Linear Algebra with Applications

Ninth Edition

Steven J. Leon

Linear Algebra with Applications

Ninth Edition

Steven J. Leon

University of Massachusetts, Dartmouth

PEARSON

Boston Columbus Indianapolis New York San Francisco
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Editor in Chief: Deirdre Lynch
Acquisitions Editor: William Hoffman
Editorial Assistant: Salena Casha
Executive Marketing Manager: Jeff Weidenaar
Marketing Assistant: Brooke Smith
Project Management Team Lead: Christina Lepre
Production Project Manager: Mary Sanger
Procurement Manager: Carol Melville
Senior Author Support/Technology Specialist: Joe Vetere
Cover Designer: Heather Scott
Image Manager: Diahanne Lucas
Cover Image: Ola Dusegard/E+/Getty Images
Production Coordination, Composition, Illustrations: Integra
Printer/Binder: Edwards Brothers
Cover Printer: Phoenix Color

Library of Congress Cataloging-in-Publication Data
Leon, Steven J.
Linear algebra with applications / Steven J. Leon, University of Massachusetts, Dartmouth. – Ninth edition.
p. cm.
Includes bibliographical references.
ISBN-13: 978-0-321-96221-8 (hardcover)
ISBN-10: 0-321-96221-4 (hardcover)
1. Algebras, Linear–Textbooks. I. Title.
QA184.2.L46 2015
512'.5–dc23
2013037914

10 9 8 7 6 5 4 3 2 1—EBM—18 17 16 15 14

PEARSON

www.pearsonhighered.com

ISBN-10: 0-321-96221-4
ISBN-13: 978-0-321-96221-8

To the memories of

Florence and Rudolph Leon,
devoted and loving parents

and to the memories of

Gene Golub, Germund Dahlquist, and Jim Wilkinson,
friends, mentors, and role models

This page intentionally left blank

Contents

* Online: The supplemental Chapters 8 and 9 can be downloaded from the Internet. See the section of the Preface on supplementary materials.

This page intentionally left blank

Preface

I am pleased to see the text reach its ninth edition. The continued support and enthusiasm of the many users has been most gratifying. Linear algebra is more exciting now than at almost any time in the past. Its applications continue to spread to more and more fields. Largely due to the computer revolution of the last 75 years, linear algebra has risen to a role of prominence in the mathematical curriculum rivaling that of calculus. Modern software has also made it possible to dramatically improve the way the course is taught.

The first edition of this book was published in 1980. Many significant changes were made for the second edition (1986), most notably the exercise sets were greatly expanded and the linear transformations chapter of the book was completely revised. Each of the following editions has seen significant modifications including the addition of comprehensive sets of MATLAB computer exercises, a dramatic increase in the number of applications, and many revisions in the various sections of the book. I have been fortunate to have had outstanding reviewers and their suggestions have led to many important improvements in the book. For the ninth edition we have given special attention to Chapter 7 as it is the only chapter that has not seen major revisions in any of the previous editions. The following is an outline of the most significant revisions that were made for the ninth edition.

What's New in the Ninth Edition?

1. New Subsection Added to Chapter 3
Section 2 of Chapter 3 deals with the topic of subspaces. One important example of a subspace occurs when we find all solutions to a homogeneous system of linear equations. This type of subspace is referred to as a *null space*. A new subsection has been added to show how the null space is also useful in finding the solution set to a nonhomogeneous linear system. The subsection contains a new theorem and a new figure that provides a geometric illustration of the theorem. Three related problems have been added to the exercises at the end of Section 2.

2. New Applications Added to Chapters 1, 5, 6, and 7
In Chapter 1, we introduce an important application to the field of Management Science. Management decisions often involve making choices between a number of alternatives. We assume that the choices are to be made with a fixed goal in mind and should be based on a set of evaluation criteria. These decisions often involve a number of human judgments that may not always be completely consistent. The analytic hierarchy process is a technique for rating the various alternatives based on a chart consisting of weighted criteria and ratings that measure how well each alternative satisfies each of the criteria.

In Chapter 1, we see how to set up such a chart or decision tree for the process. After weights and ratings have been assigned to each entry in the chart, an overall ranking of the alternatives is calculated using simple matrix-vector operations. In Chapters 5 and 6, we revisit the application and discuss how to use advanced matrix techniques to determine appropriate weights and ratings for the decision process. Finally in Chapter 7, we present a numerical algorithm for computing the weight vectors used in the decision process.

3. Section 1 of Chapter 7 Revised and Two Subsections Added
   Section 7.1 has been revised and modernized. A new subsection on IEEE floating-point representation of numbers and a second subsection on accuracy and stability of numerical algorithms have been added. New examples and additional exercises on these topics are also included.
4. Section 5 of Chapter 7 Revised
   The discussion of Householder transformations has been revised and expanded. A new subsection has been added, which discusses the practicalities of using QR factorizations for solving linear systems. New exercises have also been added to this section.
5. Section 7 of Chapter 7 Revised
   Section 7.7 deals with numerical methods for solving least squares problems. The section has been revised and a new subsection on using the modified Gram–Schmidt process to solve least squares problems has been added. The subsection contains one new algorithm.

Overview of Text

This book is suitable for either a sophomore-level course or for a junior/senior level course. The student should have some familiarity with the basics of differential and integral calculus. This prerequisite can be met by either one semester or two quarters of elementary calculus.

If the text is used for a sophomore-level course, the instructor should probably spend more time on the early chapters and omit many of the sections in the later chapters. For more advanced courses, a quick review of the topics in the first two chapters and then a more complete coverage of the later chapters would be appropriate. The explanations in the text are given in sufficient detail so that beginning students should have little trouble reading and understanding the material. To further aid the student, a large number of examples have been worked out completely. Additionally, computer exercises at the end of each chapter give students the opportunity to perform numerical experiments and try to generalize the results. Applications are presented throughout the book. These applications can be used to motivate new material or to illustrate the relevance of material that has already been covered.

The text contains all the topics recommended by the National Science Foundation (NSF) sponsored Linear Algebra Curriculum Study Group (LACSG) and much more. Although there is more material than can be covered in a one-quarter or one-semester course, it is my feeling that it is easier for an instructor to leave out or skip material than it is to supplement a book with outside material. Even if many topics are omitted, the book should still provide students with a feeling for the overall scope of the subject matter. Furthermore, students may use the book later as a reference and consequently may end up learning omitted topics on their own.

In the next section of this preface, a number of outlines are provided for one-semester courses at either the sophomore level or the junior/senior level and with either a matrix-oriented emphasis or a slightly more theoretical emphasis.

Ideally, the entire book could be covered in a two-quarter or two-semester sequence. Although two semesters of linear algebra has been recommended by the LACSG, it is still not practical at many universities and colleges. At present there is no universal agreement on a core syllabus for a second course. Indeed, if all of the topics that instructors would like to see in a second course were included in a single volume, it would be a weighty book. An effort has been made in this text to cover all of the basic linear algebra topics that are necessary for modern applications. Furthermore, two additional chapters for a second course are available for downloading from the special Pearson Web site developed for this book:

http://pearsonhighered.com/leon

Suggested Course Outlines

I. Two-Semester Sequence: In a two-semester sequence, it is possible to cover all 40 sections of the book. When the author teaches the course, he also includes an extra lecture demonstrating how to use the MATLAB software.

II. One-Semester Sophomore-Level Course

A. A Basic Sophomore-Level Course

Chapter 1	Sections 1–6		7 lectures
Chapter 2	Sections 1–2		2 lectures
Chapter 3	Sections 1–6		9 lectures
Chapter 4	Sections 1–3		4 lectures
Chapter 5	Sections 1–6		9 lectures
Chapter 6	Sections 1–3		4 lectures
		Total	35 lectures

B. The LACSG Matrix Oriented Course: The core course recommended by the Linear Algebra Curriculum Study Group involves only the Euclidean vector spaces. Consequently, for this course you should omit Section 1 of Chapter 3 (on general vector spaces) and all references and exercises involving function spaces in Chapters 3 to 6. All of the topics in the LACSG core syllabus are included in the text. It is not necessary to introduce any supplementary materials. The LACSG recommended 28 lectures to cover the core material. This is possible if the class is taught in lecture format with an additional recitation section meeting once a week. If the course is taught without recitations, it is my feeling that the following schedule of 35 lectures is perhaps more reasonable.

Chapter 1	Sections 1–6	7 lectures
Chapter 2	Sections 1–2	2 lectures
Chapter 3	Sections 2–6	7 lectures
Chapter 4	Sections 1–3	2 lectures
Chapter 5	Sections 1–6	9 lectures
Chapter 6	Sections 1,3–5	8 lectures
	Total	35 lectures

III. One-Semester Junior/Senior Level Courses: The coverage in an upper division course is dependent on the background of the students. Below are two possible courses with 35 lectures each.

A. Course 1

Chapter 1	Sections 1–6	6 lectures
Chapter 2	Sections 1–2	2 lectures
Chapter 3	Sections 1–6	7 lectures
Chapter 5	Sections 1–6	9 lectures
Chapter 6	Sections 1–7	10 lectures
	Section 8 if time allows	
Chapter 7	Section 4	1 lecture

B. Course 2

Review of Topics in Chapters 1–3		5 lectures
Chapter 4	Sections 1–3	2 lectures
Chapter 5	Sections 1–6	10 lectures
Chapter 6	Sections 1–7	11 lectures
	Section 8 if time allows	
Chapter 7	Sections 4–7	7 lectures
	If time allows, Sections 1–3	

Computer Exercises

This edition contains a section of computing exercises at the end of each chapter. These exercises are based on the software package MATLAB. The MATLAB Appendix in the book explains the basics of using the software. MATLAB has the advantage that it is a powerful tool for matrix computations and yet it is easy to learn. After reading the Appendix, students should be able to do the computing exercises without having to refer to any other software books or manuals. To help students get started, we recommend one 50-minute classroom demonstration of the software. The assignments can be done either as ordinary homework assignments or as part of a formally scheduled computer laboratory course.

Another source of MATLAB exercises for linear algebra is the ATLAST book, which is available as a companion manual to supplement this book. (See the list of supplementary materials in the next section of this preface.)

While the course can be taught without any reference to the computer, we believe that computer exercises can greatly enhance student learning and provide a new dimension to linear algebra education. One of the recommendations of the Linear Algebra Curriculum Study Group is that technology should be used in a first course in linear algebra. That recommendation has been widely accepted, and it is now common to see mathematical software packages used in linear algebra courses.

Supplementary Materials

Web Supplements and Additional Chapters

Two supplemental chapters for this book may be downloaded using links from the author's home page:

http://www.umassd.edu/cas/math/people/facultyandstaff/steveleon

or from the Pearson Web site for this book:

http://pearsonhighered.com/leon

The additional chapters are:

- Chapter 8. Iterative Methods
- Chapter 9. Canonical Forms

The Pearson Web site for this book contains materials for students and instructors including links to online exercises for each of the original seven chapters of the book. The author's home page contains a link to the errata list for this textbook. Please send any additional errata items that you discover to the author so that the list can be updated and corrections can be made in later printings of the book.

Companion Books

A *Student Study Guide* has been developed to accompany this textbook. A number of MATLAB and Maple computer manuals are also available as companion books. Instructors wishing to use one of the companion manuals along with the textbook can order both the book and the manual for their classes and have each pair bundled together in a shrink-wrapped package. These packages are available for classes at special rates that are comparable to the price of ordering the textbook alone. Thus, when students buy the textbook, they get the manual at little or no extra cost. To obtain information about the companion packages available, instructors should either consult their Pearson sales representative or search the instructor section of the Pearson higher education Web site (*www.pearsonhighered.com*). The following is a list of some of the companion books being offered as bundles with this textbook:

- *Student Study Guide for Linear Algebra with Applications*. The manual is available to students as a study tool to accompany this textbook. The manual summarizes important theorems, definitions, and concepts presented in the textbook. It provides solutions to some of the exercises and hints and suggestions on many other exercises.

- *ATLAST Computer Exercises for Linear Algebra, Second Edition*. ATLAST (Augmenting the Teaching of Linear Algebra through the use of Software Tools) was an NSF-sponsored project to encourage and facilitate the use of software in the teaching of linear algebra. During a five-year period, 1992–1997, the ATLAST Project conducted 18 faculty workshops using the MATLAB software package. Participants in those workshops designed computer exercises, projects, and lesson plans for software-based teaching of linear algebra. A selection of these materials was first published as a manual in 1997. That manual was greatly expanded for the second edition published in 2003. Each of the eight chapters in the second edition contains a section of short exercises and a section of longer projects.

  The collection of software tools (M-files) developed to accompany the ATLAST book may be downloaded from the ATLAST Web site:

  www1.umassd.edu/specialprograms/atlast

  Additionally, Mathematica users can download the collection of *ATLAST Mathematica Notebooks* that has been developed by Richard Neidinger.

- *Linear Algebra Labs with MATLAB: 3rd ed.* by David Hill and David Zitarelli
- *Visualizing Linear Algebra using Maple*, by Sandra Keith
- *A Maple Supplement for Linear Algebra*, by John Maloney
- *Understanding Linear Algebra Using MATLAB*, by Erwin and Margaret Kleinfeld

Acknowledgments

I would like to express my gratitude to the long list of reviewers that have contributed so much to all previous editions of this book. Special thanks is due to the reviewers of the ninth edition:

Mark Arnold, University of Arkansas
J'Lee Bumpus, Austin College
Michael Cranston, University of California Irvine
Matthias Kawski, Arizona State University

Thanks also to the many users who have sent in comments and suggestions. In particular the author would like to thank LeSheng Jin for suggesting the inclusion of the analytic hierarchy process application.

Special thanks to Pearson Production Project Manager Mary Sanger and Editorial Assistant Salena Casha. I am grateful to Tom Wegleitner for doing the accuracy checking for the book and the associated manuals. Thanks to the entire editorial, production, and sales staff at Pearson for all their efforts. Thanks also to Integra Software Services Project Manager Abinaya Rajendran.

I would like to acknowledge the contributions of Gene Golub and Jim Wilkinson. Most of the first edition of the book was written in 1977–1978 while I was a Visiting Scholar at Stanford University. During that period, I attended courses and lectures on numerical linear algebra given by Gene Golub and J. H. Wilkinson. Those lectures

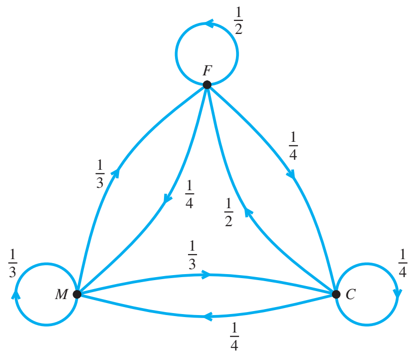
have greatly influenced me in writing this book. Finally, I would like to express my gratitude to Germund Dahlquist for his helpful suggestions on earlier editions of the book. Although Gene Golub, Jim Wilkinson, and Germund Dahlquist are no longer with us, they continue to live on in the memories of their friends.

Steven J. Leon
sleon@umassd.edu

This page intentionally left blank

CHAPTER

1



Matrices and Systems of Equations

Probably the most important problem in mathematics is that of solving a system of linear equations. Well over 75 percent of all mathematical problems encountered in scientific or industrial applications involve solving a linear system at some stage. By using the methods of modern mathematics, it is often possible to take a sophisticated problem and reduce it to a single system of linear equations. Linear systems arise in applications to such areas as business, economics, sociology, ecology, demography, genetics, electronics, engineering, and physics. Therefore, it seems appropriate to begin this book with a section on linear systems.

1.1 Systems of Linear Equations

A *linear equation in n unknowns* is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are real numbers and $x_1, x_2, \ldots, x_n$ are variables. A *linear system* of m equations in n unknowns is then a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1}$$

where the a_{ij}'s and the b_i's are all real numbers. We will refer to systems of the form (1) as $m \times n$ linear systems. The following are examples of linear systems:

(a) $\begin{aligned} x_1 + 2x_2 &= 5 \\ 2x_1 + 3x_2 &= 8 \end{aligned}$ **(b)** $\begin{aligned} x_1 - x_2 + x_3 &= 2 \\ 2x_1 + x_2 - x_3 &= 4 \end{aligned}$ **(c)** $\begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 1 \\ x_1 \quad &= 4 \end{aligned}$

System **(a)** is a 2×2 system, **(b)** is a 2×3 system, and **(c)** is a 3×2 system.

By a solution of an $m \times n$ system, we mean an ordered n-tuple of numbers $(x_1, x_2, \ldots, x_n)$ that satisfies all the equations of the system. For example, the ordered pair $(1, 2)$ is a solution of system **(a)**, since

$$\begin{aligned} 1 \cdot (1) + 2 \cdot (2) &= 5 \\ 2 \cdot (1) + 3 \cdot (2) &= 8 \end{aligned}$$

The ordered triple $(2, 0, 0)$ is a solution of system **(b)**, since

$$\begin{aligned} 1 \cdot (2) - 1 \cdot (0) + 1 \cdot (0) &= 2 \\ 2 \cdot (2) + 1 \cdot (0) - 1 \cdot (0) &= 4 \end{aligned}$$

Actually, system **(b)** has many solutions. If α is any real number, it is easily seen that the ordered triple $(2, \alpha, \alpha)$ is a solution. However, system **(c)** has no solution. It follows from the third equation that the first coordinate of any solution would have to be 4. Using $x_1 = 4$ in the first two equations, we see that the second coordinate must satisfy

$$\begin{aligned} 4 + x_2 &= 2 \\ 4 - x_2 &= 1 \end{aligned}$$

Since there is no real number that satisfies both of these equations, the system has no solution. If a linear system has no solution, we say that the system is *inconsistent*. If the system has at least one solution, we say that it is *consistent*. Thus system **(c)** is inconsistent, while systems **(a)** and **(b)** are both consistent.

The set of all solutions of a linear system is called the *solution set* of the system. If a system is inconsistent, its solution set is empty. A consistent system will have a nonempty solution set. To solve a consistent system, we must find its solution set.

2×2 Systems

Let us examine geometrically a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

Each equation can be represented graphically as a line in the plane. The ordered pair (x_1, x_2) will be a solution of the system if and only if it lies on both lines. For example, consider the three systems

$$\textbf{(i)}\ \begin{aligned} x_1 + x_2 &= 2 \\ x_1 - x_2 &= 2 \end{aligned} \qquad \textbf{(ii)}\ \begin{aligned} x_1 + x_2 &= 2 \\ x_1 + x_2 &= 1 \end{aligned} \qquad \textbf{(iii)}\ \begin{aligned} x_1 + x_2 &= 2 \\ -x_1 - x_2 &= -2 \end{aligned}$$

The two lines in system (i) intersect at the point $(2, 0)$. Thus, $\{(2, 0)\}$ is the solution set of (i). In system (ii) the two lines are parallel. Therefore, system (ii) is inconsistent and hence its solution set is empty. The two equations in system (iii) both represent the same line. Any point on this line will be a solution of the system (see Figure 1.1.1).

In general, there are three possibilities: the lines intersect at a point, they are parallel, or both equations represent the same line. The solution set then contains either one, zero, or infinitely many points.

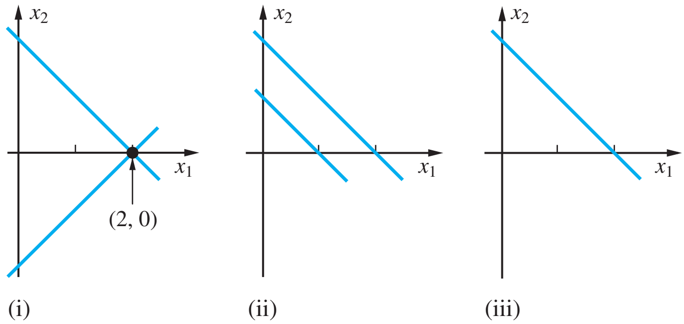


Figure 1.1.1.

The situation is the same for $m \times n$ systems. An $m \times n$ system may or may not be consistent. If it is consistent, it must have either exactly one solution or infinitely many solutions. These are the only possibilities. We will see why this is so in Section 1.2 when we study the row echelon form. Of more immediate concern is the problem of finding all solutions of a given system. To tackle this problem, we introduce the notion of *equivalent systems*.

Equivalent Systems

Consider the two systems

$$\textbf{(a)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ x_2 &= 3 \\ 2x_3 &= 4 \end{aligned} \qquad\qquad \textbf{(b)}\quad \begin{aligned} 3x_1 + 2x_2 - x_3 &= -2 \\ -3x_1 - x_2 + x_3 &= 5 \\ 3x_1 + 2x_2 + x_3 &= 2 \end{aligned}$$

System **(a)** is easy to solve because it is clear from the last two equations that $x_2 = 3$ and $x_3 = 2$. Using these values in the first equation, we get

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

Thus, the solution of the system is $(-2, 3, 2)$. System **(b)** seems to be more difficult to solve. Actually, system **(b)** has the same solution as system **(a)**. To see this, add the first two equations of the system:

$$\begin{array}{r} 3x_1 + 2x_2 - x_3 = -2 \\ -3x_1 - x_2 + x_3 = 5 \\ \hline x_2 = 3 \end{array}$$

If (x_1, x_2, x_3) is any solution of **(b)**, it must satisfy all the equations of the system. Thus, it must satisfy any new equation formed by adding two of its equations. Therefore, x_2 must equal 3. Similarly, (x_1, x_2, x_3) must satisfy the new equation formed by subtracting the first equation from the third:

$$\begin{array}{r} 3x_1 + 2x_2 + x_3 = 2 \\ 3x_1 + 2x_2 - x_3 = -2 \\ \hline 2x_3 = 4 \end{array}$$

Therefore, any solution of system **(b)** must also be a solution of system **(a)**. By a similar argument, it can be shown that any solution of **(a)** is also a solution of **(b)**. This can be done by subtracting the first equation from the second:

$$\begin{array}{rcl} x_2 \qquad\quad & = & 3 \\ 3x_1 + 2x_2 - x_3 & = & -2 \\ \hline -3x_1 - x_2 + x_3 & = & 5 \end{array}$$

Then add the first and third equations:

$$\begin{array}{rcl} 3x_1 + 2x_2 - x_3 & = & -2 \\ 2x_3 & = & 4 \\ \hline 3x_1 + 2x_2 + x_3 & = & 2 \end{array}$$

Thus, (x_1, x_2, x_3) is a solution of system **(b)** if and only if it is a solution of system **(a)**. Therefore, both systems have the same solution set, $\{(-2, 3, 2)\}$.

Definition

> Two systems of equations involving the same variables are said to be **equivalent** if they have the same solution set.

Clearly, if we interchange the order in which two equations of a system are written, this will have no effect on the solution set. The reordered system will be equivalent to the original system. For example, the systems

$$\begin{array}{rcl} x_1 + 2x_2 & = & 4 \\ 3x_1 - x_2 & = & 2 \\ 4x_1 + x_2 & = & 6 \end{array} \qquad \text{and} \qquad \begin{array}{rcl} 4x_1 + x_2 & = & 6 \\ 3x_1 - x_2 & = & 2 \\ x_1 + 2x_2 & = & 4 \end{array}$$

both involve the same three equations and, consequently, they must have the same solution set.

If one equation of a system is multiplied through by a nonzero real number, this will have no effect on the solution set, and the new system will be equivalent to the original system. For example, the systems

$$\begin{array}{rcl} x_1 + x_2 + x_3 & = & 3 \\ -2x_1 - x_2 + 4x_3 & = & 1 \end{array} \qquad \text{and} \qquad \begin{array}{rcl} 2x_1 + 2x_2 + 2x_3 & = & 6 \\ -2x_1 - x_2 + 4x_3 & = & 1 \end{array}$$

are equivalent.

If a multiple of one equation is added to another equation, the new system will be equivalent to the original system. This follows since the n-tuple $(x_1, \ldots, x_n)$ will satisfy the two equations

$$\begin{array}{rcl} a_{i1}x_1 + \cdots + a_{in}x_n & = & b_i \\ a_{j1}x_1 + \cdots + a_{jn}x_n & = & b_j \end{array}$$

if and only if it satisfies the equations

$$\begin{array}{rcl} a_{i1}x_1 + \cdots + a_{in}x_n & = & b_i \\ (a_{j1} + \alpha a_{i1})x_1 + \cdots + (a_{jn} + \alpha a_{in})x_n & = & b_j + \alpha b_i \end{array}$$

To summarize, there are three operations that can be used on a system to obtain an equivalent system:

I. The order in which any two equations are written may be interchanged.
II. Both sides of an equation may be multiplied by the same nonzero real number.
III. A multiple of one equation may be added to (or subtracted from) another.

Given a system of equations, we may use these operations to obtain an equivalent system that is easier to solve.

$n \times n$ Systems

Let us restrict ourselves to $n \times n$ systems for the remainder of this section. We will show that if an $n \times n$ system has exactly one solution, then operations **I** and **III** can be used to obtain an equivalent "strictly triangular system."

Definition

> A system is said to be in **strict triangular form** if, in the kth equation, the coefficients of the first $k-1$ variables are all zero and the coefficient of x_k is nonzero $(k = 1, \ldots, n)$.

EXAMPLE 1 The system

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

is in strict triangular form, since in the second equation the coefficients are $0, 1, -1$, respectively, and in the third equation the coefficients are $0, 0, 2$, respectively. Because of the strict triangular form, the system is easy to solve. It follows from the third equation that $x_3 = 2$. Using this value in the second equation, we obtain

$$x_2 - 2 = 2 \qquad \text{or} \qquad x_2 = 4$$

Using $x_2 = 4$, $x_3 = 2$ in the first equation, we end up with

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

Thus, the solution of the system is $(-3, 4, 2)$. ■

Any $n \times n$ strictly triangular system can be solved in the same manner as the last example. First, the nth equation is solved for the value of x_n. This value is used in the $(n-1)$st equation to solve for x_{n-1}. The values x_n and x_{n-1} are used in the $(n-2)$nd equation to solve for x_{n-2}, and so on. We will refer to this method of solving a strictly triangular system as *back substitution*.

EXAMPLE 2 Solve the system

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

Solution

Using back substitution, we obtain

$$\begin{aligned} 4x_4 &= 4 & x_4 &= 1 \\ 4x_3 + 3 \cdot 1 &= 3 & x_3 &= 0 \\ x_2 - 2 \cdot 0 + 3 \cdot 1 &= 2 & x_2 &= -1 \\ 2x_1 - (-1) + 3 \cdot 0 - 2 \cdot 1 &= 1 & x_1 &= 1 \end{aligned}$$

Thus the solution is $(1, -1, 0, 1)$. ■

In general, given a system of n linear equations in n unknowns, we will use operations **I** and **III** to try to obtain an equivalent system that is strictly triangular. (We will see in the next section of the book that it is not possible to reduce the system to strictly triangular form in the cases where the system does not have a unique solution.)

