbf{x} = \mathbf{b}$ is consistent, then the solution set of $A\mathbf{x} = \mathbf{b}$ is obtained by translating the solution set of $A\mathbf{x} = \mathbf{0}$.

25. Prove Theorem 6:

a. Suppose $\mathbf{p}$ is a solution of $A\mathbf{x} = \mathbf{b}$, so that $A\mathbf{p} = \mathbf{b}$. Let $\mathbf{v}_h$ be any solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$, and let $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$. Show that $\mathbf{w}$ is a solution of $A\mathbf{x} = \mathbf{b}$.

b. Let $\mathbf{w}$ be any solution of $A\mathbf{x} = \mathbf{b}$, and define $\mathbf{v}_h = \mathbf{w} - \mathbf{p}$. Show that $\mathbf{v}_h$ is a solution of $A\mathbf{x} = \mathbf{0}$. This shows that every solution of $A\mathbf{x} = \mathbf{b}$ has the form $\mathbf{w} = \mathbf{p} + \mathbf{v}_h$, with $\mathbf{p}$ a particular solution of $A\mathbf{x} = \mathbf{b}$ and $\mathbf{v}_h$ a solution of $A\mathbf{x} = \mathbf{0}$.

26. Suppose A is the 3×3 *zero* matrix (with all zero entries). Describe the solution set of the equation $A\mathbf{x} = \mathbf{0}$.

27. Suppose $A\mathbf{x} = \mathbf{b}$ has a solution. Explain why the solution is unique precisely when $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

In Exercises 28–31, (a) does the equation $A\mathbf{x} = \mathbf{0}$ have a nontrivial solution and (b) does the equation $A\mathbf{x} = \mathbf{b}$ have at least one solution for every possible $\mathbf{b}$?

28. A is a 3×3 matrix with three pivot positions.

29. A is a 4×4 matrix with three pivot positions.

30. A is a 2×5 matrix with two pivot positions.

31. A is a 3×2 matrix with two pivot positions.

32. If $\mathbf{b} \neq \mathbf{0}$, can the solution set of $A\mathbf{x} = \mathbf{b}$ be a plane through the origin? Explain.

33. Construct a 3×3 nonzero matrix A such that the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is a solution of $A\mathbf{x} = \mathbf{0}$.

34. Construct a 3×3 nonzero matrix A such that the vector $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is a solution of $A\mathbf{x} = \mathbf{0}$.

35. Given $A = \begin{bmatrix} -1 & -3 \\ 7 & 21 \\ -2 & -6 \end{bmatrix}$, find one nontrivial solution of $A\mathbf{x} = \mathbf{0}$ by inspection. [*Hint*: Think of the equation $A\mathbf{x} = \mathbf{0}$ written as a vector equation.]

36. Given $A = \begin{bmatrix} 3 & -2 \\ -6 & 4 \\ 12 & -8 \end{bmatrix}$, find one nontrivial solution of $A\mathbf{x} = \mathbf{0}$ by inspection.

37. Construct a 2×2 matrix A such that the solution set of the equation $A\mathbf{x} = \mathbf{0}$ is the line in $\mathbb{R}^2$ through $(4, 1)$ and the origin. Then, find a vector $\mathbf{b}$ in $\mathbb{R}^2$ such that the solution set of $A\mathbf{x} = \mathbf{b}$ is *not* a line in $\mathbb{R}^2$ parallel to the solution set of $A\mathbf{x} = \mathbf{0}$. Why does this *not* contradict Theorem 6?

38. Let A be an $m \times n$ matrix and let $\mathbf{w}$ be a vector in $\mathbb{R}^n$ that satisfies the equation $A\mathbf{x} = \mathbf{0}$. Show that for any scalar c, the vector $c\mathbf{w}$ also satisfies $A\mathbf{x} = \mathbf{0}$. [That is, show that $A(c\mathbf{w}) = \mathbf{0}$.]

39. Let A be an $m \times n$ matrix, and let $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ with the property that $A\mathbf{v} = \mathbf{0}$ and $A\mathbf{w} = \mathbf{0}$. Explain why $A(\mathbf{v} + \mathbf{w})$ must be the zero vector. Then explain why $A(c\mathbf{v} + d\mathbf{w}) = \mathbf{0}$ for each pair of scalars c and d.

40. Suppose A is a 3×3 matrix and $\mathbf{b}$ is a vector in $\mathbb{R}^3$ such that the equation $A\mathbf{x} = \mathbf{b}$ does *not* have a solution. Does there exist a vector $\mathbf{y}$ in $\mathbb{R}^3$ such that the equation $A\mathbf{x} = \mathbf{y}$ has a unique solution? Discuss.

SOLUTIONS TO PRACTICE PROBLEMS

1. Row reduce the augmented matrix:

$$\begin{bmatrix} 1 & 4 & -5 & 0 \\ 2 & -1 & 8 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & -5 & 0 \\ 0 & -9 & 18 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & -2 & -1 \end{bmatrix}$$

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

Thus $x_1 = 4 - 3x_3$, $x_2 = -1 + 2x_3$, with x_3 free. The general solution in parametric vector form is

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

The intersection of the two planes is the line through $\mathbf{p}$ in the direction of $\mathbf{v}$.

2. The augmented matrix $\begin{bmatrix} 10 & -3 & -2 & 7 \end{bmatrix}$ is row equivalent to $\begin{bmatrix} 1 & -.3 & -.2 & .7 \end{bmatrix}$, and the general solution is $x_1 = .7 + .3x_2 + .2x_3$, with x_2 and x_3 free. That is,

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

$$= \quad \mathbf{p} \quad + \quad x_2\mathbf{u} \quad + \quad x_3\mathbf{v}$$

The solution set of the nonhomogeneous equation $A\mathbf{x} = \mathbf{b}$ is the translated plane $\mathbf{p} + \text{Span}\,\{\mathbf{u}, \mathbf{v}\}$, which passes through $\mathbf{p}$ and is parallel to the solution set of the homogeneous equation in Example 2.

1.6 APPLICATIONS OF LINEAR SYSTEMS

You might expect that a real-life problem involving linear algebra would have only one solution, or perhaps no solution. The purpose of this section is to show how linear systems with many solutions can arise naturally. The applications here come from economics, chemistry, and network flow.

A Homogeneous System in Economics

WEB

The system of 500 equations in 500 variables, mentioned in this chapter's introduction, is now known as a Leontief "input–output" (or "production") model.[1] Section 2.6 will examine this model in more detail, when more theory and better notation are available. For now, we look at a simpler "exchange model," also due to Leontief.

Suppose a nation's economy is divided into many sectors, such as various manufacturing, communication, entertainment, and service industries. Suppose that for each sector we know its total output for one year and we know exactly how this output is divided or "exchanged" among the other sectors of the economy. Let the total dollar value of a sector's output be called the **price** of that output. Leontief proved the following result.

> There exist *equilibrium prices* that can be assigned to the total outputs of the various sectors in such a way that the income of each sector exactly balances its expenses.

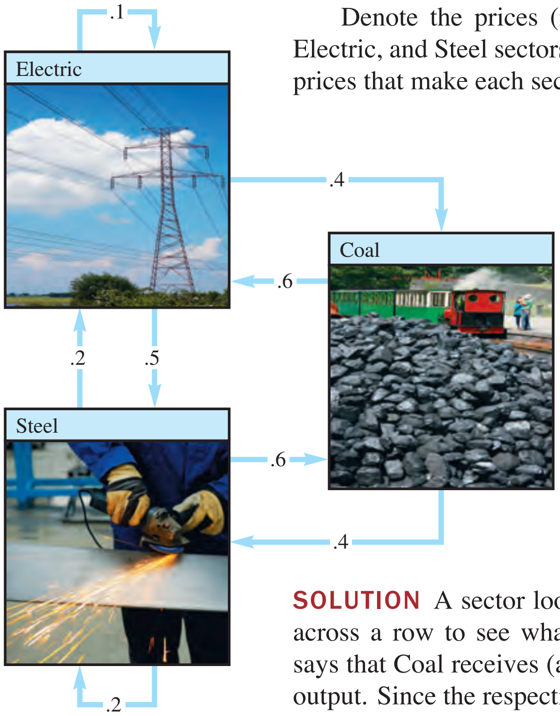
The following example shows how to find the equilibrium prices.

EXAMPLE 1 Suppose an economy consists of the Coal, Electric (power), and Steel sectors, and the output of each sector is distributed among the various sectors as shown in Table 1 on page 50, where the entries in a column represent the fractional parts of a sector's total output.

The second column of Table 1, for instance, says that the total output of the Electric sector is divided as follows: 40% to Coal, 50% to Steel, and the remaining 10% to Electric. (Electric treats this 10% as an expense it incurs in order to operate its business.) Since all output must be taken into account, the decimal fractions in each column must sum to 1.

---

[1] See Wassily W. Leontief, "Input–Output Economics," *Scientific American*, October 1951, pp. 15–21.



Denote the prices (i.e., dollar values) of the total annual outputs of the Coal, Electric, and Steel sectors by p_C, p_E, and p_S, respectively. If possible, find equilibrium prices that make each sector's income match its expenditures.

TABLE 1 A Simple Economy

Distribution of Output from:			
Coal	**Electric**	**Steel**	**Purchased by:**
.0	.4	.6	Coal
.6	.1	.2	Electric
.4	.5	.2	Steel

SOLUTION A sector looks down a column to see where its output goes, and it looks across a row to see what it needs as inputs. For instance, the first row of Table 1 says that Coal receives (and pays for) 40% of the Electric output and 60% of the Steel output. Since the respective values of the total outputs are p_E and p_S, Coal must spend $.4p_E$ dollars for its share of Electric's output and $.6p_S$ for its share of Steel's output. Thus Coal's total expenses are $.4p_E + .6p_S$. To make Coal's income, p_C, equal to its expenses, we want

$$p_C = .4p_E + .6p_S \tag{1}$$

The second row of the exchange table shows that the Electric sector spends $.6p_C$ for coal, $.1p_E$ for electricity, and $.2p_S$ for steel. Hence the income/expense requirement for Electric is

$$p_E = .6p_C + .1p_E + .2p_S \tag{2}$$

Finally, the third row of the exchange table leads to the final requirement:

$$p_S = .4p_C + .5p_E + .2p_S \tag{3}$$

To solve the system of equations (1), (2), and (3), move all the unknowns to the left sides of the equations and combine like terms. [For instance, on the left side of (2), write $p_E - .1p_E$ as $.9p_E$.]

$$\begin{aligned} p_C - .4p_E - .6p_S &= 0 \\ -.6p_C + .9p_E - .2p_S &= 0 \\ -.4p_C - .5p_E + .8p_S &= 0 \end{aligned}$$

Row reduction is next. For simplicity here, decimals are rounded to two places.

$$\begin{bmatrix} 1 & -.4 & -.6 & 0 \\ -.6 & .9 & -.2 & 0 \\ -.4 & -.5 & .8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & -.66 & .56 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & .66 & -.56 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -.4 & -.6 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -.94 & 0 \\ 0 & 1 & -.85 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The general solution is $p_C = .94p_S$, $p_E = .85p_S$, and p_S is free. The equilibrium price vector for the economy has the form

$$\mathbf{p} = \begin{bmatrix} p_C \\ p_E \\ p_S \end{bmatrix} = \begin{bmatrix} .94p_S \\ .85p_S \\ p_S \end{bmatrix} = p_S \begin{bmatrix} .94 \\ .85 \\ 1 \end{bmatrix}$$

Any (nonnegative) choice for p_S results in a choice of equilibrium prices. For instance, if we take p_S to be 100 (or \$100 million), then $p_C = 94$ and $p_E = 85$. The incomes and expenditures of each sector will be equal if the output of Coal is priced at \$94 million, that of Electric at \$85 million, and that of Steel at \$100 million. ■

Balancing Chemical Equations

Chemical equations describe the quantities of substances consumed and produced by chemical reactions. For instance, when propane gas burns, the propane (C_3H_8) combines with oxygen (O_2) to form carbon dioxide (CO_2) and water (H_2O), according to an equation of the form

$$(x_1)\mathrm{C_3H_8} + (x_2)\mathrm{O_2} \rightarrow (x_3)\mathrm{CO_2} + (x_4)\mathrm{H_2O} \tag{4}$$

To "balance" this equation, a chemist must find whole numbers $x_1, \dots, x_4$ such that the total numbers of carbon (C), hydrogen (H), and oxygen (O) atoms on the left match the corresponding numbers of atoms on the right (because atoms are neither destroyed nor created in the reaction).

A systematic method for balancing chemical equations is to set up a vector equation that describes the numbers of atoms of each type present in a reaction. Since equation (4) involves three types of atoms (carbon, hydrogen, and oxygen), construct a vector in $\mathbb{R}^3$ for each reactant and product in (4) that lists the numbers of "atoms per molecule," as follows:

$$\mathrm{C_3H_8}: \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix}, \quad \mathrm{O_2}: \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \quad \mathrm{CO_2}: \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathrm{H_2O}: \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \begin{matrix} \leftarrow \text{Carbon} \\ \leftarrow \text{Hydrogen} \\ \leftarrow \text{Oxygen} \end{matrix}$$

To balance equation (4), the coefficients $x_1, \dots, x_4$ must satisfy

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$$

To solve, move all the terms to the left (changing the signs in the third and fourth vectors):

$$x_1 \begin{bmatrix} 3 \\ 8 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Row reduction of the augmented matrix for this equation leads to the general solution

$$x_1 = \tfrac{1}{4}x_4, \; x_2 = \tfrac{5}{4}x_4, \; x_3 = \tfrac{3}{4}x_4, \; \text{ with } x_4 \text{ free}$$

Since the coefficients in a chemical equation must be integers, take $x_4 = 4$, in which case $x_1 = 1$, $x_2 = 5$, and $x_3 = 3$. The balanced equation is

$$\mathrm{C_3H_8} + 5\mathrm{O_2} \rightarrow 3\mathrm{CO_2} + 4\mathrm{H_2O}$$

The equation would also be balanced if, for example, each coefficient were doubled. For most purposes, however, chemists prefer to use a balanced equation whose coefficients are the smallest possible whole numbers.

Network Flow

WEB

Systems of linear equations arise naturally when scientists, engineers, or economists study the flow of some quantity through a network. For instance, urban planners and traffic engineers monitor the pattern of traffic flow in a grid of city streets. Electrical engineers calculate current flow through electrical circuits. And economists analyze the distribution of products from manufacturers to consumers through a network of wholesalers and retailers. For many networks, the systems of equations involve hundreds or even thousands of variables and equations.

A *network* consists of a set of points called *junctions*, or *nodes*, with lines or arcs called *branches* connecting some or all of the junctions. The direction of flow in each branch is indicated, and the flow amount (or rate) is either shown or is denoted by a variable.

The basic assumption of network flow is that the total flow into the network equals the total flow out of the network and that the total flow into a junction equals the total flow out of the junction. For example, Fig. 1 shows 30 units flowing into a junction through one branch, with x_1 and x_2 denoting the flows out of the junction through other branches. Since the flow is "conserved" at each junction, we must have $x_1 + x_2 = 30$. In a similar fashion, the flow at each junction is described by a linear equation. The problem of network analysis is to determine the flow in each branch when partial information (such as the flow into and out of the network) is known.

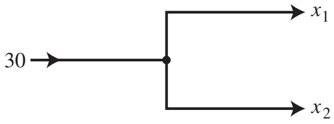


FIGURE 1
A junction, or node.

EXAMPLE 2 The network in Fig. 2 shows the traffic flow (in vehicles per hour) over several one-way streets in downtown Baltimore during a typical early afternoon. Determine the general flow pattern for the network.

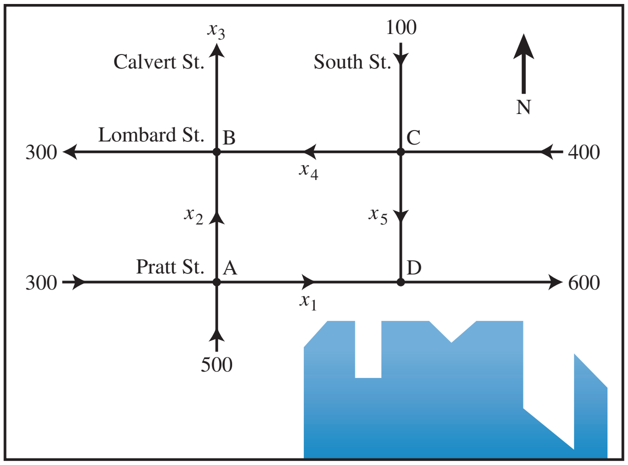


FIGURE 2 Baltimore streets.

SOLUTION Write equations that describe the flow, and then find the general solution of the system. Label the street intersections (junctions) and the unknown flows in the branches, as shown in Fig. 2. At each intersection, set the flow in equal to the flow out.

Intersection	Flow in		Flow out
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

Also, the total flow into the network $(500 + 300 + 100 + 400)$ equals the total flow out of the network $(300 + x_3 + 600)$, which simplifies to $x_3 = 400$. Combine this equation with a rearrangement of the first four equations to obtain the following system of equations:

$$\begin{aligned} x_1 + x_2 \qquad\qquad\qquad &= 800 \\ x_2 - x_3 + x_4 \qquad &= 300 \\ x_4 + x_5 &= 500 \\ x_1 \qquad\qquad\qquad + x_5 &= 600 \\ x_3 \qquad\qquad &= 400 \end{aligned}$$

Row reduction of the associated augmented matrix leads to

$$\begin{aligned} x_1 \qquad\qquad + x_5 &= 600 \\ x_2 \qquad - x_5 &= 200 \\ x_3 \qquad\quad &= 400 \\ x_4 + x_5 &= 500 \end{aligned}$$

The general flow pattern for the network is described by

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ is free} \end{cases}$$

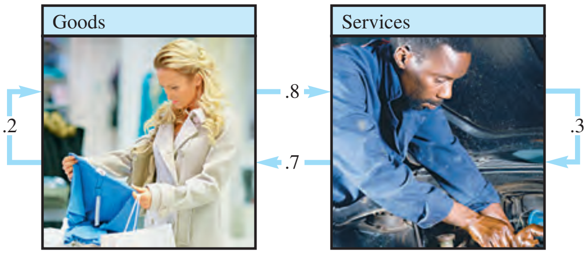
A negative flow in a network branch corresponds to flow in the direction opposite to that shown on the model. Since the streets in this problem are one-way, none of the variables here can be negative. This fact leads to certain limitations on the possible values of the variables. For instance, $x_5 \leq 500$ because x_4 cannot be negative. Other constraints on the variables are considered in Practice Problem 2. ■

PRACTICE PROBLEMS

1. Suppose an economy has three sectors: Agriculture, Mining, and Manufacturing. Agriculture sells 5% of its output to Mining and 30% to Manufacturing, and retains the rest. Mining sells 20% of its output to Agriculture and 70% to Manufacturing, and retains the rest. Manufacturing sells 20% of its output to Agriculture and 30% to Mining, and retains the rest. Determine the exchange table for this economy, where the columns describe how the output of each sector is exchanged among the three sectors.
2. Consider the network flow studied in Example 2. Determine the possible range of values of x_1 and x_2. [*Hint:* The example showed that $x_5 \leq 500$. What does this imply about x_1 and x_2? Also, use the fact that $x_5 \geq 0$.]

1.6 EXERCISES

1. Suppose an economy has only two sectors: Goods and Services. Each year, Goods sells 80% of its output to Services and keeps the rest, while Services sells 70% of its output to Goods and retains the rest. Find equilibrium prices for the annual outputs of the Goods and Services sectors that make each sector's income match its expenditures.



2. Find another set of equilibrium prices for the economy in Example 1. Suppose the same economy used Japanese yen instead of dollars to measure the values of the various sectors' outputs. Would this change the problem in any way? Discuss.

3. Consider an economy with three sectors: Fuels and Power, Manufacturing, and Services. Fuels and Power sells 80% of its output to Manufacturing, 10% to Services, and retains the rest. Manufacturing sells 10% of its output to Fuels and Power, 80% to Services, and retains the rest. Services sells 20% to Fuels and Power, 40% to Manufacturing, and retains the rest.
   a. Construct the exchange table for this economy.
   b. Develop a system of equations that leads to prices at which each sector's income matches its expenses. Then write the augmented matrix that can be row reduced to find these prices.
   c. [**M**] Find a set of equilibrium prices when the price for the Services output is 100 units.

4. Suppose an economy has four sectors: Mining, Lumber, Energy, and Transportation. Mining sells 10% of its output to Lumber, 60% to Energy, and retains the rest. Lumber sells 15% of its output to Mining, 50% to Energy, 20% to Transportation, and retains the rest. Energy sells 20% of its output to Mining, 15% to Lumber, 20% to Transportation, and retains the rest. Transportation sells 20% of its output to Mining, 10% to Lumber, 50% to Energy, and retains the rest.
   a. Construct the exchange table for this economy.
   b. [**M**] Find a set of equilibrium prices for the economy.