EXAMPLE 3 Solve the system

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

Solution

Subtracting 3 times the first row from the second row yields

$$-7x_2 - 6x_3 = -10$$

Subtracting 2 times the first row from the third row yields

$$-x_2 - x_3 = -2$$

If the second and third equations of our system, respectively, are replaced by these new equations, we obtain the equivalent system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -x_2 - x_3 &= -2 \end{aligned}$$

If the third equation of this system is replaced by the sum of the third equation and $-\frac{1}{7}$ times the second equation, we end up with the following strictly triangular system:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 3 \\ -7x_2 - 6x_3 &= -10 \\ -\tfrac{1}{7}x_3 &= -\tfrac{4}{7} \end{aligned}$$

Using back substitution, we get

$$x_3 = 4, \qquad x_2 = -2, \qquad x_1 = 3$$

Let us look back at the system of equations in the last example. We can associate with that system a 3×3 array of numbers whose entries are the coefficients of the x_i's:

$$\begin{Bmatrix} 1 & 2 & 1 \\ 3 & -1 & -3 \\ 2 & 3 & 1 \end{Bmatrix}$$

We will refer to this array as the *coefficient matrix* of the system. The term *matrix* means simply a rectangular array of numbers. A matrix having m rows and n columns is said to be $m \times n$. A matrix is said to be *square* if it has the same number of rows and columns, that is, if $m = n$.

If we attach to the coefficient matrix an additional column whose entries are the numbers on the right-hand side of the system, we obtain the new matrix

$$\left\{\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right\}$$

We will refer to this new matrix as the *augmented matrix*. In general, when an $m \times r$ matrix B is attached to an $m \times n$ matrix A in this way, the augmented matrix is denoted by $(A|B)$. Thus, if

$$A = \begin{Bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{Bmatrix}, \qquad B = \begin{Bmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \vdots & & & \\ b_{m1} & b_{m2} & \cdots & b_{mr} \end{Bmatrix}$$

then

$$(A|B) = \left\{\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & b_{11} & \cdots & b_{1r} \\ \vdots & & & \vdots & & \\ a_{m1} & \cdots & a_{mn} & b_{m1} & \cdots & b_{mr} \end{array}\right\}$$

With each system of equations we may associate an augmented matrix of the form

$$\left\{\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right\}$$

The system can be solved by performing operations on the augmented matrix. The x_i's are placeholders that can be omitted until the end of the computation. Corresponding to the three operations used to obtain equivalent systems, the following row operations may be applied to the augmented matrix:

Elementary Row Operations

I. Interchange two rows.
II. Multiply a row by a nonzero real number.
III. Replace a row by its sum with a multiple of another row.

Returning to the example, we find that the first row is used to eliminate the elements in the first column of the remaining rows. We refer to the first row as the *pivotal row*. For emphasis, the entries in the pivotal row are all in bold type and the entire row is color shaded. The first nonzero entry in the pivotal row is called the *pivot*.

$$\begin{array}{r} (\text{pivot } a_{11} = 1) \\ \left.\begin{array}{r}\text{entries to be eliminated} \\ a_{21} = 3 \text{ and } a_{31} = 2\end{array}\right\} \rightarrow \end{array} \left[\begin{array}{rrr|r} \mathbf{1} & \mathbf{2} & \mathbf{1} & \mathbf{3} \\ 3 & -1 & -3 & -1 \\ 2 & 3 & 1 & 4 \end{array}\right] \begin{array}{l} \leftarrow \text{pivotal row} \\ \\ \\ \end{array}$$

By using row operation III, 3 times the first row is subtracted from the second row and 2 times the first row is subtracted from the third. When this is done, we end up with the matrix

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ \mathbf{0} & \mathbf{-7} & \mathbf{-6} & \mathbf{-10} \\ 0 & -1 & -1 & -2 \end{array}\right] \begin{array}{l} \\ \leftarrow \text{pivotal row} \\ \\ \end{array}$$

At this step we choose the second row as our new pivotal row and apply row operation III to eliminate the last element in the second column. This time the pivot is -7 and the quotient $\frac{-1}{-7} = \frac{1}{7}$ is the multiple of the pivotal row that is subtracted from the third row. We end up with the matrix

$$\left[\begin{array}{rrr|r} 1 & 2 & 1 & 3 \\ 0 & -7 & -6 & -10 \\ 0 & 0 & -\frac{1}{7} & -\frac{4}{7} \end{array}\right]$$

This is the augmented matrix for the strictly triangular system, which is equivalent to the original system. The solution of the system is easily obtained by back substitution.

EXAMPLE 4 Solve the system

$$\begin{aligned} -x_2 - x_3 + x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 6 \\ 2x_1 + 4x_2 + x_3 - 2x_4 &= -1 \\ 3x_1 + x_2 - 2x_3 + 2x_4 &= 3 \end{aligned}$$

Solution

The augmented matrix for this system is

$$\left[\begin{array}{rrrr|r} 0 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 6 \\ 2 & 4 & 1 & -2 & -1 \\ 3 & 1 & -2 & 2 & 3 \end{array}\right]$$

Since it is not possible to eliminate any entries by using 0 as a pivot element, we will use row operation I to interchange the first two rows of the augmented matrix. The new first row will be the pivotal row and the pivot element will be 1:

$$\begin{array}{r} (\text{pivot } a_{11} = 1) \\ \\ \\ \\ \end{array} \left[\begin{array}{rrrr|r} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{6} \\ 0 & -1 & -1 & 1 & 0 \\ 2 & 4 & 1 & -2 & -1 \\ 3 & 1 & -2 & 2 & 3 \end{array}\right] \begin{array}{l} \leftarrow \text{pivotal row} \\ \\ \\ \\ \end{array}$$

Row operation III is then used twice to eliminate the two nonzero entries in the first column:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ \mathbf{0} & \mathbf{-1} & \mathbf{-1} & \mathbf{1} & \mathbf{0} \\ 0 & 2 & -1 & -4 & -13 \\ 0 & -2 & -5 & -1 & -15 \end{array}\right]$$

Next, the second row is used as the pivotal row to eliminate the entries in the second column below the pivot element -1:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{-3} & \mathbf{-2} & \mathbf{-13} \\ 0 & 0 & -3 & -3 & -15 \end{array}\right]$$

Finally, the third row is used as the pivotal row to eliminate the last element in the third column:

$$\left[\begin{array}{rrrr|r} 1 & 1 & 1 & 1 & 6 \\ 0 & -1 & -1 & 1 & 0 \\ 0 & 0 & -3 & -2 & -13 \\ 0 & 0 & 0 & -1 & -2 \end{array}\right]$$

This augmented matrix represents a strictly triangular system. Solving by back substitution, we obtain the solution $(2, -1, 3, 2)$. ■

In general, if an $n \times n$ linear system can be reduced to strictly triangular form, then it will have a unique solution that can be obtained by performing back substitution on the triangular system. We can think of the reduction process as an algorithm involving $n - 1$ steps. At the first step, a pivot element is chosen from among the nonzero entries in the first column of the matrix. The row containing the pivot element is called the *pivotal row*. We interchange rows (if necessary) so that the pivotal row is the new first row. Multiples of the pivotal row are then subtracted from each of the remaining $n - 1$ rows so as to obtain 0's in the first entries of rows 2 through n. At the second step, a pivot element is chosen from the nonzero entries in column 2, rows 2 through n, of the matrix. The row containing the pivot is then interchanged with the second row of the matrix and is used as the new pivotal row. Multiples of the pivotal row are then subtracted from the remaining $n - 2$ rows so as to eliminate all entries below the pivot in the second column. The same procedure is repeated for columns 3 through $n - 1$. Note that at the second step row 1 and column 1 remain unchanged, at the third step the first two rows and first two columns remain unchanged, and so on. At each step, the overall dimensions of the system are effectively reduced by 1 (see Figure 1.1.2).

If the elimination process can be carried out as described, we will arrive at an equivalent strictly triangular system after $n - 1$ steps. However, the procedure will break down if, at any step, all possible choices for a pivot element are equal to 0. When this happens, the alternative is to reduce the system to certain special echelon, or staircase-shaped, forms. These echelon forms will be studied in the next section. They will also be used for $m \times n$ systems, where $m \neq n$.

Step 1 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right)$

Step 2 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right)$

Step 3 $\left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & x & x & x \end{array}\right) \rightarrow \left(\begin{array}{cccc|c} x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{array}\right)$

Figure 1.1.2.

SECTION 1.1 EXERCISES

1. Use back substitution to solve each of the following systems of equations:

(a) $\begin{aligned} x_1 - 3x_2 &= 2 \\ 2x_2 &= 6 \end{aligned}$

(b) $\begin{aligned} x_1 + x_2 + x_3 &= 8 \\ 2x_2 + x_3 &= 5 \\ 3x_3 &= 9 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= 5 \\ 3x_2 + x_3 - 2x_4 &= 1 \\ -x_3 + 2x_4 &= -1 \\ 4x_4 &= 4 \end{aligned}$

(d) $\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 5 \\ 2x_2 + x_3 - 2x_4 + x_5 &= 1 \\ 4x_3 + x_4 - 2x_5 &= 1 \\ x_4 - 3x_5 &= 0 \\ 2x_5 &= 2 \end{aligned}$

2. Write out the coefficient matrix for each of the systems in Exercise 1.

3. In each of the following systems, interpret each equation as a line in the plane. For each system, graph the lines and determine geometrically the number of solutions.

(a) $\begin{aligned} x_1 + x_2 &= 4 \\ x_1 - x_2 &= 2 \end{aligned}$

(b) $\begin{aligned} x_1 + 2x_2 &= 4 \\ -2x_1 - 4x_2 &= 4 \end{aligned}$

(c) $\begin{aligned} 2x_1 - x_2 &= 3 \\ -4x_1 + 2x_2 &= -6 \end{aligned}$

(d) $\begin{aligned} x_1 + x_2 &= 1 \\ x_1 - x_2 &= 1 \\ -x_1 + 3x_2 &= 3 \end{aligned}$

4. Write an augmented matrix for each of the systems in Exercise 3.

5. Write out the system of equations that corresponds to each of the following augmented matrices:

(a) $\left[\begin{array}{cc|c} 3 & 2 & 8 \\ 1 & 5 & 7 \end{array}\right]$

(b) $\left[\begin{array}{ccc|c} 5 & -2 & 1 & 3 \\ 2 & 3 & -4 & 0 \end{array}\right]$

(c) $\left[\begin{array}{ccc|c} 2 & 1 & 4 & -1 \\ 4 & -2 & 3 & 4 \\ 5 & 2 & 6 & -1 \end{array}\right]$

(d) $\left[\begin{array}{cccc|c} 4 & -3 & 1 & 2 & 4 \\ 3 & 1 & -5 & 6 & 5 \\ 1 & 1 & 2 & 4 & 8 \\ 5 & 1 & 3 & -2 & 7 \end{array}\right]$

6. Solve each of the following systems.

(a) $\begin{aligned} x_1 - 2x_2 &= 5 \\ 3x_1 + x_2 &= 1 \end{aligned}$

(b) $\begin{aligned} 2x_1 + x_2 &= 8 \\ 4x_1 - 3x_2 &= 6 \end{aligned}$

(c) $\begin{aligned} 4x_1 + 3x_2 &= 4 \\ \tfrac{2}{3}x_1 + 4x_2 &= 3 \end{aligned}$

(d) $\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 - x_2 + x_3 &= 3 \\ -x_1 + 2x_2 + 3x_3 &= 7 \end{aligned}$

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

(f) $\begin{aligned} 3x_1 + 2x_2 + x_3 &= 0 \\ -2x_1 + x_2 - x_3 &= 2 \\ 2x_1 - x_2 + 2x_3 &= -1 \end{aligned}$

(g) $\begin{aligned} \tfrac{1}{3}x_1 + \tfrac{2}{3}x_2 + 2x_3 &= -1 \\ x_1 + 2x_2 + \tfrac{3}{2}x_3 &= \tfrac{3}{2} \\ \tfrac{1}{2}x_1 + 2x_2 + \tfrac{12}{5}x_3 &= \tfrac{1}{10} \end{aligned}$

(h) $\begin{aligned} x_2 + x_3 + x_4 &= 0 \\ 3x_1 \qquad + 3x_3 - 4x_4 &= 7 \\ x_1 + x_2 + x_3 + 2x_4 &= 6 \\ 2x_1 + 3x_2 + x_3 + 3x_4 &= 6 \end{aligned}$

7. The two systems

$$\begin{aligned} 2x_1 + x_2 &= 3 \\ 4x_1 + 3x_2 &= 5 \end{aligned} \quad \text{and} \quad \begin{aligned} 2x_1 + x_2 &= -1 \\ 4x_1 + 3x_2 &= 1 \end{aligned}$$

have the same coefficient matrix but different right-hand sides. Solve both systems simultaneously by eliminating the first entry in the second row of the augmented matrix

$$\left[\begin{array}{cc|cc} 2 & 1 & 3 & -1 \\ 4 & 3 & 5 & 1 \end{array}\right]$$

and then performing back substitutions for each of the columns corresponding to the right-hand sides.

8. Solve the two systems

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= 9 \end{aligned} \qquad \begin{aligned} x_1 + 2x_2 - 2x_3 &= 9 \\ 2x_1 + 5x_2 + x_3 &= 9 \\ x_1 + 3x_2 + 4x_3 &= -2 \end{aligned}$$

by doing elimination on a 3×5 augmented matrix and then performing two back substitutions.

9. Given a system of the form

$$\begin{aligned} -m_1x_1 + x_2 &= b_1 \\ -m_2x_1 + x_2 &= b_2 \end{aligned}$$

where m_1, m_2, b_1, and b_2 are constants:

(a) Show that the system will have a unique solution if $m_1 \neq m_2$.

(b) Show that if $m_1 = m_2$, then the system will be consistent only if $b_1 = b_2$.

(c) Give a geometric interpretation of parts (a) and (b).

10. Consider a system of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

where a_{11}, a_{12}, a_{21}, and a_{22} are constants. Explain why a system of this form must be consistent.

11. Give a geometrical interpretation of a linear equation in three unknowns. Give a geometrical description of the possible solution sets for a 3×3 linear system.

1.2 Row Echelon Form

In Section 1.1 we learned a method for reducing an $n \times n$ linear system to strict triangular form. However, this method will fail if, at any stage of the reduction process, all the possible choices for a pivot element in a given column are 0.

EXAMPLE 1 Consider the system represented by the augmented matrix

$$\left[\begin{array}{rrrrr|r} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 1 & 1 & 2 & 2 & 4 & 1 \end{array}\right] \begin{array}{l} \leftarrow \text{pivotal row} \\ \\ \\ \\ \\ \end{array}$$

If row operation III is used to eliminate the nonzero entries in the last four rows of the first column, the resulting matrix will be

$$\left[\begin{array}{rrrrr|r} 1 & 1 & 1 & 1 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{2} & \mathbf{0} \\ 0 & 0 & 2 & 2 & 5 & 3 \\ 0 & 0 & 1 & 1 & 3 & -1 \\ 0 & 0 & 1 & 1 & 3 & 0 \end{array}\right] \begin{array}{l} \\ \leftarrow \text{pivotal row} \\ \\ \\ \\ \end{array}$$

At this stage, the reduction to strict triangular form breaks down. All four possible choices for the pivot element in the second column are 0. How do we proceed from here? Since our goal is to simplify the system as much as possible, it seems natural to move over to the third column and eliminate the last three entries:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{3} \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

In the fourth column, all the choices for a pivot element are 0; so again we move on to the next column. If we use the third row as the pivotal row, the last two entries in the fifth column are eliminated and we end up with the matrix

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 & -3 \end{array}\right]$$

The coefficient matrix that we end up with is not in strict triangular form; it is in staircase, or echelon, form. The horizontal and vertical line segments in the array for the coefficient matrix indicate the structure of the staircase form. Note that the vertical drop is 1 for each step, but the horizontal span for a step can be more than 1.

The equations represented by the last two rows are

$$\begin{aligned} 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= -4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= -3 \end{aligned}$$

Since there are no 5-tuples that could satisfy these equations, the system is inconsistent. ■

Suppose now that we change the right-hand side of the system in the last example so as to obtain a consistent system. For example, if we start with

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -2 & -2 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 1 & 3 & 3 \\ 1 & 1 & 2 & 2 & 4 & 4 \end{array}\right]$$

then the reduction process will yield the echelon-form augmented matrix

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The last two equations of the reduced system will be satisfied for any 5-tuple. Thus the solution set will be the set of all 5-tuples satisfying the first three equations.

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 1 \\ x_3 + x_4 + 2x_5 &= 0 \\ x_5 &= 3 \end{aligned} \tag{1}$$

The variables corresponding to the first nonzero elements in each row of the reduced matrix will be referred to as *lead variables*. Thus x_1, x_3, and x_5 are the lead variables. The remaining variables corresponding to the columns skipped in the reduction process will be referred to as *free variables*. Hence, x_2 and x_4 are the free variables. If we transfer the free variables over to the right-hand side in (1), we obtain the system

$$\begin{aligned} x_1 + x_3 + x_5 &= 1 - x_2 - x_4 \\ x_3 + 2x_5 &= -x_4 \\ x_5 &= 3 \end{aligned} \tag{2}$$

System (2) is strictly triangular in the unknowns x_1, x_3, and x_5. Thus, for each pair of values assigned to x_2 and x_4, there will be a unique solution. For example, if $x_2 = x_4 = 0$, then $x_5 = 3$, $x_3 = -6$, and $x_1 = 4$, and hence $(4, 0, -6, 0, 3)$ is a solution of the system.

Definition

A matrix is said to be in **row echelon form** if

(i) The first nonzero entry in each nonzero row is 1.
(ii) If row k does not consist entirely of zeros, the number of leading zero entries in row $k + 1$ is greater than the number of leading zero entries in row k.
(iii) If there are rows whose entries are all zero, they are below the rows having nonzero entries.

EXAMPLE 2 The following matrices are in row echelon form:

$$\begin{bmatrix} 1 & 4 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 & 1 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

■

EXAMPLE 3 The following matrices are not in row echelon form:

$$\begin{bmatrix} 2 & 4 & 6 \\ 0 & 3 & 5 \\ 0 & 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The first matrix does not satisfy condition (i). The second matrix fails to satisfy condition (iii), and the third matrix fails to satisfy condition (ii). ■

Definition

The process of using row operations I, II, and III to transform a linear system into one whose augmented matrix is in row echelon form is called **Gaussian elimination**.

Note that row operation II is necessary in order to scale the rows so that the leading coefficients are all 1. If the row echelon form of the augmented matrix contains a row of the form

$$\left(\begin{array}{cccc|c} 0 & 0 & \cdots & 0 & 1 \end{array}\right)$$

the system is inconsistent. Otherwise, the system will be consistent. If the system is consistent and the nonzero rows of the row echelon form of the matrix form a strictly triangular system, the system will have a unique solution.

Overdetermined Systems

A linear system is said to be *overdetermined* if there are more equations than unknowns. Overdetermined systems are *usually* (but not always) inconsistent.

EXAMPLE 4 Solve each of the following overdetermined systems:

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

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

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

Solution

By now the reader should be familiar enough with the elimination process that we can omit the intermediate steps in reducing each of these systems. Thus, we may write

$$\text{System (a):}\quad \left(\begin{array}{rr|r} 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 2 & -2 \end{array}\right) \rightarrow \left(\begin{array}{rr|r} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{array}\right)$$

It follows from the last row of the reduced matrix that the system is inconsistent. The three equations in system (a) represent lines in the plane. The first two lines intersect at the point $(2, -1)$. However, the third line does not pass through this point. Thus, there are no points that lie on all three lines (see Figure 1.2.1).

$$\text{System (b):}\quad \left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 2 & -1 & 3 & 5 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Using back substitution, we see that system (b) has exactly one solution $(0.1, -0.3, 1.5)$. The solution is unique because the nonzero rows of the reduced matrix

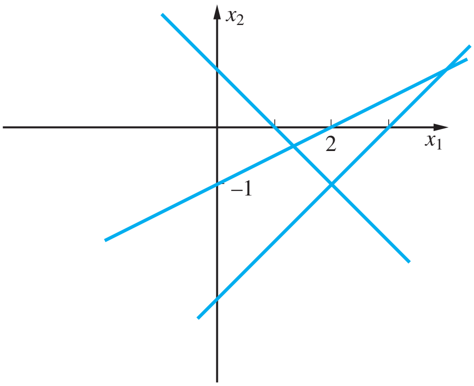


Figure 1.2.1.

form a strictly triangular system.

$$\text{System (c):} \quad \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 2 & -1 & 1 & 2 \\ 4 & 3 & 3 & 4 \\ 3 & 1 & 2 & 3 \end{array}\right] \rightarrow \left[\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Solving for x_2 and x_1 in terms of x_3, we obtain

$$\begin{aligned} x_2 &= -0.2x_3 \\ x_1 &= 1 - 2x_2 - x_3 = 1 - 0.6x_3 \end{aligned}$$

It follows that the solution set is the set of all ordered triples of the form $(1 - 0.6\alpha, -0.2\alpha, \alpha)$, where α is a real number. This system is consistent and has infinitely many solutions because of the free variable x_3. ■

Underdetermined Systems

A system of m linear equations in n unknowns is said to be *underdetermined* if there are fewer equations than unknowns ($m < n$). Although it is possible for underdetermined systems to be inconsistent, they are usually consistent with infinitely many solutions. It is not possible for an underdetermined system to have a unique solution. The reason for this is that any row echelon form of the coefficient matrix will involve $r \leq m$ nonzero rows. Thus there will be r lead variables and $n - r$ free variables, where $n - r \geq n - m > 0$. If the system is consistent, we can assign the free variables arbitrary values and solve for the lead variables. Therefore, a consistent underdetermined system will have infinitely many solutions.

EXAMPLE 5 Solve the following underdetermined systems:

(a)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 4x_2 + 2x_3 &= 3 \end{aligned}$$

(b)
$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 &= 2 \\ x_1 + x_2 + x_3 + 2x_4 + 2x_5 &= 3 \\ x_1 + x_2 + x_3 + 2x_4 + 3x_5 &= 2 \end{aligned}$$

Solution

$$\text{System (a):} \quad \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 4 & 2 & 3 \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Clearly, system (a) is inconsistent. We can think of the two equations in system (a) as representing planes in 3-space. Usually, two planes intersect in a line; however, in this case the planes are parallel.

$$\text{System (b):} \quad \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 3 & 2 \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

System (b) is consistent, and since there are two free variables, the system will have infinitely many solutions. In cases such as these it is convenient to continue the elimination process and simplify the form of the reduced matrix even further. We continue eliminating until all the terms above each leading 1 are eliminated. Thus, for system (b), we will continue and eliminate the first two entries in the fifth column and then the first element in the fourth column.

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} \end{array}\right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 0 & 3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{2} \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$
$$\rightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right)$$

If we put the free variables over on the right-hand side, it follows that

$$\begin{aligned} x_1 &= 1 - x_2 - x_3 \\ x_4 &= 2 \\ x_5 &= -1 \end{aligned}$$

Thus, for any real numbers α and β, the 5-tuple

$$(1 - \alpha - \beta, \alpha, \beta, 2, -1)$$

is a solution of the system. ■

In the case where the row echelon form of a consistent system has free variables, the standard procedure is to continue the elimination process until all the entries above each leading 1 have been eliminated, as in system (b) of the previous example. The resulting reduced matrix is said to be in *reduced row echelon form*.

Reduced Row Echelon Form

Definition

A matrix is said to be in **reduced row echelon form** if

(i) The matrix is in row echelon form.

(ii) The first nonzero entry in each row is the only nonzero entry in its column.

The following matrices are in reduced row echelon form:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The process of using elementary row operations to transform a matrix into reduced row echelon form is called *Gauss–Jordan reduction*.

EXAMPLE 6 Use Gauss–Jordan reduction to solve the system

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

Solution

$$\left[\begin{array}{rrrr|r} \mathbf{-1} & \mathbf{1} & \mathbf{-1} & \mathbf{3} & \mathbf{0} \\ 3 & 1 & -1 & -1 & 0 \\ 2 & -1 & -2 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} -1 & 1 & -1 & 3 & 0 \\ \mathbf{0} & \mathbf{4} & \mathbf{-4} & \mathbf{8} & \mathbf{0} \\ 0 & 1 & -4 & 5 & 0 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrrr|r} -1 & 1 & -1 & 3 & 0 \\ 0 & 4 & -4 & 8 & 0 \\ 0 & 0 & -3 & 3 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 1 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{-1} & \mathbf{0} \end{array}\right] \quad \text{row echelon form}$$

$$\rightarrow \left[\begin{array}{rrrr|r} 1 & -1 & 0 & -2 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ 0 & 0 & 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array}\right] \quad \text{reduced row echelon form}$$

If we set x_4 equal to any real number α, then $x_1 = \alpha$, $x_2 = -\alpha$, and $x_3 = \alpha$. Thus, all ordered 4-tuples of the form $(\alpha, -\alpha, \alpha, \alpha)$ are solutions of the system. ■

APPLICATION 1 Traffic Flow

In the downtown section of a certain city, two sets of one-way streets intersect as shown in Figure 1.2.2. The average hourly volume of traffic entering and leaving this section during rush hour is given in the diagram. Determine the amount of traffic between each of the four intersections.

Solution

At each intersection the number of automobiles entering must be the same as the number leaving. For example, at intersection A, the number of automobiles entering is $x_1 + 450$ and the number leaving is $x_2 + 610$. Thus

$$x_1 + 450 = x_2 + 610 \qquad \text{(intersection } A\text{)}$$

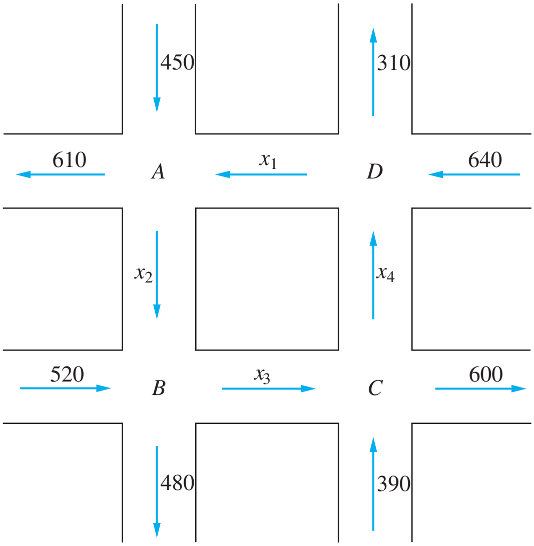


Figure 1.2.2.