5. An economy has four sectors: Agriculture, Manufacturing, Services, and Transportation. Agriculture sells 20% of its output to Manufacturing, 30% to Services, 30% to Transportation, and retains the rest. Manufacturing sells 35% of its output to Agriculture, 35% to Services, 20% to Transportation, and retains the rest. Services sells 10% of its output to Agriculture, 20% to Manufacturing, 20% to Transportation, and retains the rest. Transportation sells 20% of its output to Agriculture, 30% to Manufacturing, 20% to Services, and retains the rest.
   a. Construct the exchange table for this economy.
   b. [**M**] Find a set of equilibrium prices for the economy if the value of Transportation is \$10.00 per unit.
   c. The Services sector launches a successful "eat farm fresh" campaign, and increases its share of the output from the Agricultural sector to 40%, whereas the share of Agricultural production going to Manufacturing falls to 10%. Construct the exchange table for this new economy.
   d. [**M**] Find a set of equilibrium prices for this new economy if the value of Transportation is still \$10.00 per unit. What effect has the "eat farm fresh" campaign had on the equilibrium prices for the sectors in this economy?

Balance the chemical equations in Exercises 6–11 using the vector equation approach discussed in this section.

6. Aluminum oxide and carbon react to create elemental aluminum and carbon dioxide:

   $Al_2O_3 + C \rightarrow Al + CO_2$

   [For each compound, construct a vector that lists the numbers of atoms of aluminum, oxygen, and carbon.]

7. Alka-Seltzer contains sodium bicarbonate ($NaHCO_3$) and citric acid ($H_3C_6H_5O_7$). When a tablet is dissolved in water, the following reaction produces sodium citrate, water, and carbon dioxide (gas):

   $NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + H_2O + CO_2$

8. Limestone, $CaCO_3$, neutralizes the acid, H_3O, in acid rain by the following unbalanced equation:

   $H_3O + CaCO_3 \rightarrow H_2O + Ca + CO_2$

9. Boron sulfide reacts violently with water to form boric acid and hydrogen sulfide gas (the smell of rotten eggs). The unbalanced equation is

   $B_2S_3 + H_2O \rightarrow H_3BO_3 + H_2S$

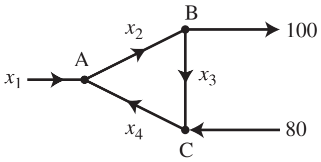
10. [**M**] If possible, use exact arithmetic or a rational format for calculations in balancing the following chemical reaction:

    $PbN_6 + CrMn_2O_8 \rightarrow Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$

11. [**M**] The chemical reaction below can be used in some industrial processes, such as the production of arsene (AsH_3). Use exact arithmetic or a rational format for calculations to balance this equation.

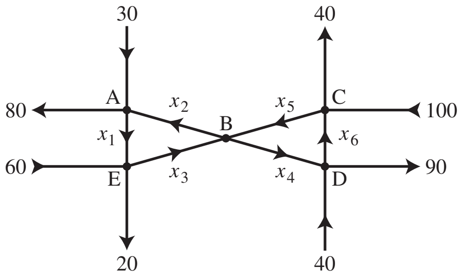
    $$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \\ \rightarrow HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

12. Find the general flow pattern of the network shown in the figure. Assuming that the flows are all nonnegative, what is the smallest possible value for x_4?



13. a. Find the general flow pattern of the network shown in the figure.

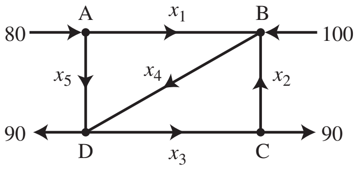
b. Assuming that the flow must be in the directions indicated, find the minimum flows in the branches denoted by x_2, x_3, x_4, and x_5.



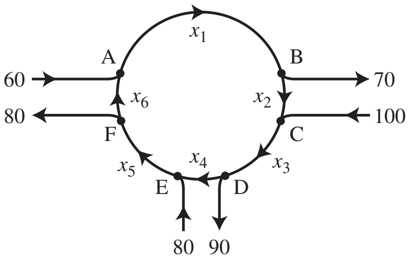
14. a. Find the general traffic pattern of the freeway network shown in the figure. (Flow rates are in cars/minute.)

b. Describe the general traffic pattern when the road whose flow is x_5 is closed.

c. When $x_5 = 0$, what is the minimum value of x_4?



15. Intersections in England are often constructed as one-way "roundabouts," such as the one shown in the figure. Assume that traffic must travel in the directions shown. Find the general solution of the network flow. Find the smallest possible value for x_6.



SOLUTIONS TO PRACTICE PROBLEMS

1. Write the percentages as decimals. Since all output must be taken into account, each column must sum to 1. This fact helps to fill in any missing entries.

Distribution of Output from:			
Agriculture	**Mining**	**Manufacturing**	**Purchased by:**
.65	.20	.20	Agriculture
.05	.10	.30	Mining
.30	.70	.50	Manufacturing

2. Since $x_5 \le 500$, the equations D and A for x_1 and x_2 imply that $x_1 \ge 100$ and $x_2 \le 700$. The fact that $x_5 \ge 0$ implies that $x_1 \le 600$ and $x_2 \ge 200$. So, $100 \le x_1 \le 600$, and $200 \le x_2 \le 700$.

1.7 LINEAR INDEPENDENCE

The homogeneous equations in Section 1.5 can be studied from a different perspective by writing them as vector equations. In this way, the focus shifts from the unknown solutions of $A\mathbf{x} = \mathbf{0}$ to the vectors that appear in the vector equations.

For instance, consider the equation

$$x_1\begin{bmatrix}1\\2\\3\end{bmatrix}+x_2\begin{bmatrix}4\\5\\6\end{bmatrix}+x_3\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix} \tag{1}$$

This equation has a trivial solution, of course, where $x_1 = x_2 = x_3 = 0$. As in Section 1.5, the main issue is whether the trivial solution is the *only one*.

DEFINITION

An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{2}$$

Equation (2) is called a **linear dependence relation** among $\mathbf{v}_1, \dots, \mathbf{v}_p$ when the weights are not all zero. An indexed set is linearly dependent if and only if it is not linearly independent. For brevity, we may say that $\mathbf{v}_1, \dots, \mathbf{v}_p$ are linearly dependent when we mean that $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a linearly dependent set. We use analogous terminology for linearly independent sets.

EXAMPLE 1 Let $\mathbf{v}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix}2\\1\\0\end{bmatrix}$.

a. Determine if the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

b. If possible, find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

SOLUTION

a. We must determine if there is a nontrivial solution of equation (1) above. Row operations on the associated augmented matrix show that

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

Clearly, x_1 and x_2 are basic variables, and x_3 is free. Each nonzero value of x_3 determines a nontrivial solution of (1). Hence $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent (and not linearly independent).

b. To find a linear dependence relation among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, completely row reduce the augmented matrix and write the new system:

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

Thus $x_1 = 2x_3$, $x_2 = -x_3$, and x_3 is free. Choose any nonzero value for x_3—say, $x_3 = 5$. Then $x_1 = 10$ and $x_2 = -5$. Substitute these values into equation (1) and obtain

$$10\mathbf{v}_1 - 5\mathbf{v}_2 + 5\mathbf{v}_3 = \mathbf{0}$$

This is one (out of infinitely many) possible linear dependence relations among $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. ■

Linear Independence of Matrix Columns

Suppose that we begin with a matrix $A = [\,\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n\,]$ instead of a set of vectors. The matrix equation $A\mathbf{x} = \mathbf{0}$ can be written as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0}$$

Each linear dependence relation among the columns of A corresponds to a nontrivial solution of $A\mathbf{x} = \mathbf{0}$. Thus we have the following important fact.

> The columns of a matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has *only* the trivial solution. (3)

EXAMPLE 2 Determine if the columns of the matrix $A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$ are linearly independent.

SOLUTION To study $A\mathbf{x} = \mathbf{0}$, row reduce the augmented matrix:

$$\begin{bmatrix} 0 & 1 & 4 & 0 \\ 1 & 2 & -1 & 0 \\ 5 & 8 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & -2 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 13 & 0 \end{bmatrix}$$

At this point, it is clear that there are three basic variables and no free variables. So the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, and the columns of A are linearly independent. ■

Sets of One or Two Vectors

A set containing only one vector—say, $\mathbf{v}$—is linearly independent if and only if $\mathbf{v}$ is not the zero vector. This is because the vector equation $x_1\mathbf{v} = \mathbf{0}$ has only the trivial solution when $\mathbf{v} \neq \mathbf{0}$. The zero vector is linearly dependent because $x_1\mathbf{0} = \mathbf{0}$ has many nontrivial solutions.

The next example will explain the nature of a linearly dependent set of two vectors.

EXAMPLE 3 Determine if the following sets of vectors are linearly independent.

a. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$ b. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$

SOLUTION

a. Notice that $\mathbf{v}_2$ is a multiple of $\mathbf{v}_1$, namely, $\mathbf{v}_2 = 2\mathbf{v}_1$. Hence $-2\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{0}$, which shows that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent.

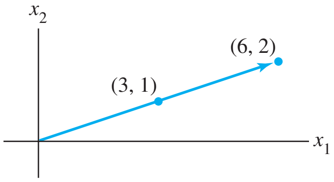
b. The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are certainly *not* multiples of one another. Could they be linearly dependent? Suppose c and d satisfy

$$c\mathbf{v}_1 + d\mathbf{v}_2 = \mathbf{0}$$

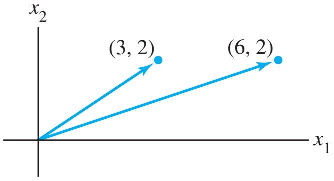
If $c \neq 0$, then we can solve for $\mathbf{v}_1$ in terms of $\mathbf{v}_2$, namely, $\mathbf{v}_1 = (-d/c)\mathbf{v}_2$. This result is impossible because $\mathbf{v}_1$ is *not* a multiple of $\mathbf{v}_2$. So c must be zero. Similarly, d must also be zero. Thus $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set. ■

The arguments in Example 3 show that you can always decide *by inspection* when a set of two vectors is linearly dependent. Row operations are unnecessary. Simply check whether at least one of the vectors is a scalar times the other. (The test applies only to sets of *two* vectors.)

A set of two vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent if at least one of the vectors is a multiple of the other. The set is linearly independent if and only if neither of the vectors is a multiple of the other.



Linearly dependent



Linearly independent

FIGURE 1

In geometric terms, two vectors are linearly dependent if and only if they lie on the same line through the origin. Figure 1 shows the vectors from Example 3.

Sets of Two or More Vectors

The proof of the next theorem is similar to the solution of Example 3. Details are given at the end of this section.

THEOREM 7

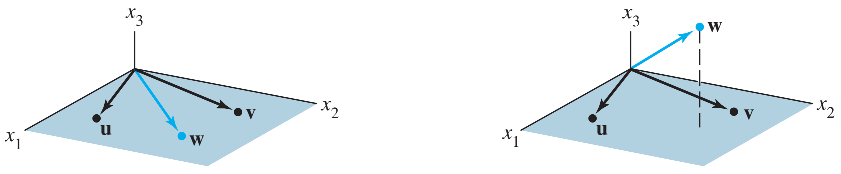
Characterization of Linearly Dependent Sets

An indexed set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others. In fact, if S is linearly dependent and $\mathbf{v}_1 \neq \mathbf{0}$, then some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.

Warning: Theorem 7 does *not* say that *every* vector in a linearly dependent set is a linear combination of the preceding vectors. A vector in a linearly dependent set may fail to be a linear combination of the other vectors. See Practice Problem 3.

EXAMPLE 4 Let $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Describe the set spanned by $\mathbf{u}$ and $\mathbf{v}$, and explain why a vector $\mathbf{w}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$ if and only if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

SOLUTION The vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent because neither vector is a multiple of the other, and so they span a plane in $\mathbb{R}^3$. (See Section 1.3.) In fact, Span $\{\mathbf{u}, \mathbf{v}\}$ is the x_1x_2-plane (with $x_3 = 0$). If $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$, then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent, by Theorem 7. Conversely, suppose that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent. By Theorem 7, some vector in $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a linear combination of the preceding vectors (since $\mathbf{u} \neq \mathbf{0}$). That vector must be $\mathbf{w}$, since $\mathbf{v}$ is not a multiple of $\mathbf{u}$. So $\mathbf{w}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$. See Fig. 2. ■



Linearly dependent,
$\mathbf{w}$ in Span$\{\mathbf{u}, \mathbf{v}\}$

Linearly independent,
$\mathbf{w}$ *not* in Span$\{\mathbf{u}, \mathbf{v}\}$

FIGURE 2 Linear dependence in $\mathbb{R}^3$.

Example 4 generalizes to any set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ in $\mathbb{R}^3$ with $\mathbf{u}$ and $\mathbf{v}$ linearly independent. The set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ will be linearly dependent if and only if $\mathbf{w}$ is in the plane spanned by $\mathbf{u}$ and $\mathbf{v}$.

The next two theorems describe special cases in which the linear dependence of a set is automatic. Moreover, Theorem 8 will be a key result for work in later chapters.

THEOREM 8

If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

FIGURE 3

If $p > n$, the columns are linearly dependent.

PROOF Let $A = [\, \mathbf{v}_1 \ \cdots \ \mathbf{v}_p \,]$. Then A is $n \times p$, and the equation $A\mathbf{x} = \mathbf{0}$ corresponds to a system of n equations in p unknowns. If $p > n$, there are more variables than equations, so there must be a free variable. Hence $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, and the columns of A are linearly dependent. See Fig. 3 for a matrix version of this theorem. ■

Warning: Theorem 8 says nothing about the case in which the number of vectors in the set does *not* exceed the number of entries in each vector.

EXAMPLE 5 The vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ are linearly dependent by Theorem 8, because there are three vectors in the set and there are only two entries in each vector. Notice, however, that none of the vectors is a multiple of one of the other vectors. See Fig. 4. ■

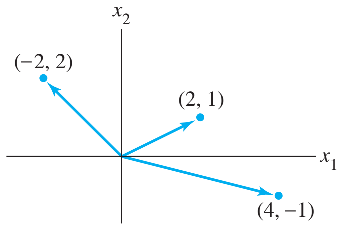


FIGURE 4

A linearly dependent set in $\mathbb{R}^2$.

THEOREM 9

If a set $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

PROOF By renumbering the vectors, we may suppose $\mathbf{v}_1 = \mathbf{0}$. Then the equation $1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p = \mathbf{0}$ shows that S is linearly dependent. ■

EXAMPLE 6 Determine by inspection if the given set is linearly dependent.

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. $\begin{bmatrix} -2 \\ 4 \\ 6 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ -9 \\ 15 \end{bmatrix}$

SOLUTION

a. The set contains four vectors, each of which has only three entries. So the set is linearly dependent by Theorem 8.

b. Theorem 8 does not apply here because the number of vectors does not exceed the number of entries in each vector. Since the zero vector is in the set, the set is linearly dependent by Theorem 9.

c. Compare the corresponding entries of the two vectors. The second vector seems to be $-3/2$ times the first vector. This relation holds for the first three pairs of entries, but fails for the fourth pair. Thus neither of the vectors is a multiple of the other, and hence they are linearly independent. ■

SG Mastering: Linear Independence 1–31

In general, you should read a section thoroughly *several* times to absorb an important concept such as linear independence. The notes in the *Study Guide* for this section will help you learn to form mental images of key ideas in linear algebra. For instance, the following proof is worth reading carefully because it shows how the definition of linear independence can be *used*.

PROOF OF THEOREM 7 (Characterization of Linearly Dependent Sets)

If some $\mathbf{v}_j$ in S equals a linear combination of the other vectors, then $\mathbf{v}_j$ can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (-1) on $\mathbf{v}_j$. [For instance, if $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, then $\mathbf{0} = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \cdots + 0\mathbf{v}_p$.] Thus S is linearly dependent.

Conversely, suppose S is linearly dependent. If $\mathbf{v}_1$ is zero, then it is a (trivial) linear combination of the other vectors in S. Otherwise, $\mathbf{v}_1 \neq \mathbf{0}$, and there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

Let j be the largest subscript for which $c_j \neq 0$. If $j = 1$, then $c_1\mathbf{v}_1 = \mathbf{0}$, which is impossible because $\mathbf{v}_1 \neq \mathbf{0}$. So $j > 1$, and

$$c_1\mathbf{v}_1 + \cdots + c_j\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_p = \mathbf{0}$$

$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \cdots - c_{j-1}\mathbf{v}_{j-1}$$

$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1} \quad ■$$

PRACTICE PROBLEMS

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \\ 7 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

1. Are the sets $\{\mathbf{u}, \mathbf{v}\}$, $\{\mathbf{u}, \mathbf{w}\}$, $\{\mathbf{u}, \mathbf{z}\}$, $\{\mathbf{v}, \mathbf{w}\}$, $\{\mathbf{v}, \mathbf{z}\}$, and $\{\mathbf{w}, \mathbf{z}\}$ each linearly independent? Why or why not?
2. Does the answer to Problem 1 imply that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ is linearly independent?
3. To determine if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ is linearly dependent, is it wise to check if, say, $\mathbf{w}$ is a linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{z}$?
4. Is $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ linearly dependent?

1.7 EXERCISES

In Exercises 1–4, determine if the vectors are linearly independent. Justify each answer.

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. $\begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -9 \end{bmatrix}$

In Exercises 5–8, determine if the columns of the matrix form a linearly independent set. Justify each answer.

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

In Exercises 9 and 10, (a) for what values of h is $\mathbf{v}_3$ in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$, and (b) for what values of h is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ linearly *dependent*? Justify each answer.

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. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ 15 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -5 \\ h \end{bmatrix}$

In Exercises 11–14, find the value(s) of h for which the vectors are linearly *dependent*. Justify each answer.

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. $\begin{bmatrix} 1 \\ -2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ h \end{bmatrix}$

Determine by inspection whether the vectors in Exercises 15–20 are linearly *independent*. Justify each answer.

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

In Exercises 21 and 22, mark each statement True or False. Justify each answer on the basis of a careful reading of the text.

21. a. The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution.

b. If S is a linearly dependent set, then each vector is a linear combination of the other vectors in S.

c. The columns of any 4×5 matrix are linearly dependent.

d. If $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, and if $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is linearly dependent, then $\mathbf{z}$ is in Span $\{\mathbf{x}, \mathbf{y}\}$.

22. a. If $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, and if $\mathbf{w}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$, then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

b. If three vectors in $\mathbb{R}^3$ lie in the same plane in $\mathbb{R}^3$, then they are linearly dependent.

c. If a set contains fewer vectors than there are entries in the vectors, then the set is linearly independent.

d. If a set in $\mathbb{R}^n$ is linearly dependent, then the set contains more than n vectors.

In Exercises 23–26, describe the possible echelon forms of the matrix. Use the notation of Example 1 in Section 1.2.

23. A is a 2×2 matrix with linearly dependent columns.

24. A is a 3×3 matrix with linearly independent columns.

25. A is a 4×2 matrix, $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, and $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$.

26. A is a 4×3 matrix, $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$, such that $\{\mathbf{a}_1, \mathbf{a}_2\}$ is linearly independent and $\mathbf{a}_3$ is not in Span $\{\mathbf{a}_1, \mathbf{a}_2\}$.

27. How many pivot columns must a 6×4 matrix have if its columns are linearly independent? Why?

28. How many pivot columns must a 4×6 matrix have if its columns span $\mathbb{R}^4$? Why?

29. Construct 3×2 matrices A and B such that $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, but $B\mathbf{x} = \mathbf{0}$ has only the trivial solution.

30. a. Fill in the blank in the following statement: "If A is an $m \times n$ matrix, then the columns of A are linearly independent if and only if A has _____ pivot columns."

b. Explain why the statement in (a) is true.

Exercises 31 and 32 should be solved *without performing row operations*. [*Hint:* Write $A\mathbf{x} = \mathbf{0}$ as a vector equation.]

31. Given $A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$, observe that the third column is the sum of the first two columns. Find a nontrivial solution of $A\mathbf{x} = \mathbf{0}$.

32. Given $A = \begin{bmatrix} 4 & 3 & -5 \\ -2 & -2 & 4 \\ -2 & -3 & 7 \end{bmatrix}$, observe that the first column minus three times the second column equals the third column. Find a nontrivial solution of $A\mathbf{x} = \mathbf{0}$.

Each statement in Exercises 33–38 is either true (in all cases) or false (for at least one example). If false, construct a specific example to show that the statement is not always true. Such an example is called a *counterexample* to the statement. If a statement is true, give a justification. (One specific example cannot explain why a statement is always true. You will have to do more work here than in Exercises 21 and 22.)

33. If $\mathbf{v}_1, \ldots, \mathbf{v}_4$ are in $\mathbb{R}^4$ and $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly dependent.

34. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^4$ and $\mathbf{v}_2$ is not a scalar multiple of $\mathbf{v}_1$, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.

35. If $\mathbf{v}_1, \ldots, \mathbf{v}_5$ are in $\mathbb{R}^5$ and $\mathbf{v}_3 = \mathbf{0}$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ is linearly dependent.

36. If $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are in $\mathbb{R}^3$ and $\mathbf{v}_3$ is *not* a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

37. If $\mathbf{v}_1, \ldots, \mathbf{v}_4$ are in $\mathbb{R}^4$ and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is also linearly dependent.

38. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ is a linearly independent set of vectors in $\mathbb{R}^4$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also linearly independent. [*Hint:* Think about $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + 0 \cdot \mathbf{v}_4 = \mathbf{0}$.]

39. Suppose A is an $m \times n$ matrix with the property that for all $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has at most one solution. Use the definition of linear independence to explain why the columns of A must be linearly independent.

40. Suppose an $m \times n$ matrix A has n pivot columns. Explain why for each $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has at most one solution. [*Hint:* Explain why $A\mathbf{x} = \mathbf{b}$ cannot have infinitely many solutions.]

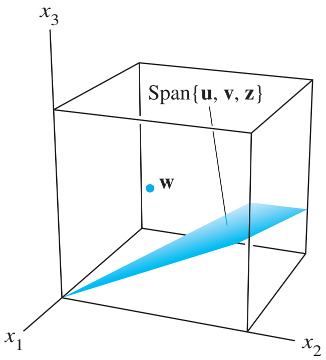
[**M**] In Exercises 41 and 42, use as many columns of A as possible to construct a matrix B with the property that the equation $B\mathbf{x} = \mathbf{0}$ has only the trivial solution. Solve $B\mathbf{x} = \mathbf{0}$ to verify your work.

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**] With A and B as in Exercise 41, select a column $\mathbf{v}$ of A that was not used in the construction of B and determine if $\mathbf{v}$ is in the set spanned by the columns of B. (Describe your calculations.)

44. [**M**] Repeat Exercise 43 with the matrices A and B from Exercise 42. Then give an explanation for what you discover, assuming that B was constructed as specified.



SOLUTIONS TO PRACTICE PROBLEMS

1. Yes. In each case, neither vector is a multiple of the other. Thus each set is linearly independent.
2. No. The observation in Practice Problem 1, by itself, says nothing about the linear independence of $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$.
3. No. When testing for linear independence, it is usually a poor idea to check if one selected vector is a linear combination of the others. It may happen that the selected vector is not a linear combination of the others and yet the whole set of vectors is linearly dependent. In this practice problem, $\mathbf{w}$ is not a linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{z}$.
4. Yes, by Theorem 8. There are more vectors (four) than entries (three) in them.

1.8 INTRODUCTION TO LINEAR TRANSFORMATIONS

The difference between a matrix equation $A\mathbf{x} = \mathbf{b}$ and the associated vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ is merely a matter of notation. However, a matrix equation $A\mathbf{x} = \mathbf{b}$ can arise in linear algebra (and in applications such as computer graphics and signal processing) in a way that is not directly connected with linear combinations of vectors. This happens when we think of the matrix A as an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$.

For instance, the equations

$$\underset{\overset{\uparrow}{A}}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\overset{\uparrow}{\mathbf{x}}}{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}} = \underset{\overset{\uparrow}{\mathbf{b}}}{\begin{bmatrix} 5 \\ 8 \end{bmatrix}} \quad \text{and} \quad \underset{\overset{\uparrow}{A}}{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}} \underset{\overset{\uparrow}{\mathbf{u}}}{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}} = \underset{\overset{\uparrow}{\mathbf{0}}}{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}$$

say that multiplication by A transforms $\mathbf{x}$ into $\mathbf{b}$ and transforms $\mathbf{u}$ into the zero vector. See Fig. 1.

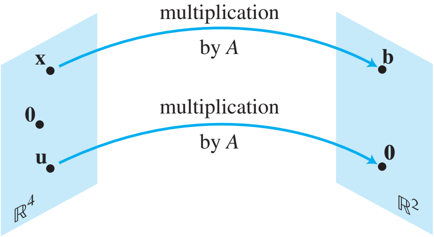


FIGURE 1 Transforming vectors via matrix multiplication.

From this new point of view, solving the equation $A\mathbf{x} = \mathbf{b}$ amounts to finding all vectors $\mathbf{x}$ in $\mathbb{R}^4$ that are transformed into the vector $\mathbf{b}$ in $\mathbb{R}^2$ under the "action" of multiplication by A.

The correspondence from $\mathbf{x}$ to $A\mathbf{x}$ is a *function* from one set of vectors to another. This concept generalizes the common notion of a function as a rule that transforms one real number into another.

A **transformation** (or **function** or **mapping**) T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of T, and $\mathbb{R}^m$ is called the **codomain** of T. The notation $T : \mathbb{R}^n \to \mathbb{R}^m$ indicates that the domain of T is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$. For $\mathbf{x}$ in $\mathbb{R}^n$, the vector $T(\mathbf{x})$ in $\mathbb{R}^m$ is called the **image** of $\mathbf{x}$ (under the action of T). The set of all images $T(\mathbf{x})$ is called the **range** of T. See Fig. 2.

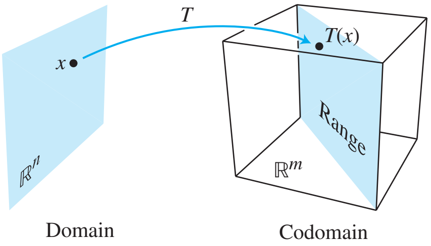


FIGURE 2 Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

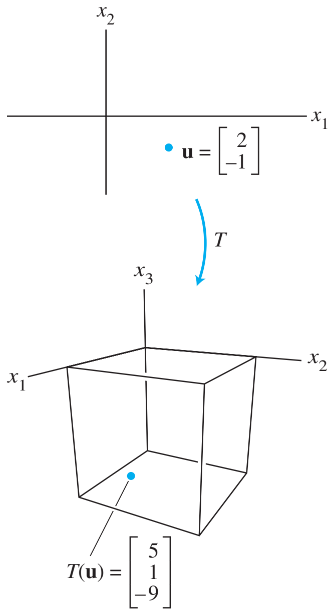
The new terminology in this section is important because a dynamic view of matrix–vector multiplication is the key to understanding several ideas in linear algebra and to building mathematical models of physical systems that evolve over time. Such *dynamical systems* will be discussed in Sections 1.10, 4.8, and 4.9 and throughout Chapter 5.

Matrix Transformations

The rest of this section focuses on mappings associated with matrix multiplication. For each $\mathbf{x}$ in $\mathbb{R}^n$, $T(\mathbf{x})$ is computed as $A\mathbf{x}$, where A is an $m \times n$ matrix. For simplicity, we sometimes denote such a *matrix transformation* by $\mathbf{x} \mapsto A\mathbf{x}$. Observe that the domain of T is $\mathbb{R}^n$ when A has n columns and the codomain of T is $\mathbb{R}^m$ when each column of A has m entries. The range of T is the set of all linear combinations of the columns of A, because each image $T(\mathbf{x})$ is of the form $A\mathbf{x}$.

EXAMPLE 1 Let $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$, and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, so that

$$T(\mathbf{x}) = A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix}$$



a. Find $T(\mathbf{u})$, the image of $\mathbf{u}$ under the transformation T.

b. Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

c. Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

d. Determine if $\mathbf{c}$ is in the range of the transformation T.

SOLUTION

a. Compute

$$T(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}$$

b. Solve $T(\mathbf{x}) = \mathbf{b}$ for $\mathbf{x}$. That is, solve $A\mathbf{x} = \mathbf{b}$, or

$$\begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -5 \end{bmatrix} \tag{1}$$

Using the method discussed in Section 1.4, row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1.5 \\ 0 & 1 & -.5 \\ 0 & 0 & 0 \end{bmatrix} \tag{2}$$

Hence $x_1 = 1.5$, $x_2 = -.5$, and $\mathbf{x} = \begin{bmatrix} 1.5 \\ -.5 \end{bmatrix}$. The image of this $\mathbf{x}$ under T is the given vector $\mathbf{b}$.

c. Any $\mathbf{x}$ whose image under T is $\mathbf{b}$ must satisfy equation (1). From (2), it is clear that equation (1) has a unique solution. So there is exactly one $\mathbf{x}$ whose image is $\mathbf{b}$.

d. The vector $\mathbf{c}$ is in the range of T if $\mathbf{c}$ is the image of some $\mathbf{x}$ in $\mathbb{R}^2$, that is, if $\mathbf{c} = T(\mathbf{x})$ for some $\mathbf{x}$. This is just another way of asking if the system $A\mathbf{x} = \mathbf{c}$ is consistent. To find the answer, row reduce the augmented matrix:

$$\begin{bmatrix} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 14 & -7 \\ 0 & 4 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 14 & -7 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & -35 \end{bmatrix}$$

The third equation, $0 = -35$, shows that the system is inconsistent. So $\mathbf{c}$ is *not* in the range of T. ■

The question in Example 1(c) is a *uniqueness* problem for a system of linear equations, translated here into the language of matrix transformations: Is $\mathbf{b}$ the image of a *unique* $\mathbf{x}$ in $\mathbb{R}^n$? Similarly, Example 1(d) is an *existence* problem: Does there *exist* an $\mathbf{x}$ whose image is $\mathbf{c}$?

The next two matrix transformations can be viewed geometrically. They reinforce the dynamic view of a matrix as something that transforms vectors into other vectors. Section 2.7 contains other interesting examples connected with computer graphics.

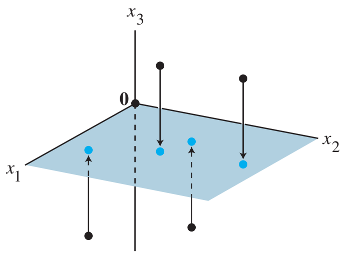


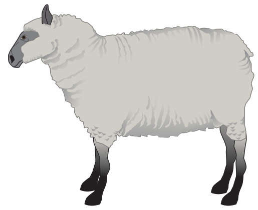
FIGURE 3
A projection transformation.

EXAMPLE 2 If $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ *projects* points in $\mathbb{R}^3$ onto the x_1x_2-plane because

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}$$

See Fig. 3. ■

EXAMPLE 3 Let $A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$. The transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is called a **shear transformation**. It can be shown that if T acts on each point in the 2×2 square shown in Fig. 4, then the set of images forms the shaded parallelogram. The key idea is to show that T maps line segments onto line segments (as shown in Exercise 27) and then to check that the corners of the square map onto the vertices of the parallelogram. For instance, the image of the point $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$ is $T(\mathbf{u}) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \end{bmatrix}$, and the image of $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is $\begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$. T deforms the square as if the top of the square were pushed to the right while the base is held fixed. Shear transformations appear in physics, geology, and crystallography. ■



sheep

sheared sheep

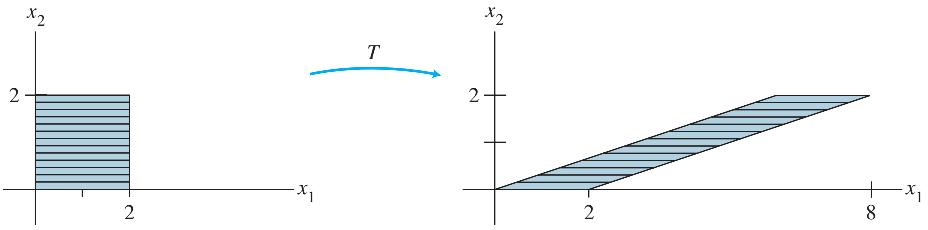


FIGURE 4 A shear transformation.

Linear Transformations

Theorem 5 in Section 1.4 shows that if A is $m \times n$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ has the properties

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \quad \text{and} \quad A(c\mathbf{u}) = cA\mathbf{u}$$

for all $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$ and all scalars c. These properties, written in function notation, identify the most important class of transformations in linear algebra.

DEFINITION

A transformation (or mapping) T is **linear** if:

(i) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;
(ii) $T(c\mathbf{u}) = cT(\mathbf{u})$ for all scalars c and all $\mathbf{u}$ in the domain of T.

Every matrix transformation is a linear transformation. Important examples of linear transformations that are not matrix transformations will be discussed in Chapters 4 and 5.

Linear transformations *preserve the operations of vector addition and scalar multiplication*. Property (i) says that the result $T(\mathbf{u}+\mathbf{v})$ of first adding $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and then applying T is the same as first applying T to $\mathbf{u}$ and to $\mathbf{v}$ and then adding $T(\mathbf{u})$ and $T(\mathbf{v})$ in $\mathbb{R}^m$. These two properties lead easily to the following useful facts.

If T is a linear transformation, then

$$T(\mathbf{0}) = \mathbf{0} \tag{3}$$

and

$$T(c\mathbf{u} + d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v}) \tag{4}$$

for all vectors $\mathbf{u}$, $\mathbf{v}$ in the domain of T and all scalars c, d.

Property (3) follows from condition (ii) in the definition, because $T(\mathbf{0}) = T(0\mathbf{u}) = 0T(\mathbf{u}) = \mathbf{0}$. Property (4) requires both (i) and (ii):

$$T(c\mathbf{u} + d\mathbf{v}) = T(c\mathbf{u}) + T(d\mathbf{v}) = cT(\mathbf{u}) + dT(\mathbf{v})$$

Observe that *if a transformation satisfies* (4) *for all* $\mathbf{u}$, $\mathbf{v}$ *and* c, d, *it must be linear*. (Set $c = d = 1$ for preservation of addition, and set $d = 0$ for preservation of scalar multiplication.) Repeated application of (4) produces a useful generalization:

$$T(c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \cdots + c_pT(\mathbf{v}_p) \tag{5}$$

In engineering and physics, (5) is referred to as a *superposition principle*. Think of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ as signals that go into a system and $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)$ as the responses of that system to the signals. The system satisfies the superposition principle if whenever an input is expressed as a linear combination of such signals, the system's response is *the same* linear combination of the responses to the individual signals. We will return to this idea in Chapter 4.

EXAMPLE 4 Given a scalar r, define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = r\mathbf{x}$. T is called a **contraction** when $0 \le r \le 1$ and a **dilation** when $r > 1$. Let $r = 3$, and show that T is a linear transformation.

SOLUTION Let $\mathbf{u}$, $\mathbf{v}$ be in $\mathbb{R}^2$ and let c, d be scalars. Then

$$\begin{aligned}
T(c\mathbf{u} + d\mathbf{v}) &= 3(c\mathbf{u} + d\mathbf{v}) && \text{Definition of } T \\
&= 3c\mathbf{u} + 3d\mathbf{v} && \Big\} \text{ Vector arithmetic} \\
&= c(3\mathbf{u}) + d(3\mathbf{v}) && \\
&= cT(\mathbf{u}) + dT(\mathbf{v})
\end{aligned}$$

Thus T is a linear transformation because it satisfies (4). See Fig. 5. ■

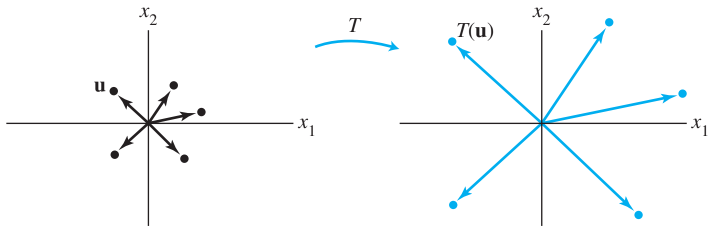


FIGURE 5 A dilation transformation.

EXAMPLE 5 Define a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T(\mathbf{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

Find the images under T of $\mathbf{u} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

SOLUTION

$$T(\mathbf{u}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \qquad T(\mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix},$$

$$T(\mathbf{u} + \mathbf{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

Note that $T(\mathbf{u} + \mathbf{v})$ is obviously equal to $T(\mathbf{u}) + T(\mathbf{v})$. It appears from Fig. 6 that T rotates $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} + \mathbf{v}$ counterclockwise about the origin through 90°. In fact, T transforms the entire parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$ into the one determined by $T(\mathbf{u})$ and $T(\mathbf{v})$. (See Exercise 28.) ■

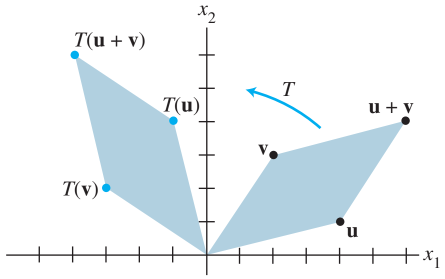


FIGURE 6 A rotation transformation.

The final example is not geometrical; instead, it shows how a linear mapping can transform one type of data into another.

EXAMPLE 6 A company manufactures two products, B and C. Using data from Example 7 in Section 1.3, we construct a "unit cost" matrix, $U = [\,\mathbf{b} \quad \mathbf{c}\,]$, whose columns describe the "costs per dollar of output" for the products:

$$U = \begin{array}{c} \text{Product} \\ \begin{array}{cc} \text{B} & \text{C} \end{array} \\ \begin{bmatrix} .45 & .40 \\ .25 & .35 \\ .15 & .15 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \text{Materials} \\ \text{Labor} \\ \text{Overhead} \end{array}$$

Let $\mathbf{x} = (x_1, x_2)$ be a "production" vector, corresponding to x_1 dollars of product B and x_2 dollars of product C, and define $T : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$T(\mathbf{x}) = U\mathbf{x} = x_1 \begin{bmatrix} .45 \\ .25 \\ .15 \end{bmatrix} + x_2 \begin{bmatrix} .40 \\ .30 \\ .15 \end{bmatrix} = \begin{bmatrix} \text{Total cost of materials} \\ \text{Total cost of labor} \\ \text{Total cost of overhead} \end{bmatrix}$$

The mapping T transforms a list of production quantities (measured in dollars) into a list of total costs. The linearity of this mapping is reflected in two ways:

1. If production is increased by a factor of, say, 4, from $\mathbf{x}$ to $4\mathbf{x}$, then the costs will increase by the same factor, from $T(\mathbf{x})$ to $4T(\mathbf{x})$.

2. If $\mathbf{x}$ and $\mathbf{y}$ are production vectors, then the total cost vector associated with the combined production $\mathbf{x}+\mathbf{y}$ is precisely the sum of the cost vectors $T(\mathbf{x})$ and $T(\mathbf{y})$. ■

PRACTICE PROBLEMS

1. Suppose $T : \mathbb{R}^5 \to \mathbb{R}^2$ and $T(\mathbf{x}) = A\mathbf{x}$ for some matrix A and for each $\mathbf{x}$ in $\mathbb{R}^5$. How many rows and columns does A have?
2. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Give a geometric description of the transformation $\mathbf{x} \mapsto A\mathbf{x}$.
3. The line segment from $\mathbf{0}$ to a vector $\mathbf{u}$ is the set of points of the form $t\mathbf{u}$, where $0 \le t \le 1$. Show that a linear transformation T maps this segment into the segment between $\mathbf{0}$ and $T(\mathbf{u})$.

1.8 EXERCISES

1. Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, and define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$. Find the images under T of $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

2. Let $A = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 3 \\ 6 \\ -9 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$. Find $T(\mathbf{u})$ and $T(\mathbf{v})$.

In Exercises 3–6, with T defined by $T(\mathbf{x}) = A\mathbf{x}$, find a vector $\mathbf{x}$ whose image under T is $\mathbf{b}$, and determine whether $\mathbf{x}$ is unique.

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. Let A be a 6×5 matrix. What must a and b be in order to define $T : \mathbb{R}^a \to \mathbb{R}^b$ by $T(\mathbf{x}) = A\mathbf{x}$?

8. How many rows and columns must a matrix A have in order to define a mapping from $\mathbb{R}^5$ into $\mathbb{R}^7$ by the rule $T(\mathbf{x}) = A\mathbf{x}$?

For Exercises 9 and 10, find all $\mathbf{x}$ in $\mathbb{R}^4$ that are mapped into the zero vector by the transformation $\mathbf{x} \mapsto A\mathbf{x}$ for the given matrix A.

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. Let $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, and let A be the matrix in Exercise 9. Is $\mathbf{b}$ in the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$? Why or why not?

12. Let $\mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 4 \end{bmatrix}$, and let A be the matrix in Exercise 10. Is $\mathbf{b}$ in the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$? Why or why not?

In Exercises 13–16, use a rectangular coordinate system to plot $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, and their images under the given transformation T. (Make a separate and reasonably large sketch for each exercise.) Describe geometrically what T does to each vector $\mathbf{x}$ in $\mathbb{R}^2$.

13. $T(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

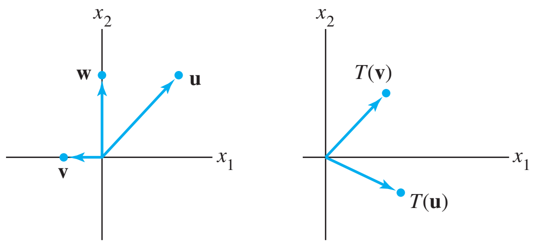
14. $T(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

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. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ into $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and maps $\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ into $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$. Use the fact that T is linear to find the images under T of $2\mathbf{u}$, $3\mathbf{v}$, and $2\mathbf{u}+3\mathbf{v}$.

18. The figure shows vectors **u**, **v**, and **w**, along with the images $T(\mathbf{u})$ and $T(\mathbf{v})$ under the action of a linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$. Copy this figure carefully, and draw the image $T(\mathbf{w})$ as accurately as possible. [*Hint:* First, write **w** as a linear combination of **u** and **v**.]



19. Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y}_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, and $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$, and let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps $\mathbf{e}_1$ into $\mathbf{y}_1$ and maps $\mathbf{e}_2$ into $\mathbf{y}_2$. Find the images of $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

20. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, and let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps **x** into $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$. Find a matrix A such that $T(\mathbf{x})$ is $A\mathbf{x}$ for each **x**.

In Exercises 21 and 22, mark each statement True or False. Justify each answer.

21. a. A linear transformation is a special type of function.

b. If A is a 3×5 matrix and T is a transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then the domain of T is $\mathbb{R}^3$.

c. If A is an $m \times n$ matrix, then the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is $\mathbb{R}^m$.

d. Every linear transformation is a matrix transformation.

e. A transformation T is linear if and only if

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

for all $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain of T and for all scalars c_1 and c_2.