Similarly,

$$x_2 + 520 = x_3 + 480 \qquad \text{(intersection } B\text{)}$$
$$x_3 + 390 = x_4 + 600 \qquad \text{(intersection } C\text{)}$$
$$x_4 + 640 = x_1 + 310 \qquad \text{(intersection } D\text{)}$$

The augmented matrix for the system is

$$\left[\begin{array}{rrrr|r} 1 & -1 & 0 & 0 & 160 \\ 0 & 1 & -1 & 0 & -40 \\ 0 & 0 & 1 & -1 & 210 \\ -1 & 0 & 0 & 1 & -330 \end{array}\right]$$

The reduced row echelon form for this matrix is

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & -1 & 330 \\ 0 & 1 & 0 & -1 & 170 \\ 0 & 0 & 1 & -1 & 210 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The system is consistent, and since there is a free variable, there are many possible solutions. The traffic flow diagram does not give enough information to determine x_1, x_2, x_3, and x_4 uniquely. If the amount of traffic were known between any pair of intersections, the traffic on the remaining arteries could easily be calculated. For example, if the amount of traffic between intersections C and D averages 200 automobiles per hour, then $x_4 = 200$. Using this value, we can then solve for x_1, x_2, and x_3:

$$x_1 = x_4 + 330 = 530$$
$$x_2 = x_4 + 170 = 370$$
$$x_3 = x_4 + 210 = 410$$

APPLICATION 2 Electrical Networks

In an electrical network, it is possible to determine the amount of current in each branch in terms of the resistances and the voltages. An example of a typical circuit is given in Figure 1.2.3.

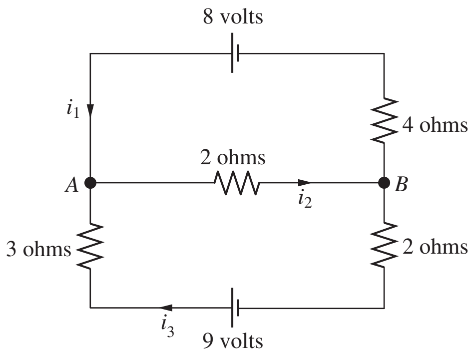
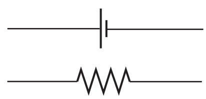


Figure 1.2.3.

The symbols in the figure have the following meanings:



A path along which current may flow

An electrical source

A resistor

The electrical source is usually a battery with a voltage (measured in volts) that drives a charge and produces a current. The current will flow out from the terminal of the battery that is represented by the longer vertical line. The resistances are measured in ohms. The letters represent nodes and the i's represent the currents between the nodes. The currents are measured in amperes. The arrows show the direction of the currents. If, however, one of the currents, say, i_2, turns out to be negative, this would mean that the current along that branch is in the direction opposite that of the arrow.

To determine the currents, the following rules are used:

Kirchhoff's Laws

1. At every node the sum of the incoming currents equals the sum of the outgoing currents.
2. Around every closed loop, the algebraic sum of the voltage gains must equal the algebraic sum of the voltage drops.

The voltage drops E for each resistor are given by *Ohm's law*:

$$E = iR$$

where i represents the current in amperes and R the resistance in ohms.

Let us find the currents in the network pictured in Figure 1.2.3. From the first law, we have

$$\begin{aligned} i_1 - i_2 + i_3 &= 0 \qquad &\text{(node } A\text{)} \\ -i_1 + i_2 - i_3 &= 0 \qquad &\text{(node } B\text{)} \end{aligned}$$

By the second law,

$$\begin{aligned} 4i_1 + 2i_2 &= 8 \qquad \text{(top loop)} \\ 2i_2 + 5i_3 &= 9 \qquad \text{(bottom loop)} \end{aligned}$$

The network can be represented by the augmented matrix

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 4 & 2 & 0 & 8 \\ 0 & 2 & 5 & 9 \end{array}\right]$$

This matrix is easily reduced to the row echelon form

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{2}{3} & \frac{4}{3} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Solving by back substitution, we see that $i_1 = 1$, $i_2 = 2$, and $i_3 = 1$.

Homogeneous Systems

A system of linear equations is said to be *homogeneous* if the constants on the right-hand side are all zero. Homogeneous systems are always consistent. It is a trivial matter to find a solution; just set all the variables equal to zero. Thus, if an $m \times n$ homogeneous system has a unique solution, it must be the trivial solution $(0, 0, \ldots, 0)$. The homogeneous system in Example 6 consisted of $m = 3$ equations in $n = 4$ unknowns. In the case that $n > m$, there will always be free variables and, consequently, additional nontrivial solutions. This result has essentially been proved in our discussion of underdetermined systems, but, because of its importance, we state it as a theorem.

Theorem 1.2.1 *An $m \times n$ homogeneous system of linear equations has a nontrivial solution if $n > m$.*

Proof A homogeneous system is always consistent. The row echelon form of the matrix can have at most m nonzero rows. Thus there are at most m lead variables. Since there are n variables altogether and $n > m$, there must be some free variables. The free variables can be assigned arbitrary values. For each assignment of values to the free variables, there is a solution of the system. ∎

APPLICATION 3 Chemical Equations

In the process of photosynthesis, plants use radiant energy from sunlight to convert carbon dioxide (CO_2) and water (H_2O) into glucose ($C_6H_{12}O_6$) and oxygen (O_2). The chemical equation of the reaction is of the form

$$x_1CO_2 + x_2H_2O \rightarrow x_3O_2 + x_4C_6H_{12}O_6$$

To balance the equation, we must choose x_1, x_2, x_3, and x_4 so that the numbers of carbon, hydrogen, and oxygen atoms are the same on each side of the equation. Since carbon dioxide contains one carbon atom and glucose contains six, to balance the carbon atoms we require that

$$x_1 = 6x_4$$

Similarly, to balance the oxygen, we need

$$2x_1 + x_2 = 2x_3 + 6x_4$$

and finally, to balance the hydrogen, we need

$$2x_2 = 12x_4$$

If we move all the unknowns to the left-hand sides of the three equations, we end up with the homogeneous linear system

$$\begin{aligned} x_1 \qquad\qquad\qquad - \;\; 6x_4 &= 0 \\ 2x_1 + \;\; x_2 - 2x_3 - \;\; 6x_4 &= 0 \\ 2x_2 \qquad\quad - 12x_4 &= 0 \end{aligned}$$

By Theorem 1.2.1, the system has nontrivial solutions. To balance the equation, we must find solutions (x_1, x_2, x_3, x_4) whose entries are nonnegative integers. If we solve the system in the usual way, we see that x_4 is a free variable and

$$x_1 = x_2 = x_3 = 6x_4$$

In particular, if we take $x_4 = 1$, then $x_1 = x_2 = x_3 = 6$ and the equation takes the form

$$6CO_2 + 6H_2O \rightarrow 6O_2 + C_6H_{12}O_6$$

APPLICATION 4 Economic Models for Exchange of Goods

Suppose that in a primitive society the members of a tribe are engaged in three occupations: farming, manufacturing of tools and utensils, and weaving and sewing of clothing. Assume that initially the tribe has no monetary system and that all goods and services are bartered. Let us denote the three groups by F, M, and C, and suppose that the directed graph in Figure 1.2.4 indicates how the bartering system works in practice.

The figure indicates that the farmers keep half of their produce and give one-fourth of their produce to the manufacturers and one-fourth to the clothing producers. The manufacturers divide the goods evenly among the three groups, one-third going to each group. The group producing clothes gives half of the clothes to the farmers and divides the other half evenly between the manufacturers and themselves. The result is summarized in the following table:

	F	M	C
F	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{2}$
M	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$
C	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{4}$

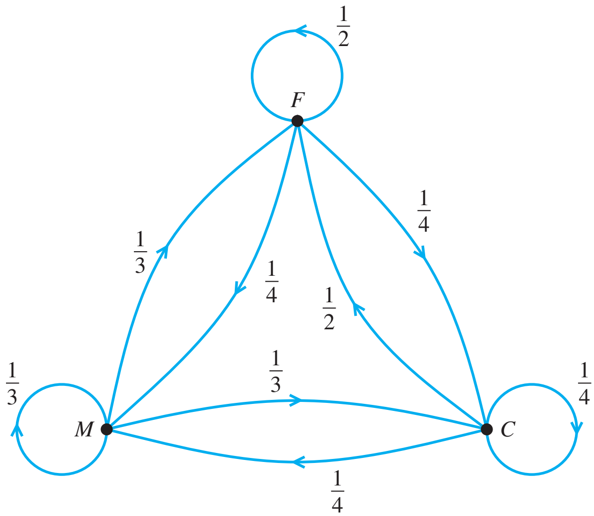


Figure 1.2.4.

The first column of the table indicates the distribution of the goods produced by the farmers, the second column indicates the distribution of the manufactured goods, and the third column indicates the distribution of the clothing.

As the size of the tribe grows, the system of bartering becomes too cumbersome and, consequently, the tribe decides to institute a monetary system of exchange. For this simple economic system, we assume that there will be no accumulation of capital or debt and that the prices for each of the three types of goods will reflect the values of the existing bartering system. The question is how to assign values to the three types of goods that fairly represent the current bartering system.

The problem can be turned into a linear system of equations using an economic model that was originally developed by the Nobel Prize-winning economist Wassily Leontief. For this model, we will let x_1 be the monetary value of the goods produced by the farmers, x_2 be the value of the manufactured goods, and x_3 be the value of all the clothing produced. According to the first row of the table, the value of the goods received by the farmers amounts to half the value of the farm goods produced, plus one-third the value of the manufactured products, and half the value of the clothing goods. Thus the total value of goods received by the farmer is $\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3$. If the system is fair, the total value of goods received by the farmers should equal x_1, the total value of the farm goods produced. Hence, we have the linear equation

$$\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3 = x_1$$

Using the second row of the table and equating the value of the goods produced and received by the manufacturers, we obtain a second equation:

$$\frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = x_2$$

Finally, using the third row of the table, we get

$$\frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = x_3$$

These equations can be rewritten as a homogeneous system:

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

The reduced row echelon form of the augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{5}{3} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

There is one free variable: x_3. Setting $x_3 = 3$, we obtain the solution $(5, 3, 3)$, and the general solution consists of all multiples of $(5, 3, 3)$. It follows that the variables x_1, x_2, and x_3 should be assigned values in the ratio

$$x_1 : x_2 : x_3 = 5 : 3 : 3$$

This simple system is an example of the closed Leontief input-output model. Leontief's models are fundamental to our understanding of economic systems. Modern applications would involve thousands of industries and lead to very large linear systems. The Leontief models will be studied in greater detail later in Section 6.8 of the book.

SECTION 1.2 EXERCISES

1. Which of the matrices that follow are in row echelon form? Which are in reduced row echelon form?

(a) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ **(b)** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

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

(e) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ **(f)** $\begin{bmatrix} 1 & 4 & 6 \\ 0 & 0 & 1 \\ 0 & 1 & 3 \end{bmatrix}$

(g) $\begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 & 4 \\ 0 & 0 & 1 & 3 & 6 \end{bmatrix}$ **(h)** $\begin{bmatrix} 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

2. The augmented matrices that follow are in row echelon form. For each case, indicate whether the corresponding linear system is consistent. If the system has a unique solution, find it.

(a) $\left[\begin{array}{cc|c} 1 & 2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{array}\right]$ **(b)** $\left[\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right]$

(c) $\left[\begin{array}{ccc|c} 1 & -2 & 4 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(d) $\left[\begin{array}{ccc|c} 1 & -2 & 2 & -2 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$

(e) $\left[\begin{array}{ccc|c} 1 & 3 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{array}\right]$

(f) $\left[\begin{array}{rrr|r} 1 & -1 & 3 & 8 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$

3. The augmented matrices that follow are in reduced row echelon form. In each case, find the solution set to the corresponding linear system.

(a) $\left[\begin{array}{rrr|r} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right]$ (b) $\left[\begin{array}{rrr|r} 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right]$

(c) $\left[\begin{array}{rrr|r} 1 & -3 & 0 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(d) $\left[\begin{array}{rrrr|r} 1 & 2 & 0 & 1 & 5 \\ 0 & 0 & 1 & 3 & 4 \end{array}\right]$

(e) $\left[\begin{array}{rrrr|r} 1 & 5 & -2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

(f) $\left[\begin{array}{rrr|r} 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$

4. For each of the systems in Exercise 3, make a list of the lead variables and a second list of the free variables.

5. For each of the systems of equations that follow, use Gaussian elimination to obtain an equivalent system whose coefficient matrix is in row echelon form. Indicate whether the system is consistent. If the system is consistent and involves no free variables, use back substitution to find the unique solution. If the system is consistent and there are free variables, transform it to reduced row echelon form and find all solutions.

(a) $\begin{aligned} x_1 - 2x_2 &= 3 \\ 2x_1 - x_2 &= 9 \end{aligned}$ (b) $\begin{aligned} 2x_1 - 3x_2 &= 5 \\ -4x_1 + 6x_2 &= 8 \end{aligned}$

(c) $\begin{aligned} x_1 + x_2 &= 0 \\ 2x_1 + 3x_2 &= 0 \\ 3x_1 - 2x_2 &= 0 \end{aligned}$ (d) $\begin{aligned} 3x_1 + 2x_2 - x_3 &= 4 \\ x_1 - 2x_2 + 2x_3 &= 1 \\ 11x_1 + 2x_2 + x_3 &= 14 \end{aligned}$

(e) $\begin{aligned} 2x_1 + 3x_2 + x_3 &= 1 \\ x_1 + x_2 + x_3 &= 3 \\ 3x_1 + 4x_2 + 2x_3 &= 4 \end{aligned}$

(f) $\begin{aligned} x_1 - x_2 + 2x_3 &= 4 \\ 2x_1 + 3x_2 - x_3 &= 1 \\ 7x_1 + 3x_2 + 4x_3 &= 7 \end{aligned}$

(g) $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ 2x_1 + 3x_2 - x_3 - x_4 &= 2 \\ 3x_1 + 2x_2 + x_3 + x_4 &= 5 \\ 3x_1 + 6x_2 - x_3 - x_4 &= 4 \end{aligned}$ (h) $\begin{aligned} x_1 - 2x_2 &= 3 \\ 2x_1 + x_2 &= 1 \\ -5x_1 + 8x_2 &= 4 \end{aligned}$

(i) $\begin{aligned} -x_1 + 2x_2 - x_3 &= 2 \\ -2x_1 + 2x_2 + x_3 &= 4 \\ 3x_1 + 2x_2 + 2x_3 &= 5 \\ -3x_1 + 8x_2 + 5x_3 &= 17 \end{aligned}$

(j) $\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 1 \\ -x_1 - x_2 + 4x_3 - x_4 &= 6 \\ -2x_1 - 4x_2 + 7x_3 - x_4 &= 1 \end{aligned}$

(k) $\begin{aligned} x_1 + 3x_2 + x_3 + x_4 &= 3 \\ 2x_1 - 2x_2 + x_3 + 2x_4 &= 8 \\ x_1 - 5x_2 + x_4 &= 5 \end{aligned}$

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

6. Use Gauss–Jordan reduction to solve each of the following systems.

(a) $\begin{aligned} x_1 + x_2 &= -1 \\ 4x_1 - 3x_2 &= 3 \end{aligned}$

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

(c) $\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0 \end{aligned}$

(d) $\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ 2x_1 + x_2 - x_3 + 3x_4 &= 0 \\ x_1 - 2x_2 + x_3 + x_4 &= 0 \end{aligned}$

7. Give a geometric explanation of why a homogeneous linear system consisting of two equations in three unknowns must have infinitely many solutions. What are the possible numbers of solutions of a nonhomogeneous 2×3 linear system? Give a geometric explanation of your answer.

8. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 1 \\ -1 & 4 & 3 & 2 \\ 2 & -2 & a & 3 \end{array}\right)$$

For what values of a will the system have a unique solution?

9. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{rrr|r} 1 & 2 & 1 & 0 \\ 2 & 5 & 3 & 0 \\ -1 & 1 & \beta & 0 \end{array}\right)$$

(a) Is it possible for the system to be inconsistent? Explain.

(b) For what values of β will the system have infinitely many solutions?

10. Consider a linear system whose augmented matrix is of the form

$$\left(\begin{array}{rrr|r} 1 & 1 & 3 & 2 \\ 1 & 2 & 4 & 3 \\ 1 & 3 & a & b \end{array}\right)$$

(a) For what values of a and b will the system have infinitely many solutions?

(b) For what values of a and b will the system be inconsistent?

11. Given the linear systems

(i) $x_1 + 2x_2 = 2$ **(ii)** $x_1 + 2x_2 = 1$

$3x_1 + 7x_2 = 8$ $3x_1 + 7x_2 = 7$

solve both systems by incorporating the right-hand sides into a 2×2 matrix B and computing the reduced row echelon form of

$$(A|B) = \left(\begin{array}{rr|rr} 1 & 2 & 2 & 1 \\ 3 & 7 & 8 & 7 \end{array}\right)$$

12. Given the linear systems

(i)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= 2 \\ -x_1 - x_2 + 2x_3 &= 3 \\ 2x_1 + 3x_2 &= 0 \end{aligned}$$

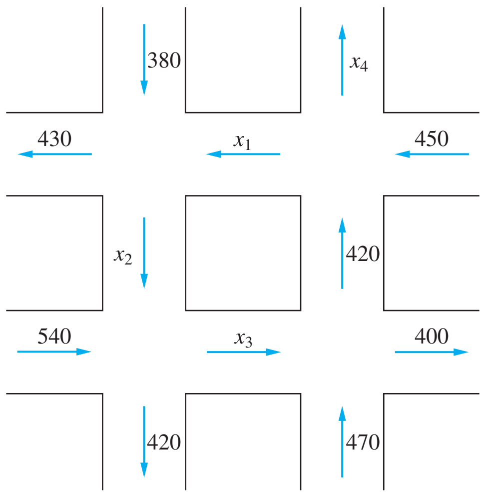
(ii)
$$\begin{aligned} x_1 + 2x_2 + x_3 &= -1 \\ -x_1 - x_2 + 2x_3 &= 2 \\ 2x_1 + 3x_2 &= -2 \end{aligned}$$

solve both systems by computing the row echelon form of an augmented matrix $(A|B)$ and performing back substitution twice.

13. Given a homogeneous system of linear equations, if the system is overdetermined, what are the possibilities as to the number of solutions? Explain.

14. Given a nonhomogeneous system of linear equations, if the system is underdetermined, what are the possibilities as to the number of solutions? Explain.

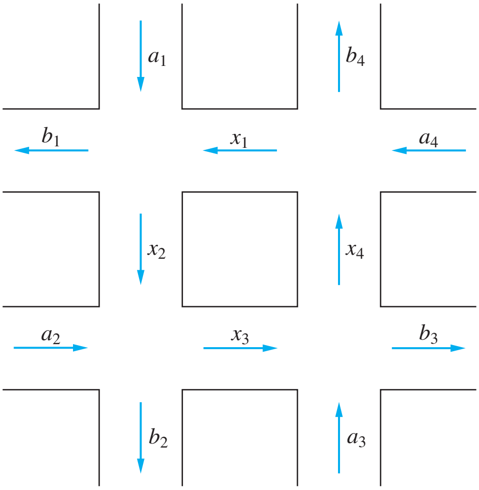
15. Determine the values of x_1, x_2, x_3, x_4 for the following traffic flow diagram.



16. Consider the traffic flow diagram that follows, where $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4$ are fixed positive integers. Set up a linear system in the unknowns x_1, x_2, x_3, x_4 and show that the system will be consistent if and only if

$$a_1 + a_2 + a_3 + a_4 = b_1 + b_2 + b_3 + b_4$$

What can you conclude about the number of automobiles entering and leaving the traffic network?



17. Let (c_1, c_2) be a solution of the 2×2 system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= 0 \\ a_{21}x_1 + a_{22}x_2 &= 0 \end{aligned}$$

Show that for any real number α the ordered pair $(\alpha c_1, \alpha c_2)$ is also a solution.

18. In Application 3 the solution $(6, 6, 6, 1)$ was obtained by setting the free variable $x_4 = 1$.

(a) Determine the solution corresponding to $x_4 = 0$. What information, if any, does this solution give about the chemical reaction? Is the term "trivial solution" appropriate in this case?

(b) Choose some other values of x_4, such as 2, 4, or 5, and determine the corresponding solutions. How are these nontrivial solutions related?

19. Liquid benzene burns in the atmosphere. If a cold object is placed directly over the benzene, water will condense on the object and a deposit of soot (carbon) will also form on the object. The chemical equation for this reaction is of the form

$$x_1 C_6H_6 + x_2 O_2 \rightarrow x_3 C + x_4 H_2O$$

Determine values of x_1, x_2, x_3, and x_4 to balance the equation.

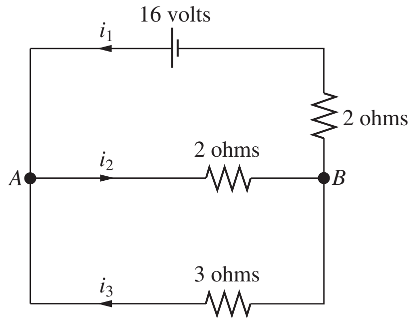
20. Nitric acid is prepared commercially by a series of three chemical reactions. In the first reaction, nitrogen (N_2) is combined with hydrogen (H_2) to form ammonia (NH_3). Next the ammonia is combined with oxygen (O_2) to form nitrogen dioxide (NO_2) and water. Finally, the NO_2 reacts with some of the water to form nitric acid (HNO_3) and nitric oxide (NO). The amounts of each of the components of these reactions are measured in moles (a standard unit of measurement for chemical reactions). How many moles of nitrogen, hydrogen, and oxygen are necessary to produce 8 moles of nitric acid?

21. In Application 4, determine the relative values of x_1, x_2, and x_3 if the distribution of goods is as described in the following table.

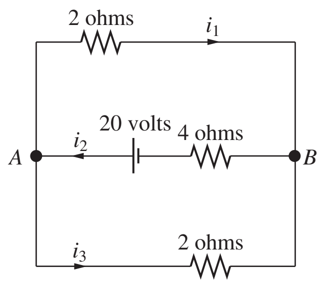
	F	M	C
F	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
M	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$
C	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{2}$

22. Determine the amount of each current for the following networks:

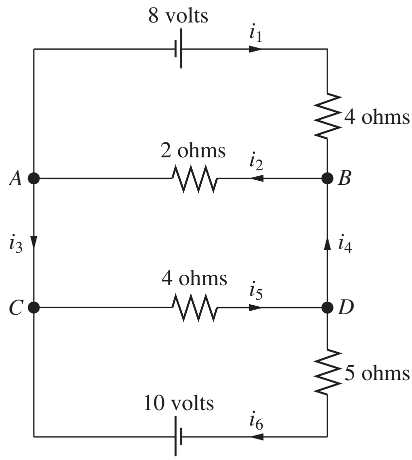
(a)



(b)



(c)



1.3 Matrix Arithmetic

In this section, we introduce the standard notations used for matrices and vectors and define arithmetic operations (addition, subtraction, and multiplication) with matrices. We will also introduce two additional operations: *scalar multiplication* and *transposition*. We will see how to represent linear systems as equations involving matrices and vectors and then derive a theorem characterizing when a linear system is consistent.

The entries of a matrix are called *scalars*. They are usually either real or complex numbers. For the most part, we will be working with matrices whose entries are real numbers. Throughout the first five chapters of the book, the reader may assume that the term *scalar* refers to a real number. However, in Chapter 6 there will be occasions when we will use the set of complex numbers as our scalar field.

Matrix Notation

If we wish to refer to matrices without specifically writing out all their entries, we will use capital letters A, B, C, and so on. In general, a_{ij} will denote the entry of the matrix A that is in the ith row and the jth column. We will refer to this entry as the (i, j) entry of A. Thus, if A is an $m \times n$ matrix, then

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

We will sometimes shorten this to $A = (a_{ij})$. Similarly, a matrix B may be referred to as (b_{ij}), a matrix C as (c_{ij}), and so on.

Vectors

Matrices that have only one row or one column are of special interest, since they are used to represent solutions of linear systems. A solution of a system of m linear equations in n unknowns is an n-tuple of real numbers. We will refer to an n-tuple of real numbers as a *vector*. If an n-tuple is represented in terms of a $1 \times n$ matrix, then we will refer to it as a *row vector*. Alternatively, if the n-tuple is represented by an $n \times 1$ matrix, then we will refer to it as a *column vector*. For example, the solution of the linear system

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

can be represented by the row vector $(2, 1)$ or the column vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

In working with matrix equations, it is generally more convenient to represent the solutions in terms of column vectors ($n \times 1$ matrices). The set of all $n \times 1$ matrices of real numbers is called *Euclidean n-space* and is usually denoted by $\mathbb{R}^n$. Since we will be working almost exclusively with column vectors in the future, we will generally omit the word "column" and refer to the elements of $\mathbb{R}^n$ as simply *vectors*, rather than as column vectors. The standard notation for a column vector is a boldface lowercase letter, as in

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

For row vectors, there is no universal standard notation. In this book, we will represent both row and column vectors with boldface lower case letters and to distinguish a row vector from a column vector we will place a horizontal arrow above the letter. Thus, the horizontal arrow indicates an horizontal array (row vector) rather than a vertical array (column vector). For example,

$$\vec{\mathbf{x}} = (x_1, x_2, x_3, x_4) \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

are row and column vectors, respectively, with four entries each.

Given an $m \times n$ matrix A, it is often necessary to refer to a particular row or column. The standard notation for the jth column vector of A is $\mathbf{a}_j$. There is no universally accepted standard notation for the ith row vector of a matrix A. In this book, since we use horizontal arrows to indicate row vectors, we denote the ith row vector of A by $\vec{\mathbf{a}}_i$.

If A is an $m \times n$ matrix, then the row vectors of A are given by

$$\vec{\mathbf{a}}_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \qquad i = 1, \ldots, m$$

and the column vectors are given by

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \qquad j = 1, \ldots, n$$

The matrix A can be represented in terms of either its column vectors or its row vectors:

$$A = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) \qquad \text{or} \qquad A = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$$

Similarly, if B is an $n \times r$ matrix, then

$$B = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_r) = \begin{bmatrix} \vec{\mathbf{b}}_1 \\ \vec{\mathbf{b}}_2 \\ \vdots \\ \vec{\mathbf{b}}_n \end{bmatrix}$$

EXAMPLE 1 If

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

then

$$\mathbf{a}_1 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

and

$$\vec{\mathbf{a}}_1 = (3,\ 2,\ 5), \quad \vec{\mathbf{a}}_2 = (-1,\ 8,\ 4)$$

■

Equality

For two matrices to be equal, they must have the same dimensions and their corresponding entries must agree.

Definition

Two $m \times n$ matrices A and B are said to be **equal** if $a_{ij} = b_{ij}$ for each i and j.

Scalar Multiplication

If A is a matrix and α is a scalar, then αA is the matrix formed by multiplying each of the entries of A by α.

Definition

If A is an $m \times n$ matrix and α is a scalar, then αA is the $m \times n$ matrix whose (i, j) entry is αa_{ij}.

For example, if

$$A = \begin{bmatrix} 4 & 8 & 2 \\ 6 & 8 & 10 \end{bmatrix}$$

then

$$\frac{1}{2}A = \begin{bmatrix} 2 & 4 & 1 \\ 3 & 4 & 5 \end{bmatrix} \quad \text{and} \quad 3A = \begin{bmatrix} 12 & 24 & 6 \\ 18 & 24 & 30 \end{bmatrix}$$

Matrix Addition

Two matrices with the same dimensions can be added by adding their corresponding entries.

Definition

If $A = (a_{ij})$ and $B = (b_{ij})$ are both $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose (i, j) entry is $a_{ij} + b_{ij}$ for each ordered pair (i, j).

For example,

$$\begin{pmatrix} 3 & 2 & 1 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 2 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 3 \\ 5 & 7 & 9 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \\ 8 \end{pmatrix} + \begin{pmatrix} -8 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -6 \\ 4 \\ 10 \end{pmatrix}$$

If we define $A - B$ to be $A + (-1)B$, then it turns out that $A - B$ is formed by subtracting the corresponding entry of B from each entry of A. Thus,

$$\begin{aligned}
\begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} - \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} &= \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + (-1)\begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 4 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} -4 & -5 \\ -2 & -3 \end{pmatrix} \\
&= \begin{pmatrix} 2-4 & 4-5 \\ 3-2 & 1-3 \end{pmatrix} \\
&= \begin{pmatrix} -2 & -1 \\ 1 & -2 \end{pmatrix}
\end{aligned}$$

If O represents the matrix, with the same dimensions as A, whose entries are all 0, then

$$A + O = O + A = A$$

We will refer to O as the *zero matrix*. It acts as an additive identity on the set of all $m \times n$ matrices. Furthermore, each $m \times n$ matrix A has an additive inverse. Indeed,

$$A + (-1)A = O = (-1)A + A$$

It is customary to denote the additive inverse by $-A$. Thus,

$$-A = (-1)A$$

Matrix Multiplication and Linear Systems

We have yet to define the most important operation: the multiplication of two matrices. Much of the motivation behind the definition comes from the applications to linear systems of equations. If we have a system of one linear equation in one unknown, it can be written in the form

$$ax = b \tag{2}$$

We generally think of a, x, and b as being scalars; however, they could also be treated as 1×1 matrices. Our goal now is to generalize equation (2) so that we can represent an $m \times n$ linear system by a single matrix equation of the form

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

where A is an $m \times n$ matrix, $\mathbf{x}$ is an unknown vector in $\mathbb{R}^n$, and $\mathbf{b}$ is in $\mathbb{R}^m$. We consider first the case of one equation in several unknowns.

Case 1. One Equation in Several Unknowns
Let us begin by examining the case of one equation in several variables. Consider, for example, the equation

$$3x_1 + 2x_2 + 5x_3 = 4$$

If we set

$$A = \begin{pmatrix} 3 & 2 & 5 \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = \begin{pmatrix} 3 & 2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 3x_1 + 2x_2 + 5x_3$$

then the equation $3x_1 + 2x_2 + 5x_3 = 4$ can be written as the matrix equation

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

For a linear equation with n unknowns of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

if we let

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

then the system can be written in the form $A\mathbf{x} = \mathbf{b}$.

For example, if

$$A = \begin{pmatrix} 2 & 1 & -3 & 4 \end{pmatrix} \qquad \text{and} \qquad \mathbf{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ -2 \end{pmatrix}$$

then

$$A\mathbf{x} = 2 \cdot 3 + 1 \cdot 2 + (-3) \cdot 1 + 4 \cdot (-2) = -3$$

Note that the result of multiplying a row vector on the left by a column vector on the right is a scalar. Consequently, this type of multiplication is often referred to as a *scalar product*.

Case 2. *M* Equations in *N* Unknowns
Consider now an $m \times n$ linear system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{3}$$

It is desirable to write the system (3) in a form similar to (2), that is, as a matrix equation

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

where $A = (a_{ij})$ is known, $\mathbf{x}$ is an $n \times 1$ matrix of unknowns, and $\mathbf{b}$ is an $m \times 1$ matrix representing the right-hand side of the system. Thus, if we set

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

and define the product $A\mathbf{x}$ by

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} \tag{5}$$

then the linear system of equations (3) is equivalent to the matrix equation (4).

Given an $m \times n$ matrix A and a vector $\mathbf{x}$ in $\mathbb{R}^n$, it is possible to compute a product $A\mathbf{x}$ by (5). The product $A\mathbf{x}$ will be an $m \times 1$ matrix, that is, a vector in $\mathbb{R}^m$. The rule for determining the ith entry of $A\mathbf{x}$ is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$$

which is equal to $\vec{\mathbf{a}}_i\mathbf{x}$, the scalar product of the ith row vector of A and the column vector $\mathbf{x}$. Thus,

$$A\mathbf{x} = \begin{bmatrix} \vec{\mathbf{a}}_1\mathbf{x} \\ \vec{\mathbf{a}}_2\mathbf{x} \\ \vdots \\ \vec{\mathbf{a}}_n\mathbf{x} \end{bmatrix}$$

EXAMPLE 2

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 5 & 3 & 7 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} 4x_1 + 2x_2 + x_3 \\ 5x_1 + 3x_2 + 7x_3 \end{bmatrix}$$

■

EXAMPLE 3

$$A = \begin{bmatrix} -3 & 1 \\ 2 & 5 \\ 4 & 2 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} -3 \cdot 2 + 1 \cdot 4 \\ 2 \cdot 2 + 5 \cdot 4 \\ 4 \cdot 2 + 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 24 \\ 16 \end{bmatrix}$$ ■

EXAMPLE 4 Write the following system of equations as a matrix equation of the form $A\mathbf{x} = \mathbf{b}$:

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

Solution

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

An alternative way to represent the linear system (3) as a matrix equation is to express the product $A\mathbf{x}$ as a sum of column vectors:

$$A\mathbf{x} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Thus, we have

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \tag{6}$$

Using this formula, we can represent the system of equations (3) as a matrix equation of the form

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{7}$$

EXAMPLE 5 The linear system

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

can be written as a matrix equation

$$x_1 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ -4 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$ ■

Definition

If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ are vectors in $\mathbb{R}^m$ and $c_1, c_2, \ldots, c_n$ are scalars, then a sum of the form

$$c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n$$

is said to be a **linear combination** of the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$.

It follows from equation (6) that the product $A\mathbf{x}$ is a linear combination of the column vectors of A. Some books even use this linear combination representation as the definition of matrix vector multiplication.

If A is an $m \times n$ matrix and $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then

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

EXAMPLE 6 If we choose $x_1 = 2$, $x_2 = 3$, and $x_3 = 4$ in Example 5, then

$$\begin{Bmatrix} 5 \\ 6 \end{Bmatrix} = 2\begin{Bmatrix} 2 \\ 5 \end{Bmatrix} + 3\begin{Bmatrix} 3 \\ -4 \end{Bmatrix} + 4\begin{Bmatrix} -2 \\ 2 \end{Bmatrix}$$

Thus, the vector $\begin{Bmatrix} 5 \\ 6 \end{Bmatrix}$ is a linear combination of the three column vectors of the coefficient matrix. It follows that the linear system in Example 5 is consistent and

$$\mathbf{x} = \begin{Bmatrix} 2 \\ 3 \\ 4 \end{Bmatrix}$$

is a solution of the system. ■

The matrix equation (7) provides a nice way of characterizing whether a linear system of equations is consistent. Indeed, the following theorem is a direct consequence of (7).

Theorem 1.3.1 Consistency Theorem for Linear Systems

A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be written as a linear combination of the column vectors of A.

EXAMPLE 7 The linear system

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ 2x_1 + 4x_2 &= 1 \end{aligned}$$

is inconsistent since the vector $\begin{Bmatrix} 1 \\ 1 \end{Bmatrix}$ cannot be written as a linear combination of the column vectors $\begin{Bmatrix} 1 \\ 2 \end{Bmatrix}$ and $\begin{Bmatrix} 2 \\ 4 \end{Bmatrix}$. Note that any linear combination of these vectors would be of the form

$$x_1\begin{Bmatrix} 1 \\ 2 \end{Bmatrix} + x_2\begin{Bmatrix} 2 \\ 4 \end{Bmatrix} = \begin{Bmatrix} x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{Bmatrix}$$

and hence the second entry of the vector must be double the first entry. ■

Matrix Multiplication

More generally, it is possible to multiply a matrix A times a matrix B if the number of columns of A equals the number of rows of B. The first column of the product is determined by the first column of B; that is, the first column of AB is $A\mathbf{b}_1$, the second column of AB is $A\mathbf{b}_2$, and so on. Thus the product AB is the matrix whose columns are $A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_n$.

$$AB = (A\mathbf{b}_1, A\mathbf{b}_2, \ldots, A\mathbf{b}_n)$$

The (i, j) entry of AB is the ith entry of the column vector $A\mathbf{b}_j$. It is determined by multiplying the ith row vector of A times the jth column vector of B.

Definition

If $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times r$ matrix, then the product $AB = C = (c_{ij})$ is the $m \times r$ matrix whose entries are defined by

$$c_{ij} = \vec{\mathbf{a}}_i \mathbf{b}_j = \sum_{k=1}^{n} a_{ik} b_{kj}$$

EXAMPLE 8 If

$$A = \begin{bmatrix} 3 & -2 \\ 2 & 4 \\ 1 & -3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -2 & 1 & 3 \\ 4 & 1 & 6 \end{bmatrix}$$

then

$$\begin{aligned} AB &= \begin{bmatrix} 3 & -2 \\ \mathbf{2} & \mathbf{4} \\ 1 & -3 \end{bmatrix} \begin{bmatrix} -2 & 1 & \mathbf{3} \\ 4 & 1 & \mathbf{6} \end{bmatrix} \\ &= \begin{bmatrix} 3 \cdot (-2) - 2 \cdot 4 & 3 \cdot 1 - 2 \cdot 1 & 3 \cdot 3 - 2 \cdot 6 \\ 2 \cdot (-2) + 4 \cdot 4 & 2 \cdot 1 + 4 \cdot 1 & \mathbf{2 \cdot 3 + 4 \cdot 6} \\ 1 \cdot (-2) - 3 \cdot 4 & 1 \cdot 1 - 3 \cdot 1 & 1 \cdot 3 - 3 \cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} -14 & 1 & -3 \\ 12 & 6 & \mathbf{30} \\ -14 & -2 & -15 \end{bmatrix} \end{aligned}$$

The shading indicates how the $(2, 3)$ entry of the product AB is computed as a scalar product of the second row vector of A and the third column vector of B. It is also possible to multiply B times A; however, the resulting matrix BA is not equal to AB. In fact, AB and BA do not even have the same dimensions.

$$\begin{aligned} BA &= \begin{bmatrix} -2 \cdot 3 + 1 \cdot 2 + 3 \cdot 1 & -2 \cdot (-2) + 1 \cdot 4 + 3 \cdot (-3) \\ 4 \cdot 3 + 1 \cdot 2 + 6 \cdot 1 & 4 \cdot (-2) + 1 \cdot 4 + 6 \cdot (-3) \end{bmatrix} \\ &= \begin{bmatrix} -1 & -1 \\ 20 & -22 \end{bmatrix} \end{aligned}$$

■

EXAMPLE 9 If

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix}$$

then it is impossible to multiply A times B, since the number of columns of A does not equal the number of rows of B. However, it is possible to multiply B times A.

$$BA = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 8 \\ 17 & 26 \\ 15 & 24 \end{bmatrix}$$

■

If A and B are both $n \times n$ matrices, then AB and BA will also be $n \times n$ matrices, but, in general, they will not be equal. *Multiplication of matrices is not commutative*.

EXAMPLE 10 If

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

then

$$AB = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 0 \end{bmatrix}$$

and

$$BA = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

Hence $AB \neq BA$. ■

APPLICATION 1 Production Costs

A company manufactures three products. Its production expenses are divided into three categories. In each category, an estimate is given for the cost of producing a single item of each product. An estimate is also made of the amount of each product to be produced per quarter. These estimates are given in Tables 1 and 2. At its stockholders' meeting the company would like to present a single table showing the total costs for each quarter in each of the three categories: raw materials, labor, and overhead.

Table 1 Production Costs per Item (dollars)

	Product		
Expenses	**A**	**B**	**C**
Raw materials	0.10	0.30	0.15
Labor	0.30	0.40	0.25
Overhead and miscellaneous	0.10	0.20	0.15

Table 2 Amount Produced per Quarter

	Season			
Product	**Summer**	**Fall**	**Winter**	**Spring**
A	4000	4500	4500	4000
B	2000	2600	2400	2200
C	5800	6200	6000	6000

Solution

Let us consider the problem in terms of matrices. Each of the two tables can be represented by a matrix, namely,

$$M = \begin{bmatrix} 0.10 & 0.30 & 0.15 \\ 0.30 & 0.40 & 0.25 \\ 0.10 & 0.20 & 0.15 \end{bmatrix}$$

and

$$P = \begin{bmatrix} 4000 & 4500 & 4500 & 4000 \\ 2000 & 2600 & 2400 & 2200 \\ 5800 & 6200 & 6000 & 6000 \end{bmatrix}$$

If we form the product MP, the first column of MP will represent the costs for the summer quarter:

$$\begin{array}{ll} \text{Raw materials:} & (0.10)(4000) + (0.30)(2000) + (0.15)(5800) = 1870 \\ \text{Labor:} & (0.30)(4000) + (0.40)(2000) + (0.25)(5800) = 3450 \\ \text{Overhead and} & \\ \text{miscellaneous:} & (0.10)(4000) + (0.20)(2000) + (0.15)(5800) = 1670 \end{array}$$

The costs for the fall quarter are given in the second column of MP:

$$\begin{array}{ll} \text{Raw materials:} & (0.10)(4500) + (0.30)(2600) + (0.15)(6200) = 2160 \\ \text{Labor:} & (0.30)(4500) + (0.40)(2600) + (0.25)(6200) = 3940 \\ \text{Overhead and} & \\ \text{miscellaneous:} & (0.10)(4500) + (0.20)(2600) + (0.15)(6200) = 1900 \end{array}$$

Columns 3 and 4 of MP represent the costs for the winter and spring quarters.

$$MP = \begin{bmatrix} 1870 & 2160 & 2070 & 1960 \\ 3450 & 3940 & 3810 & 3580 \\ 1670 & 1900 & 1830 & 1740 \end{bmatrix}$$

The entries in row 1 of MP represent the total cost of raw materials for each of the four quarters. The entries in rows 2 and 3 represent the total cost for labor and overhead, respectively, for each of the four quarters. The yearly expenses in each category may be obtained by adding the entries in each row. The numbers in each of the columns may be added to obtain the total production costs for each quarter. Table 3 summarizes the total production costs. ■

Table 3

	Season				
	Summer	**Fall**	**Winter**	**Spring**	**Year**
Raw materials	1870	2160	2070	1960	8060
Labor	3450	3940	3810	3580	14,780
Overhead and miscellaneous	1670	1900	1830	1740	7140
Total production costs	6990	8000	7710	7280	29,980

APPLICATION 2 Management Science—Analytic Hierarchy Process

The analytic hierarchy process (AHP) is a common technique that is used for analyzing complex decisions. The technique was developed by T. L. Saaty during the 1970s. AHP is used in a wide variety of areas including business, industry, government, education, and health care. The technique is applied to problems with a specific goal and a fixed number of alternatives for achieving the goal. The decision as to which alternative to pick is based on a list of evaluation criteria. In the case of more complex decisions, each evaluation criterion could have a list of subcritera and these in turn could also have subcriteria, and so on. Thus for complex decisions one could have a multilayered hierarchy of decision criteria.

To illustrate how AHP actually works we consider a simple example. A search and screen committee in the Mathematics Department of a state university is conducting a screening process to fill a full professor position in the department. The committee does a preliminary round of screening and narrows the pool down to three candidates: Dr. Gauss, Dr. O'Leary, and Dr. Taussky. After interviewing the finalists the committee must pick the candidate best qualified for the position. To do this they must evaluate each of the candidates in terms of the following criteria: Research, Teaching Ability, and Professional Activities. The hierarchal structural of the decision-making process is illustrated in Figure 1.3.1.

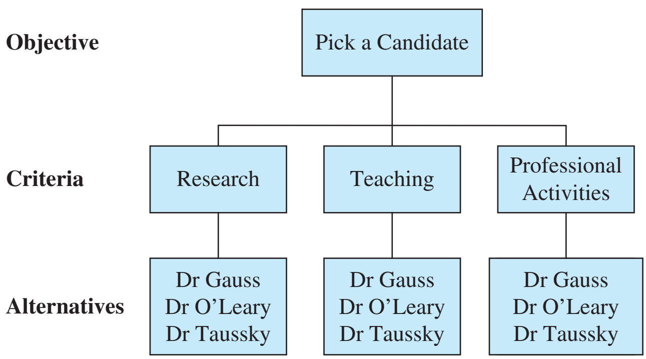


Figure 1.3.1. Analytic Hierarchy Process

The first step of the AHP process is to determine the relative importance of the three areas of evaluation. This can be done using pairwise comparisons. Suppose, for example, that the committee decides that Research and Teaching should be given equal importance and that both of these categories are twice as important as the category of Professional Activities. These relative ratings can be expressed mathematically by assigning the weights 0.40, 0.40, and 0.20 to the respective categories of evaluation. Note that the weights of the first two evaluation criteria are equal and have double the weight of the third. Note also that the weights are chosen so that they all add up to 1. The weight vector

$$\mathbf{w} = \begin{Bmatrix} 0.40 \\ 0.40 \\ 0.20 \end{Bmatrix}$$

provides a numerical representation of the relative importance of the search criteria.

The next step in the process is to assign relative ratings or weights to the three candidates for each of the criteria in our list. Methods for assigning these weights may be either quantitative or qualitative. For example, one could do a quantitative evaluation of research using weights based on the total number of pages published by the candidates in research journals. Thus if Gauss has published 500 pages, O'Leary 250 pages, and Taussky 250 pages, then one could obtain weights by dividing each of these page counts by 1000 (the combined page count for all three individuals). Thus the quantitative weights produced in this manner would be 0.50, 0.25, and 0.25. The quantitative method does not factor in differences in the quality of the publications. Determining qualitative weights involves making some judgments, but the process need not be entirely subjective. Later in the text (in Chapters 5 and 6) we will revisit this example and discuss how to determine qualitative weights. The methods we will consider involve making pairwise comparisons and then using advanced matrix techniques to assign weights based on those comparisons.

Another way the committee could refine the search process would be to break up the research criteria up into two subclasses, quantitative research and qualitative research. In this case one would add a subcriteria row to Figure 1.3.1 directly below the row for criteria. We will incorporate this refinement later when we revisit the AHP application in Section 3 of Chapter 5.

For now, let us assume that the search committee has determined the relative weights for each of the three criteria and that those weights are specified in Figure 1.3.2. The relative ratings for the candidates for research, teaching, and professional activities are given by the vectors

$$\mathbf{a}_1 = \begin{Bmatrix} 0.50 \\ 0.25 \\ 0.25 \end{Bmatrix}, \quad \mathbf{a}_2 = \begin{Bmatrix} 0.20 \\ 0.50 \\ 0.30 \end{Bmatrix}, \quad \mathbf{a}_3 = \begin{Bmatrix} 0.25 \\ 0.50 \\ 0.25 \end{Bmatrix}$$

To determine the overall ranking for the candidates we multiply each of these vectors by the corresponding weights w_1, w_2, w_3 and add.

$$\mathbf{r} = w_1\mathbf{a}_1 + w_2\mathbf{a}_2 + w_3\mathbf{a}_3 = 0.40 \begin{Bmatrix} 0.50 \\ 0.25 \\ 0.25 \end{Bmatrix} + 0.40 \begin{Bmatrix} 0.20 \\ 0.50 \\ 0.30 \end{Bmatrix} + 0.20 \begin{Bmatrix} 0.25 \\ 0.50 \\ 0.25 \end{Bmatrix} = \begin{Bmatrix} 0.33 \\ 0.40 \\ 0.27 \end{Bmatrix}$$

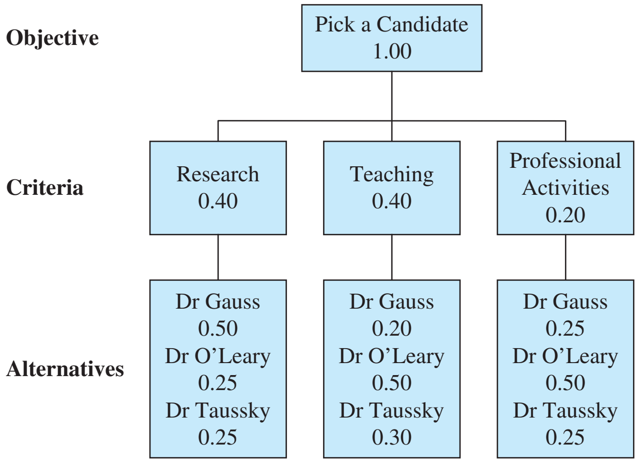


Figure 1.3.2. APH Diagram with Weights

Note that if we set $A = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix}$, then the vector $\mathbf{r}$ of relative ratings is determined by multiplying the matrix A times the vector $\mathbf{w}$.

$$\mathbf{r} = A\mathbf{w} = \begin{pmatrix} 0.50 & 0.20 & 0.25 \\ 0.25 & 0.50 & 0.50 \\ 0.25 & 0.30 & 0.25 \end{pmatrix} \begin{pmatrix} 0.40 \\ 0.40 \\ 0.20 \end{pmatrix} = \begin{pmatrix} 0.33 \\ 0.40 \\ 0.27 \end{pmatrix}$$

In this example the second candidate has the highest relative rating, so the committee eliminates Gauss and Taussky and offers the position to O'Leary. If O'Leary refuses the offer, then next in line is Gauss, the candidate with the second highest rating.

References

1. Saaty, T. L., *The Analytic Hierarchy Process*, McGraw Hill, 1980

Notational Rules

Just as in ordinary algebra, if an expression involves both multiplication and addition and there are no parentheses to indicate the order of the operations, multiplications are carried out before additions. This is true for both scalar and matrix multiplications. For example, if

$$A = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}$$

then

$$A + BC = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} + \begin{pmatrix} 7 & 7 \\ -1 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 11 \\ 0 & 6 \end{pmatrix}$$

and

$$3A + B = \begin{pmatrix} 9 & 12 \\ 3 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 15 \\ 5 & 7 \end{pmatrix}$$

The Transpose of a Matrix

Given an $m \times n$ matrix A, it is often useful to form a new $n \times m$ matrix whose columns are the rows of A.

Definition

The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix B defined by

$$b_{ji} = a_{ij} \tag{8}$$

for $j = 1, \ldots, n$ and $i = 1, \ldots, m$. The transpose of A is denoted by A^T.

It follows from (8) that the jth row of A^T has the same entries, respectively, as the jth column of A, and the ith column of A^T has the same entries, respectively, as the ith row of A.

EXAMPLE 11

(a) If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

(b) If $B = \begin{bmatrix} -3 & 2 & 1 \\ 4 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$, then $B^T = \begin{bmatrix} -3 & 4 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 5 \end{bmatrix}$.

(c) If $C = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, then $C^T = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. ■

The matrix C in Example 11 is its own transpose. This frequently happens with matrices that arise in applications.

Definition

An $n \times n$ matrix A is said to be **symmetric** if $A^T = A$.

The following are some examples of symmetric matrices:

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

APPLICATION 3 Information Retrieval

The growth of digital libraries on the Internet has led to dramatic improvements in the storage and retrieval of information. Modern retrieval methods are based on matrix theory and linear algebra.

In a typical situation, a database consists of a collection of documents and we wish to search the collection and find the documents that best match some particular search conditions. Depending on the type of database, we could search for such items as research articles in journals, Web pages on the Internet, books in a library, or movies in a film collection.

To see how the searches are done, let us assume that our database consists of m documents and that there are n dictionary words that can be used as keywords for searches. Not all words are allowable since it would not be practical to search for common words such as articles or prepositions. If the key dictionary words are ordered alphabetically, then we can represent the database by an $m \times n$ matrix A. Each document is represented by a column of the matrix. The first entry in the jth column of A would be a number representing the relative frequency of the first key dictionary word in the jth document. The entry a_{2j} represents the relative frequency of the second word in the jth document, and so on. The list of keywords to be used in the search is represented by a vector $\mathbf{x}$ in $\mathbb{R}^m$. The ith entry of $\mathbf{x}$ is taken to be 1 if the ith word in the list of keywords is on our search list; otherwise, we set $x_i = 0$. To carry out the search, we simply multiply A^T times $\mathbf{x}$.

Simple Matching Searches

The simplest type of search determines how many of the key search words are in each document; it does not take into account the relative frequencies of the words. Suppose, for example, that our database consists of these book titles:

B1. *Applied Linear Algebra*
B2. *Elementary Linear Algebra*
B3. *Elementary Linear Algebra with Applications*
B4. *Linear Algebra and Its Applications*
B5. *Linear Algebra with Applications*
B6. *Matrix Algebra with Applications*
B7. *Matrix Theory*

The collection of keywords is given by the following alphabetical list:

algebra, application, elementary, linear, matrix, theory

For a simple matching search, we just use 0's and 1's, rather than relative frequencies, for the entries of the database matrix. Thus, the (i, j) entry of the matrix will be 1 if the ith word appears in the title of the jth book and 0 if it does not. We will assume that our search engine is sophisticated enough to equate various forms of a word. So, for example, in our list of titles the words *applied* and *applications* are both counted as forms of the word *application*. The database matrix for our list of books is the array defined by Table 4.

If the words we are searching for are *applied, linear*, and *algebra*, then the database matrix and search vector are, respectively, given by

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Table 4 Array Representation for Database of Linear Algebra Books

	Books						
Key Words	**B1**	**B2**	**B3**	**B4**	**B5**	**B6**	**B7**
algebra	1	1	1	1	1	1	0
application	1	0	1	1	1	1	0
elementary	0	1	1	0	0	0	0
linear	1	1	1	1	1	0	0
matrix	0	0	0	0	0	1	1
theory	0	0	0	0	0	0	1

If we set $\mathbf{y} = A^T\mathbf{x}$, then

$$\mathbf{y} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 3 \\ 3 \\ 3 \\ 2 \\ 0 \end{bmatrix}$$

The value of y_1 is the number of search word matches in the title of the first book, the value of y_2 is the number of matches in the second book title, and so on. Since $y_1 = y_3 = y_4 = y_5 = 3$, the titles of books B1, B3, B4, and B5 must contain all three search words. If the search is set up to find titles matching all search words, then the search engine will report the titles of the first, third, fourth, and fifth books.