22. a. The range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of all linear combinations of the columns of A.

b. Every matrix transformation is a linear transformation.

c. If $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and if **c** is in $\mathbb{R}^m$, then a uniqueness question is "Is **c** in the range of T?"

d. A linear transformation preserves the operations of vector addition and scalar multiplication.

e. A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ always maps the origin of $\mathbb{R}^n$ to the origin of $\mathbb{R}^m$.

23. Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = mx + b$.

a. Show that f is a linear transformation when $b = 0$.

b. Find a property of a linear transformation that is violated when $b \neq 0$.

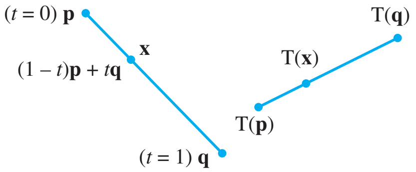
c. Why is f called a linear function?

24. An *affine transformation* $T: \mathbb{R}^n \to \mathbb{R}^m$ has the form $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, with A an $m \times n$ matrix and **b** in $\mathbb{R}^m$. Show that T is *not* a linear transformation when $\mathbf{b} \neq \mathbf{0}$. (Affine transformations are important in computer graphics.)

25. Given $\mathbf{v} \neq \mathbf{0}$ and **p** in $\mathbb{R}^n$, the line through **p** in the direction of **v** has the parametric equation $\mathbf{x} = \mathbf{p} + t\mathbf{v}$. Show that a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ maps this line onto another line or onto a single point (a *degenerate line*).

26. a. Show that the line through vectors **p** and **q** in $\mathbb{R}^n$ may be written in the parametric form $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$. (Refer to the figure with Exercises 21 and 22 in Section 1.5.)

b. The line segment from **p** to **q** is the set of points of the form $(1 - t)\mathbf{p} + t\mathbf{q}$ for $0 \leq t \leq 1$ (as shown in the figure below). Show that a linear transformation T maps this line segment onto a line segment or onto a single point.



27. Let **u** and **v** be linearly independent vectors in $\mathbb{R}^3$, and let P be the plane through **u**, **v**, and **0**. The parametric equation of P is $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (with s, t in $\mathbb{R}$). Show that a linear transformation $T: \mathbb{R}^3 \to \mathbb{R}^3$ maps P onto a plane through **0**, or onto a line through **0**, or onto just the origin in $\mathbb{R}^3$. What must be true about $T(\mathbf{u})$ and $T(\mathbf{v})$ in order for the image of the plane P to be a plane?

28. Let **u** and **v** be vectors in $\mathbb{R}^n$. It can be shown that the set P of all points in the parallelogram determined by **u** and **v** has the form $a\mathbf{u} + b\mathbf{v}$, for $0 \leq a \leq 1, 0 \leq b \leq 1$. Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Explain why the image of a point in P under the transformation T lies in the parallelogram determined by $T(\mathbf{u})$ and $T(\mathbf{v})$.

29. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that reflects each point through the x_2-axis. Make two sketches similar to Fig. 6 that illustrate properties (i) and (ii) of a linear transformation.

30. Suppose vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ span $\mathbb{R}^n$, and let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Suppose $T(\mathbf{v}_i) = \mathbf{0}$ for $i = 1, \ldots, p$. Show that T is the zero transformation. That is, show that if **x** is any vector in $\mathbb{R}^n$, then $T(\mathbf{x}) = \mathbf{0}$.

31. Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly dependent set in $\mathbb{R}^n$. Explain why the set $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ is linearly dependent.

In Exercises 32–36, column vectors are written as rows, such as $\mathbf{x} = (x_1, x_2)$, and $T(\mathbf{x})$ is written as $T(x_1, x_2)$.

32. Show that the transformation T defined by $T(x_1, x_2) = (x_1 - 2|x_2|, x_1 - 4x_2)$ is not linear.

33. Show that the transformation T defined by $T(x_1, x_2) = (x_1 - 2x_2, x_1 - 3, 2x_1 - 5x_2)$ is not linear.

34. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the transformation that reflects each vector $\mathbf{x} = (x_1, x_2, x_3)$ through the plane $x_3 = 0$ onto $T(\mathbf{x}) = (x_1, x_2, -x_3)$. Show that T is a linear transformation. [See Example 4 for ideas.]

35. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the transformation that projects each vector $\mathbf{x} = (x_1, x_2, x_3)$ onto the plane $x_2 = 0$, so $T(\mathbf{x}) = (x_1, 0, x_3)$. Show that T is a linear transformation.

36. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Suppose $\{\mathbf{u}, \mathbf{v}\}$ is a linearly independent set, but $\{T(\mathbf{u}), T(\mathbf{v})\}$ is a linearly dependent set. Show that $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution. [*Hint:* Use the fact that $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$ for some weights c_1 and c_2, not both zero.]

[**M**] In Exercises 37 and 38, the given matrix determines a linear transformation T. Find all $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{0}$.

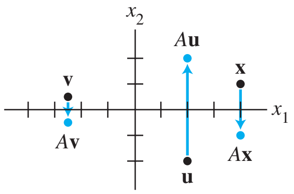
37. $\begin{bmatrix} 2 & 3 & 5 & -5 \\ -7 & 7 & 0 & 0 \\ -3 & 4 & 1 & 3 \\ -9 & 3 & -6 & -4 \end{bmatrix}$ **38.** $\begin{bmatrix} 3 & 4 & -7 & 0 \\ 5 & -8 & 7 & 4 \\ 6 & -8 & 6 & 4 \\ 9 & -7 & -2 & 0 \end{bmatrix}$

39. [**M**] Let $\mathbf{b} = \begin{bmatrix} 8 \\ 7 \\ 5 \\ -3 \end{bmatrix}$ and let A be the matrix in Exercise 37. Is $\mathbf{b}$ in the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$? If so, find an $\mathbf{x}$ whose image under the transformation is $\mathbf{b}$.

40. [**M**] Let $\mathbf{b} = \begin{bmatrix} 4 \\ -4 \\ -4 \\ -7 \end{bmatrix}$ and let A be the matrix in Exercise 38. Is $\mathbf{b}$ in the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$? If so, find an $\mathbf{x}$ whose image under the transformation is $\mathbf{b}$.

SG Mastering: Linear Transformations 1–34

SOLUTIONS TO PRACTICE PROBLEMS

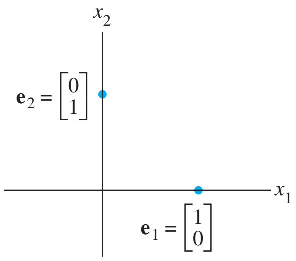


The transformation $\mathbf{x} \mapsto A\mathbf{x}$.

1. A must have five columns for $A\mathbf{x}$ to be defined. A must have two rows for the codomain of T to be $\mathbb{R}^2$.
2. Plot some random points (vectors) on graph paper to see what happens. A point such as $(4, 1)$ maps into $(4, -1)$. The transformation $\mathbf{x} \mapsto A\mathbf{x}$ reflects points through the x-axis (or x_1-axis).
3. Let $\mathbf{x} = t\mathbf{u}$ for some t such that $0 \le t \le 1$. Since T is linear, $T(t\mathbf{u}) = t\,T(\mathbf{u})$, which is a point on the line segment between $\mathbf{0}$ and $T(\mathbf{u})$.

1.9 THE MATRIX OF A LINEAR TRANSFORMATION

Whenever a linear transformation T arises geometrically or is described in words, we usually want a “formula” for $T(\mathbf{x})$. The discussion that follows shows that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is actually a matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ and that important properties of T are intimately related to familiar properties of A. The key to finding A is to observe that T is completely determined by what it does to the columns of the $n \times n$ identity matrix I_n.



EXAMPLE 1 The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ such that

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

With no additional information, find a formula for the image of an arbitrary $\mathbf{x}$ in $\mathbb{R}^2$.

SOLUTION Write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 \tag{1}$$

Since T is a *linear* transformation,

$$\begin{aligned} T(\mathbf{x}) &= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) && (2) \\ &= x_1\begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 5x_1 - 3x_2 \\ -7x_1 + 8x_2 \\ 2x_1 + 0 \end{bmatrix} \end{aligned}$$

■

The step from equation (1) to equation (2) explains why knowledge of $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ is sufficient to determine $T(\mathbf{x})$ for any $\mathbf{x}$. Moreover, since (2) expresses $T(\mathbf{x})$ as a linear combination of vectors, we can put these vectors into the columns of a matrix A and write (2) as

$$T(\mathbf{x}) = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A\mathbf{x}$$

THEOREM 10

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique matrix A such that

$$T(\mathbf{x}) = A\mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

In fact, A is the $m \times n$ matrix whose jth column is the vector $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth column of the identity matrix in $\mathbb{R}^n$:

$$A = \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix} \tag{3}$$

PROOF Write $\mathbf{x} = I_n\mathbf{x} = [\,\mathbf{e}_1 \ \cdots \ \mathbf{e}_n\,]\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, and use the linearity of T to compute

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + \cdots + x_nT(\mathbf{e}_n) \\ &= \begin{bmatrix} T(\mathbf{e}_1) & \cdots & T(\mathbf{e}_n) \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x} \end{aligned}$$

The uniqueness of A is treated in Exercise 33. ■

The matrix A in (3) is called the **standard matrix for the linear transformation T**.

We know now that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be viewed as a matrix transformation, and vice versa. The term *linear transformation* focuses on a property of a mapping, while *matrix transformation* describes how such a mapping is implemented, as Examples 2 and 3 illustrate.

EXAMPLE 2 Find the standard matrix A for the dilation transformation $T(\mathbf{x}) = 3\mathbf{x}$, for $\mathbf{x}$ in $\mathbb{R}^2$.

SOLUTION Write

$$T(\mathbf{e}_1) = 3\mathbf{e}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = 3\mathbf{e}_2 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

■

EXAMPLE 3 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that rotates each point in $\mathbb{R}^2$ about the origin through an angle φ, with counterclockwise rotation for a positive angle. We could show geometrically that such a transformation is linear. (See Fig. 6 in Section 1.8.) Find the standard matrix A of this transformation.

SOLUTION $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ rotates into $\begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$, and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ rotates into $\begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$. See Fig. 1. By Theorem 10,

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

Example 5 in Section 1.8 is a special case of this transformation, with $\varphi = \pi/2$. ■

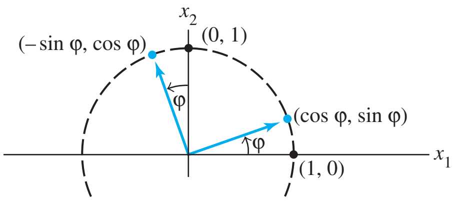


FIGURE 1 A rotation transformation.

Geometric Linear Transformations of $\mathbb{R}^2$

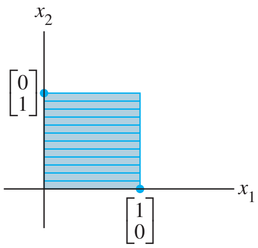


FIGURE 2
The unit square.

Examples 2 and 3 illustrate linear transformations that are described geometrically. Tables 1–4 illustrate other common geometric linear transformations of the plane. Because the transformations are linear, they are determined completely by what they do to the columns of I_2. Instead of showing only the images of $\mathbf{e}_1$ and $\mathbf{e}_2$, the tables show what a transformation does to the unit square (Fig. 2).

Other transformations can be constructed from those listed in Tables 1–4 by applying one transformation after another. For instance, a horizontal shear could be followed by a reflection in the x_2-axis. Section 2.1 will show that such a *composition* of linear transformations is linear. (Also, see Exercise 34.)

Existence and Uniqueness Questions

The concept of a linear transformation provides a new way to understand the existence and uniqueness questions asked earlier. The two definitions following Tables 1–4 give the appropriate terminology for transformations.

TABLE 1 **Reflections**

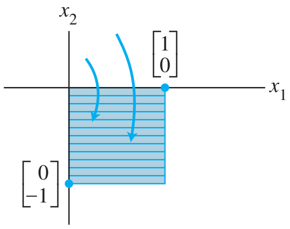
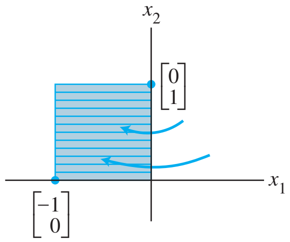
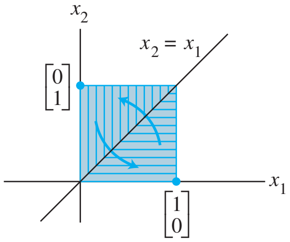
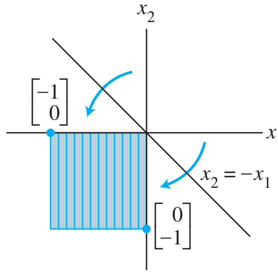
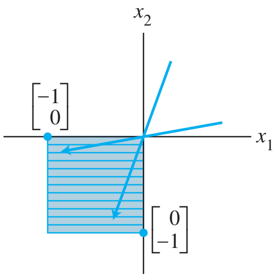
Transformation	Image of the Unit Square	Standard Matrix
Reflection through the x_1-axis		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection through the x_2-axis		$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection through the line $x_2 = x_1$		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Reflection through the line $x_2 = -x_1$		$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$
Reflection through the origin		$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

TABLE 2 Contractions and Expansions

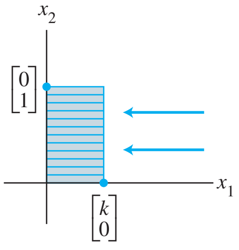
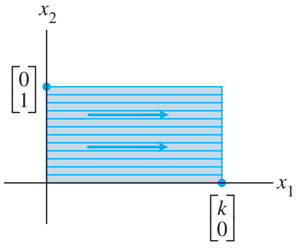
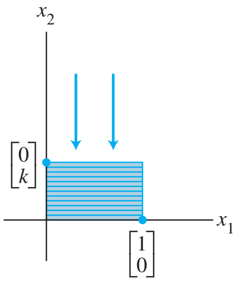
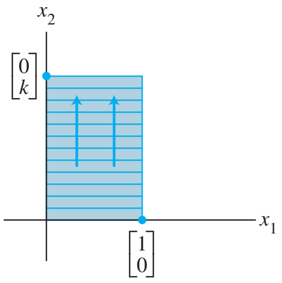
Transformation	Image of the Unit Square		Standard Matrix
Horizontal contraction and expansion	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}k & 0\\0 & 1\end{bmatrix}$
Vertical contraction and expansion	 $0 < k < 1$	 $k > 1$	$\begin{bmatrix}1 & 0\\0 & k\end{bmatrix}$

TABLE 3 Shears

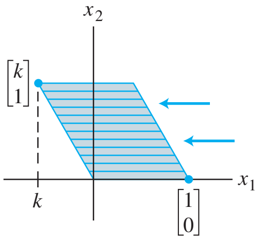
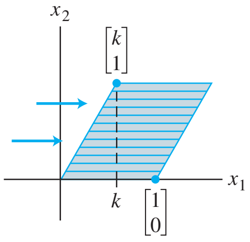
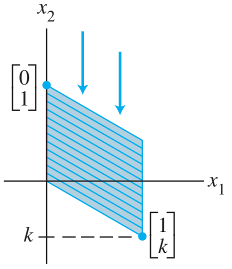
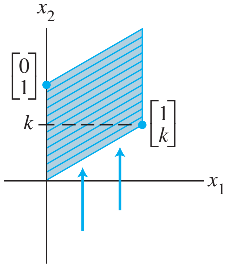
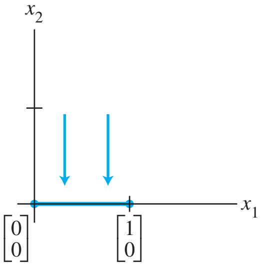
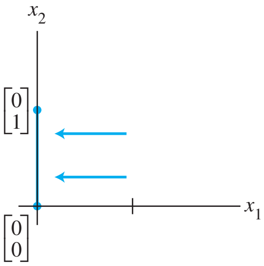
Transformation	Image of the Unit Square		Standard Matrix
Horizontal shear	 $k < 0$	 $k > 0$	$\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$
Vertical shear	 $k < 0$	 $k > 0$	$\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$

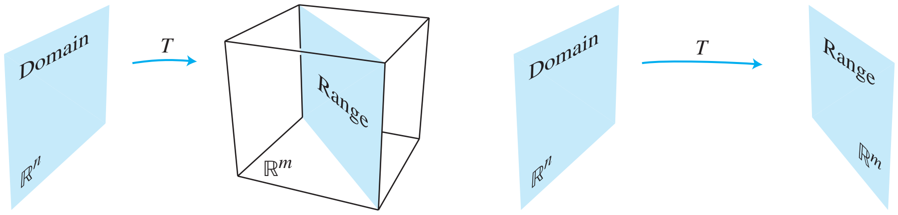
TABLE 4 Projections

Transformation	**Image of the Unit Square**	**Standard Matrix**
Projection onto the x_1-axis		$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$
Projection onto the x_2-axis		$\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$

DEFINITION

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **onto** $\mathbb{R}^m$ if each **b** in $\mathbb{R}^m$ is the image of *at least one* **x** in $\mathbb{R}^n$.

Equivalently, T is onto $\mathbb{R}^m$ when the range of T is all of the codomain $\mathbb{R}^m$. That is, T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if, for each **b** in the codomain $\mathbb{R}^m$, there exists at least one solution of $T(\mathbf{x}) = \mathbf{b}$. "Does T map $\mathbb{R}^n$ onto $\mathbb{R}^m$?" is an existence question. The mapping T is *not* onto when there is some **b** in $\mathbb{R}^m$ for which the equation $T(\mathbf{x}) = \mathbf{b}$ has no solution. See Fig. 3.



T is *not* onto $\mathbb{R}^m$

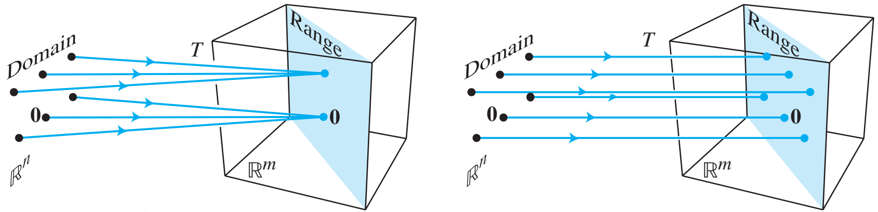
T *is* onto $\mathbb{R}^m$

FIGURE 3 Is the range of T all of $\mathbb{R}^m$?

DEFINITION

A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if each **b** in $\mathbb{R}^m$ is the image of *at most one* **x** in $\mathbb{R}^n$.

Equivalently, T is one-to-one if, for each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $T(\mathbf{x}) = \mathbf{b}$ has either a unique solution or none at all. "Is T one-to-one?" is a uniqueness question. The mapping T is *not* one-to-one when some $\mathbf{b}$ in $\mathbb{R}^m$ is the image of more than one vector in $\mathbb{R}^n$. If there is no such $\mathbf{b}$, then T is one-to-one. See Fig. 4.



T is *not* one-to-one T *is* one-to-one

FIGURE 4 Is every $\mathbf{b}$ the image of at most one vector?

SG Mastering: Existence and Uniqueness 1–39

The projection transformations shown in Table 4 are *not* one-to-one and do *not* map $\mathbb{R}^2$ onto $\mathbb{R}^2$. The transformations in Tables 1, 2, and 3 are one-to-one *and* do map $\mathbb{R}^2$ onto $\mathbb{R}^2$. Other possibilities are shown in the two examples below.

Example 4 and the theorems that follow show how the function properties of being one-to-one and mapping onto are related to important concepts studied earlier in this chapter.

EXAMPLE 4 Let T be the linear transformation whose standard matrix is

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

Does T map $\mathbb{R}^4$ onto $\mathbb{R}^3$? Is T a one-to-one mapping?

SOLUTION Since A happens to be in echelon form, we can see at once that A has a pivot position in each row. By Theorem 4 in Section 1.4, for each $\mathbf{b}$ in $\mathbb{R}^3$, the equation $A\mathbf{x} = \mathbf{b}$ is consistent. In other words, the linear transformation T maps $\mathbb{R}^4$ (its domain) onto $\mathbb{R}^3$. However, since the equation $A\mathbf{x} = \mathbf{b}$ has a free variable (because there are four variables and only three basic variables), each $\mathbf{b}$ is the image of more than one $\mathbf{x}$. That is, T is *not* one-to-one. ■

THEOREM 11

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is one-to-one if and only if the equation $T(\mathbf{x}) = \mathbf{0}$ has only the trivial solution.

PROOF Since T is linear, $T(\mathbf{0}) = \mathbf{0}$. If T is one-to-one, then the equation $T(\mathbf{x}) = \mathbf{0}$ has at most one solution and hence only the trivial solution. If T is not one-to-one, then there is a $\mathbf{b}$ that is the image of at least two different vectors in $\mathbb{R}^n$—say, $\mathbf{u}$ and $\mathbf{v}$. That is, $T(\mathbf{u}) = \mathbf{b}$ and $T(\mathbf{v}) = \mathbf{b}$. But then, since T is linear,

$$T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

The vector $\mathbf{u} - \mathbf{v}$ is not zero, since $\mathbf{u} \neq \mathbf{v}$. Hence the equation $T(\mathbf{x}) = \mathbf{0}$ has more than one solution. So, either the two conditions in the theorem are both true or they are both false. ■

THEOREM 12

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;

b. T is one-to-one if and only if the columns of A are linearly independent.

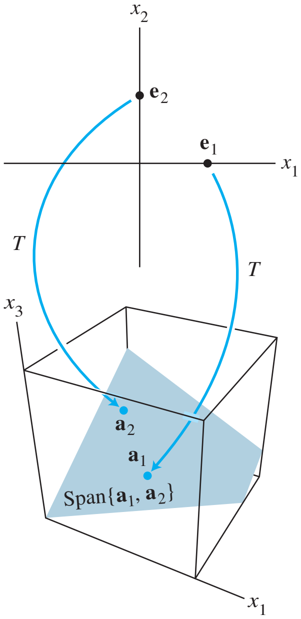
PROOF

a. By Theorem 4 in Section 1.4, the columns of A span $\mathbb{R}^m$ if and only if for each $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ is consistent—in other words, if and only if for every $\mathbf{b}$, the equation $T(\mathbf{x}) = \mathbf{b}$ has at least one solution. This is true if and only if T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

b. The equations $T(\mathbf{x}) = \mathbf{0}$ and $A\mathbf{x} = \mathbf{0}$ are the same except for notation. So, by Theorem 11, T is one-to-one if and only if $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. This happens if and only if the columns of A are linearly independent, as was already noted in the boxed statement (3) in Section 1.7. ■

Statement (a) in Theorem 12 is equivalent to the statement "T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if every vector in $\mathbb{R}^m$ is a linear combination of the columns of A." See Theorem 4 in Section 1.4.

In the next example and in some exercises that follow, column vectors are written in rows, such as $\mathbf{x} = (x_1, x_2)$, and $T(\mathbf{x})$ is written as $T(x_1, x_2)$ instead of the more formal $T((x_1, x_2))$.



The transformation T is not onto $\mathbb{R}^3$.

EXAMPLE 5 Let $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Show that T is a one-to-one linear transformation. Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?

SOLUTION When $\mathbf{x}$ and $T(\mathbf{x})$ are written as column vectors, you can determine the standard matrix of T by inspection, visualizing the row–vector computation of each entry in $A\mathbf{x}$.

$$T(\mathbf{x}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4}$$

So T is indeed a linear transformation, with its standard matrix A shown in (4). The columns of A are linearly independent because they are not multiples. By Theorem 12(b), T is one-to-one. To decide if T is onto $\mathbb{R}^3$, examine the span of the columns of A. Since A is 3×2, the columns of A span $\mathbb{R}^3$ if and only if A has 3 pivot positions, by Theorem 4. This is impossible, since A has only 2 columns. So the columns of A do not span $\mathbb{R}^3$, and the associated linear transformation is not onto $\mathbb{R}^3$. ■

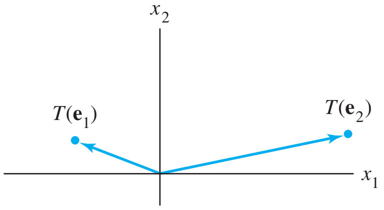
PRACTICE PROBLEM

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that first performs a horizontal shear that maps $\mathbf{e}_2$ into $\mathbf{e}_2 - .5\mathbf{e}_1$ (but leaves $\mathbf{e}_1$ unchanged) and then reflects the result through the x_2-axis. Assuming that T is linear, find its standard matrix. [*Hint:* Determine the final location of the images of $\mathbf{e}_1$ and $\mathbf{e}_2$.]

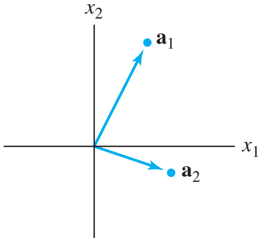
1.9 EXERCISES

In Exercises 1–10, assume that T is a linear transformation. Find the standard matrix of T.

1. $T : \mathbb{R}^2 \to \mathbb{R}^4$, $T(\mathbf{e}_1) = (3, 1, 3, 1)$, and $T(\mathbf{e}_2) = (-5, 2, 0, 0)$, where $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$.
2. $T : \mathbb{R}^3 \to \mathbb{R}^2$, $T(\mathbf{e}_1) = (1, 4)$, $T(\mathbf{e}_2) = (-2, 9)$, and $T(\mathbf{e}_3) = (3, -8)$, where $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the columns of the 3×3 identity matrix.
3. $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a vertical shear transformation that maps $\mathbf{e}_1$ into $\mathbf{e}_1 - 3\mathbf{e}_2$, but leaves $\mathbf{e}_2$ unchanged.
4. $T : \mathbb{R}^2 \to \mathbb{R}^2$ is a horizontal shear transformation that leaves $\mathbf{e}_1$ unchanged and maps $\mathbf{e}_2$ into $\mathbf{e}_2 + 2\mathbf{e}_1$.
5. $T : \mathbb{R}^2 \to \mathbb{R}^2$ rotates points (about the origin) through $\pi/2$ radians (counterclockwise).
6. $T : \mathbb{R}^2 \to \mathbb{R}^2$ rotates points (about the origin) through $-3\pi/2$ radians (clockwise).
7. $T : \mathbb{R}^2 \to \mathbb{R}^2$ first rotates points through $-3\pi/4$ radians (clockwise) and then reflects points through the horizontal x_1-axis. [*Hint:* $T(\mathbf{e}_1) = (-1/\sqrt{2}, 1/\sqrt{2})$.]
8. $T : \mathbb{R}^2 \to \mathbb{R}^2$ first performs a horizontal shear that transforms $\mathbf{e}_2$ into $\mathbf{e}_2 + 2\mathbf{e}_1$ (leaving $\mathbf{e}_1$ unchanged) and then reflects points through the line $x_2 = -x_1$.
9. $T : \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the horizontal x_1-axis and then rotates points $-\pi/2$ radians.
10. $T : \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the horizontal x_1-axis and then reflects points through the line $x_2 = x_1$.
11. A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the x_1-axis and then reflects points through the x_2-axis. Show that T can also be described as a linear transformation that rotates points about the origin. What is the angle of that rotation?
12. Show that the transformation in Exercise 10 is merely a rotation about the origin. What is the angle of the rotation?
13. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation such that $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ are the vectors shown in the figure. Using the figure, sketch the vector $T(2, 1)$.



14. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with standard matrix $A = [\,\mathbf{a}_1 \quad \mathbf{a}_2\,]$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are shown in the figure at the top of column 2. Using the figure, draw the image of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ under the transformation T.



In Exercises 15 and 16, fill in the missing entries of the matrix, assuming that the equation holds for all values of the 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. $$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 - 2x_2 \\ x_1 + 4x_2 \\ x_2 \end{bmatrix}$$

In Exercises 17–20, show that T is a linear transformation by finding a matrix that implements the mapping. Note that $x_1, x_2, \ldots$ are not vectors but are entries in vectors.

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. $T(x_1, x_2, x_3, x_4) = 3x_1 + 4x_3 - 2x_4$ (Notice: $T : \mathbb{R}^4 \to \mathbb{R}$)
21. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that $T(x_1, x_2) = (x_1 + x_2, 4x_1 + 5x_2)$. Find $\mathbf{x}$ such that $T(\mathbf{x}) = (3, 8)$.
22. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation with $T(x_1, x_2) = (2x_1 - x_2, -3x_1 + x_2, 2x_1 - 3x_2)$. Find $\mathbf{x}$ such that $T(\mathbf{x}) = (0, -1, -4)$.

In Exercises 23 and 24, mark each statement True or False. Justify each answer.

23. a. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is completely determined by its effect on the columns of the $n \times n$ identity matrix.
    b. If $T : \mathbb{R}^2 \to \mathbb{R}^2$ rotates vectors about the origin through an angle φ, then T is a linear transformation.
    c. When two linear transformations are performed one after another, the combined effect may not always be a linear transformation.
    d. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is onto $\mathbb{R}^m$ if every vector $\mathbf{x}$ in $\mathbb{R}^n$ maps onto some vector in $\mathbb{R}^m$.
    e. If A is a 3×2 matrix, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ cannot be one-to-one.
24. a. If A is a 4×3 matrix, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^3$ onto $\mathbb{R}^4$.

b. Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation.

c. The columns of the standard matrix for a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ are the images of the columns of the $n \times n$ identity matrix under T.

d. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each vector in $\mathbb{R}^n$ maps onto a unique vector in $\mathbb{R}^m$.

e. The standard matrix of a horizontal shear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ has the form $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$, where a and d are ±1.

In Exercises 25–28, determine if the specified linear transformation is (a) one-to-one and (b) onto. Justify each answer.

25. The transformation in Exercise 17

26. The transformation in Exercise 2

27. The transformation in Exercise 19

28. The transformation in Exercise 14

In Exercises 29 and 30, describe the possible echelon forms of the standard matrix for a linear transformation T. Use the notation of Example 1 in Section 1.2.

29. $T : \mathbb{R}^3 \to \mathbb{R}^4$ is one-to-one. **30.** $T : \mathbb{R}^4 \to \mathbb{R}^3$ is onto.

31. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, with A its standard matrix. Complete the following statement to make it true: "T is one-to-one if and only if A has ____ pivot columns." Explain why the statement is true. [*Hint:* Look in the exercises for Section 1.7.]

32. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, with A its standard matrix. Complete the following statement to make it true: "T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if A has ____ pivot columns." Find some theorems that explain why the statement is true.

33. Verify the uniqueness of A in Theorem 10. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation such that $T(\mathbf{x}) = B\mathbf{x}$ for some $m \times n$ matrix B. Show that if A is the standard matrix for T, then $A = B$. [*Hint:* Show that A and B have the same columns.]

34. Let $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations. Show that the mapping $\mathbf{x} \mapsto T(S(\mathbf{x}))$ is a linear transformation (from $\mathbb{R}^p$ to $\mathbb{R}^m$). [*Hint:* Compute $T(S(c\mathbf{u} + d\mathbf{v}))$ for $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^p$ and scalars c and d. Justify each step of the computation, and explain why this computation gives the desired conclusion.]

35. If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$, can you give a relation between m and n? If T is one-to-one, what can you say about m and n?

36. Why is the question "Is the linear transformation T onto?" an existence question?

[**M**] In Exercises 37–40, let T be the linear transformation whose standard matrix is given. In Exercises 37 and 38, decide if T is a one-to-one mapping. In Exercises 39 and 40, decide if T maps $\mathbb{R}^5$ onto $\mathbb{R}^5$. Justify your answers.

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. $\begin{bmatrix} 9 & 43 & 5 & 6 & -1 \\ 14 & 15 & -7 & -5 & 4 \\ -8 & -6 & 12 & -5 & -9 \\ -5 & -6 & -4 & 9 & 8 \\ 13 & 14 & 15 & 3 & 11 \end{bmatrix}$

SOLUTION TO PRACTICE PROBLEM

WEB

Follow what happens to $\mathbf{e}_1$ and $\mathbf{e}_2$. See Fig. 5. First, $\mathbf{e}_1$ is unaffected by the shear and then is reflected into $-\mathbf{e}_1$. So $T(\mathbf{e}_1) = -\mathbf{e}_1$. Second, $\mathbf{e}_2$ goes to $\mathbf{e}_2 - .5\mathbf{e}_1$ by the shear

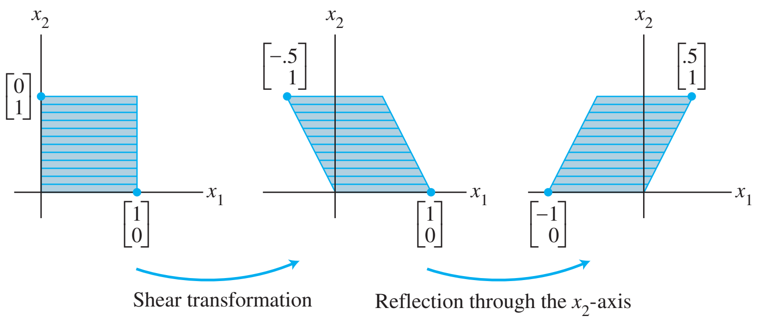


FIGURE 5 The composition of two transformations.

transformation. Since reflection through the x_2-axis changes $\mathbf{e}_1$ into $-\mathbf{e}_1$ and leaves $\mathbf{e}_2$ unchanged, the vector $\mathbf{e}_2 - .5\mathbf{e}_1$ goes to $\mathbf{e}_2 + .5\mathbf{e}_1$. So $T(\mathbf{e}_2) = \mathbf{e}_2 + .5\mathbf{e}_1$. Thus the standard matrix of T is

$$\begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} -\mathbf{e}_1 & \mathbf{e}_2 + .5\mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} -1 & .5 \\ 0 & 1 \end{bmatrix}$$

1.10 LINEAR MODELS IN BUSINESS, SCIENCE, AND ENGINEERING

The mathematical models in this section are all *linear*; that is, each describes a problem by means of a linear equation, usually in vector or matrix form. The first model concerns nutrition but actually is representative of a general technique in linear programming problems. The second model comes from electrical engineering. The third model introduces the concept of a *linear difference equation*, a powerful mathematical tool for studying dynamic processes in a wide variety of fields such as engineering, ecology, economics, telecommunications, and the management sciences. Linear models are important because natural phenomena are often linear or nearly linear when the variables involved are held within reasonable bounds. Also, linear models are more easily adapted for computer calculation than are complex nonlinear models.

As you read about each model, pay attention to how its linearity reflects some property of the system being modeled.

Constructing a Nutritious Weight-Loss Diet

WEB

The formula for the Cambridge Diet, a popular diet in the 1980s, was based on years of research. A team of scientists headed by Dr. Alan H. Howard developed this diet at Cambridge University after more than eight years of clinical work with obese patients.[1] The very low-calorie powdered formula diet combines a precise balance of carbohydrate, high-quality protein, and fat, together with vitamins, minerals, trace elements, and electrolytes. Millions of persons have used the diet to achieve rapid and substantial weight loss.

To achieve the desired amounts and proportions of nutrients, Dr. Howard had to incorporate a large variety of foodstuffs in the diet. Each foodstuff supplied several of the required ingredients, but not in the correct proportions. For instance, nonfat milk was a major source of protein but contained too much calcium. So soy flour was used for part of the protein because soy flour contains little calcium. However, soy flour contains proportionally too much fat, so whey was added since it supplies less fat in relation to calcium. Unfortunately, whey contains too much carbohydrate. . . .

The following example illustrates the problem on a small scale. Listed in Table 1 are three of the ingredients in the diet, together with the amounts of certain nutrients supplied by 100 grams (g) of each ingredient.[2]

EXAMPLE 1 If possible, find some combination of nonfat milk, soy flour, and whey to provide the exact amounts of protein, carbohydrate, and fat supplied by the diet in one day (Table 1).

---

[1] The first announcement of this rapid weight-loss regimen was given in the *International Journal of Obesity* (1978) **2**, 321–332.

[2] Ingredients in the diet as of 1984; nutrient data for ingredients adapted from USDA Agricultural Handbooks No. 8-1 and 8-6, 1976.

TABLE 1

Amounts (g) Supplied per 100 g of Ingredient				Amounts (g) Supplied by Cambridge Diet in One Day
Nutrient	Nonfat milk	Soy flour	Whey	
Protein	36	51	13	33
Carbohydrate	52	34	74	45
Fat	0	7	1.1	3

SOLUTION Let x_1, x_2, and x_3, respectively, denote the number of units (100 g) of these foodstuffs. One approach to the problem is to derive equations for each nutrient separately. For instance, the product

$$\left\{\begin{matrix} x_1 \text{ units of} \\ \text{nonfat milk} \end{matrix}\right\} \cdot \left\{\begin{matrix} \text{protein per unit} \\ \text{of nonfat milk} \end{matrix}\right\}$$

gives the amount of protein supplied by x_1 units of nonfat milk. To this amount, we would then add similar products for soy flour and whey and set the resulting sum equal to the amount of protein we need. Analogous calculations would have to be made for each nutrient.

A more efficient method, and one that is conceptually simpler, is to consider a "nutrient vector" for each foodstuff and build just one vector equation. The amount of nutrients supplied by x_1 units of nonfat milk is the scalar multiple

$$\overset{\text{Scalar}}{\left\{\begin{matrix} x_1 \text{ units of} \\ \text{nonfat milk} \end{matrix}\right\}} \cdot \overset{\text{Vector}}{\left\{\begin{matrix} \text{nutrients per unit} \\ \text{of nonfat milk} \end{matrix}\right\}} = x_1\mathbf{a}_1 \tag{1}$$

where $\mathbf{a}_1$ is the first column in Table 1. Let $\mathbf{a}_2$ and $\mathbf{a}_3$ be the corresponding vectors for soy flour and whey, respectively, and let $\mathbf{b}$ be the vector that lists the total nutrients required (the last column of the table). Then $x_2\mathbf{a}_2$ and $x_3\mathbf{a}_3$ give the nutrients supplied by x_2 units of soy flour and x_3 units of whey, respectively. So the relevant equation is

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 = \mathbf{b} \tag{2}$$

Row reduction of the augmented matrix for the corresponding system of equations shows that

$$\begin{bmatrix} 36 & 51 & 13 & 33 \\ 52 & 34 & 74 & 45 \\ 0 & 7 & 1.1 & 3 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & 0 & .277 \\ 0 & 1 & 0 & .392 \\ 0 & 0 & 1 & .233 \end{bmatrix}$$

To three significant digits, the diet requires .277 units of nonfat milk, .392 units of soy flour, and .233 units of whey in order to provide the desired amounts of protein, carbohydrate, and fat. ■

It is important that the values of x_1, x_2, and x_3 found above are nonnegative. This is necessary for the solution to be physically feasible. (How could you use $-.233$ units of whey, for instance?) With a large number of nutrient requirements, it may be necessary to use a larger number of foodstuffs in order to produce a system of equations with a "nonnegative" solution. Thus many, many different combinations of foodstuffs may need to be examined in order to find a system of equations with such a solution. In fact, the manufacturer of the Cambridge Diet was able to supply 31 nutrients in precise amounts using only 33 ingredients.

The diet construction problem leads to the *linear* equation (2) because the amount of nutrients supplied by each foodstuff can be written as a scalar multiple of a vector, as in (1). That is, the nutrients supplied by a foodstuff are *proportional* to the amount of the foodstuff added to the diet mixture. Also, each nutrient in the mixture is the *sum* of the amounts from the various foodstuffs.

Problems of formulating specialized diets for humans and livestock occur frequently. Usually they are treated by linear programming techniques. Our method of constructing vector equations often simplifies the task of formulating such problems.

Linear Equations and Electrical Networks

WEB

Current flow in a simple electrical network can be described by a system of linear equations. A voltage source such as a battery forces a current of electrons to flow through the network. When the current passes through a resistor (such as a lightbulb or motor), some of the voltage is "used up"; by Ohm's law, this "voltage drop" across a resistor is given by

$$V = RI$$

where the voltage V is measured in *volts*, the resistance R in *ohms* (denoted by Ω), and the current flow I in *amperes* (*amps*, for short).

The network in Fig. 1 contains three closed loops. The currents flowing in loops 1, 2, and 3 are denoted by I_1, I_2, and I_3, respectively. The designated directions of such *loop currents* are arbitrary. If a current turns out to be negative, then the actual direction of current flow is opposite to that chosen in the figure. If the current direction shown is away from the positive (longer) side of a battery (⊣⊢) around to the negative (shorter) side, the voltage is positive; otherwise, the voltage is negative.

Current flow in a loop is governed by the following rule.

KIRCHHOFF'S VOLTAGE LAW

The algebraic sum of the RI voltage drops in one direction around a loop equals the algebraic sum of the voltage sources in the same direction around the loop.

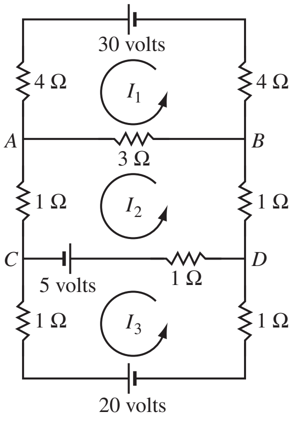


FIGURE 1

EXAMPLE 2 Determine the loop currents in the network in Fig. 1.

SOLUTION For loop 1, the current I_1 flows through three resistors, and the sum of the RI voltage drops is

$$4I_1 + 4I_1 + 3I_1 = (4 + 4 + 3)I_1 = 11I_1$$

Current from loop 2 also flows in part of loop 1, through the short *branch* between A and B. The associated RI drop there is $3I_2$ volts. However, the current direction for the branch AB in loop 1 is opposite to that chosen for the flow in loop 2, so the algebraic sum of all RI drops for loop 1 is $11I_1 - 3I_2$. Since the voltage in loop 1 is $+30$ volts, Kirchhoff's voltage law implies that

$$11I_1 - 3I_2 = 30$$

The equation for loop 2 is

$$-3I_1 + 6I_2 - I_3 = 5$$

The term $-3I_1$ comes from the flow of the loop-1 current through the branch AB (with a negative voltage drop because the current flow there is opposite to the flow in loop 2). The term $6I_2$ is the sum of all resistances in loop 2, multiplied by the loop current. The term $-I_3 = -1 \cdot I_3$ comes from the loop-3 current flowing through the 1-ohm resistor in branch CD, in the direction opposite to the flow in loop 2. The loop-3 equation is

$$-I_2 + 3I_3 = -25$$

Note that the 5-volt battery in branch CD is counted as part of both loop 2 and loop 3, but it is -5 volts for loop 3 because of the direction chosen for the current in loop 3. The 20-volt battery is negative for the same reason.