Relative Frequency Searches

Searches of noncommercial databases generally find all documents containing the key search words and then order the documents based on the relative frequency. In this case, the entries of the database matrix should represent the relative frequencies of the keywords in the documents. For example, suppose that in the dictionary of all key words of the database the 6th word is *algebra* and the 8th word is *applied*, where all words are listed alphabetically. If, say, document 9 in the database contains a total of 200 occurrences of keywords from the dictionary and if the word *algebra* occurred 10 times in the document and the word *applied* occurred 6 times, then the relative frequencies for these words would be $\frac{10}{200}$ and $\frac{6}{200}$, and the corresponding entries in the database matrix would be

$$a_{69} = 0.05 \quad \text{and} \quad a_{89} = 0.03$$

To search for these two words, we take our search vector $\mathbf{x}$ to be the vector whose entries x_6 and x_8 are both equal to 1 and whose remaining entries are all 0. We then compute

$$\mathbf{y} = A^T\mathbf{x}$$

The entry of $\mathbf{y}$ corresponding to document 9 is

$$y_9 = a_{69} \cdot 1 + a_{89} \cdot 1 = 0.08$$

Note that 16 of the 200 words (8 percent of the words) in document 9 match the key search words. If y_j is the largest entry of $\mathbf{y}$, this would indicate that the jth document in the database is the one that contains the keywords with the greatest relative frequencies.

Advanced Search Methods

A search for the keywords *linear* and *algebra* could easily turn up hundreds of documents, some of which may not even be about linear algebra. If we were to increase the number of search words and require that all search words be matched, then we would run the risk of excluding some crucial linear algebra documents. Rather than match all words of the expanded search list, our database search should give priority to those documents which match most of the keywords with high relative frequencies. To accomplish this, we need to find the columns of the database matrix A that are "closest" to the search vector $\mathbf{x}$. One way to measure how close two vectors are is to define *the angle between the vectors*. We will do this later in Section 5.1 of the book.

We will also revisit the information retrieval application after we have learned about the *singular value decomposition* (Chapter 6, Section 5). This decomposition can be used to find a simpler approximation to the database matrix, which will speed up the searches dramatically. Often it has the added advantage of filtering out *noise*; that is, using the approximate version of the database matrix may automatically have the effect of eliminating documents that use keywords in unwanted contexts. For example, a dental student and a mathematics student could both use *calculus* as one of their search words. Since the list of mathematics search words does not contain any other dental terms, a mathematics search using an approximate database matrix is likely to eliminate all documents relating to dentistry. Similarly, the mathematics documents would be filtered out in the dental student's search.

Web Searches and Page Ranking

Modern Web searches could easily involve billions of documents with hundreds of thousands of keywords. Indeed, as of July 2008, there were more than 1 trillion Web pages on the Internet, and it is not uncommon for search engines to acquire or update as many as 10 million Web pages in a single day. Although the database matrix for pages on the Internet is extremely large, searches can be simplified dramatically, since the matrices and search vectors are *sparse*; that is, most of the entries in any column are 0's.

For Internet searches, the better search engines will do simple matching searches to find all pages matching the keywords, but they will not order them on the basis of the relative frequencies of the keywords. Because of the commercial nature of the Internet, people who want to sell products may deliberately make repeated use of keywords to ensure that their Web site is highly ranked in any relative-frequency search. In fact, it is easy to surreptitiously list a keyword hundreds of times. If the font color of the word matches the background color of the page, then the viewer will not be aware that the word is listed repeatedly.

For Web searches, a more sophisticated algorithm is necessary for ranking the pages that contain all of the key search words. In Chapter 6 we will study a special type of matrix model for assigning probabilities in certain random processes. This type of model is referred to as a *Markov process* or a *Markov chain*. In Section 6.3 we will see how to use Markov chains to model Web surfing and obtain rankings of Web pages.

References

1. Berry, Michael W., and Murray Browne, *Understanding Search Engines: Mathematical Modeling and Text Retrieval*, SIAM, Philadelphia, 1999.
2. Langville, Amy N., and Carl D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings*, Princeton University Press, 2012.

SECTION 1.3 EXERCISES

1. If

$$A = \begin{bmatrix} 3 & 1 & 4 \\ -2 & 0 & 1 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 1 & 1 \\ 2 & -4 & 1 \end{bmatrix}$$

compute

(a) $2A$ **(b)** $A + B$

(c) $2A - 3B$ **(d)** $(2A)^T - (3B)^T$

(e) AB **(f)** BA

(g) $A^T B^T$ **(h)** $(BA)^T$

2. For each of the pairs of matrices that follow, determine whether it is possible to multiply the first matrix times the second. If it is possible, perform the multiplication.

(a) $\begin{bmatrix} 3 & 5 & 1 \\ -2 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \\ 4 & 1 \end{bmatrix}$

(b) $\begin{bmatrix} 4 & -2 \\ 6 & -4 \\ 8 & -6 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 1 \\ 4 & 5 \end{bmatrix}$

(d) $\begin{bmatrix} 4 & 6 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 5 \\ 4 & 1 & 6 \end{bmatrix}$

(e) $\begin{bmatrix} 4 & 6 & 1 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 5 \\ 4 & 1 & 6 \end{bmatrix}$

(f) $\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \begin{bmatrix} 3 & 2 & 4 & 5 \end{bmatrix}$

3. For which of the pairs in Exercise 2 is it possible to multiply the second matrix times the first, and what would the dimension of the product matrix be?

4. Write each of the following systems of equations as a matrix equation:

(a) $\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 2x_1 - 3x_2 &= 5 \end{aligned}$

(b) $\begin{aligned} x_1 + x_2 \phantom{{}-x_3} &= 5 \\ 2x_1 + x_2 - x_3 &= 6 \\ 3x_1 - 2x_2 + 2x_3 &= 7 \end{aligned}$

(c) $\begin{aligned} 2x_1 + x_2 + x_3 &= 4 \\ x_1 - x_2 + 2x_3 &= 2 \\ 3x_1 - 2x_2 - x_3 &= 0 \end{aligned}$

5. If

$$A = \begin{bmatrix} 3 & 4 \\ 1 & 1 \\ 2 & 7 \end{bmatrix}$$

verify that

(a) $5A = 3A + 2A$ **(b)** $6A = 3(2A)$

(c) $(A^T)^T = A$

6. If

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

verify that

(a) $A + B = B + A$

(b) $3(A + B) = 3A + 3B$

(c) $(A + B)^T = A^T + B^T$

7. If

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \\ -2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}$$

verify that

(a) $3(AB) = (3A)B = A(3B)$,

(b) $(AB)^T = B^T A^T$

8. If

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 1 \\ 0 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$$

verify that

(a) $(A + B) + C = A + (B + C)$

(b) $(AB)C = A(BC)$

(c) $A(B + C) = AB + AC$

(d) $(A + B)C = AC + BC$

9. Let

$$A = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} -3 \\ -2 \end{pmatrix}$$

(a) Write $\mathbf{b}$ as a linear combination of the column vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

(b) Use the result from part (a) to determine a solution of the linear system $A\mathbf{x} = \mathbf{b}$. Does the system have any other solutions? Explain.

(c) Write $\mathbf{c}$ as a linear combination of the column vectors $\mathbf{a}_1$ and $\mathbf{a}_2$.

10. For each of the choices of A and $\mathbf{b}$ that follow, determine whether the system $A\mathbf{x} = \mathbf{b}$ is consistent by examining how $\mathbf{b}$ relates to the column vectors of A. Explain your answers in each case.

(a) $A = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$

(b) $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5 \\ 5 \end{pmatrix}$

(c) $A = \begin{pmatrix} 3 & 2 & 1 \\ 3 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$

11. Let A be a 5×3 matrix. If

$$\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_2 + \mathbf{a}_3$$

then what can you conclude about the number of solutions of the linear system $A\mathbf{x} = \mathbf{b}$? Explain.

12. Let A be a 3×4 matrix. If

$$\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4$$

then what can you conclude about the number of solutions to the linear system $A\mathbf{x} = \mathbf{b}$? Explain.

13. Let $A\mathbf{x} = \mathbf{b}$ be a linear system whose augmented matrix $(A|\mathbf{b})$ has reduced row echelon form

$$\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 1 & -2 \\ 0 & 0 & 1 & 2 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

(a) Find all solutions to the system.

(b) If

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 4 \end{pmatrix} \quad \text{and} \quad \mathbf{a}_3 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 3 \end{pmatrix}$$

determine $\mathbf{b}$.

14. Suppose in the search and screen example in Application 2 the committee decides that research is actually 1.5 times as important as teaching and 3 times as important as professional activities. The committee still rates teaching twice as important as professional activities. Determine a new weight vector $\mathbf{w}$ that reflects these revised priorities. Determine also a new rating vector $\mathbf{r}$. Will the new weights have any effect on the overall rankings of the candidates?

15. Let A be an $m \times n$ matrix. Explain why the matrix multiplications $A^T A$ and AA^T are possible.

16. A matrix A is said to be *skew symmetric* if $A^T = -A$. Show that if a matrix is skew symmetric, then its diagonal entries must all be 0.

17. In Application 3, suppose that we are searching the database of seven linear algebra books for the search words *elementary, matrix, algebra*. Form a search vector $\mathbf{x}$, and then compute a vector $\mathbf{y}$ that represents the results of the search. Explain the significance of the entries of the vector $\mathbf{y}$.

18. Let A be a 2×2 matrix with $a_{11} \neq 0$ and let $\alpha = a_{21}/a_{11}$. Show that A can be factored into a product of the form

$$\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ 0 & b \end{pmatrix}$$

What is the value of b?

1.4 Matrix Algebra

The algebraic rules used for real numbers may or may not work when matrices are used. For example, if a and b are real numbers, then

$$a + b = b + a \quad \text{and} \quad ab = ba$$

For real numbers, the operations of addition and multiplication are both commutative. The first of these algebraic rules works when we replace a and b by square matrices A and B, that is,

$$A + B = B + A$$

However, we have already seen that matrix multiplication is not commutative. This fact deserves special emphasis.

> **Warning:** In general, $AB \neq BA$. Matrix multiplication is *not* commutative.

In this section we examine which algebraic rules work for matrices and which do not.

Algebraic Rules

The following theorem provides some useful rules for doing matrix algebra.

Theorem 1.4.1 *Each of the following statements is valid for any scalars α and β and for any matrices A, B, and C for which the indicated operations are defined.*

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $(AB)C = A(BC)$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AC + BC$
6. $(\alpha\beta)A = \alpha(\beta A)$
7. $\alpha(AB) = (\alpha A)B = A(\alpha B)$
8. $(\alpha + \beta)A = \alpha A + \beta A$
9. $\alpha(A + B) = \alpha A + \alpha B$

We will prove two of the rules and leave the rest for the reader to verify.

Proof of Rule 4 Assume that $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ and $C = (c_{ij})$ are both $n \times r$ matrices. Let $D = A(B + C)$ and $E = AB + AC$. It follows that

$$d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

and

$$e_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

But

$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}$$

so that $d_{ij} = e_{ij}$ and hence $A(B + C) = AB + AC$. ■

Proof of Rule 3 Let A be an $m \times n$ matrix, B an $n \times r$ matrix, and C an $r \times s$ matrix. Let $D = AB$ and $E = BC$. We must show that $DC = AE$. By the definition of matrix multiplication,

$$d_{il} = \sum_{k=1}^{n} a_{ik}b_{kl} \qquad \text{and} \qquad e_{kj} = \sum_{l=1}^{r} b_{kl}c_{lj}$$

The ijth term of DC is

$$\sum_{l=1}^{r} d_{il}c_{lj} = \sum_{l=1}^{r}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj}$$

and the (i,j) entry of AE is

$$\sum_{k=1}^{n} a_{ik}e_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{r} b_{kl}c_{lj}\right)$$

Since

$$\sum_{l=1}^{r}\left(\sum_{k=1}^{n} a_{ik}b_{kl}\right)c_{lj} = \sum_{l=1}^{r}\left(\sum_{k=1}^{n} a_{ik}b_{kl}c_{lj}\right) = \sum_{k=1}^{n} a_{ik}\left(\sum_{l=1}^{r} b_{kl}c_{lj}\right)$$

it follows that

$$(AB)C = DC = AE = A(BC)$$ ■

The algebraic rules given in Theorem 1.4.1 seem quite natural, since they are similar to the rules that we use with real numbers. However, there are important differences between the rules for matrix algebra and the algebraic rules for real numbers. Some of these differences are illustrated in Exercises 1 through 5 at the end of this section.

EXAMPLE 1 If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 \\ -3 & 2 \end{bmatrix}, \qquad \text{and} \qquad C = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$$

verify that $A(BC) = (AB)C$ and $A(B+C) = AB + AC$.

Solution

$$\begin{aligned} A(BC) &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 16 & 11 \end{bmatrix} \\ (AB)C &= \begin{bmatrix} -4 & 5 \\ -6 & 11 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 16 & 11 \end{bmatrix} \end{aligned}$$

Thus

$$A(BC) = \begin{bmatrix} 6 & 5 \\ 16 & 11 \end{bmatrix} = (AB)C$$

$$A(B+C) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ 5 & 15 \end{bmatrix}$$

$$AB + AC = \begin{bmatrix} -4 & 5 \\ -6 & 11 \end{bmatrix} + \begin{bmatrix} 5 & 2 \\ 11 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 7 \\ 5 & 15 \end{bmatrix}$$

Therefore,

$$A(B+C) = AB + AC$$

■

Notation

Since $(AB)C = A(BC)$, we may simply omit the parentheses and write ABC. The same is true for a product of four or more matrices. In the case where an $n \times n$ matrix is multiplied by itself a number of times, it is convenient to use exponential notation. Thus, if k is a positive integer, then

$$A^k = \underbrace{AA\cdots A}_{k \text{ times}}$$

EXAMPLE 2 If

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

then

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$$

$$A^3 = AAA = AA^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

and in general

$$A^n = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix}$$

■

APPLICATION 1 A Simple Model for Marital Status Computations

In a certain town, 30 percent of the married women get divorced each year and 20 percent of the single women get married each year. There are 8000 married women and 2000 single women. Assuming that the total population of women remains constant, how many married women and how many single women will there be after one year? After two years?

Solution

Form a matrix A as follows: The entries in the first row of A will be the percentages of married and single women, respectively, who are married after one year. The entries in the second row will be the percentages of women who are single after one year. Thus,

$$A = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}$$

If we let $\mathbf{x} = \begin{bmatrix} 8000 \\ 2000 \end{bmatrix}$, the number of married and single women after one year can be computed by multiplying A times $\mathbf{x}$.

$$A\mathbf{x} = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 8000 \\ 2000 \end{bmatrix} = \begin{bmatrix} 6000 \\ 4000 \end{bmatrix}$$

After one year, there will be 6000 married women and 4000 single women. To find the number of married and single women after two years, compute

$$A^2\mathbf{x} = A(A\mathbf{x}) = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 6000 \\ 4000 \end{bmatrix} = \begin{bmatrix} 5000 \\ 5000 \end{bmatrix}$$

After two years, half of the women will be married and half will be single. In general, the number of married and single women after n years can be determined by computing $A^n\mathbf{x}$. ■

APPLICATION 2 Ecology: Demographics of the Loggerhead Sea Turtle

The management and preservation of many wildlife species depends on our ability to model population dynamics. A standard modeling technique is to divide the life cycle of a species into a number of stages. The models assume that the population sizes for each stage depend only on the female population and that the probability of survival of an individual female from one year to the next depends only on the stage of the life cycle and not on the actual age of an individual. For example, let us consider a four-stage model for analyzing the population dynamics of the loggerhead sea turtle (see Figure 1.4.1).

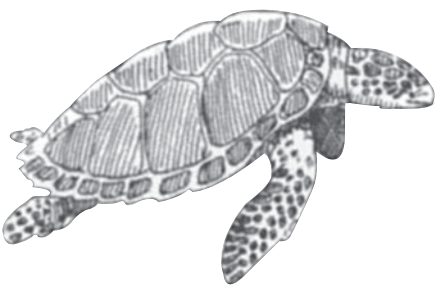


Figure 1.4.1. Loggerhead Sea Turtle

At each stage, we estimate the probability of survival over a one-year period. We also estimate the ability to reproduce in terms of the expected number of eggs laid in

Table 1 Four-Stage Model for Loggerhead Sea Turtle Demographics

Stage Number	Description (age in years)	Annual survivorship	Eggs laid per year
1	Eggs, hatchlings (<1)	0.67	0
2	Juveniles and subadults (1–21)	0.74	0
3	Novice breeders (22)	0.81	127
4	Mature breeders (23–54)	0.81	79

a given year. The results are summarized in Table 1. The approximate ages for each stage are listed in parentheses next to the stage description.

If d_i represents the duration of the ith stage and s_i is the annual survivorship rate for that stage, then it can be shown that the proportion remaining in stage i the following year will be

$$p_i = \left(\frac{1 - s_i^{d_i - 1}}{1 - s_i^{d_i}} \right) s_i \tag{1}$$

and the proportion of the population that will survive and move into stage $i + 1$ the following year will be

$$q_i = \frac{s_i^{d_i}(1 - s_i)}{1 - s_i^{d_i}} \tag{2}$$

If we let e_i denote the average number of eggs laid by a member of stage i ($i = 2, 3, 4$) in 1 year and form the matrix

$$L = \begin{pmatrix} p_1 & e_2 & e_3 & e_4 \\ q_1 & p_2 & 0 & 0 \\ 0 & q_2 & p_3 & 0 \\ 0 & 0 & q_3 & p_4 \end{pmatrix} \tag{3}$$

then L can be used to predict the turtle populations at each stage in future years. A matrix of the form (3) is called a *Leslie matrix*, and the corresponding population model is sometimes referred to as a *Leslie population model*. Using the figures from Table 1, the Leslie matrix for our model is

$$L = \begin{pmatrix} 0 & 0 & 127 & 79 \\ 0.67 & 0.7394 & 0 & 0 \\ 0 & 0.0006 & 0 & 0 \\ 0 & 0 & 0.81 & 0.8097 \end{pmatrix}$$

Suppose that the initial populations at each stage were 200,000, 300,000, 500, and 1500, respectively. If we represent these initial populations by a vector $\mathbf{x}_0$, the populations at each stage after one year are determined by computing

$$\mathbf{x}_1 = L\mathbf{x}_0 = \begin{pmatrix} 0 & 0 & 127 & 79 \\ 0.67 & 0.7394 & 0 & 0 \\ 0 & 0.0006 & 0 & 0 \\ 0 & 0 & 0.81 & 0.8097 \end{pmatrix} \begin{pmatrix} 200{,}000 \\ 300{,}000 \\ 500 \\ 1500 \end{pmatrix} = \begin{pmatrix} 182{,}000 \\ 355{,}820 \\ 180 \\ 1620 \end{pmatrix}$$

(The computations have been rounded to the nearest integer.) To determine the population vector after two years, we multiply again by the matrix L.

$$\mathbf{x}_2 = L\mathbf{x}_1 = L^2\mathbf{x}_0$$

In general, the population after k years is determined by computing $\mathbf{x}_k = L^k\mathbf{x}_0$. To see longer-range trends, we compute $\mathbf{x}_{10}$, $\mathbf{x}_{25}$, $\mathbf{x}_{50}$, and $\mathbf{x}_{100}$. The results are summarized in Table 2. The model predicts that the total number of breeding-age turtles will decrease by approximately 95 percent over a 100-year period.

Table 2 Loggerhead Sea Turtle Population Projections

Stage Number	Initial population	10 years	25 years	50 years	100 years
1	200,000	115,403	75,768	37,623	9276
2	300,000	331,274	217,858	108,178	26,673
3	500	215	142	70	17
4	1500	1074	705	350	86

A seven-stage model describing the population dynamics is presented in reference [1] to follow. We will use the seven-stage model in the computer exercises at the end of this chapter. Reference [2] is the original paper by Leslie.

References

1. Crouse, Deborah T., Larry B. Crowder, and Hal Caswell, "A Stage-Based Population Model for Loggerhead Sea Turtles and Implications for Conservation," *Ecology*, 68(5), 1987.
2. Leslie, P. H., "On the Use of Matrices in Certain Population Mathematics," *Biometrika*, 33, 1945.

The Identity Matrix

Just as the number 1 acts as an identity for the multiplication of real numbers, there is a special matrix I that acts as an identity for matrix multiplication; that is,

$$IA = AI = A \tag{4}$$

for any $n \times n$ matrix A. It is easy to verify that, if we define I to be an $n \times n$ matrix with 1's on the main diagonal and 0's elsewhere, then I satisfies equation (4) for any $n \times n$ matrix A. More formally, we have the following definition:

Definition

The $n \times n$ **identity matrix** is the matrix $I = (\delta_{ij})$, where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

As an example, let us verify equation (4) in the case $n = 3$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix}$$

and

$$\begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 6 & 3 \\ 0 & 1 & 8 \end{bmatrix}$$

In general, if B is any $m \times n$ matrix and C is any $n \times r$ matrix, then

$$BI = B \qquad \text{and} \qquad IC = C$$

The column vectors of the $n \times n$ identity matrix I are the standard vectors used to define a coordinate system in Euclidean n-space. The standard notation for the jth column vector of I is $\mathbf{e}_j$, rather than the usual $\mathbf{i}_j$. Thus, the $n \times n$ identity matrix can be written

$$I = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$$

Matrix Inversion

A real number a is said to have a multiplicative inverse if there exists a number b such that $ab = 1$. Any nonzero number a has a multiplicative inverse $b = \frac{1}{a}$. We generalize the concept of multiplicative inverses to matrices with the following definition.

Definition

An $n \times n$ matrix A is said to be **nonsingular** or **invertible** if there exists a matrix B such that $AB = BA = I$. The matrix B is said to be a **multiplicative inverse** of A.

If B and C are both multiplicative inverses of A, then

$$B = BI = B(AC) = (BA)C = IC = C$$

Thus, a matrix can have at most one multiplicative inverse. We will refer to the multiplicative inverse of a nonsingular matrix A as simply the *inverse* of A and denote it by A^{-1}.

EXAMPLE 3 The matrices

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix}$$

are inverses of each other, since

$$\begin{bmatrix} 2 & 4 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$\begin{Bmatrix} -\frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & -\frac{1}{5} \end{Bmatrix} \begin{Bmatrix} 2 & 4 \\ 3 & 1 \end{Bmatrix} = \begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix}$$

■

EXAMPLE 4 The 3×3 matrices

$$\begin{Bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{Bmatrix} \quad \text{and} \quad \begin{Bmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{Bmatrix}$$

are inverses, since

$$\begin{Bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{Bmatrix} \begin{Bmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{Bmatrix} = \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix}$$

and

$$\begin{Bmatrix} 1 & -2 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{Bmatrix} \begin{Bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{Bmatrix} = \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{Bmatrix}$$

■

EXAMPLE 5 The matrix

$$A = \begin{Bmatrix} 1 & 0 \\ 0 & 0 \end{Bmatrix}$$

has no inverse. Indeed, if B is any 2×2 matrix, then

$$BA = \begin{Bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{Bmatrix} \begin{Bmatrix} 1 & 0 \\ 0 & 0 \end{Bmatrix} = \begin{Bmatrix} b_{11} & 0 \\ b_{21} & 0 \end{Bmatrix}$$

Thus, BA cannot equal I. ■

Definition

An $n \times n$ matrix is said to be **singular** if it does not have a multiplicative inverse.

Note

Only square matrices have multiplicative inverses. One should not use the terms *singular* and *nonsingular* when referring to nonsquare matrices.

Often we will be working with products of nonsingular matrices. It turns out that any product of nonsingular matrices is nonsingular. The following theorem characterizes how the inverse of the product of a pair of nonsingular matrices A and B is related to the inverses of A and B:

Theorem 1.4.2 *If A and B are nonsingular $n \times n$ matrices, then AB is also nonsingular and $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}B = I$$
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I$$

It follows by induction that, if $A_1, \ldots, A_k$ are all nonsingular $n \times n$ matrices, then the product $A_1A_2 \cdots A_k$ is nonsingular and

$$(A_1A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1}A_1^{-1}$$

In the next section we will learn how to determine whether a matrix has a multiplicative inverse. We will also learn a method for computing the inverse of a nonsingular matrix.

Algebraic Rules for Transposes

There are four basic algebraic rules involving transposes.

Algebraic Rules for Transposes

1. $(A^T)^T = A$
2. $(\alpha A)^T = \alpha A^T$
3. $(A + B)^T = A^T + B^T$
4. $(AB)^T = B^TA^T$

The first three rules are straightforward. We leave it to the reader to verify that they are valid. To prove the fourth rule, we need only show that the (i, j) entries of $(AB)^T$ and B^TA^T are equal. If A is an $m \times n$ matrix, then, for the multiplications to be possible, B must have n rows. The (i, j) entry of $(AB)^T$ is the (j, i) entry of AB. It is computed by multiplying the jth row vector of A times the ith column vector of B:

$$\vec{\mathbf{a}}_j\mathbf{b}_i = (a_{j1}, a_{j2}, \ldots, a_{jn})\begin{bmatrix} b_{1i} \\ b_{2i} \\ \vdots \\ b_{ni} \end{bmatrix} = a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jn}b_{ni} \tag{5}$$

The (i, j) entry of B^TA^T is computed by multiplying the ith row of B^T times the jth column of A^T. Since the ith row of B^T is the transpose of the ith column of B and the jth column of A^T is the transpose of the jth row of A, it follows that the (i, j) entry of B^TA^T is given by

$$\mathbf{b}_i^T\vec{\mathbf{a}}_j^T = (b_{1i}, b_{2i}, \ldots, b_{ni})\begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jn} \end{bmatrix} = b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{ni}a_{jn} \tag{6}$$

It follows from (5) and (6) that the (i, j) entries of $(AB)^T$ and B^TA^T are equal.

The next example illustrates the idea behind the last proof.

EXAMPLE 6 Let

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 5 \\ 2 & 4 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \\ 5 & 4 & 1 \end{bmatrix}$$

Note that, on the one hand, the (3, 2) entry of AB is computed taking the scalar product of the third row of A and the second column of B.

$$AB = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 3 & 5 \\ \mathbf{2} & \mathbf{4} & \mathbf{1} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} & 2 \\ 2 & \mathbf{1} & 1 \\ 5 & \mathbf{4} & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 & 5 \\ 34 & 23 & 14 \\ 15 & \mathbf{8} & 9 \end{bmatrix}$$

When the product is transposed, the (3, 2) entry of AB becomes the (2, 3) entry of $(AB)^T$.

$$(AB)^T = \begin{bmatrix} 10 & 34 & 15 \\ 6 & 23 & \mathbf{8} \\ 5 & 14 & 9 \end{bmatrix}$$

On the other hand, the (2, 3) entry of B^TA^T is computed taking the scalar product of the second row of B^T and the third column of A^T.

$$B^TA^T = \begin{bmatrix} 1 & 2 & 5 \\ \mathbf{0} & \mathbf{1} & \mathbf{4} \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & \mathbf{2} \\ 2 & 3 & \mathbf{4} \\ 1 & 5 & \mathbf{1} \end{bmatrix} = \begin{bmatrix} 10 & 34 & 15 \\ 6 & 23 & \mathbf{8} \\ 5 & 14 & 9 \end{bmatrix}$$

In both cases, the arithmetic for computing the (2, 3) entry is the same. ■

Symmetric Matrices and Networks

Recall that a matrix A is symmetric if $A^T = A$. One type of application that leads to symmetric matrices is problems involving networks. These problems are often solved using the techniques of an area of mathematics called *graph theory*.