The loop currents are found by solving the system

$$\begin{array}{rcrcrcr} 11I_1 & - & 3I_2 & & & = & 30 \\ -3I_1 & + & 6I_2 & - & I_3 & = & 5 \\ & - & I_2 & + & 3I_3 & = & -25 \end{array} \tag{3}$$

Row operations on the augmented matrix lead to the solution: $I_1 = 3$ amps, $I_2 = 1$ amp, and $I_3 = -8$ amps. The negative value of I_3 indicates that the actual current in loop 3 flows in the direction opposite to that shown in Fig. 1. ■

It is instructive to look at system (3) as a vector equation:

$$I_1 \underset{\substack{\uparrow \\ \mathbf{r}_1}}{\begin{bmatrix} 11 \\ -3 \\ 0 \end{bmatrix}} + I_2 \underset{\substack{\uparrow \\ \mathbf{r}_2}}{\begin{bmatrix} -3 \\ 6 \\ -1 \end{bmatrix}} + I_3 \underset{\substack{\uparrow \\ \mathbf{r}_3}}{\begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}} = \underset{\substack{\uparrow \\ \mathbf{v}}}{\begin{bmatrix} 30 \\ 5 \\ -25 \end{bmatrix}} \tag{4}$$

The first entry of each vector concerns the first loop, and similarly for the second and third entries. The first resistor vector $\mathbf{r}_1$ lists the resistance in the various loops through which current I_1 flows. A resistance is written negatively when I_1 flows against the flow direction in another loop. Examine Fig. 1 and see how to compute the entries in $\mathbf{r}_1$; then do the same for $\mathbf{r}_2$ and $\mathbf{r}_3$. The matrix form of equation (4),

$$R\mathbf{i} = \mathbf{v}, \quad \text{where} \quad R = [\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3\,] \quad \text{and} \quad \mathbf{i} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}$$

provides a matrix version of Ohm's law. If all loop currents are chosen in the same direction (say, counterclockwise), then all entries off the main diagonal of R will be negative.

The matrix equation $R\mathbf{i} = \mathbf{v}$ makes the linearity of this model easy to see at a glance. For instance, if the voltage vector is doubled, then the current vector must double. Also, a *superposition principle* holds. That is, the solution of equation (4) is the sum of the solutions of the equations

$$R\mathbf{i} = \begin{bmatrix} 30 \\ 0 \\ 0 \end{bmatrix}, \qquad R\mathbf{i} = \begin{bmatrix} 0 \\ 5 \\ 0 \end{bmatrix}, \quad \text{and} \quad R\mathbf{i} = \begin{bmatrix} 0 \\ 0 \\ -25 \end{bmatrix}$$

Each equation here corresponds to the circuit with only one voltage source (the other sources being replaced by wires that close each loop). The model for current flow is *linear* precisely because Ohm's law and Kirchhoff's law are linear: The voltage drop across a resistor is *proportional* to the current flowing through it (Ohm), and the *sum* of the voltage drops in a loop equals the sum of the voltage sources in the loop (Kirchhoff).

Loop currents in a network can be used to determine the current in any branch of the network. If only one loop current passes through a branch, such as from B to D in Fig. 1, the branch current equals the loop current. If more than one loop current passes through a branch, such as from A to B, the branch current is the algebraic sum of the loop currents in the branch (*Kirchhoff's current law*). For instance, the current in branch AB is $I_1 - I_2 = 3 - 1 = 2$ amps, in the direction of I_1. The current in branch CD is $I_2 - I_3 = 9$ amps.

Difference Equations

In many fields such as ecology, economics, and engineering, a need arises to model mathematically a dynamic system that changes over time. Several features of the system are each measured at discrete time intervals, producing a sequence of vectors $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2, \ldots$. The entries in $\mathbf{x}_k$ provide information about the *state* of the system at the time of the kth measurement.

If there is a matrix A such that $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1$, and, in general,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots \tag{5}$$

then (5) is called a **linear difference equation** (or **recurrence relation**). Given such an equation, one can compute $\mathbf{x}_1$, $\mathbf{x}_2$, and so on, provided $\mathbf{x}_0$ is known. Sections 4.8 and 4.9, and several sections in Chapter 5, will develop formulas for $\mathbf{x}_k$ and describe what can happen to $\mathbf{x}_k$ as k increases indefinitely. The discussion below illustrates how a difference equation might arise.

A subject of interest to demographers is the movement of populations or groups of people from one region to another. The simple model here considers the changes in the population of a certain city and its surrounding suburbs over a period of years.

Fix an initial year—say, 2000—and denote the populations of the city and suburbs that year by r_0 and s_0, respectively. Let $\mathbf{x}_0$ be the population vector

$$\mathbf{x}_0 = \begin{bmatrix} r_0 \\ s_0 \end{bmatrix} \begin{matrix} \text{City population, 2000} \\ \text{Suburban population, 2000} \end{matrix}$$

For 2001 and subsequent years, denote the populations of the city and suburbs by the vectors

$$\mathbf{x}_1 = \begin{bmatrix} r_1 \\ s_1 \end{bmatrix}, \qquad \mathbf{x}_2 = \begin{bmatrix} r_2 \\ s_2 \end{bmatrix}, \qquad \mathbf{x}_3 = \begin{bmatrix} r_3 \\ s_3 \end{bmatrix}, \ldots$$

Our goal is to describe mathematically how these vectors might be related.

Suppose demographic studies show that each year about 5% of the city's population moves to the suburbs (and 95% remains in the city), while 3% of the suburban population moves to the city (and 97% remains in the suburbs). See Fig. 2.

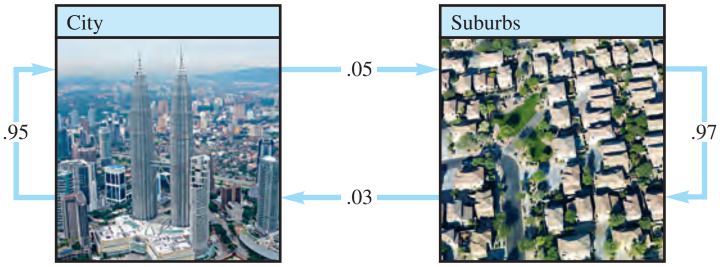


FIGURE 2 Annual percentage migration between city and suburbs.

After 1 year, the original r_0 persons in the city are now distributed between city and suburbs as

$$\begin{bmatrix} .95r_0 \\ .05r_0 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} \begin{matrix} \text{Remain in city} \\ \text{Move to suburbs} \end{matrix} \tag{6}$$

The s_0 persons in the suburbs in 2000 are distributed 1 year later as

$$s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} \begin{matrix} \text{Move to city} \\ \text{Remain in suburbs} \end{matrix} \tag{7}$$

The vectors in (6) and (7) account for all of the population in 2001.[3] Thus

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} + s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

That is,

$$\mathbf{x}_1 = M\mathbf{x}_0 \tag{8}$$

where M is the **migration matrix** determined by the following table:

$$\begin{array}{ccl} \multicolumn{2}{c}{\text{From:}} & \\ \text{City} & \text{Suburbs} & \text{To:} \\ \left[\begin{matrix} .95 \\ .05 \end{matrix}\right. & \left.\begin{matrix} .03 \\ .97 \end{matrix}\right] & \begin{matrix} \text{City} \\ \text{Suburbs} \end{matrix} \end{array}$$

Equation (8) describes how the population changes from 2000 to 2001. If the migration percentages remain constant, then the change from 2001 to 2002 is given by

$$\mathbf{x}_2 = M\mathbf{x}_1$$

and similarly for 2002 to 2003 and subsequent years. In general,

$$\mathbf{x}_{k+1} = M\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots \tag{9}$$

The sequence of vectors $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ describes the population of the city/suburban region over a period of years.

EXAMPLE 3 Compute the population of the region just described for the years 2001 and 2002, given that the population in 2000 was 600,000 in the city and 400,000 in the suburbs.

SOLUTION The initial population in 2000 is $\mathbf{x}_0 = \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$. For 2001,

$$\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

For 2002,

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix} = \begin{bmatrix} 565{,}440 \\ 434{,}560 \end{bmatrix}$$ ■

The model for population movement in (9) is *linear* because the correspondence $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ is a linear transformation. The linearity depends on two facts: the number of people who chose to move from one area to another is *proportional* to the number of people in that area, as shown in (6) and (7), and the cumulative effect of these choices is found by *adding* the movement of people from the different areas.

PRACTICE PROBLEM

Find a matrix A and vectors $\mathbf{x}$ and $\mathbf{b}$ such that the problem in Example 1 amounts to solving the equation $A\mathbf{x} = \mathbf{b}$.

---

[3] For simplicity, we ignore other influences on the population such as births, deaths, and migration into and out of the city/suburban region.

1.10 EXERCISES

1. The container of a breakfast cereal usually lists the number of calories and the amounts of protein, carbohydrate, and fat contained in one serving of the cereal. The amounts for two common cereals are given below. Suppose a mixture of these two cereals is to be prepared that contains exactly 295 calories, 9 g of protein, 48 g of carbohydrate, and 8 g of fat.
   a. Set up a vector equation for this problem. Include a statement of what the variables in your equation represent.
   b. Write an equivalent matrix equation, and then determine if the desired mixture of the two cereals can be prepared.

Nutrition Information per Serving

Nutrient	General Mills Cheerios®	Quaker® 100% Natural Cereal
Calories	110	130
Protein (g)	4	3
Carbohydrate (g)	20	18
Fat (g)	2	5

2. One serving of Shredded Wheat supplies 160 calories, 5 g of protein, 6 g of fiber, and 1 g of fat. One serving of Crispix® supplies 110 calories, 2 g of protein, .1 g of fiber, and .4 g of fat.
   a. Set up a matrix B and a vector $\mathbf{u}$ such that $B\mathbf{u}$ gives the amounts of calories, protein, fiber, and fat contained in a mixture of three servings of Shredded Wheat and two servings of Crispix.
   b. **[M]** Suppose that you want a cereal with more fiber than Crispix but fewer calories than Shredded Wheat. Is it possible for a mixture of the two cereals to supply 130 calories, 3.20 g of protein, 2.46 g of fiber, and .64 g of fat? If so, what is the mixture?

3. After taking a nutrition class, a big Annie's® Mac and Cheese fan decides to improve the levels of protein and fiber in her favorite lunch by adding broccoli and canned chicken. The nutritional information for the foods referred to in this exercise are given in the table below.

Nutrition Information per Serving

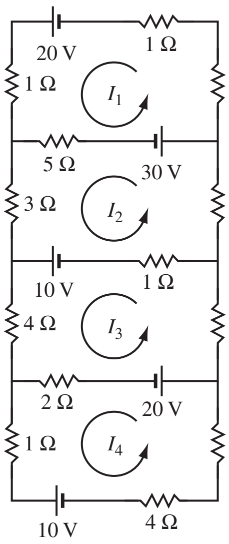
Nutrient	Mac and Cheese	Broccoli	Chicken	Shells
Calories	270	51	70	260
Protein (g)	10	5.4	15	9
Fiber (g)	2	5.2	0	5

   a. **[M]** If she wants to limit her lunch to 400 calories but get 30 g of protein and 10 g of fiber, what proportions of servings of Mac and Cheese, broccoli, and chicken should she use?
   b. **[M]** She found that there was too much broccoli in the proportions from part (a), so she decided to switch from classical Mac and Cheese to Annie's® Whole Wheat Shells and White Cheddar. What proportions of servings of each food should she use to meet the same goals as in part (a)?

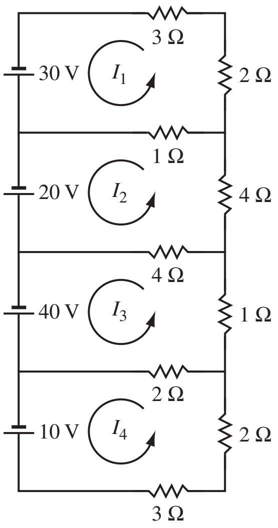
4. The Cambridge Diet supplies .8 g of calcium per day, in addition to the nutrients listed in the Table 1 for Example 1. The amounts of calcium per unit (100 g) supplied by the three ingredients in the Cambridge Diet are as follows: 1.26 g from nonfat milk, .19 g from soy flour, and .8 g from whey. Another ingredient in the diet mixture is isolated soy protein, which provides the following nutrients in each unit: 80 g of protein, 0 g of carbohydrate, 3.4 g of fat, and .18 g of calcium.
   a. Set up a matrix equation whose solution determines the amounts of nonfat milk, soy flour, whey, and isolated soy protein necessary to supply the precise amounts of protein, carbohydrate, fat, and calcium in the Cambridge Diet. State what the variables in the equation represent.
   b. **[M]** Solve the equation in (a) and discuss your answer.

In Exercises 5–8, write a matrix equation that determines the loop currents. **[M]** If MATLAB or another matrix program is available, solve the system for the loop currents.

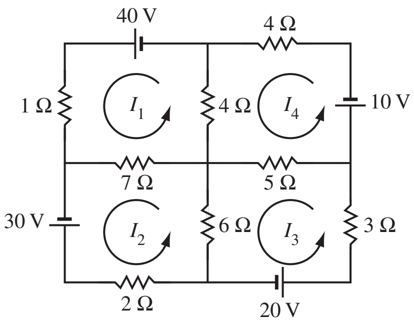
5.



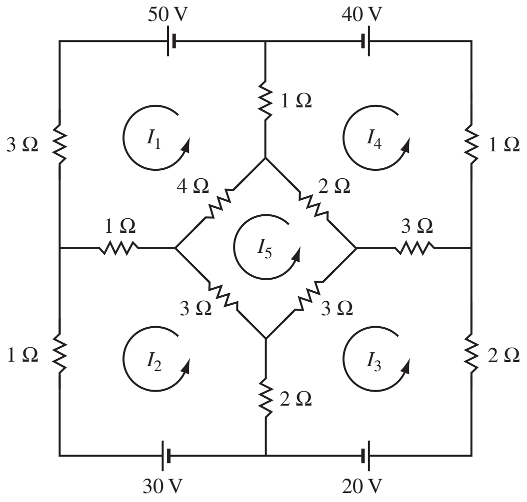
6.



7.



8.



9. In a certain region, about 7% of a city's population moves to the surrounding suburbs each year, and about 5% of the suburban population moves into the city. In 2010, there were 800,000 residents in the city and 500,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2010. Then estimate the populations in the city and in the suburbs two years later, in 2012. (Ignore other factors that might influence the population sizes.)

10. In a certain region, about 6% of a city's population moves to the surrounding suburbs each year, and about 4% of the suburban population moves into the city. In 2010, there were 10,000,000 residents in the city and 800,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2010. Then estimate the populations in the city and in the suburbs two years later, in 2012.

11. In 1994, the population of California was 31,524,000, and the population living in the United States but *outside* California was 228,680,000. During the year, it is estimated that 516,100 persons moved from California to elsewhere in the United States, while 381,262 persons moved to California from elsewhere in the United States.[4]

a. Set up the migration matrix for this situation, using five decimal places for the migration rates into and out of California. Let your work show how you produced the migration matrix.

b. [**M**] Compute the projected populations in the year 2000 for California and elsewhere in the United States, assuming that the migration rates did not change during the 6-year period. (These calculations do not take into account births, deaths, or the substantial migration of persons into California and elsewhere in the United States from other countries.)

12. [**M**] Budget® Rent A Car in Wichita, Kansas has a fleet of about 500 cars, at three locations. A car rented at one location may be returned to any of the three locations. The various fractions of cars returned to the three locations are shown in the matrix below. Suppose that on Monday there are 295 cars at the airport (or rented from there), 55 cars at the east side office, and 150 cars at the west side office. What will be the approximate distribution of cars on Wednesday?

$$\begin{array}{cl}
\text{Cars Rented From:} & \\
\begin{array}{ccc}\text{Airport} & \text{East} & \text{West}\end{array} & \text{Returned To:} \\
\begin{bmatrix} .97 & .05 & .10 \\ .00 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix} & \begin{array}{l}\text{Airport} \\ \text{East} \\ \text{West}\end{array}
\end{array}$$

13. [**M**] Let M and $\mathbf{x}_0$ be as in Example 3.

a. Compute the population vectors $\mathbf{x}_k$ for $k = 1, \ldots, 20$. Discuss what you find.

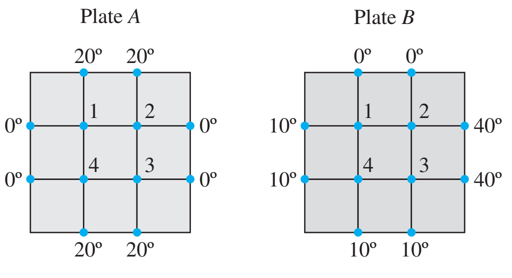
b. Repeat part (a) with an initial population of 350,000 in the city and 650,000 in the suburbs. What do you find?

14. [**M**] Study how changes in boundary temperatures on a steel plate affect the temperatures at interior points on the plate.

a. Begin by estimating the temperatures T_1, T_2, T_3, T_4 at each of the sets of four points on the steel plate shown in the figure. In each case, the value of T_k is approximated by the average of the temperatures at the four closest points. See Exercises 33 and 34 in Section 1.1, where the values (in degrees) turn out to be (20, 27.5, 30, 22.5). How is this list of values related to your results for the points in set (a) and set (b)?

b. Without making any computations, guess the interior temperatures in (a) when the boundary temperatures are all multipled by 3. Check your guess.

c. Finally, make a general conjecture about the correspondence from the list of eight boundary temperatures to the list of four interior temperatures.



(a) (b)

---

[4] Migration data supplied by the Demographic Research Unit of the California State Department of Finance.

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

CHAPTER 1 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. (If true, cite appropriate facts or theorems. If false, explain why or give a counterexample that shows why the statement is not true in every case.

a. Every matrix is row equivalent to a unique matrix in echelon form.

b. Any system of n linear equations in n variables has at most n solutions.

c. If a system of linear equations has two different solutions, it must have infinitely many solutions.

d. If a system of linear equations has no free variables, then it has a unique solution.

e. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ is transformed into $[\,C \quad \mathbf{d}\,]$ by elementary row operations, then the equations $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have exactly the same solution sets.

f. If a system $A\mathbf{x} = \mathbf{b}$ has more than one solution, then so does the system $A\mathbf{x} = \mathbf{0}$.

g. If A is an $m \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some $\mathbf{b}$, then the columns of A span $\mathbb{R}^m$.

h. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ can be transformed by elementary row operations into reduced echelon form, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent.

i. If matrices A and B are row equivalent, they have the same reduced echelon form.

j. The equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution if and only if there are no free variables.

k. If A is an $m \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$, then A has m pivot columns.

l. If an $m \times n$ matrix A has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ in $\mathbb{R}^m$.

m. If an $n \times n$ matrix A has n pivot positions, then the reduced echelon form of A is the $n \times n$ identity matrix.

n. If 3×3 matrices A and B each have three pivot positions, then A can be transformed into B by elementary row operations.

o. If A is an $m \times n$ matrix, if the equation $A\mathbf{x} = \mathbf{b}$ has at least two different solutions, and if the equation $A\mathbf{x} = \mathbf{c}$ is consistent, then the equation $A\mathbf{x} = \mathbf{c}$ has many solutions.

p. If A and B are row equivalent $m \times n$ matrices and if the columns of A span $\mathbb{R}^m$, then so do the columns of B.

q. If none of the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$ is a multiple of one of the other vectors, then S is linearly independent.

r. If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, then $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are not in $\mathbb{R}^2$.

s. In some cases, it is possible for four vectors to span $\mathbb{R}^5$.

t. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^m$, then $-\mathbf{u}$ is in Span$\{\mathbf{u}, \mathbf{v}\}$.

u. If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^2$, then $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

v. If $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, then $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$.

w. Suppose that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are in $\mathbb{R}^5$, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$, and $\mathbf{v}_3$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

x. A linear transformation is a function.

y. If A is a 6×5 matrix, the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ cannot map $\mathbb{R}^5$ onto $\mathbb{R}^6$.

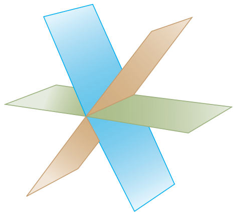
z. If A is an $m \times n$ matrix with m pivot columns, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one mapping.

2. Let a and b represent real numbers. Describe the possible solution sets of the (linear) equation $ax = b$. [*Hint:* The number of solutions depends upon a and b.]

3. The solutions (x, y, z) of a single linear equation

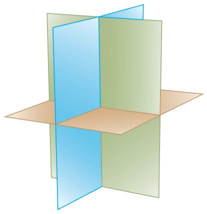
$$ax + by + cz = d$$

form a plane in $\mathbb{R}^3$ when a, b, and c are not all zero. Construct sets of three linear equations whose graphs (a) intersect in a single line, (b) intersect in a single point, and (c) have no points in common. Typical graphs are illustrated in the figure.



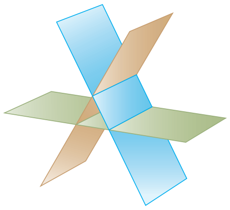
Three planes intersecting in a line

(a)



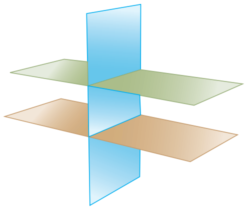
Three planes intersecting in a point

(b)



Three planes with no intersection

(c)



Three planes with no intersection

(c')

4. Suppose the coefficient matrix of a linear system of three equations in three variables has a pivot position in each column. Explain why the system has a unique solution.

5. Determine h and k such that the solution set of the system (i) is empty, (ii) contains a unique solution, and (iii) contains infinitely many solutions.

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. Consider the problem of determining whether the following system of equations is consistent:

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

a. Define appropriate vectors, and restate the problem in terms of linear combinations. Then solve that problem.

b. Define an appropriate matrix, and restate the problem using the phrase "columns of A."

c. Define an appropriate linear transformation T using the matrix in (b), and restate the problem in terms of T.

7. Consider the problem of determining whether the following system of equations is consistent for all 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. Define appropriate vectors, and restate the problem in terms of Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Then solve that problem.

b. Define an appropriate matrix, and restate the problem using the phrase "columns of A."

c. Define an appropriate linear transformation T using the matrix in (b), and restate the problem in terms of T.

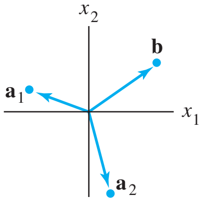
8. Describe the possible echelon forms of the matrix A. Use the notation of Example 1 in Section 1.2.

a. A is a 2×3 matrix whose columns span $\mathbb{R}^2$.

b. A is a 3×3 matrix whose columns span $\mathbb{R}^3$.

9. Write the vector $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ as the sum of two vectors, one on the line $\{(x, y) : y = 2x\}$ and one on the line $\{(x, y) : y = x/2\}$.

10. Let $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}$ be the vectors in $\mathbb{R}^2$ shown in the figure, and let $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$. Does the equation $A\mathbf{x} = \mathbf{b}$ have a solution? If so, is the solution unique? Explain.



11. Construct a 2×3 matrix A, not in echelon form, such that the solution of $A\mathbf{x} = \mathbf{0}$ is a line in $\mathbb{R}^3$.

12. Construct a 2×3 matrix A, not in echelon form, such that the solution of $A\mathbf{x} = \mathbf{0}$ is a plane in $\mathbb{R}^3$.

13. Write the *reduced* echelon form of a 3×3 matrix A such that the first two columns of A are pivot columns and $A\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

14. Determine the value(s) of a such that $\left\{\begin{bmatrix} 1 \\ a \end{bmatrix}, \begin{bmatrix} a \\ a+2 \end{bmatrix}\right\}$ is linearly independent.

15. In (a) and (b), suppose the vectors are linearly independent. What can you say about the numbers $a, \ldots, f$? Justify your answers. [*Hint:* Use a theorem for (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. Use Theorem 7 in Section 1.7 to explain why the columns of the matrix A are linearly independent.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$$

17. Explain why a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{R}^5$ must be linearly independent when $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and $\mathbf{v}_4$ is *not* in Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

18. Suppose $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set in $\mathbb{R}^n$. Show that $\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2\}$ is also linearly independent.

19. Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are distinct points on one line in $\mathbb{R}^3$. The line need not pass through the origin. Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.

20. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and suppose $T(\mathbf{u}) = \mathbf{v}$. Show that $T(-\mathbf{u}) = -\mathbf{v}$.

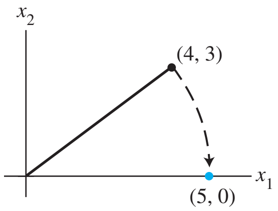
21. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation that reflects each vector through the plane $x_2 = 0$. That is, $T(x_1, x_2, x_3) = (x_1, -x_2, x_3)$. Find the standard matrix of T.

22. Let A be a 3×3 matrix with the property that the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^3$ onto $\mathbb{R}^3$. Explain why the transformation must be one-to-one.

23. A *Givens rotation* is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ used in computer programs to create a zero entry in a vector (usually a column of a matrix). The standard matrix of a Givens rotation in $\mathbb{R}^2$ has the form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \qquad a^2 + b^2 = 1$$

Find a and b such that $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ is rotated into $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$.



A Givens rotation in $\mathbb{R}^2$.

WEB

24. The following equation describes a Givens rotation in $\mathbb{R}^3$. Find a and 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. A large apartment building is to be built using modular construction techniques. The arrangement of apartments on any particular floor is to be chosen from one of three basic floor plans. Plan A has 18 apartments on one floor, including 3 three-bedroom units, 7 two-bedroom units, and 8 one-bedroom units. Each floor of plan B includes 4 three-bedroom units, 4 two-bedroom units, and 8 one-bedroom units. Each floor of plan C includes 5 three-bedroom units, 3 two-bedroom units, and 9 one-bedroom units. Suppose the building contains a total of x_1 floors of plan A, x_2 floors of plan B, and x_3 floors of plan C.

a. What interpretation can be given to the vector $x_1 \begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix}$?

b. Write a formal linear combination of vectors that expresses the total numbers of three-, two-, and one-bedroom apartments contained in the building.

c. [**M**] Is it possible to design the building with exactly 66 three-bedroom units, 74 two-bedroom units, and 136 one-bedroom units? If so, is there more than one way to do it? Explain your answer.

2 Matrix Algebra

INTRODUCTORY EXAMPLE

Computer Models in Aircraft Design

To design the next generation of commercial and military aircraft, engineers at Boeing's Phantom Works use 3D modeling and computational fluid dynamics (CFD). They study the airflow around a virtual airplane to answer important design questions before physical models are created. This has drastically reduced design cycle times and cost—and linear algebra plays a crucial role in the process.

The virtual airplane begins as a mathematical "wire-frame" model that exists only in computer memory and on graphics display terminals. (A model of a Boeing 777 is shown.) This mathematical model organizes and influences each step of the design and manufacture of the airplane—both the exterior and interior. The CFD analysis concerns the exterior surface.

Although the finished skin of a plane may seem smooth, the geometry of the surface is complicated. In addition to wings and a fuselage, an aircraft has nacelles, stabilizers, slats, flaps, and ailerons. The way air flows around these structures determines how the plane moves through the sky. Equations that describe the airflow are complicated, and they must account for engine intake, engine exhaust, and the wakes left by the wings of the plane. To study the airflow, engineers need a highly refined description of the plane's surface.

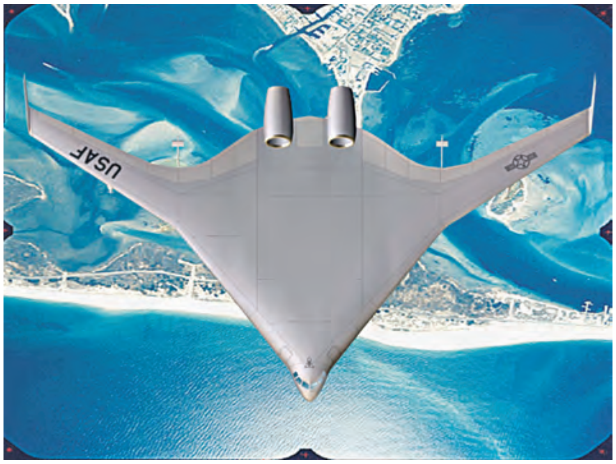
A computer creates a model of the surface by first superimposing a three-dimensional grid of "boxes" on the original wire-frame model. Boxes in this grid lie either completely inside or completely outside the plane, or they intersect the surface of the plane. The computer selects the boxes that intersect the surface and subdivides them, retaining only the smaller boxes that still intersect the surface. The subdividing process is repeated until the grid is extremely fine. A typical grid can include over 400,000 boxes.

The process for finding the airflow around the plane involves repeatedly solving a system of linear equations $A\mathbf{x} = \mathbf{b}$ that may involve up to 2 million equations and variables. The vector $\mathbf{b}$ changes each time, based on data from the grid and solutions of previous equations. Using the fastest computers available commercially, a Phantom Works team can spend from a few hours to several days setting up and solving a single airflow problem. After the team analyzes the solution, they may make small changes to the airplane surface and begin the whole process again. Thousands of CFD runs may be required.

This chapter presents two important concepts that assist in the solution of such massive systems of equations:

- *Partitioned matrices:* A typical CFD system of equations has a "sparse" coefficient matrix with mostly zero entries. Grouping the variables correctly leads to a partitioned matrix with many zero blocks. Section 2.4 introduces such matrices and describes some of their applications.

- *Matrix factorizations:* Even when written with partitioned matrices, the system of equations is complicated. To further simplify the computations, the CFD software at Boeing uses what is called an LU factorization of the coefficient matrix. Section 2.5 discusses LU and other useful matrix factorizations. Further details about factorizations appear at several points later in the text.



Modern CFD has revolutionized wing design. The Boeing Blended Wing Body is in design for the year 2020 or sooner.

To analyze a solution of an airflow system, engineers want to visualize the airflow over the surface of the plane. They use computer graphics, and linear algebra provides the engine for the graphics. The wire-frame model of the plane's surface is stored as data in many matrices. Once the image has been rendered on a computer screen, engineers can change its scale, zoom in or out of small regions, and rotate the image to see parts that may be hidden from view. Each of these operations is accomplished by appropriate matrix multiplications. Section 2.7 explains the basic ideas.

WEB

Our ability to analyze and solve equations will be greatly enhanced when we can perform algebraic operations with matrices. Furthermore, the definitions and theorems in this chapter provide some basic tools for handling the many applications of linear algebra that involve two or more matrices. For square matrices, the Invertible Matrix Theorem in Section 2.3 ties together most of the concepts treated earlier in the text. Sections 2.4 and 2.5 examine partitioned matrices and matrix factorizations, which appear in most modern uses of linear algebra. Sections 2.6 and 2.7 describe two interesting applications of matrix algebra, to economics and to computer graphics.

2.1 MATRIX OPERATIONS

If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A. See Fig. 1. For instance, the $(3, 2)$-entry is the number a_{32} in the third row, second column. Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$. Often, these columns are denoted by $\mathbf{a}_1, \dots, \mathbf{a}_n$, and the matrix A is written as

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$$

Observe that the number a_{ij} is the ith entry (from the top) of the jth column vector $\mathbf{a}_j$.

The **diagonal entries** in an $m \times n$ matrix $A = [\,a_{ij}\,]$ are $a_{11}, a_{22}, a_{33}, \dots$, and they form the **main diagonal** of A. A **diagonal matrix** is a square $n \times n$ matrix whose nondiagonal entries are zero. An example is the $n \times n$ identity matrix, I_n. An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0. The size of a zero matrix is usually clear from the context.

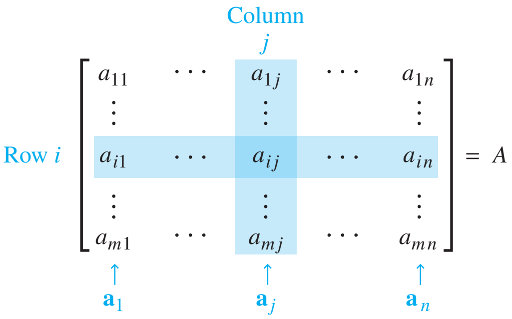


FIGURE 1 Matrix notation.

Sums and Scalar Multiples

The arithmetic for vectors described earlier has a natural extension to matrices. We say that two matrices are **equal** if they have the same size (i.e., the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal. If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B. Since vector addition of the columns is done entrywise, each entry in $A + B$ is the sum of the corresponding entries in A and B. The sum $A + B$ is defined only when A and B are the same size.

EXAMPLE 1 Let

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

Then

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

but $A + C$ is not defined because A and C have different sizes. ■

If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A. As with vectors, $-A$ stands for $(-1)A$, and $A - B$ is the same as $A + (-1)B$.

EXAMPLE 2 If A and B are the matrices in Example 1, then

$$2B = 2\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix}$$

$$A - 2B = \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix}$$ ■

It was unnecessary in Example 2 to compute $A - 2B$ as $A + (-1)2B$ because the usual rules of algebra apply to sums and scalar multiples of matrices, as the following theorem shows.

THEOREM 1

Let A, B, and C be matrices of the same size, and let r and s be scalars.

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. $r(sA) = (rs)A$

Each equality in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because A, B, and C are equal in size. The equality of columns follows immediately from analogous properties of vectors. For instance, if the jth columns of A, B, and C are $\mathbf{a}_j$, $\mathbf{b}_j$, and $\mathbf{c}_j$, respectively, then the jth columns of $(A+B)+C$ and $A+(B+C)$ are

$$(\mathbf{a}_j+\mathbf{b}_j)+\mathbf{c}_j \quad \text{and} \quad \mathbf{a}_j+(\mathbf{b}_j+\mathbf{c}_j)$$

respectively. Since these two vector sums are equal for each j, property (b) is verified.

Because of the associative property of addition, we can simply write $A+B+C$ for the sum, which can be computed either as $(A+B)+C$ or as $A+(B+C)$. The same applies to sums of four or more matrices.

Matrix Multiplication

When a matrix B multiplies a vector $\mathbf{x}$, it transforms $\mathbf{x}$ into the vector $B\mathbf{x}$. If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(B\mathbf{x})$. See Fig. 2.

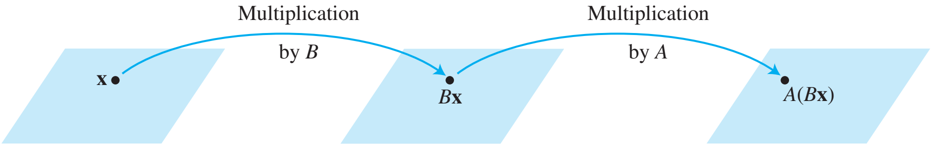


FIGURE 2 Multiplication by B and then A.

Thus $A(B\mathbf{x})$ is produced from $\mathbf{x}$ by a *composition* of mappings—the linear transformations studied in Section 1.8. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB, so that

$$A(B\mathbf{x}) = (AB)\mathbf{x} \tag{1}$$

See Fig. 3.

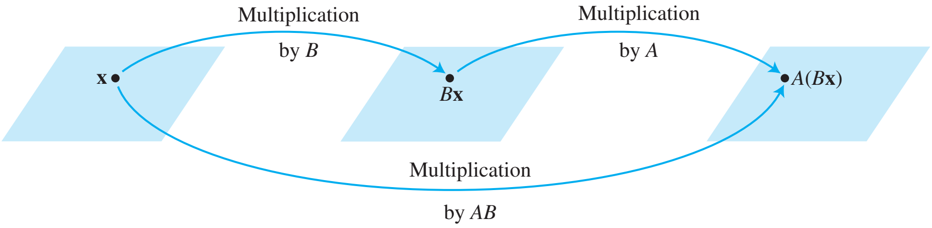


FIGURE 3 Multiplication by AB.

If A is $m\times n$, B is $n\times p$, and $\mathbf{x}$ is in $\mathbb{R}^p$, denote the columns of B by $\mathbf{b}_1,\ldots,\mathbf{b}_p$ and the entries in $\mathbf{x}$ by $x_1,\ldots,x_p$. Then

$$B\mathbf{x} = x_1\mathbf{b}_1+\cdots+x_p\mathbf{b}_p$$

By the linearity of multiplication by A,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1)+\cdots+A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1+\cdots+x_pA\mathbf{b}_p \end{aligned}$$

The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \dots, A\mathbf{b}_p$, using the entries in $\mathbf{x}$ as weights. In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = [\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p \,]\mathbf{x}$$

Thus multiplication by $[\, A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_p \,]$ transforms $\mathbf{x}$ into $A(B\mathbf{x})$. We have found the matrix we sought!

DEFINITION

If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$. That is,

$$AB = A\begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_p \end{bmatrix}$$

This definition makes equation (1) true for all $\mathbf{x}$ in $\mathbb{R}^p$. Equation (1) proves that the composite mapping in Fig. 3 is a linear transformation and that its standard matrix is AB. *Multiplication of matrices corresponds to composition of linear transformations.*

EXAMPLE 3 Compute AB, where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

SOLUTION Write $B = [\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,]$, and compute:

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix}, \quad A\mathbf{b}_3 = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}\begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

$$= \begin{bmatrix} 11 \\ -1 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 0 \\ 13 \end{bmatrix} \qquad\qquad = \begin{bmatrix} 21 \\ -9 \end{bmatrix}$$

Then

$$AB = A[\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,] = \begin{array}{c} \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{array} \end{array}$$

■

Notice that since the first column of AB is $A\mathbf{b}_1$, this column is a linear combination of the columns of A using the entries in $\mathbf{b}_1$ as weights. A similar statement is true for each column of AB.

Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B.

Obviously, the number of columns of A must match the number of rows in B in order for a linear combination such as $A\mathbf{b}_1$ to be defined. Also, the definition of AB shows that *AB has the same number of rows as A and the same number of columns as B.*

EXAMPLE 4 If A is a 3×5 matrix and B is a 5×2 matrix, what are the sizes of AB and BA, if they are defined?

SOLUTION Since A has 5 columns and B has 5 rows, the product AB is defined and is a 3×2 matrix:

$$\begin{array}{cccc} A & B & & AB \\ \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} & = & \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \end{array}$$

$$\underbrace{3 \times \underbrace{5 \qquad 5}_{\text{Match}} \times 2}_{\text{Size of } AB} \qquad 3 \times 2$$

The product BA is *not* defined because the 2 columns of B do not match the 3 rows of A. ■

The definition of AB is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in AB when working small problems by hand.

ROW–COLUMN RULE FOR COMPUTING *AB*

If the product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B. If $(AB)_{ij}$ denotes the (i, j)-entry in AB, and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

To verify this rule, let $B = [\,\mathbf{b}_1 \ \cdots \ \mathbf{b}_p\,]$. Column j of AB is $A\mathbf{b}_j$, and we can compute $A\mathbf{b}_j$ by the row–vector rule for computing $A\mathbf{x}$ from Section 1.4. The ith entry in $A\mathbf{b}_j$ is the sum of the products of corresponding entries from row i of A and the vector $\mathbf{b}_j$, which is precisely the computation described in the rule for computing the (i, j)-entry of AB.

EXAMPLE 5 Use the row–column rule to compute two of the entries in AB for the matrices in Example 3. An inspection of the numbers involved will make it clear how the two methods for calculating AB produce the same matrix.

SOLUTION To find the entry in row 1 and column 3 of AB, consider row 1 of A and column 3 of B. Multiply corresponding entries and add the results, as shown below:

$$AB = \begin{matrix} \rightarrow \\ \end{matrix}\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & \overset{\downarrow}{6} \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

For the entry in row 2 and column 2 of AB, use row 2 of A and column 2 of B:

$$\begin{matrix} \\ \rightarrow \end{matrix}\begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & \overset{\downarrow}{3} & 6 \\ 1 & -2 & 3 \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$

■

EXAMPLE 6 Find the entries in the second row of AB, where

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

SOLUTION By the row–column rule, the entries of the second row of AB come from row 2 of A (and the columns of B):

$$\begin{array}{r} \\ \rightarrow \\ \\ \\ \end{array}\begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} \overset{\downarrow}{4} & \overset{\downarrow}{-6} \\ 7 & 1 \\ 3 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} \square & \square \\ -4 + 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix} \qquad \blacksquare$$

Notice that since Example 6 requested only the second row of AB, we could have written just the second row of A to the left of B and computed

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

This observation about rows of AB is true in general and follows from the row–column rule. Let $\text{row}_i(A)$ denote the ith row of a matrix A. Then

$$\boxed{\text{row}_i(AB) = \text{row}_i(A) \cdot B} \tag{2}$$

Properties of Matrix Multiplication

The following theorem lists the standard properties of matrix multiplication. Recall that I_m represents the $m \times m$ identity matrix and $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

THEOREM 2

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B + C) = AB + AC$ (left distributive law)
c. $(B + C)A = BA + CA$ (right distributive law)
d. $r(AB) = (rA)B = A(rB)$
   for any scalar r
e. $I_mA = A = AI_n$ (identity for matrix multiplication)

PROOF Properties (b)–(e) are considered in the exercises. Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative. Here is another proof of (a) that rests on the "column definition" of the product of two matrices. Let

$$C = [\,\mathbf{c}_1 \;\; \cdots \;\; \mathbf{c}_p\,]$$

By the definition of matrix multiplication,

$$\begin{aligned} BC &= [\, B\mathbf{c}_1 \quad \cdots \quad B\mathbf{c}_p \,] \\ A(BC) &= [\, A(B\mathbf{c}_1) \quad \cdots \quad A(B\mathbf{c}_p) \,] \end{aligned}$$

Recall from equation (1) that the definition of AB makes $A(B\mathbf{x}) = (AB)\mathbf{x}$ for all $\mathbf{x}$, so

$$A(BC) = [\, (AB)\mathbf{c}_1 \quad \cdots \quad (AB)\mathbf{c}_p \,] = (AB)C \qquad \blacksquare$$

The associative and distributive laws in Theorems 1 and 2 say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write ABC for the product, which can be computed either as $A(BC)$ or as $(AB)C$.[1] Similarly, a product $ABCD$ of four matrices can be computed as $A(BCD)$ or $(ABC)D$ or $A(BC)D$, and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because AB and BA are usually not the same. This is not surprising, because the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B. The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A. If $AB = BA$, we say that A and B **commute** with one another.

EXAMPLE 7 Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Show that these matrices do not commute. That is, verify that $AB \neq BA$.

SOLUTION

$$\begin{aligned} AB &= \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix} \\ BA &= \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix} \end{aligned} \qquad \blacksquare$$

Example 7 illustrates the first of the following list of important differences between matrix algebra and the ordinary algebra of real numbers. See Exercises 9–12 for examples of these situations.