APPLICATION 3 Networks and Graphs

Graph theory is an important area of applied mathematics. It is used to model problems in virtually all the applied sciences. Graph theory is particularly useful in applications involving communications networks.

A *graph* is defined to be a set of points called *vertices*, together with a set of unordered pairs of vertices, which are referred to as *edges*. Figure 1.4.2 gives a geometrical representation of a graph. We can think of the vertices V_1, V_2, V_3, V_4, and V_5 as corresponding to the nodes in a communications network.

The line segments joining the vertices correspond to the edges:

$$\{V_1, V_2\},\ \{V_2, V_5\},\ \{V_3, V_4\},\ \{V_3, V_5\},\ \{V_4, V_5\}$$

Each edge represents a direct communications link between two nodes of the network.

An actual communications network could involve a large number of vertices and edges. Indeed, if there are millions of vertices, a graphical picture of the network would

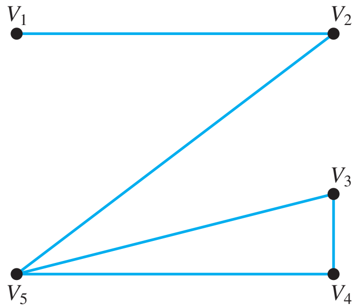


Figure 1.4.2.

be quite confusing. An alternative is to use a matrix representation for the network. If the graph contains a total of n vertices, we can define an $n \times n$ matrix A by

$$a_{ij} = \begin{cases} 1 & \text{if } \{V_i, V_j\} \text{ is an edge of the graph} \\ 0 & \text{if there is no edge joining } V_i \text{ and } V_j \end{cases}$$

The matrix A is called the *adjacency matrix* of the graph. The adjacency matrix for the graph in Figure 1.4.2 is given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$$

Note that the matrix A is symmetric. Indeed, any adjacency matrix must be symmetric, for if $\{V_i, V_j\}$ is an edge of the graph, then $a_{ij} = a_{ji} = 1$ and $a_{ij} = a_{ji} = 0$ if there is no edge joining V_i and V_j. In either case, $a_{ij} = a_{ji}$.

We can think of a *walk* in a graph as a sequence of edges linking one vertex to another. For example, in Figure 1.4.2 the edges $\{V_1, V_2\}$, $\{V_2, V_5\}$ represent a walk from vertex V_1 to vertex V_5. The length of the walk is said to be 2 since it consists of two edges. A simple way to describe the walk is to indicate the movement between vertices by arrows. Thus, $V_1 \to V_2 \to V_5$ denotes a walk of length 2 from V_1 to V_5. Similarly, $V_4 \to V_5 \to V_2 \to V_1$ represents a walk of length 3 from V_4 to V_1. It is possible to traverse the same edges more than once in a walk. For example, $V_5 \to V_3 \to V_5 \to V_3$ is a walk of length 3 from V_5 to V_3. In general, by taking powers of the adjacency matrix we can determine the number of walks of any specified length between two vertices.

Theorem 1.4.3 *If A is an $n \times n$ adjacency matrix of a graph and $a_{ij}^{(k)}$ represents the (i,j) entry of A^k, then $a_{ij}^{(k)}$ is equal to the number of walks of length k from V_i to V_j.*

Proof The proof is by mathematical induction. In the case $k = 1$, it follows from the definition of the adjacency matrix that a_{ij} represents the number of walks of length 1 from V_i to V_j. Assume for some m that each entry of A^m is equal to the number of walks of length m between the corresponding vertices. Thus $a_{il}^{(m)}$ is the number of walks of length m

from V_i to V_l. Now on the one hand, if there is an edge $\{V_l, V_j\}$, then $a_{il}^{(m)} a_{lj} = a_{il}^{(m)}$ is the number of walks of length $m + 1$ from V_i to V_j of the form

$$V_i \to \cdots \to V_l \to V_j$$

On the other hand, if $\{V_l, V_j\}$ is not an edge, then there are no walks of length $m + 1$ of this form from V_i to V_j and

$$a_{il}^{(m)} a_{lj} = a_{il}^{(m)} \cdot 0 = 0$$

It follows that the total number of walks of length $m + 1$ from V_i to V_j is given by

$$a_{i1}^{(m)} a_{1j} + a_{i2}^{(m)} a_{2j} + \cdots + a_{in}^{(m)} a_{nj}$$

But this is just the (i, j) entry of A^{m+1}. ■

EXAMPLE 7 To determine the number of walks of length 3 between any two vertices of the graph in Figure 1.4.2, we need only compute

$$A^3 = \begin{Bmatrix} 0 & 2 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 4 \\ 1 & 1 & 2 & 3 & 4 \\ 1 & 1 & 3 & 2 & 4 \\ 0 & 4 & 4 & 4 & 2 \end{Bmatrix}$$

Thus, the number of walks of length 3 from V_3 to V_5 is $a_{35}^{(3)} = 4$. Note that the matrix A^3 is symmetric. This reflects the fact that there are the same number of walks of length 3 from V_i to V_j as there are from V_j to V_i. ■

SECTION 1.4 EXERCISES

1. Explain why each of the following algebraic rules will not work in general when the real numbers a and b are replaced by $n \times n$ matrices A and B.

(a) $(a + b)^2 = a^2 + 2ab + b^2$

(b) $(a + b)(a - b) = a^2 - b^2$

2. Will the rules in Exercise 1 work if a is replaced by an $n \times n$ matrix A and b is replaced by the $n \times n$ identity matrix I?

3. Find nonzero 2×2 matrices A and B such that $AB = O$.

4. Find nonzero matrices A, B, and C such that

$$AC = BC \qquad \text{and} \qquad A \neq B$$

5. The matrix

$$A = \begin{Bmatrix} 1 & -1 \\ 1 & -1 \end{Bmatrix}$$

has the property that $A^2 = O$. Is it possible for a nonzero symmetric 2×2 matrix to have this property? Prove your answer.

6. Prove the associative law of multiplication for 2×2 matrices; that is, let

$$A = \begin{Bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Bmatrix}, \quad B = \begin{Bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{Bmatrix},$$

$$C = \begin{Bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{Bmatrix}$$

and show that

$$(AB)C = A(BC)$$

7. Let

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Compute A^2 and A^3 . What will A^n turn out to be?

8. Let

$$A = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

Compute A^2 and A^3. What will A^{2n} and A^{2n+1} turn out to be?

9. Let

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Show that $A^n = O$ for $n \geq 4$.

10. Let A and B be symmetric $n \times n$ matrices. For each of the following, determine whether the given matrix must be symmetric or could be nonsymmetric:

(a) $C = A + B$ **(b)** $D = A^2$
(c) $E = AB$ **(d)** $F = ABA$
(e) $G = AB + BA$ **(f)** $H = AB - BA$

11. Let C be a nonsymmetric $n \times n$ matrix. For each of the following, determine whether the given matrix must necessarily be symmetric or could possibly be nonsymmetric:

(a) $A = C + C^T$ **(b)** $B = C - C^T$
(c) $D = C^T C$ **(d)** $E = C^T C - CC^T$
(e) $F = (I + C)(I + C^T)$
(f) $G = (I + C)(I - C^T)$

12. Let

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Show that if $d = a_{11}a_{22} - a_{21}a_{12} \neq 0$, then

$$A^{-1} = \frac{1}{d} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

13. Use the result from Exercise 12 to find the inverse of each of the following matrices:

(a) $\begin{pmatrix} 7 & 2 \\ 3 & 1 \end{pmatrix}$ **(b)** $\begin{pmatrix} 3 & 5 \\ 2 & 3 \end{pmatrix}$ **(c)** $\begin{pmatrix} 4 & 3 \\ 2 & 2 \end{pmatrix}$

14. Let A and B are $n \times n$ matrices. Show that if

$$AB = A \quad \text{and} \quad B \neq I$$

then A must be singular.

15. Let A be a nonsingular matrix. Show that A^{-1} is also nonsingular and $(A^{-1})^{-1} = A$.

16. Prove that if A is nonsingular then A^T is nonsingular and

$$(A^T)^{-1} = (A^{-1})^T$$

Hint: $(AB)^T = B^T A^T$.

17. Let A be an $n \times n$ matrix and let $\mathbf{x}$ and $\mathbf{y}$ be vectors in $\mathbb{R}^n$. Show that if $A\mathbf{x} = A\mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$, then the matrix A must be singular.

18. Let A be a nonsingular $n \times n$ matrix. Use mathematical induction to prove that A^m is nonsingular and

$$(A^m)^{-1} = (A^{-1})^m$$

for $m = 1, 2, 3, \ldots$.

19. Let A be an $n \times n$ matrix. Show that if $A^2 = O$, then $I - A$ is nonsingular and $(I - A)^{-1} = I + A$.

20. Let A be an $n \times n$ matrix. Show that if $A^{k+1} = O$, then $I - A$ is nonsingular and

$$(I - A)^{-1} = I + A + A^2 + \cdots + A^k$$

21. Given

$$R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

show that R is nonsingular and $R^{-1} = R^T$.

22. An $n \times n$ matrix A is said to be an *involution* if $A^2 = I$. Show that if G is any matrix of the form

$$G = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

then G is an involution.

23. Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$ (i.e., $\mathbf{u}^T\mathbf{u} = 1$) and let $H = I - 2\mathbf{u}\mathbf{u}^T$. Show that H is an involution.

24. A matrix A is said to be *idempotent* if $A^2 = A$. Show that each of the following matrices are idempotent.

(a) $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

(c) $\begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$

25. Let A be an idempotent matrix.

(a) Show that $I - A$ is also idempotent.

(b) Show that $I + A$ is nonsingular and $(I + A)^{-1} = I - \frac{1}{2}A$

26. Let D be an $n \times n$ diagonal matrix whose diagonal entries are either 0 or 1.

(a) Show that D is idempotent.

(b) Show that if X is a nonsingular matrix and $A = XDX^{-1}$, then A is idempotent.

27. Let A be an involution matrix and let

$$B = \frac{1}{2}(I + A) \quad \text{and} \quad C = \frac{1}{2}(I - A)$$

Show that B and C are both idempotent and $BC = O$.

28. Let A be an $m \times n$ matrix. Show that A^TA and AA^T are both symmetric.

29. Let A and B be symmetric $n \times n$ matrices. Prove that $AB = BA$ if and only if AB is also symmetric.

30. Let A be an $n \times n$ matrix and let

$$B = A + A^T \qquad \text{and} \qquad C = A - A^T$$

(a) Show that B is symmetric and C is skew symmetric.

(b) Show that every $n \times n$ matrix can be represented as a sum of a symmetric matrix and a skew-symmetric matrix.

31. In Application 1, how many married women and how many single women will there be after 3 years?

32. Consider the matrix

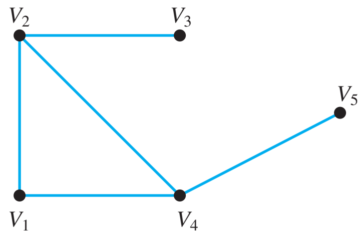
$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

(a) Draw a graph that has A as its adjacency matrix. Be sure to label the vertices of the graph.

(b) By inspecting the graph, determine the number of walks of length 2 from V_2 to V_3 and from V_2 to V_5.

(c) Compute the second row of A^3 and use it to determine the number of walks of length 3 from V_2 to V_3 and from V_2 to V_5.

33. Consider the graph



(a) Determine the adjacency matrix A of the graph.

(b) Compute A^2. What do the entries in the first row of A^2 tell you about walks of length 2 that start from V_1?

(c) Compute A^3. How many walks of length 3 are there from V_2 to V_4? How many walks of length less than or equal to 3 are there from V_2 to V_4?

For each of the conditional statements that follow, answer true *if the statement is always true and answer* false *otherwise. In the case of a true statement, explain or prove your answer. In the case of a false statement, give an example to show that the statement is not always true.*

34. If $A\mathbf{x} = B\mathbf{x}$ for some nonzero vector $\mathbf{x}$, then the matrices A and B must be equal.

35. If A and B are singular $n \times n$ matrices, then $A + B$ is also singular.

36. If A and B are nonsingular matrices, then $(AB)^T$ is nonsingular and

$$((AB)^T)^{-1} = (A^{-1})^T(B^{-1})^T$$

1.5 Elementary Matrices

In this section, we view the process of solving a linear system in terms of matrix multiplications rather than row operations. Given a linear system $A\mathbf{x} = \mathbf{b}$, we can multiply both sides by a sequence of special matrices to obtain an equivalent system in row echelon form. The special matrices we will use are called *elementary matrices*. We will use them to see how to compute the inverse of a nonsingular matrix and also to obtain an important matrix factorization. We begin by considering the effects of multiplying both sides of a linear system by a nonsingular matrix.

Equivalent Systems

Given an $m \times n$ linear system $A\mathbf{x} = \mathbf{b}$, we can obtain an equivalent system by multiplying both sides of the equation by a nonsingular $m \times m$ matrix M:

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

$$MA\mathbf{x} = M\mathbf{b} \tag{2}$$

Clearly, any solution of (1) will also be a solution of (2). However, if $\hat{\mathbf{x}}$ is a solution of (2), then

$$\begin{aligned} M^{-1}(MA\hat{\mathbf{x}}) &= M^{-1}(M\mathbf{b}) \\ A\hat{\mathbf{x}} &= \mathbf{b} \end{aligned}$$

and it follows that the two systems are equivalent.

To obtain an equivalent system that is easier to solve, we can apply a sequence of nonsingular matrices $E_1, \ldots, E_k$ to both sides of the equation $A\mathbf{x} = \mathbf{b}$ to obtain a simpler system of the form

$$U\mathbf{x} = \mathbf{c}$$

where $U = E_k \cdots E_1 A$ and $\mathbf{c} = E_k \cdots E_2 E_1 \mathbf{b}$. The new system will be equivalent to the original, provided that $M = E_k \cdots E_1$ is nonsingular. However, M is nonsingular since it is a product of nonsingular matrices.

We will show next that any of the three elementary row operations can be accomplished by multiplying A on the left by a nonsingular matrix.

Elementary Matrices

If we start with the identity matrix I and then perform exactly one elementary row operation, the resulting matrix is called an *elementary* matrix.

There are three types of elementary matrices corresponding to the three types of elementary row operations.

Type I An elementary matrix of type I is a matrix obtained by interchanging two rows of I.

EXAMPLE 1 The matrix

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix of type I since it was obtained by interchanging the first two rows of I. If A is a 3×3 matrix, then

$$E_1 A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$AE_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{bmatrix}$$

Multiplying A on the left by E_1 interchanges the first and second rows of A. Right multiplication of A by E_1 is equivalent to the elementary column operation of interchanging the first and second columns. ■

Type II An elementary matrix of type II is a matrix obtained by multiplying a row of I by a nonzero constant.

EXAMPLE 2

$$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is an elementary matrix of type II. If A is a 3×3 matrix, then

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 3a_{31} & 3a_{32} & 3a_{33} \end{bmatrix}$$

$$AE_2 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & 3a_{13} \\ a_{21} & a_{22} & 3a_{23} \\ a_{31} & a_{32} & 3a_{33} \end{bmatrix}$$

Multiplication on the left by E_2 performs the elementary row operation of multiplying the third row by 3, while multiplication on the right by E_2 performs the elementary column operation of multiplying the third column by 3. ■

Type III An elementary matrix of type III is a matrix obtained from I by adding a multiple of one row to another row.

EXAMPLE 3

$$E_3 = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix of type III. If A is a 3×3 matrix, then

$$E_3A = \begin{bmatrix} a_{11} + 3a_{31} & a_{12} + 3a_{32} & a_{13} + 3a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$AE_3 = \begin{bmatrix} a_{11} & a_{12} & 3a_{11} + a_{13} \\ a_{21} & a_{22} & 3a_{21} + a_{23} \\ a_{31} & a_{32} & 3a_{31} + a_{33} \end{bmatrix}$$

Multiplication on the left by E_3 adds 3 times the third row to the first row. Multiplication on the right adds 3 times the first column to the third column. ■

In general, suppose that E is an $n \times n$ elementary matrix. We can think of E as being obtained from I by either a row operation or a column operation. If A is an $n \times r$

matrix, *premultiplying A by E has the effect of performing that same row operation on A*. If B is an $m \times n$ matrix, *postmultiplying B by E is equivalent to performing that same column operation on B*.

Theorem 1.5.1 *If E is an elementary matrix, then E is nonsingular and E^{-1} is an elementary matrix of the same type.*

Proof If E is the elementary matrix of type I formed from I by interchanging the ith and jth rows, then E can be transformed back into I by interchanging these same rows again. Therefore $EE = I$ and hence E is its own inverse. If E is the elementary matrix of type II formed by multiplying the ith row of I by a nonzero scalar α, then E can be transformed into the identity matrix by multiplying either its ith row or its ith column by $1/\alpha$. Thus,

$$E^{-1} = \left[\begin{array}{ccccccc} 1 & & & & & & \\ & \ddots & & & & O & \\ & & 1 & & & & \\ & & & 1/\alpha & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{array}\right] \begin{array}{l} \\ \\ \\ i\text{th row} \\ \\ \\ \\ \end{array}$$

Finally, if E is the elementary matrix of type III formed from I by adding m times the ith row to the jth row, that is,

$$E = \left[\begin{array}{ccccccc} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{array}\right] \begin{array}{l} \\ \\ i\text{th row} \\ \\ j\text{th row} \\ \\ \\ \end{array}$$

then E can be transformed back into I either by subtracting m times the ith row from the jth row or by subtracting m times the jth column from the ith column. Thus,

$$E^{-1} = \left[\begin{array}{ccccccc} 1 & & & & & & \\ \vdots & \ddots & & & & O & \\ 0 & \cdots & 1 & & & & \\ \vdots & & & \ddots & & & \\ 0 & \cdots & -m & \cdots & 1 & & \\ \vdots & & & & & \ddots & \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{array}\right] \qquad \blacksquare$$

Definition A matrix B is **row equivalent** to a matrix A if there exists a finite sequence $E_1, E_2, \ldots, E_k$ of elementary matrices such that

$$B = E_k E_{k-1} \cdots E_1 A$$

In other words, B is row equivalent to A if B can be obtained from A by a finite number of row operations. In particular, if two augmented matrices $(A \,|\, \mathbf{b})$ and $(B \,|\, \mathbf{c})$ are row equivalent, then $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ are equivalent systems.

The following properties of row equivalent matrices are easily established.

I. If A is row equivalent to B, then B is row equivalent to A.

II. If A is row equivalent to B, and B is row equivalent to C, then A is row equivalent to C.

Property (I) can be proved using Theorem 1.5.1. The details of the proofs of (I) and (II) are left as an exercise for the reader.

Theorem 1.5.2 Equivalent Conditions for Nonsingularity

Let A be an $n \times n$ matrix. The following are equivalent:

(a) *A is nonsingular.*

(b) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution* $\mathbf{0}$.

(c) *A is row equivalent to I.*

Proof We prove first that statement (a) implies statement (b). If A is nonsingular and $\hat{\mathbf{x}}$ is a solution of $A\mathbf{x} = \mathbf{0}$, then

$$\hat{\mathbf{x}} = I\hat{\mathbf{x}} = (A^{-1}A)\hat{\mathbf{x}} = A^{-1}(A\hat{\mathbf{x}}) = A^{-1}\mathbf{0} = \mathbf{0}$$

Thus, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Next, we show that statement (b) implies statement (c). If we use elementary row operations, the system can be transformed into the form $U\mathbf{x} = \mathbf{0}$, where U is in row echelon form. If one of the diagonal elements of U were 0, the last row of U would consist entirely of 0's. But then $A\mathbf{x} = \mathbf{0}$ would be equivalent to a system with more unknowns than equations and hence, by Theorem 1.2.1, would have a nontrivial solution. Thus, U must be a strictly triangular matrix with diagonal elements all equal to 1. It then follows that I is the reduced row echelon form of A and hence A is row equivalent to I.

Finally, we will show that statement (c) implies statement (a). If A is row equivalent to I, there exist elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$A = E_k E_{k-1} \cdots E_1 I = E_k E_{k-1} \cdots E_1$$

But since E_i is invertible, $i = 1, \ldots, k$, the product $E_k E_{k-1} \cdots E_1$ is also invertible. Hence, A is nonsingular and

$$A^{-1} = (E_k E_{k-1} \cdots E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1} \qquad \blacksquare$$

Corollary 1.5.3 *The system $A\mathbf{x} = \mathbf{b}$ of n linear equations in n unknowns has a unique solution if and only if A is nonsingular.*

Proof If A is nonsingular, and $\hat{\mathbf{x}}$ is any solution of $A\mathbf{x} = \mathbf{b}$, then

$$A\hat{\mathbf{x}} = \mathbf{b}$$

Multiplying both sides of this equation by A^{-1}, we see that $\hat{\mathbf{x}}$ must be equal to $A^{-1}\mathbf{b}$.

Conversely, if $A\mathbf{x} = \mathbf{b}$ has a unique solution $\hat{\mathbf{x}}$, then we claim that A cannot be singular. Indeed, if A were singular, then the equation $A\mathbf{x} = \mathbf{0}$ would have a solution $\mathbf{z} \neq \mathbf{0}$. But this would imply that $\mathbf{y} = \hat{\mathbf{x}} + \mathbf{z}$ is a second solution of $A\mathbf{x} = \mathbf{b}$, since

$$A\mathbf{y} = A(\hat{\mathbf{x}} + \mathbf{z}) = A\hat{\mathbf{x}} + A\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

Therefore if $A\mathbf{x} = \mathbf{b}$ has a unique solution, then A must be nonsingular. ■

If A is nonsingular then A is row equivalent to I and hence there exist elementary matrices $E_1, \ldots, E_k$ such that

$$E_k E_{k-1} \cdots E_1 A = I$$

Multiplying both sides of this equation on the right by A^{-1}, we obtain

$$E_k E_{k-1} \cdots E_1 I = A^{-1}$$

Thus the same series of elementary row operations that transforms a nonsingular matrix A into I will transform I into A^{-1}. This gives us a method for computing A^{-1}. If we augment A by I and perform the elementary row operations that transform A into I on the augmented matrix, then I will be transformed into A^{-1}. That is, the reduced row echelon form of the augmented matrix $(A|I)$ will be $(I|A^{-1})$.

EXAMPLE 4 Compute A^{-1} if

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

Solution

$$\left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ -1 & -2 & 0 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & -6 & -3 & -2 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 3 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 4 & 0 & \frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 6 & 1 & 3 & 1 \end{array}\right] \rightarrow \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{array}\right]$$

Thus

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix}$$ ■

EXAMPLE 5 Solve the system

$$\begin{aligned} x_1 + 4x_2 + 3x_3 &= 12 \\ -x_1 - 2x_2 &= -12 \\ 2x_1 + 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

The coefficient matrix of this system is the matrix A of the last example. The solution of the system is then

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 12 \\ -12 \\ 8 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -\frac{8}{3} \end{bmatrix}$$ ■

Diagonal and Triangular Matrices

An $n \times n$ matrix A is said to be *upper triangular* if $a_{ij} = 0$ for $i > j$ and *lower triangular* if $a_{ij} = 0$ for $i < j$. Also, A is said to be *triangular* if it is either upper triangular or lower triangular. For example, the 3×3 matrices

$$\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 \\ 6 & 0 & 0 \\ 1 & 4 & 3 \end{bmatrix}$$

are both triangular. The first is upper triangular and the second is lower triangular.

A triangular matrix may have 0's on the diagonal. However, for a linear system $A\mathbf{x} = \mathbf{b}$ to be in strict triangular form, the coefficient matrix A must be upper triangular with nonzero diagonal entries.

An $n \times n$ matrix A is *diagonal* if $a_{ij} = 0$ whenever $i \neq j$. The matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are all diagonal. A diagonal matrix is both upper triangular and lower triangular.

Triangular Factorization

If an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then it is possible to represent the reduction process in terms of a matrix factorization. We illustrate how this is done in the next example.

EXAMPLE 6 Let

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

and let us use only row operation III to carry out the reduction process. At the first step, we subtract $\frac{1}{2}$ times the first row from the second and then we subtract twice the first row from the third.

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

To keep track of the multiples of the first row that were subtracted, we set $l_{21} = \frac{1}{2}$ and $l_{31} = 2$. We complete the elimination process by eliminating the -9 in the (3,2) position.

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

Let $l_{32} = -3$, the multiple of the second row subtracted from the third row. If we call the resulting matrix U and set

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix}$$

then it is easily verified that

$$LU = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 2 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 8 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 2 \\ 1 & 5 & 2 \\ 4 & -1 & 9 \end{bmatrix} = A$$ ■

The matrix L in the previous example is lower triangular with 1's on the diagonal. We say that L is *unit lower triangular*. The factorization of the matrix A into a product of a unit lower triangular matrix L times a strictly upper triangular matrix U is often referred to as an *LU factorization*.

To see why the factorization in Example 6 works, let us view the reduction process in terms of elementary matrices. The three row operations that were applied to the matrix A can be represented in terms of multiplications by elementary matrices

$$E_3E_2E_1A = U \tag{3}$$

where

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix}$$

correspond to the row operations in the reduction process. Since each of the elementary matrices is nonsingular, we can multiply equation (3) by their inverses.

$$A = E_1^{-1}E_2^{-1}E_3^{-1}U$$

[We multiply in reverse order because $(E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}$.] However, when the inverses are multiplied in this order, the multipliers l_{21}, l_{31}, l_{32} fill in below the diagonal in the product:

$$E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix} = L$$

In general, if an $n \times n$ matrix A can be reduced to strict upper triangular form using only row operation III, then A has an LU factorization. The matrix L is unit lower triangular, and if $i > j$, then l_{ij} is the multiple of the jth row subtracted from the ith row during the reduction process.

The LU factorization is a very useful way of viewing the elimination process. We will find it particularly useful in Chapter 7 when we study computer methods for solving linear systems. Many of the major topics in linear algebra can be viewed in terms of matrix factorizations. We will study other interesting and important factorizations in Chapters 5 through 7.

SECTION 1.5 EXERCISES

1. Which of the matrices that follow are elementary matrices? Classify each elementary matrix by type.

(a) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$ **(d)** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. Find the inverse of each matrix in Exercise 1. For each elementary matrix, verify that its inverse is an elementary matrix of the same type.