WARNINGS:

1. In general, $AB \neq BA$.
2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)
3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

Powers of a Matrix

WEB

If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k

---

[1] When B is square and C has fewer columns than A has rows, it is more efficient to compute $A(BC)$ than $(AB)C$.

copies of A:

$$A^k = \underbrace{A \cdots A}_{k}$$

If A is nonzero and if $\mathbf{x}$ is in $\mathbb{R}^n$, then $A^k\mathbf{x}$ is the result of left-multiplying $\mathbf{x}$ by A repeatedly k times. If $k = 0$, then $A^0\mathbf{x}$ should be $\mathbf{x}$ itself. Thus A^0 is interpreted as the identity matrix. Matrix powers are useful in both theory and applications (Sections 2.6, 4.9, and later in the text).

The Transpose of a Matrix

Given an $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

EXAMPLE 8 Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} -5 & 2 \\ 1 & -3 \\ 0 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ -3 & 5 & -2 & 7 \end{bmatrix}$$

Then

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix}$$ ■

THEOREM 3

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. For any scalar r, $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

Proofs of (a)–(c) are straightforward and are omitted. For (d), see Exercise 33. Usually, $(AB)^T$ is not equal to $A^T B^T$, even when A and B have sizes such that the product $A^T B^T$ is defined.

The generalization of Theorem 3(d) to products of more than two factors can be stated in words as follows:

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

The exercises contain numerical examples that illustrate properties of transposes.

NUMERICAL NOTES

1. The fastest way to obtain AB on a computer depends on the way in which the computer stores matrices in its memory. The standard high-performance algorithms, such as in LAPACK, calculate AB by columns, as in our definition of the product. (A version of LAPACK written in C++ calculates AB by rows.)
2. The definition of AB lends itself well to parallel processing on a computer. The columns of B are assigned individually or in groups to different processors, which independently and hence simultaneously compute the corresponding columns of AB.

PRACTICE PROBLEMS

1. Since vectors in $\mathbb{R}^n$ may be regarded as $n \times 1$ matrices, the properties of transposes in Theorem 3 apply to vectors, too. Let
$$A = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$
Compute $(A\mathbf{x})^T$, $\mathbf{x}^T A^T$, $\mathbf{x}\mathbf{x}^T$, and $\mathbf{x}^T\mathbf{x}$. Is $A^T\mathbf{x}^T$ defined?
2. Let A be a 4×4 matrix and let $\mathbf{x}$ be a vector in $\mathbb{R}^4$. What is the fastest way to compute $A^2\mathbf{x}$? Count the multiplications.

2.1 EXERCISES

In Exercises 1 and 2, compute each matrix sum or product if it is defined. If an expression is undefined, explain why. Let

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 4 & -5 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -5 & 1 \\ 1 & -4 & -3 \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

1. $-2A$, $B - 2A$, AC, CD
2. $A + 3B$, $2C - 3E$, DB, EC

In the rest of this exercise set and in those to follow, assume that each matrix expression is defined. That is, the sizes of the matrices (and vectors) involved "match" appropriately.

3. Let $A = \begin{bmatrix} 2 & -5 \\ 3 & -2 \end{bmatrix}$. Compute $3I_2 - A$ and $(3I_2)A$.
4. Compute $A - 5I_3$ and $(5I_3)A$, where
$$A = \begin{bmatrix} 5 & -1 & 3 \\ -4 & 3 & -6 \\ -3 & 1 & 2 \end{bmatrix}.$$

In Exercises 5 and 6, compute the product AB in two ways: (a) by the definition, where $A\mathbf{b}_1$ and $A\mathbf{b}_2$ are computed separately, and (b) by the row–column rule for computing 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. If a matrix A is 5×3 and the product AB is 5×7, what is the size of B?
8. How many rows does B have if BC is a 5×4 matrix?
9. Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 9 \\ -3 & k \end{bmatrix}$. What value(s) of k, if any, will make $AB = BA$?
10. Let $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.
11. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Compute AD and DA. Explain how the columns or rows of A change when A is multiplied by D on the right or on the left. Find a 3×3 matrix B, not the identity matrix or the zero matrix, such that $AB = BA$.
12. Let $A = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}$. Construct a 2×2 matrix B such that AB is the zero matrix. Use two different nonzero columns for B.

13. Let $\mathbf{r}_1, \dots, \mathbf{r}_p$ be vectors in $\mathbb{R}^n$, and let Q be an $m \times n$ matrix. Write the matrix $[\, Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_p \,]$ as a *product* of two matrices (neither of which is an identity matrix).

14. Let U be the 3×2 cost matrix described in Example 6 in Section 1.8. The first column of U lists the costs per dollar of output for manufacturing product B, and the second column lists the costs per dollar of output for product C. (The costs are categorized as materials, labor, and overhead.) Let $\mathbf{q}_1$ be a vector in $\mathbb{R}^2$ that lists the output (measured in dollars) of products B and C manufactured during the first quarter of the year, and let $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{q}_4$ be the analogous vectors that list the amounts of products B and C manufactured in the second, third, and fourth quarters, respectively. Give an economic description of the data in the matrix UQ, where $Q = [\, \mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3 \quad \mathbf{q}_4 \,]$.

Exercises 15 and 16 concern arbitrary matrices A, B, and C for which the indicated sums and products are defined. Mark each statement True or False. Justify each answer.

15. a. If A and B are 2×2 matrices with columns $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}_1$, $\mathbf{b}_2$, respectively, then $AB = [\, \mathbf{a}_1\mathbf{b}_1 \quad \mathbf{a}_2\mathbf{b}_2 \,]$.

b. Each column of AB is a linear combination of the columns of B using weights from the corresponding column of A.

c. $AB + AC = A(B + C)$

d. $A^T + B^T = (A + B)^T$

e. The transpose of a product of matrices equals the product of their transposes in the same order.

16. a. The first row of AB is the first row of A multiplied on the right by B.

b. If A and B are 3×3 matrices and $B = [\, \mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3 \,]$, then $AB = [\, A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3 \,]$.

c. If A is an $n \times n$ matrix, then $(A^2)^T = (A^T)^2$

d. $(ABC)^T = C^T A^T B^T$

e. The transpose of a sum of matrices equals the sum of their transposes.

17. If $A = \begin{bmatrix} 1 & -3 \\ -3 & 5 \end{bmatrix}$ and $AB = \begin{bmatrix} -3 & -11 \\ 1 & 17 \end{bmatrix}$, determine the first and second columns of B.

18. Suppose the third column of B is all zeros. What can be said about the third column of AB?

19. Suppose the third column of B is the sum of the first two columns. What can be said about the third column of AB? Why?

20. Suppose the first two columns, $\mathbf{b}_1$ and $\mathbf{b}_2$, of B are equal. What can be said about the columns of AB? Why?

21. Suppose the last column of AB is entirely zeros but B itself has no column of zeros. What can be said about the columns of A?

22. Show that if the columns of B are linearly dependent, then so are the columns of AB.

23. Suppose $CA = I_n$ (the $n \times n$ identity matrix). Show that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A cannot have more columns than rows.

24. Suppose A is a $3 \times n$ matrix whose columns span $\mathbb{R}^3$. Explain how to construct an $n \times 3$ matrix D such that $AD = I_3$.

25. Suppose A is an $m \times n$ matrix and there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. Prove that $m = n$ and $C = D$. [*Hint:* Think about the product CAD.]

26. Suppose $AD = I_m$ (the $m \times m$ identity matrix). Show that for any $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution. [*Hint:* Think about the equation $AD\mathbf{b} = \mathbf{b}$.] Explain why A cannot have more rows than columns.

In Exercises 27 and 28, view vectors in $\mathbb{R}^n$ as $n \times 1$ matrices. For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, called the **scalar product**, or **inner product**, of $\mathbf{u}$ and $\mathbf{v}$. It is usually written as a single real number without brackets. The matrix product $\mathbf{u}\mathbf{v}^T$ is an $n \times n$ matrix, called the **outer product** of $\mathbf{u}$ and $\mathbf{v}$. The products $\mathbf{u}^T\mathbf{v}$ and $\mathbf{u}\mathbf{v}^T$ will appear later in the text.

27. Let $\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ -5 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Compute $\mathbf{u}^T\mathbf{v}$, $\mathbf{v}^T\mathbf{u}$, $\mathbf{u}\mathbf{v}^T$, and $\mathbf{v}\mathbf{u}^T$.

28. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^n$, how are $\mathbf{u}^T\mathbf{v}$ and $\mathbf{v}^T\mathbf{u}$ related? How are $\mathbf{u}\mathbf{v}^T$ and $\mathbf{v}\mathbf{u}^T$ related?

29. Prove Theorem 2(b) and 2(c). Use the row–column rule. The (i, j)-entry in $A(B + C)$ can be written as

$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$

or

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

30. Prove Theorem 2(d). [*Hint:* The (i, j)-entry in $(rA)B$ is $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$.]

31. Show that $I_m A = A$ where A is an $m \times n$ matrix. Assume $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

32. Show that $AI_n = A$ when A is an $m \times n$ matrix. [*Hint:* Use the (column) definition of AI_n.]

33. Prove Theorem 3(d). [*Hint:* Consider the jth row of $(AB)^T$.]

34. Give a formula for $(AB\mathbf{x})^T$, where $\mathbf{x}$ is a vector and A and B are matrices of appropriate sizes.

35. [**M**] Read the documentation for your matrix program, and write the commands that will produce the following matrices (without keying in each entry of the matrix).

a. A 4×5 matrix of zeros

b. A 5×3 matrix of ones

c. The 5×5 identity matrix

d. A 4×4 diagonal matrix, with diagonal entries 3, 4, 2, 5

A useful way to test new ideas in matrix algebra, or to make conjectures, is to make calculations with matrices selected at random. Checking a property for a few matrices does not prove that the property holds in general, but it makes the property more believable. Also, if the property is actually false, making a few calculations may help to discover this.

36. [**M**] Write the command(s) that will create a 5×6 matrix with random entries. In what range of numbers do the entries lie? Tell how to create a 4×4 matrix with random integer entries between -9 and 9. [*Hint:* If x is a random number such that $0 < x < 1$, then $-9.5 < 19(x - .5) < 9.5$.]

37. [**M**] Construct random 4×4 matrices A and B to test whether $AB = BA$. The best way to do this is to compute $AB - BA$ and check whether this difference is the zero matrix. Then test $AB - BA$ for three more pairs of random 4×4 matrices. Report your conclusions.

38. [**M**] Construct a random 5×5 matrix A and test whether $(A + I)(A - I) = A^2 - I$. The best way to do this is to compute $(A + I)(A - I) - (A^2 - I)$ and verify that this difference is the zero matrix. Do this for three random matrices. Then test $(A + B)(A - B) = A^2 - B^2$ the same way for three pairs of random 4×4 matrices. Report your conclusions.

39. [**M**] Use at least three pairs of random 4×4 matrices A and B to test the equalities $(A + B)^T = A^T + B^T$ and $(AB)^T = B^TA^T$, as well as $(AB)^T = A^TB^T$. (See Exercise 37.) Report your conclusions. [*Note*: Most matrix programs use A' for A^T.]

40. [**M**] Let

$$S = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Compute S^k for $k = 2, \ldots, 6$.

41. [**M**] Describe in words what happens when A^5, A^{10}, A^{20}, and A^{30} are computed for

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \\ 1/4 & 1/6 & 7/12 \end{bmatrix}$$

SOLUTIONS TO PRACTICE PROBLEMS

1. $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. So $(A\mathbf{x})^T = \begin{bmatrix} -4 & 2 \end{bmatrix}$. Also,

$$\mathbf{x}^TA^T = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -4 & 2 \end{bmatrix}.$$

The quantities $(A\mathbf{x})^T$ and $\mathbf{x}^TA^T$ are equal, by Theorem 3(d). Next,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 25 + 9 \end{bmatrix} = 34$$

A 1×1 matrix such as $\mathbf{x}^T\mathbf{x}$ is usually written without the brackets. Finally, $A^T\mathbf{x}^T$ is not defined, because $\mathbf{x}^T$ does not have two rows to match the two columns of A^T.

2. The fastest way to compute $A^2\mathbf{x}$ is to compute $A(A\mathbf{x})$. The product $A\mathbf{x}$ requires 16 multiplications, 4 for each entry, and $A(A\mathbf{x})$ requires 16 more. In contrast, the product A^2 requires 64 multiplications, 4 for each of the 16 entries in A^2. After that, $A^2\mathbf{x}$ takes 16 more multiplications, for a total of 80.

2.2 THE INVERSE OF A MATRIX

Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers. This section investigates the matrix analogue of the reciprocal, or multiplicative inverse, of a nonzero number.

Recall that the multiplicative inverse of a number such as 5 is 1/5 or 5^{-1}. This inverse satisfies the equations

$$5^{-1} \cdot 5 = 1 \quad \text{and} \quad 5 \cdot 5^{-1} = 1$$

The matrix generalization requires *both* equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative. Furthermore, a full generalization is possible only if the matrices involved are square.[1]

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A. In fact, C is uniquely determined by A, because if B were another inverse of A, then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

EXAMPLE 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}\begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

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

Thus $C = A^{-1}$. ■

Here is a simple formula for the inverse of a 2×2 matrix, along with a test to tell if the inverse exists.

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A, and we write

$$\det A = ad - bc$$

Theorem 4 says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

---

[1] One could say that an $m \times n$ matrix A is invertible if there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. However, these equations imply that A is square and $C = D$. Thus A is invertible as defined above. See Exercises 23–25 in Section 2.1.

EXAMPLE 2 Find the inverse of $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

SOLUTION Since $\det A = 3(6) - 4(5) = -2 \neq 0$, A is invertible, and

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \quad ■$$

Invertible matrices are indispensable in linear algebra—mainly for algebraic calculations and formula derivations, as in the next theorem. There are also occasions when an inverse matrix provides insight into a mathematical model of a real-life situation, as in Example 3, below.

THEOREM 5

If A is an invertible $n \times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

PROOF Take any $\mathbf{b}$ in $\mathbb{R}^n$. A solution exists because if $A^{-1}\mathbf{b}$ is substituted for $\mathbf{x}$, then $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. So $A^{-1}\mathbf{b}$ is a solution. To prove that the solution is unique, show that if $\mathbf{u}$ is any solution, then $\mathbf{u}$, in fact, must be $A^{-1}\mathbf{b}$. Indeed, if $A\mathbf{u} = \mathbf{b}$, we can multiply both sides by A^{-1} and obtain

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \quad I\mathbf{u} = A^{-1}\mathbf{b}, \quad \text{and} \quad \mathbf{u} = A^{-1}\mathbf{b} \quad ■$$

EXAMPLE 3 A horizontal elastic beam is supported at each end and is subjected to forces at points 1, 2, 3, as shown in Fig. 1. Let $\mathbf{f}$ in $\mathbb{R}^3$ list the forces at these points, and let $\mathbf{y}$ in $\mathbb{R}^3$ list the amounts of deflection (that is, movement) of the beam at the three points. Using Hooke's law from physics, it can be shown that

$$\mathbf{y} = D\mathbf{f}$$

where D is a *flexibility matrix*. Its inverse is called the *stiffness matrix*. Describe the physical significance of the columns of D and D^{-1}.

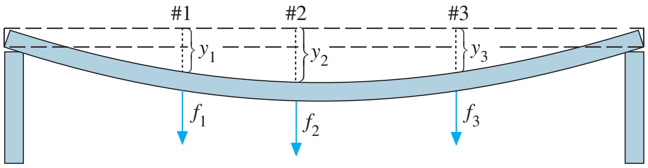


FIGURE 1 Deflection of an elastic beam.

SOLUTION Write $I_3 = [\,\mathbf{e}_1 \;\; \mathbf{e}_2 \;\; \mathbf{e}_3\,]$ and observe that

$$D = DI_3 = [\,D\mathbf{e}_1 \;\; D\mathbf{e}_2 \;\; D\mathbf{e}_3\,]$$

Interpret the vector $\mathbf{e}_1 = (1, 0, 0)$ as a unit force applied downward at point 1 on the beam (with zero force at the other two points). Then $D\mathbf{e}_1$, the first column of D, lists the beam deflections due to a unit force at point 1. Similar descriptions apply to the second and third columns of D.

To study the stiffness matrix D^{-1}, observe that the equation $\mathbf{f} = D^{-1}\mathbf{y}$ computes a force vector $\mathbf{f}$ when a deflection vector $\mathbf{y}$ is given. Write

$$D^{-1} = D^{-1}I_3 = [\,D^{-1}\mathbf{e}_1 \;\; D^{-1}\mathbf{e}_2 \;\; D^{-1}\mathbf{e}_3\,]$$

Now interpret $\mathbf{e}_1$ as a deflection vector. Then $D^{-1}\mathbf{e}_1$ lists the forces that create the deflection. That is, the first column of D^{-1} lists the forces that must be applied at the three points to produce a unit deflection at point 1 and zero deflections at the other points. Similarly, columns 2 and 3 of D^{-1} list the forces required to produce unit deflections at points 2 and 3, respectively. In each column, one or two of the forces must be negative (point upward) to produce a unit deflection at the desired point and zero deflections at the other two points. If the flexibility is measured, for example, in inches of deflection per pound of load, then the stiffness matrix entries are given in pounds of load per inch of deflection. ■

The formula in Theorem 5 is seldom used to solve an equation $A\mathbf{x} = \mathbf{b}$ numerically because row reduction of $[\,A \quad \mathbf{b}\,]$ is nearly always faster. (Row reduction is usually more accurate, too, when computations involve rounding off numbers.) One possible exception is the 2×2 case. In this case, mental computations to solve $A\mathbf{x} = \mathbf{b}$ are sometimes easier using the formula for A^{-1}, as in the next example.

EXAMPLE 4 Use the inverse of the matrix A in Example 2 to solve the system

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

SOLUTION This system is equivalent to $A\mathbf{x} = \mathbf{b}$, so

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}$$ ■

The next theorem provides three useful facts about invertible matrices.

THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and

$$(A^{-1})^{-1} = A$$

b. If A and B are $n \times n$ invertible matrices, then so is AB, and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. If A is an invertible matrix, then so is A^T, and the inverse of A^T is the transpose of A^{-1}. That is,

$$(A^T)^{-1} = (A^{-1})^T$$

PROOF To verify statement (a), find a matrix C such that

$$A^{-1}C = I \quad \text{and} \quad CA^{-1} = I$$

In fact, these equations are satisfied with A in place of C. Hence A^{-1} is invertible, and A is its inverse. Next, to prove statement (b), compute:

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$. For statement (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. Similarly, $A^T(A^{-1})^T = I^T = I$. Hence A^T is invertible, and its inverse is $(A^{-1})^T$. ■

The following generalization of Theorem 6(b) is needed later.

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.

There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by *watching the row reduction of* A *to* I.

Elementary Matrices

An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix. The next example illustrates the three kinds of elementary matrices.

EXAMPLE 5 Let

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compute E_1A, E_2A, and E_3A, and describe how these products can be obtained by elementary row operations on A.

SOLUTION Verify that

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g-4a & h-4b & i-4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

Addition of -4 times row 1 of A to row 3 produces E_1A. (This is a row replacement operation.) An interchange of rows 1 and 2 of A produces E_2A, and multiplication of row 3 of A by 5 produces E_3A. ■

Left-multiplication (that is, multiplication on the left) by E_1 in Example 5 has the same effect on any $3 \times n$ matrix. It adds -4 times row 1 to row 3. In particular, since $E_1 \cdot I = E_1$, we see that E_1 *itself* is produced by this same row operation on the identity. Thus Example 5 illustrates the following general fact about elementary matrices. See Exercises 27 and 28.

> If an elementary row operation is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is created by performing the same row operation on I_m.

Since row operations are reversible, as shown in Section 1.1, elementary matrices are invertible, for if E is produced by a row operation on I, then there is another row operation of the same type that changes E back into I. Hence there is an elementary matrix F such that $FE = I$. Since E and F correspond to reverse operations, $EF = I$, too.

Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

EXAMPLE 6 Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

SOLUTION To transform E_1 into I, add $+4$ times row 1 to row 3. The elementary matrix that does this is

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$

■

The following theorem provides the best way to "visualize" an invertible matrix, and the theorem leads immediately to a method for finding the inverse of a matrix.

THEOREM 7

An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n, and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.

PROOF Suppose that A is invertible. Then, since the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ (Theorem 5), A has a pivot position in every row (Theorem 4 in Section 1.4). Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n. That is, $A \sim I_n$.

Now suppose, conversely, that $A \sim I_n$. Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices $E_1, \ldots, E_p$ such that

$$A \sim E_1 A \sim E_2(E_1 A) \sim \cdots \sim E_p(E_{p-1} \cdots E_1 A) = I_n$$

That is,

$$E_p \cdots E_1 A = I_n \tag{1}$$

Since the product $E_p \cdots E_1$ of invertible matrices is invertible, (1) leads to

$$\begin{aligned} (E_p \cdots E_1)^{-1}(E_p \cdots E_1)A &= (E_p \cdots E_1)^{-1} I_n \\ A &= (E_p \cdots E_1)^{-1} \end{aligned}$$

Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,

$$A^{-1} = [\,(E_p \cdots E_1)^{-1}\,]^{-1} = E_p \cdots E_1$$

Then $A^{-1} = E_p \cdots E_1 \cdot I_n$, which says that A^{-1} results from applying $E_1, \ldots, E_p$ successively to I_n. This is the same sequence in (1) that reduced A to I_n. ■

An Algorithm for Finding A^{-1}

If we place A and I side-by-side to form an augmented matrix $[\,A \quad I\,]$, then row operations on this matrix produce identical operations on A and on I. By Theorem 7, either there are row operations that transform A to I_n and I_n to A^{-1} or else A is not invertible.