3. For each of the following pairs of matrices, find an elementary matrix E such that $EA = B$.

(a) $A = \begin{bmatrix} 2 & -1 \\ 5 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -4 & 2 \\ 5 & 3 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 1 & 3 \\ -2 & 4 & 5 \\ 3 & 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ -2 & 4 & 5 \end{bmatrix}$

(c) $A = \begin{bmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ -2 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -2 & 3 \\ 1 & 0 & 2 \\ 0 & 3 & 5 \end{bmatrix}$

4. For each of the following pairs of matrices, find an elementary matrix E such that $AE = B$.

(a) $A = \begin{bmatrix} 4 & 1 & 3 \\ 2 & 1 & 4 \\ 1 & 3 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 & 4 \\ 4 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & 4 \\ 1 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -2 \\ 1 & 3 \end{bmatrix}$

(c) $A = \begin{bmatrix} 4 & -2 & 3 \\ -2 & 4 & 2 \\ 6 & 1 & -2 \end{bmatrix}$,

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

5. Let

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 1 & 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 1 & 3 \\ 2 & 2 & 6 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & -3 \\ 2 & 2 & 6 \end{bmatrix}$$

(a) Find an elementary matrix E such that $EA = B$.

(b) Find an elementary matrix F such that $FB = C$.

(c) Is C row equivalent to A? Explain.

6. Let

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

(a) Find elementary matrices E_1, E_2, E_3 such that

$$E_3E_2E_1A = U$$

where U is an upper triangular matrix.

(b) Determine the inverses of E_1, E_2, E_3 and set $L = E_1^{-1}E_2^{-1}E_3^{-1}$. What type of matrix is L? Verify that $A = LU$.

7. Let

$$A = \begin{bmatrix} 2 & 1 \\ 6 & 4 \end{bmatrix}$$

(a) Express A^{-1} as a product of elementary matrices.

(b) Express A as a product of elementary matrices.

8. Compute the LU factorization of each of the following matrices.

(a) $\begin{bmatrix} 3 & 1 \\ 9 & 5 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 4 \\ -2 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 6 \\ -2 & 2 & 7 \end{bmatrix}$ **(d)** $\begin{bmatrix} -2 & 1 & 2 \\ 4 & 1 & -2 \\ -6 & -3 & 4 \end{bmatrix}$

9. Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 3 & 4 \\ 2 & 2 & 3 \end{bmatrix}$$

(a) Verify that

$$A^{-1} = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 1 & -1 \\ 0 & -2 & 3 \end{bmatrix}$$

(b) Use A^{-1} to solve $A\mathbf{x} = \mathbf{b}$ for the following choices of $\mathbf{b}$.

(i) $\mathbf{b} = (1, 1, 1)^T$ **(ii)** $\mathbf{b} = (1, 2, 3)^T$

(iii) $\mathbf{b} = (-2, 1, 0)^T$

10. Find the inverse of each of the following matrices.

(a) $\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$ **(b)** $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 6 \\ 3 & 8 \end{bmatrix}$ **(d)** $\begin{bmatrix} 3 & 0 \\ 9 & 3 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ **(f)** $\begin{bmatrix} 2 & 0 & 5 \\ 0 & 3 & 0 \\ 1 & 0 & 3 \end{bmatrix}$

(g) $\begin{bmatrix} -1 & -3 & -3 \\ 2 & 6 & 1 \\ 3 & 8 & 3 \end{bmatrix}$ **(h)** $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -2 & -3 \end{bmatrix}$

11. Given

$$A = \begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

compute A^{-1} and use it to:

(a) Find a 2×2 matrix X such that $AX = B$.

(b) Find a 2×2 matrix Y such that $YA = B$.

12. Let

$$A = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}, B = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix}, C = \begin{bmatrix} 4 & -2 \\ -6 & 3 \end{bmatrix}$$

Solve each of the following matrix equations.

(a) $AX + B = C$ **(b)** $XA + B = C$

(c) $AX + B = X$ **(d)** $XA + C = X$

13. Is the transpose of an elementary matrix an elementary matrix of the same type? Is the product of two elementary matrices an elementary matrix?

14. Let U and R be $n \times n$ upper triangular matrices and set $T = UR$. Show that T is also upper triangular and that $t_{jj} = u_{jj}r_{jj}$ for $j = 1, \dots, n$.

15. Let A be a 3×3 matrix and suppose that

$$2\mathbf{a}_1 + \mathbf{a}_2 - 4\mathbf{a}_3 = \mathbf{0}$$

How many solutions will the system $A\mathbf{x} = \mathbf{0}$ have? Explain. Is A nonsingular? Explain.

16. Let A be a 3×3 matrix and suppose that

$$\mathbf{a}_1 = 3\mathbf{a}_2 - 2\mathbf{a}_3$$

Will the system $A\mathbf{x} = \mathbf{0}$ have a nontrivial solution? Is A nonsingular? Explain your answers.

17. Let A and B be $n \times n$ matrices and let $C = A - B$. Show that if $A\mathbf{x}_0 = B\mathbf{x}_0$ and $\mathbf{x}_0 \neq \mathbf{0}$, then C must be singular.

18. Let A and B be $n \times n$ matrices and let $C = AB$. Prove that if B is singular then C must be singular. *Hint*: Use Theorem 1.5.2.

19. Let U be an $n \times n$ upper triangular matrix with nonzero diagonal entries.

(a) Explain why U must be nonsingular.

(b) Explain why U^{-1} must be upper triangular.

20. Let A be a nonsingular $n \times n$ matrix and let B be an $n \times r$ matrix. Show that the reduced row echelon form of $(A|B)$ is $(I|C)$, where $C = A^{-1}B$.

21. In general, matrix multiplication is not commutative (i.e., $AB \neq BA$). However, in certain special cases the commutative property does hold. Show that

(a) if D_1 and D_2 are $n \times n$ diagonal matrices, then $D_1D_2 = D_2D_1$.

(b) if A is an $n \times n$ matrix and

$$B = a_0I + a_1A + a_2A^2 + \cdots + a_kA^k$$

where $a_0, a_1, \ldots, a_k$ are scalars, then $AB = BA$.

22. Show that if A is a symmetric nonsingular matrix then A^{-1} is also symmetric.

23. Prove that if A is row equivalent to B then B is row equivalent to A.

24. **(a)** Prove that if A is row equivalent to B and B is row equivalent to C, then A is row equivalent to C.

(b) Prove that any two nonsingular $n \times n$ matrices are row equivalent.

25. Let A and B be an $m \times n$ matrices. Prove that if B is row equivalent to A and U is any row echelon form of A, then B is row equivalent to U.

26. Prove that B is row equivalent to A if and only if there exists a nonsingular matrix M such that $B = MA$.

27. Is it possible for a singular matrix B to be row equivalent to a nonsingular matrix A? Explain.

28. Given a vector $\mathbf{x} \in \mathbb{R}^{n+1}$, the $(n+1)\times(n+1)$ matrix V defined by

$$v_{ij} = \begin{cases} 1 & \text{if } j = 1 \\ x_i^{j-1} & \text{for } j = 2, \ldots, n+1 \end{cases}$$

is called the Vandermonde matrix.

(a) Show that if

$$V\mathbf{c} = \mathbf{y}$$

and

$$p(x) = c_1 + c_2x + \cdots + c_{n+1}x^n$$

then

$$p(x_i) = y_i, \qquad i = 1, 2, \ldots, n+1$$

(b) Suppose that $x_1, x_2, \ldots, x_{n+1}$ are all distinct. Show that if $\mathbf{c}$ is a solution of $V\mathbf{x} = \mathbf{0}$ then the coefficients $c_1, c_2, \ldots, c_n$ must all be zero, and hence V must be nonsingular.

For each of following, answer true *if the statement is always true and answer* false *otherwise. In the case of a true statement, explain or prove your answer. In the case of a false statement, give an example to show that the statement is not always true.*

29. If A is row equivalent to I and $AB = AC$, then B must equal C.

30. If E and F are elementary matrices and $G = EF$, then G is nonsingular.

31. If A is a 4×4 matrix and $\mathbf{a}_1 + \mathbf{a}_2 = \mathbf{a}_3 + 2\mathbf{a}_4$, then A must be singular.

32. If A is row equivalent to both B and C, then A is row equivalent to $B + C$.

1.6 Partitioned Matrices

Often it is useful to think of a matrix as being composed of a number of submatrices. A matrix C can be partitioned into smaller matrices by drawing horizontal lines between the rows and vertical lines between the columns. The smaller matrices are often referred to as *blocks*. For example, let

$$C = \begin{pmatrix} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{pmatrix}$$

If lines are drawn between the second and third rows and between the third and fourth columns, then C will be divided into four submatrices, C_{11}, C_{12}, C_{21}, and C_{22}.

$$\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \left(\begin{array}{ccc|cc} 1 & -2 & 4 & 1 & 3 \\ 2 & 1 & 1 & 1 & 1 \\ \hline 3 & 3 & 2 & -1 & 2 \\ 4 & 6 & 2 & 2 & 4 \end{array}\right)$$

One useful way of partitioning a matrix is to partition it into columns. For example, if

$$B = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{pmatrix}$$

we can partition B into three column submatrices:

$$B = (\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3) = \left(\begin{array}{c|c|c} -1 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 4 & 1 \end{array}\right)$$

Suppose that we are given a matrix A with three columns; then the product AB can be viewed as a block multiplication. Each block of B is multiplied by A and the result is a matrix with three blocks: $A\mathbf{b}_1$, $A\mathbf{b}_2$, and $A\mathbf{b}_3$; that is,

$$AB = A(\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3) = \begin{pmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3 \end{pmatrix}$$

For example, if

$$A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 1 & -2 \end{pmatrix}$$

then

$$A\mathbf{b}_1 = \begin{pmatrix} 6 \\ -2 \end{pmatrix}, \quad A\mathbf{b}_2 = \begin{pmatrix} 15 \\ -1 \end{pmatrix}, \quad A\mathbf{b}_3 = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$$

and hence

$$A(\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3) = \left(\begin{array}{c|c|c} 6 & 15 & 5 \\ -2 & -1 & 1 \end{array}\right)$$

In general, if A is an $m \times n$ matrix and B is an $n \times r$ matrix that has been partitioned into columns $\begin{pmatrix} \mathbf{b}_1 & \cdots & \mathbf{b}_r \end{pmatrix}$, then the block multiplication of A times B is given by

$$\boxed{AB = (A\mathbf{b}_1\ A\mathbf{b}_2\ \cdots\ A\mathbf{b}_r)}$$

In particular,

$$(\mathbf{a}_1\ \cdots\ \mathbf{a}_n) = A = AI = (A\mathbf{e}_1\ \cdots\ A\mathbf{e}_n)$$

Let A be an $m \times n$ matrix. If we partition A into rows, then

$$A = \begin{bmatrix} \vec{\mathbf{a}}_1 \\ \vec{\mathbf{a}}_2 \\ \vdots \\ \vec{\mathbf{a}}_m \end{bmatrix}$$

If B is an $n \times r$ matrix, the ith row of the product AB is determined by multiplying the ith row of A times B. Thus the ith row of AB is $\vec{\mathbf{a}}_i B$. In general, the product AB can be partitioned into rows as follows:

$$AB = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B \\ \vdots \\ \vec{\mathbf{a}}_m B \end{bmatrix}$$

To illustrate this result, let us look at an example. If

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

then

$$\vec{\mathbf{a}}_1 B = \begin{bmatrix} 1 & 9 & -1 \end{bmatrix}$$

$$\vec{\mathbf{a}}_2 B = \begin{bmatrix} 5 & 10 & -5 \end{bmatrix}$$

$$\vec{\mathbf{a}}_3 B = \begin{bmatrix} -4 & 9 & 4 \end{bmatrix}$$

These are the row vectors of the product AB:

$$AB = \begin{bmatrix} \vec{\mathbf{a}}_1 B \\ \vec{\mathbf{a}}_2 B \\ \vec{\mathbf{a}}_3 B \end{bmatrix} = \left[\begin{array}{rrr} 1 & 9 & -1 \\ \hline 5 & 10 & -5 \\ \hline -4 & 9 & 4 \end{array}\right]$$

Next, we consider how to compute the product AB in terms of more general partitions of A and B.

Block Multiplication

Let A be an $m \times n$ matrix and B an $n \times r$ matrix. It is often useful to partition A and B and express the product in terms of the submatrices of A and B. Consider the following four cases.

Case 1. If $B = \begin{pmatrix} B_1 & B_2 \end{pmatrix}$, where B_1 is an $n \times t$ matrix and B_2 is an $n \times (r-t)$ matrix, then

$$\begin{aligned} AB &= A(\mathbf{b}_1, \ldots, \mathbf{b}_t, \mathbf{b}_{t+1} \ldots \mathbf{b}_r) \\ &= (A\mathbf{b}_1, \ldots, A\mathbf{b}_t, A\mathbf{b}_{t+1}, \ldots, A\mathbf{b}_r) \\ &= (A(\mathbf{b}_1 \ldots \mathbf{b}_t), A(\mathbf{b}_{t+1} \cdots \mathbf{b}_r)) \\ &= \begin{pmatrix} AB_1 & AB_2 \end{pmatrix} \end{aligned}$$

Thus,

$$\boxed{A \begin{pmatrix} B_1 & B_2 \end{pmatrix} = \begin{pmatrix} AB_1 & AB_2 \end{pmatrix}}$$

Case 2. If $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$, where A_1 is a $k \times n$ matrix and A_2 is an $(m-k) \times n$ matrix, then

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} B = \begin{pmatrix} \vec{\mathbf{a}}_1 \\ \vdots \\ \vec{\mathbf{a}}_k \\ \hline \vec{\mathbf{a}}_{k+1} \\ \vdots \\ \vec{\mathbf{a}}_m \end{pmatrix} B = \begin{pmatrix} \vec{\mathbf{a}}_1 B \\ \vdots \\ \vec{\mathbf{a}}_k B \\ \hline \vec{\mathbf{a}}_{k+1} B \\ \vdots \\ \vec{\mathbf{a}}_m B \end{pmatrix}$$

$$= \begin{pmatrix} \begin{pmatrix} \vec{\mathbf{a}}_1 \\ \vdots \\ \vec{\mathbf{a}}_k \end{pmatrix} B \\ \begin{pmatrix} \vec{\mathbf{a}}_{k+1} \\ \vdots \\ \vec{\mathbf{a}}_m \end{pmatrix} B \end{pmatrix} = \begin{pmatrix} A_1 B \\ A_2 B \end{pmatrix}$$

Thus,

$$\boxed{\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} B = \begin{pmatrix} A_1 B \\ A_2 B \end{pmatrix}}$$

Case 3. Let $A = \begin{pmatrix} A_1 & A_2 \end{pmatrix}$ and $B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}$, where A_1 is an $m \times s$ matrix, A_2 is an $m \times (n-s)$ matrix, B_1 is an $s \times r$ matrix, and B_2 is an $(n-s) \times r$ matrix. If $C = AB$, then

$$c_{ij} = \sum_{l=1}^{n} a_{il} b_{lj} = \sum_{l=1}^{s} a_{il} b_{lj} + \sum_{l=s+1}^{n} a_{il} b_{lj}$$

Thus c_{ij} is the sum of the (i, j) entry of A_1B_1 and the (i, j) entry of A_2B_2. Therefore,

$$AB = C = A_1B_1 + A_2B_2$$

and it follows that

$$\boxed{\begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1B_1 + A_2B_2}$$

Case 4. Let A and B both be partitioned as follows:

$$A = \begin{array}{c} \left(\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right) \begin{array}{l} k \\ m-k \end{array} \\ \begin{array}{cc} s & n-s \end{array} \end{array}, \qquad B = \begin{array}{c} \left(\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right) \begin{array}{l} s \\ n-s \end{array} \\ \begin{array}{cc} t & r-t \end{array} \end{array}$$

Let

$$A_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix}, \qquad A_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix},$$

$$B_1 = \begin{pmatrix} B_{11} & B_{12} \end{pmatrix}, \qquad B_2 = \begin{pmatrix} B_{21} & B_{22} \end{pmatrix}$$

It follows from case 3 that

$$AB = \begin{pmatrix} A_1 & A_2 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = A_1B_1 + A_2B_2$$

It follows from cases 1 and 2 that

$$A_1B_1 = \begin{pmatrix} A_{11} \\ A_{21} \end{pmatrix} B_1 = \begin{pmatrix} A_{11}B_1 \\ A_{21}B_1 \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{21}B_{11} & A_{21}B_{12} \end{pmatrix}$$

$$A_2B_2 = \begin{pmatrix} A_{12} \\ A_{22} \end{pmatrix} B_2 = \begin{pmatrix} A_{12}B_2 \\ A_{22}B_2 \end{pmatrix} = \begin{pmatrix} A_{12}B_{21} & A_{12}B_{22} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

In general, if the blocks have the proper dimensions, the block multiplication can be carried out in the same manner as ordinary matrix multiplication, that is, if

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1t} \\ \vdots & & \\ A_{s1} & \cdots & A_{st} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1r} \\ \vdots & & \\ B_{t1} & \cdots & B_{tr} \end{pmatrix}$$

then

$$AB = \begin{bmatrix} C_{11} & \cdots & C_{1r} \\ \vdots & & \\ C_{s1} & \cdots & C_{sr} \end{bmatrix}$$

where

$$C_{ij} = \sum_{k=1}^{t} A_{ik}B_{kj}$$

The multiplication can be carried out in this manner only if the number of columns of A_{ik} equals the number of rows of B_{kj} for each k.

EXAMPLE 1 Let

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right]$$

Partition A into four blocks and perform the block multiplication.

Solution

Since each B_{kj} has two rows, the A_{ik}'s must each have two columns. Thus, we have one of two possibilities

$$\text{(i)} \qquad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right]$$

in which case

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 1 & 1 \\ 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ \hline 10 & 9 & 6 & 7 \\ 18 & 15 & 10 & 12 \end{array}\right]$$

or

$$\text{(ii)} \qquad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right]$$

in which case

$$\left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 \\ \hline 3 & 3 & 2 & 2 \end{array}\right] \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right] = \left[\begin{array}{cc|cc} 8 & 6 & 4 & 5 \\ 10 & 9 & 6 & 7 \\ \hline 18 & 15 & 10 & 12 \end{array}\right]$$ ■

EXAMPLE 2 Let A be an $n \times n$ matrix of the form

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix}$$

where A_{11} is a $k \times k$ matrix ($k < n$). Show that A is nonsingular if and only if A_{11} and A_{22} are nonsingular.

Solution

If A_{11} and A_{22} are nonsingular, then

$$\begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

and

$$\begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = I$$

so A is nonsingular and

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix}$$

Conversely, if A is nonsingular, then let $B = A^{-1}$ and partition B in the same manner as A. Since

$$BA = I = AB$$

it follows that

$$\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = \begin{pmatrix} A_{11} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$\begin{pmatrix} B_{11}A_{11} & B_{12}A_{22} \\ B_{21}A_{11} & B_{22}A_{22} \end{pmatrix} = \begin{pmatrix} I_k & O \\ O & I_{n-k} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} \\ A_{22}B_{21} & A_{22}B_{22} \end{pmatrix}$$

Thus,

$$\begin{aligned} B_{11}A_{11} &= I_k = A_{11}B_{11} \\ B_{22}A_{22} &= I_{n-k} = A_{22}B_{22} \end{aligned}$$

Hence, A_{11} and A_{22} are both nonsingular with inverses B_{11} and B_{22}, respectively. ■

Outer Product Expansions

Given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, it is possible to perform a matrix multiplication of the vectors if we transpose one of the vectors first. The matrix product $\mathbf{x}^T\mathbf{y}$ is the product of a row vector (a $1 \times n$ matrix) and a column vector (an $n \times 1$ matrix). The result will be a 1×1 matrix, or simply a scalar:

$$\mathbf{x}^T\mathbf{y} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

This type of product is referred to as a *scalar product* or an *inner product*. The scalar product is one of the most commonly performed operations. For example, when we multiply two matrices, each entry of the product is computed as a scalar product (a row vector times a column vector).

It is also useful to multiply a column vector times a row vector. The matrix product $\mathbf{x}\mathbf{y}^T$ is the product of an $n \times 1$ matrix times a $1 \times n$ matrix. The result is a full $n \times n$ matrix.

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} = \begin{pmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & & & \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{pmatrix}$$

The product $\mathbf{x}\mathbf{y}^T$ is referred to as the *outer product* of $\mathbf{x}$ and $\mathbf{y}$. The outer product matrix has special structure in that each of its rows is a multiple of $\mathbf{y}^T$ and each of its column vectors is a multiple of $\mathbf{x}$. For example, if

$$\mathbf{x} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 3 \\ 5 \\ 2 \end{pmatrix}$$

then

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} \begin{pmatrix} 3 & 5 & 2 \end{pmatrix} = \begin{pmatrix} 12 & 20 & 8 \\ 3 & 5 & 2 \\ 9 & 15 & 6 \end{pmatrix}$$

Note that each row is a multiple of $(3, 5, 2)$ and each column is a multiple of $\mathbf{x}$.

We are now ready to generalize the idea of an outer product from vectors to matrices. Suppose that we start with an $m \times n$ matrix X and a $k \times n$ matrix Y. We can then form a matrix product XY^T. If we partition X into columns and Y^T into rows and perform the block multiplication, we see that XY^T can be represented as a sum of outer products of vectors:

$$XY^T = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{pmatrix} \begin{pmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_n^T \end{pmatrix} = \mathbf{x}_1\mathbf{y}_1^T + \mathbf{x}_2\mathbf{y}_2^T + \cdots + \mathbf{x}_n\mathbf{y}_n^T$$

This representation is referred to as an *outer product expansion*. These types of expansions play an important role in many applications. In Section 5 of Chapter 6, we will see how outer product expansions are used in digital imaging and in information retrieval applications.

EXAMPLE 3 Given

$$X = \begin{bmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 3 & 1 \end{bmatrix}$$

compute the outer product expansion of XY^T.

Solution

$$\begin{aligned} XY^T &= \begin{bmatrix} 3 & 1 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} \begin{bmatrix} 2 & 4 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 6 & 9 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 4 & 1 \\ 8 & 16 & 4 \\ 4 & 8 & 2 \end{bmatrix} \end{aligned}$$

■

SECTION 1.6 EXERCISES

1. Let A be a nonsingular $n \times n$ matrix. Perform the following multiplications:

(a) $A^{-1}\begin{bmatrix} A & I \end{bmatrix}$ **(b)** $\begin{bmatrix} A \\ I \end{bmatrix} A^{-1}$

(c) $\begin{bmatrix} A & I \end{bmatrix}^T \begin{bmatrix} A & I \end{bmatrix}$

(d) $\begin{bmatrix} A & I \end{bmatrix} \begin{bmatrix} A & I \end{bmatrix}^T$

(e) $\begin{bmatrix} A^{-1} \\ I \end{bmatrix} \begin{bmatrix} A & I \end{bmatrix}$

2. Let $B = A^TA$. Show that $b_{ij} = \mathbf{a}_i^T \mathbf{a}_j$.

3. Let

$$A = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$$

(a) Calculate $A\mathbf{b}_1$ and $A\mathbf{b}_2$.

(b) Calculate $\vec{\mathbf{a}}_1 B$ and $\vec{\mathbf{a}}_2 B$.

(c) Multiply AB and verify that its column vectors are the vectors in part (a) and its row vectors are the vectors in part (b).

4. Let

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

and

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right]$$

Perform each of the following block multiplications.

(a) $\begin{bmatrix} O & I \\ I & O \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$

(b) $\begin{bmatrix} C & O \\ O & C \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$

(c) $\begin{bmatrix} D & O \\ O & I \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$

(d) $\begin{Bmatrix} E & O \\ O & E \end{Bmatrix} \begin{Bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{Bmatrix}$

5. Perform each of the following block multiplications:

(a) $\left\{\begin{array}{ccc|c} 1 & 1 & 1 & -1 \\ 2 & 1 & 2 & -1 \end{array}\right\} \left\{\begin{array}{ccc} 4 & -2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 2 \\ \hline 1 & 2 & 3 \end{array}\right\}$

(b) $\left\{\begin{array}{cc} 4 & -2 \\ 2 & 3 \\ 1 & 1 \\ \hline 1 & 2 \end{array}\right\} \left\{\begin{array}{ccc|c} 1 & 1 & 1 & -1 \\ 2 & 1 & 2 & -1 \end{array}\right\}$

(c) $\left\{\begin{array}{cc|cc} \frac{3}{5} & -\frac{4}{5} & 0 & 0 \\ \frac{4}{5} & \frac{3}{5} & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \end{array}\right\} \left\{\begin{array}{cc|c} \frac{3}{5} & \frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right\}$

(d) $\left\{\begin{array}{ccc|cc} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right\} \left\{\begin{array}{cc} 1 & -1 \\ 2 & -2 \\ 3 & -3 \\ \hline 4 & -4 \\ 5 & -5 \end{array}\right\}$

6. Given

$$X = \begin{Bmatrix} 2 & 1 & 5 \\ 4 & 2 & 3 \end{Bmatrix}, \qquad Y = \begin{Bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \end{Bmatrix}$$

(a) Compute the outer product expansion of XY^T.

(b) Compute the outer product expansion of YX^T. How is the outer product expansion of YX^T related to the outer product expansion of XY^T?

7. Let

$$A = \begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix} \text{ and } A^T = \begin{Bmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{Bmatrix}$$

Is it possible to perform the block multiplications of AA^T and A^TA? Explain.

8. Let A be an $m \times n$ matrix, X an $n \times r$ matrix, and B an $m \times r$ matrix. Show that

$$AX = B$$

if and only if

$$A\mathbf{x}_j = \mathbf{b}_j, \quad j = 1, \ldots, r$$

9. Let A be an $n \times n$ matrix and let D be an $n \times n$ diagonal matrix.

(a) Show that $D = (d_{11}\mathbf{e}_1, d_{22}\mathbf{e}_2, \ldots, d_{nn}\mathbf{e}_n)$.

(b) Show that $AD = (d_{11}\mathbf{a}_1, d_{22}\mathbf{a}_2, \ldots, d_{nn}\mathbf{a}_n)$.

10. Let U be an $m \times m$ matrix, let V be an $n \times n$ matrix, and let

$$\Sigma = \begin{Bmatrix} \Sigma_1 \\ O \end{Bmatrix}$$

where Σ_1 is an $n \times n$ diagonal matrix with diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_n$ and O is the $(m - n) \times n$ zero matrix.

(a) Show that if $U = (U_1, U_2)$, where U_1 has n columns, then

$$U\Sigma = U_1\Sigma_1$$

(b) Show that if $A = U\Sigma V^T$, then A can be expressed as an outer product expansion of the form

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_n\mathbf{u}_n\mathbf{v}_n^T$$

11. Let

$$A = \begin{Bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{Bmatrix}$$

where all four blocks are $n \times n$ matrices.

(a) If A_{11} and A_{22} are nonsingular, show that A must also be nonsingular and that A^{-1} must be of the form

$$\left\{\begin{array}{c|c} A_{11}^{-1} & C \\ \hline O & A_{22}^{-1} \end{array}\right\}$$

(b) Determine C.

12. Let A and B be $n \times n$ matrices and let M be a block matrix of the form

$$M = \begin{Bmatrix} A & O \\ O & B \end{Bmatrix}$$

Use condition (b) of Theorem 1.5.2 to show that if either A or B is singular, then M must be singular.

13. Let

$$A = \begin{Bmatrix} O & I \\ B & O \end{Bmatrix}$$

where all four submatrices are $k \times k$. Determine A^2 and A^4.

14. Let I denote the $n \times n$ identity matrix. Find a block form for the inverse of each of the following $2n \times 2n$ matrices.

(a) $\begin{Bmatrix} O & I \\ I & O \end{Bmatrix}$ **(b)** $\begin{Bmatrix} I & O \\ B & I \end{Bmatrix}$

15. Let O be the $k \times k$ matrix whose entries are all 0, I be the $k \times k$ identity matrix, and B be a $k \times k$ matrix with the property that $B^2 = O$. If

$$A = \begin{Bmatrix} O & I \\ I & B \end{Bmatrix}$$

determine the block form of $A^{-1} + A^2 + A^3$.

16. Let A and B be $n \times n$ matrices and define $2n \times 2n$ matrices S and M by

$$S = \begin{Bmatrix} I & A \\ O & I \end{Bmatrix}, \qquad M = \begin{Bmatrix} AB & O \\ B & O \end{Bmatrix}$$

Determine the block form of S^{-1} and use it to compute the block form of the product $S^{-1}MS$.

17. Let

$$A = \begin{Bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{Bmatrix}$$

where A_{11} is a $k \times k$ nonsingular matrix. Show that A can be factored into a product

$$\begin{Bmatrix} I & O \\ B & I \end{Bmatrix} \begin{Bmatrix} A_{11} & A_{12} \\ O & C \end{Bmatrix}$$

where

$$B = A_{21}A_{11}^{-1} \qquad \text{and} \qquad C = A_{22} - A_{21}A_{11}^{-1}A_{12}$$

(Note that this problem gives a block matrix version of the factorization in Exercise 18 of Section 1.3.)

18. Let A, B, L, M, S, and T be $n \times n$ matrices with A, B, and M nonsingular and L, S, and T singular. Determine whether it is possible to find matrices X and Y such that

$$\begin{Bmatrix} O & I & O & O & O & O \\ O & O & I & O & O & O \\ O & O & O & I & O & O \\ O & O & O & O & I & O \\ O & O & O & O & O & X \\ Y & O & O & O & O & O \end{Bmatrix} \begin{Bmatrix} M \\ A \\ T \\ L \\ A \\ B \end{Bmatrix} = \begin{Bmatrix} A \\ T \\ L \\ A \\ S \\ T \end{Bmatrix}$$

If so, show how; if not, explain why.

19. Let A be an $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$.

(a) A scalar c can also be considered as a 1×1 matrix $C = (c)$, and a vector $\mathbf{b} \in \mathbb{R}^n$ can be considered as an $n \times 1$ matrix B. Although the matrix multiplication CB is not defined, show that the matrix product BC is equal to $c\mathbf{b}$, the scalar multiplication of c times $\mathbf{b}$.

(b) Partition A into columns and $\mathbf{x}$ into rows and perform the block multiplication of A times $\mathbf{x}$.

(c) Show that

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

20. If A is an $n \times n$ matrix with the property that $A\mathbf{x} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{R}^n$, show that $A = O$. *Hint*: Let $\mathbf{x} = \mathbf{e}_j$ for $j = 1, \ldots, n$.

21. Let B and C be $n \times n$ matrices with the property that $B\mathbf{x} = C\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Show that $B = C$.

22. Consider a system of the form

$$\begin{Bmatrix} A & \mathbf{a} \\ \mathbf{c}^T & \beta \end{Bmatrix} \begin{Bmatrix} \mathbf{x} \\ x_{n+1} \end{Bmatrix} = \begin{Bmatrix} \mathbf{b} \\ b_{n+1} \end{Bmatrix}$$

where A is a nonsingular $n \times n$ matrix and $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are vectors in $\mathbb{R}^n$.

(a) Multiply both sides of the system by

$$\begin{Bmatrix} A^{-1} & \mathbf{0} \\ -\mathbf{c}^T A^{-1} & 1 \end{Bmatrix}$$

to obtain an equivalent triangular system.

(b) Set $\mathbf{y} = A^{-1}\mathbf{a}$ and $\mathbf{z} = A^{-1}\mathbf{b}$. Show that if $\beta - \mathbf{c}^T\mathbf{y} \neq 0$, then the solution of the system can be determined by letting

$$x_{n+1} = \frac{b_{n+1} - \mathbf{c}^T\mathbf{z}}{\beta - \mathbf{c}^T\mathbf{y}}$$

and then setting

$$\mathbf{x} = \mathbf{z} - x_{n+1}\mathbf{y}$$

Chapter One Exercises

MATLAB EXERCISES

The exercises that follow are to be solved computationally with the software package MATLAB, which is described in the appendix of this book. The exercises also contain questions that are related to the underlying mathematical principles illustrated in the computations. Save a record of your session in a file. After editing and printing out the file, you can fill in the answers to the questions directly on the printout.

MATLAB has a help facility that explains all its operations and commands. For example, to obtain information on the MATLAB command **rand***, you need only type* `help` **rand***. The commands used in the MATLAB exercises for this chapter are* **inv***,* **floor***,* **rand***,* **tic***,* **toc***,* **rref***,* **abs***,* **max***,* **round***,* **sum***,* **eye***,* **triu***,* **ones***,* **zeros***, and* **magic***. The operations introduced are* $+$*,* $-$*,* $*$*,* $'$*, and* $\backslash$*. The* $+$ *and* $-$ *represent the usual addition and subtraction operations for both scalars and matrices. The* $*$ *corresponds to multiplication of either scalars or matrices. For matrices whose entries are all real numbers the* $'$ *operation corresponds to the transpose operation. If* A *is a nonsingular* $n \times n$ *matrix and* B *is any* $n \times r$ *matrix, the operation* $A\backslash B$ *is equivalent to computing* $A^{-1}B$*.*

1. Use MATLAB to generate random 4×4 matrices A and B. For each of the following, compute $A1$, $A2$, $A3$, and $A4$ as indicated and determine which of the matrices are equal (you can use MATLAB to test whether two matrices are equal by computing their difference).

(a) $A1 = A * B$, $A2 = B * A$, $A3 = (A' * B')'$, $A4 = (B' * A')'$

(b) $A1 = A' * B'$, $A2 = (A * B)'$, $A3 = B' * A'$, $A4 = (B * A)'$

(c) $A1 = \mathbf{inv}(A * B)$, $A2 = \mathbf{inv}(A) * \mathbf{inv}(B)$, $A3 = \mathbf{inv}(B * A)$, $A4 = \mathbf{inv}(B) * \mathbf{inv}(A)$

(d) $A1 = \mathbf{inv}((A * B)')$, $A2 = \mathbf{inv}(A' * B')$, $A3 = \mathbf{inv}(A') * \mathbf{inv}(B')$, $A4 = (\mathbf{inv}(A) * \mathbf{inv}(B))'$

2. Set $n = 200$ and generate an $n \times n$ matrix and two vectors in $\mathbb{R}^n$, both having integer entries, by setting

$$A = \mathbf{floor}(10 * \mathbf{rand}(n));$$
$$\mathbf{b} = \mathbf{sum}(A')';$$
$$\mathbf{z} = \mathbf{ones}(n, 1);$$

(Since the matrix and vectors are large, we use semicolons to suppress the printout.)

(a) The exact solution of the system $A\mathbf{x} = \mathbf{b}$ should be the vector $\mathbf{z}$. Why? Explain. One could compute the solution in MATLAB using the "\" operation or by computing A^{-1} and then multiplying A^{-1} times $\mathbf{b}$. Let us compare these two computational methods for both speed and accuracy. One can use MATLAB's **tic** and **toc** commands to measure the elapsed time for each computation. To do this, use the commands

$$\mathbf{tic},\ \mathbf{x} = A\backslash\mathbf{b};\ \mathbf{toc}$$
$$\mathbf{tic},\ \mathbf{y} = \mathbf{inv}(A) * \mathbf{b};\ \mathbf{toc}$$

Which method is faster?

To compare the accuracy of the two methods, we can measure how close the computed solutions $\mathbf{x}$ and $\mathbf{y}$ are to the exact solution $\mathbf{z}$. Do this with the commands

$$\mathbf{max}(\mathbf{abs}(\mathbf{x} - \mathbf{z}))$$
$$\mathbf{max}(\mathbf{abs}(\mathbf{y} - \mathbf{z}))$$

Which method produces the most accurate solution?

(b) Repeat part (a), using $n = 500$ and $n = 1000$.

3. Set $A = \mathbf{floor}(10 * \mathbf{rand}(6))$. By construction, the matrix A will have integer entries. Let us change the sixth column of A so as to make the matrix singular. Set

$$B = A', \qquad A(:, 6) = -\,\mathbf{sum}(B(1:5, :))'$$

(a) Set $\mathbf{x} = \mathbf{ones}(6, 1)$ and use MATLAB to compute $A\mathbf{x}$. Why do we know that A must be singular? Explain. Check that A is singular by computing its reduced row echelon form.

(b) Set

$$B = \mathbf{x} * [1:6]$$

The product AB should equal the zero matrix. Why? Explain. Verify that this is so by computing AB with the MATLAB operation $*$.

(c) Set

$$C = \mathbf{floor}(10 * \mathbf{rand}(6))$$

and

$$D = B + C$$

Although $C \neq D$, the products AC and AD should be equal. Why? Explain. Compute $A*C$ and $A*D$, and verify that they are indeed equal.

4. Construct a matrix as follows: Set

$$B = \mathbf{eye}(10) - \mathbf{triu}(\mathbf{ones}(10), 1)$$

Why do we know that B must be nonsingular? Set

$$C = \mathbf{inv}(B) \qquad \text{and} \qquad \mathbf{x} = C(:, 10)$$

Now change B slightly by setting $B(10, 1) = -1/256$. Use MATLAB to compute the product $B\mathbf{x}$. From the result of this computation, what can you conclude about the new matrix B? Is it still nonsingular? Explain. Use MATLAB to compute its reduced row echelon form.

5. Generate a matrix A by setting

$$A = \texttt{floor}(10 * \texttt{rand}(6))$$

and generate a vector $\mathbf{b}$ by setting

$$\mathbf{b} = \texttt{floor}(20 * \texttt{rand}(6,1)) - 10$$

(a) Since A was generated randomly, we would expect it to be nonsingular. The system $A\mathbf{x} = \mathbf{b}$ should have a unique solution. Find the solution using the "\" operation. Use MATLAB to compute the reduced row echelon form U of $[A \ \ \mathbf{b}]$. How does the last column of U compare with the solution $\mathbf{x}$? In exact arithmetic, they should be the same. Why? Explain. To compare the two, compute the difference $U(:,7)-\mathbf{x}$ or examine both using **format long.**

(b) Let us now change A so as to make it singular. Set

$$A(:,3) = A(:,1:2) * [\,4 \ \ 3\,]'$$

Use MATLAB to compute $\texttt{rref}([A \ \ \mathbf{b}])$. How many solutions will the system $A\mathbf{x} = \mathbf{b}$ have? Explain.

(c) Set

$$\mathbf{y} = \texttt{floor}(20 * \texttt{rand}(6,1)) - 10$$

and

$$\mathbf{c} = A * \mathbf{y}$$

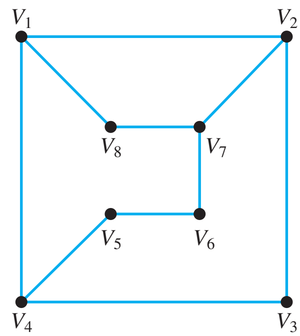
Why do we know that the system $A\mathbf{x} = \mathbf{c}$ must be consistent? Explain. Compute the reduced row echelon form U of $[\,A \ \ \mathbf{c}\,]$. How many solutions does the system $A\mathbf{x} = \mathbf{c}$ have? Explain.

(d) The free variable determined by the echelon form should be x_3. By examining the system corresponding to the matrix U, you should be able to determine the solution corresponding to $x_3 = 0$. Enter this solution into MATLAB as a column vector $\mathbf{w}$. To check that $A\mathbf{w} = \mathbf{c}$, compute the residual vector $\mathbf{c} - A\mathbf{w}$.

(e) Set $U(:,7) = \texttt{zeros}(6,1)$. The matrix U should now correspond to the reduced row echelon form of $(\,A \mid \mathbf{0}\,)$. Use U to determine the solution of the homogeneous system when the free variable $x_3 = 1$ (do this by hand) and enter your result as a vector $\mathbf{z}$. Check your answer by computing $A * \mathbf{z}$.

(f) Set $\mathbf{v} = \mathbf{w} + 3 * \mathbf{z}$. The vector $\mathbf{v}$ should be a solution of the system $A\mathbf{x} = \mathbf{c}$. Why? Explain. Verify that $\mathbf{v}$ is a solution by using MATLAB to compute the residual vector $\mathbf{c} - A\mathbf{v}$. What is the value of the free variable x_3 for this solution? How could we determine all possible solutions of the system in terms of the vectors $\mathbf{w}$ and $\mathbf{z}$? Explain.

6. Consider the graph



(a) Determine the adjacency matrix A for the graph and enter it in MATLAB.

(b) Compute A^2 and determine the number of walks of length 2 from (i) V_1 to V_7, (ii) V_4 to V_8, (iii) V_5 to V_6, and (iv) V_8 to V_3.

(c) Compute A^4, A^6, and A^8 and answer the questions in part (b) for walks of lengths 4, 6, and 8. Make a conjecture as to when there will be no walks of even length from vertex V_i to vertex V_j.

(d) Compute A^3, A^5, and A^7 and answer the questions from part (b) for walks of lengths 3, 5, and 7. Does your conjecture from part (c) hold for walks of odd length? Explain. Make a conjecture as to whether there are any walks of length k from V_i to V_j based on whether $i+j+k$ is odd or even.

(e) If we add the edges $\{V_3, V_6\}$, $\{V_5, V_8\}$ to the graph, the adjacency matrix B for the new graph can be generated by setting $B = A$ and then setting

$$B(3,6) = 1, \qquad B(6,3) = 1,$$
$$B(5,8) = 1, \qquad B(8,5) = 1$$

Compute B^k for $k = 2, 3, 4, 5$. Is your conjecture from part (d) still valid for the new graph?

(f) Add the edge $\{V_6, V_8\}$ to the figure and construct the adjacency matrix C for the resulting graph. Compute powers of C to determine whether your conjecture from part (d) will still hold for this new graph.

7. In Application 1 of Section 1.4, the numbers of married and single women after 1 and 2 years were determined by computing the products AX and A^2X for the given matrices A and X. Use **`format long`** and enter these matrices in MATLAB. Compute A^k and A^kX for $k = 5, 10, 15, 20$. What is happening to A^k as k gets large? What is the long-run distribution of married and single women in the town?

8. The following table describes a seven-stage model for the life cycle of the loggerhead sea turtle.

Table 1 Seven-Stage Model for Loggerhead Sea Turtle Demographics

Stage Number	Description (age in years)	Annual survivorship	Eggs laid per year
1	Eggs, hatchlings (<1)	0.6747	0
2	Small juveniles (1–7)	0.7857	0
3	Large juveniles (8–15)	0.6758	0
4	Subadults (16–21)	0.7425	0
5	Novice breeders (22)	0.8091	127
6	First-year remigrants (23)	0.8091	4
7	Mature breeders (24–54)	0.8091	80

The corresponding Leslie matrix is

$$L = \begin{Bmatrix} 0 & 0 & 0 & 0 & 127 & 4 & 80 \\ 0.6747 & 0.7370 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.0486 & 0.6610 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0147 & 0.6907 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.0518 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.8091 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.8091 & 0.8089 \end{Bmatrix}$$

Suppose that the number of turtles in each stage of the initial turtle population is described by the vector

$$\mathbf{x}_0 = (200,000\ \ 130,000\ \ 100,000\ \ 70,000\ \ 500\ \ 400\ \ 1100)^T$$

(a) Enter L into MATLAB and then set

x0 = [200000, 130000, 100000, 70000, 500, 400, 1100]′

Use the command

x50 = round(L^50∗x0)

to compute $\mathbf{x}_{50}$. Compute also the values of $\mathbf{x}_{100}$, $\mathbf{x}_{150}$, $\mathbf{x}_{200}$, $\mathbf{x}_{250}$, and $\mathbf{x}_{300}$.

(b) Loggerhead sea turtles lay their eggs on land. Suppose that conservationists take special measures to protect these eggs and, as a result, the survival rate for eggs and hatchlings increases to 77 percent. To incorporate this change into our model, we need only change the (2,1) entry of L to 0.77. Make this modification to the matrix L and repeat part (a). Has the survival potential of the loggerhead sea turtle improved significantly?

(c) Suppose that, instead of improving the survival rate for eggs and hatchlings, we could devise a means of protecting the small juveniles so that their survival rate increases to 88 percent. Use equations (1) and (2) from Application 2 of Section 1.4 to determine the proportion of small juveniles that survive and remain in the same stage and the proportion that survive and grow to the next stage. Modify your original matrix L accordingly and repeat part (a), using the new matrix. Has the survival potential of the loggerhead sea turtle improved significantly?

9. Set $A = \texttt{magic}(8)$ and then compute its reduced row echelon form. The leading 1's should correspond to the first three variables x_1, x_2, and x_3, and the remaining five variables are all free.

(a) Set $\mathbf{c} = [1 : 8]'$ and determine whether the system $A\mathbf{x} = \mathbf{c}$ is consistent by computing the reduced row echelon form of $[A \;\; \mathbf{c}]$. Does the system turn out to be consistent? Explain.

(b) Set

$$\mathbf{b} = [8 \;\; -8 \;\; -8 \;\; 8 \;\; 8 \;\; -8 \;\; -8 \;\; 8]';$$

and consider the system $A\mathbf{x} = \mathbf{b}$. This system should be consistent. Verify that it is by computing $U = \texttt{rref}([A \;\; \mathbf{b}])$. We should be able to find a solution for any choice of the five free variables. Indeed, set $\mathbf{x2} = \texttt{floor}(10 * \texttt{rand}(5, 1))$. If $\mathbf{x2}$ represents the last five coordinates of a solution of the system, then we should be able to determine $\mathbf{x1} = (x_1, \; x_2, \; x_3)^T$ in terms of $\mathbf{x2}$. To do this, set $U = \texttt{rref}([A \;\; \mathbf{b}])$. The nonzero rows of U correspond to a linear system with block form

$$\begin{pmatrix} I & V \end{pmatrix} \begin{pmatrix} \mathbf{x1} \\ \mathbf{x2} \end{pmatrix} = \mathbf{c} \qquad (1)$$

To solve equation (1), set

$$V = U(1:3, \; 4:8), \quad \mathbf{c} = U(1:3, \; 9)$$

and use MATLAB to compute $\mathbf{x1}$ in terms of $\mathbf{x2}$, $\mathbf{c}$, and V. Set $\mathbf{x} = [\mathbf{x1}; \;\; \mathbf{x2}]$ and verify that $\mathbf{x}$ is a solution of the system.

10. Set

$$B = [-1, -1; \; 1, 1]$$

and

$$A = [\texttt{zeros}(2), \texttt{eye}(2); \; \texttt{eye}(2), B]$$

and verify that $B^2 = O$.

(a) Use MATLAB to compute A^2, A^4, A^6, and A^8. Make a conjecture as to what the block form of A^{2k} will be in terms of the submatrices I, O, and B. Use mathematical induction to prove that your conjecture is true for any positive integer k.

(b) Use MATLAB to compute A^3, A^5, A^7, and A^9. Make a conjecture as to what the block form of A^{2k-1} will be in terms of the submatrices I, O, and B. Prove your conjecture.

11. (a) The MATLAB commands

$$A = \texttt{floor}(10 * \texttt{rand}(6)), \quad B = A' * A$$

will result in a symmetric matrix with integer entries. Why? Explain. Compute B in this way and verify these claims. Next, partition B into four 3×3 submatrices. To determine the submatrices in MATLAB, set

$$B11 = B(1:3, \; 1:3), \quad B12 = B(1:3, \; 4:6)$$

and define $B21$ and $B22$ in a similar manner using rows 4 through 6 of B.

(b) Set $C = \texttt{inv}(B11)$. It should be the case that $C^T = C$ and $B21^T = B12$. Why? Explain. Use the MATLAB operation $'$ to compute the transposes and verify these claims. Next, set

$$E = B21 * C \quad \text{and} \quad F = B22 - B21 * C * B21'$$

and use the MATLAB functions **eye** and **zeros** to construct

$$L = \begin{pmatrix} I & O \\ E & I \end{pmatrix}, \quad D = \begin{pmatrix} B11 & O \\ O & F \end{pmatrix}$$

Compute $H = L*D*L'$ and compare H with B by computing $H - B$. Prove that if all computations had been done in exact arithmetic, LDL^T would equal B exactly.

CHAPTER TEST A True or False

This chapter test consists of true-or-false questions. In each case, answer *true* if the statement is always true and *false* otherwise. In the case of a true statement, explain or prove your answer. In the case of a false statement, give an example to show that the statement is not always true. For example, consider the following statements about $n \times n$ matrices A and B:

(i) $A + B = B + A$

(ii) $AB = BA$

Statement **(i)** is always *true*. Explanation: The (i,j) entry of $A + B$ is $a_{ij} + b_{ij}$ and the (i,j) entry of $B + A$ is $b_{ij} + a_{ij}$. Since $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ for each i and j, it follows that $A + B = B + A$.

The answer to statement **(ii)** is *false*. Although the statement may be true in some cases, it is not always true. To show this, we need only exhibit one instance in which equality fails to hold. For example, if

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$$

then

$$AB = \begin{Bmatrix} 4 & 5 \\ 7 & 10 \end{Bmatrix} \quad \text{and} \quad BA = \begin{Bmatrix} 11 & 7 \\ 4 & 3 \end{Bmatrix}$$

This proves that statement **(ii)** is false.

1. If the row reduced echelon form of A involves free variables, then the system $A\mathbf{x} = \mathbf{b}$ will have infinitely many solutions.

2. Every homogeneous linear system is consistent.

3. An $n \times n$ matrix A is nonsingular if and only if the reduced row echelon form of A is I (the identity matrix).

4. If A is nonsingular, then A can be factored into a product of elementary matrices.

5. If A and B are nonsingular $n \times n$ matrices, then $A + B$ is also nonsingular and $(A + B)^{-1} = A^{-1} + B^{-1}$.

6. If $A = A^{-1}$, then A must be equal to either I or $-I$.

7. If A and B are $n \times n$ matrices, then $(A - B)^2 = A^2 - 2AB + B^2$.

8. If $AB = AC$ and $A \neq O$ (the zero matrix), then $B = C$.

9. If $AB = O$, then $BA = O$.

10. If A is a 3×3 matrix and $\mathbf{a}_1 + 2\mathbf{a}_2 - \mathbf{a}_3 = \mathbf{0}$, then A must be singular.

11. If A is a 4×3 matrix and $\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_3$, then the system $A\mathbf{x} = \mathbf{b}$ must be consistent.

12. Let A be a 4×3 matrix with $\mathbf{a}_2 = \mathbf{a}_3$. If $\mathbf{b} = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3$, then the system $A\mathbf{x} = \mathbf{b}$ will have infinitely many solutions.

13. If E is an elementary matrix, then E^T is also an elementary matrix.

14. The product of two elementary matrices is an elementary matrix.

15. If $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors in $\mathbb{R}^n$ and $A = \mathbf{x}\mathbf{y}^T$, then the row echelon form of A will have exactly one nonzero row.

CHAPTER TEST B

1. Find all solutions of the linear system

$$\begin{aligned} x_1 - x_2 + 3x_3 + 2x_4 &= 1 \\ -x_1 + x_2 - 2x_3 + x_4 &= -2 \\ 2x_1 - 2x_2 + 7x_3 + 7x_4 &= 1 \end{aligned}$$

2. (a) A linear equation in two unknowns corresponds to a line in the plane. Give a similar geometric interpretation of a linear equation in three unknowns.

(b) Given a linear system consisting of two equations in three unknowns, what is the possible number of solutions? Give a geometric explanation of your answer.

(c) Given a homogeneous linear system consisting of two equations in three unknowns, how many solutions will it have? Explain.

3. Let $A\mathbf{x} = \mathbf{b}$ be a system of n linear equations in n unknowns and suppose that $\mathbf{x}_1$ and $\mathbf{x}_2$ are both solutions and $\mathbf{x}_1 \neq \mathbf{x}_2$.

(a) How many solutions will the system have? Explain.

(b) Is the matrix A nonsingular? Explain.

4. Let A be a matrix of the form

$$A = \begin{Bmatrix} \alpha & \beta \\ 2\alpha & 2\beta \end{Bmatrix}$$

where α and β are fixed scalars not both equal to 0.

(a) Explain why the system

$$A\mathbf{x} = \begin{Bmatrix} 3 \\ 1 \end{Bmatrix}$$

must be inconsistent.

(b) How can one choose a nonzero vector $\mathbf{b}$ so that the system $A\mathbf{x} = \mathbf{b}$ will be consistent? Explain.

5. Let

$$A = \begin{Bmatrix} 2 & 1 & 3 \\ 4 & 2 & 7 \\ 1 & 3 & 5 \end{Bmatrix}, \quad B = \begin{Bmatrix} 2 & 1 & 3 \\ 1 & 3 & 5 \\ 4 & 2 & 7 \end{Bmatrix},$$

$$C = \begin{Bmatrix} 0 & 1 & 3 \\ 0 & 2 & 7 \\ -5 & 3 & 5 \end{Bmatrix}$$

(a) Find an elementary matrix E such that $EA = B$.

(b) Find an elementary matrix F such that $AF = C$.

6. Let A be a 3×3 matrix and let

$$\mathbf{b} = 3\mathbf{a}_1 + \mathbf{a}_2 + 4\mathbf{a}_3$$

Will the system $A\mathbf{x} = \mathbf{b}$ be consistent? Explain.

7. Let A be a 3×3 matrix and suppose that

$$\mathbf{a}_1 - 3\mathbf{a}_2 + 2\mathbf{a}_3 = \mathbf{0} \text{ (the zero vector)}$$

Is A nonsingular? Explain.

8. Given the vector

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is it possible to find 2×2 matrices A and B so that $A \neq B$ and $A\mathbf{x}_0 = B\mathbf{x}_0$? Explain.

9. Let A and B be symmetric $n \times n$ matrices and let $C = AB$. Is C symmetric? Explain.

10. Let E and F be $n \times n$ elementary matrices and let $C = EF$. Is C nonsingular? Explain.

11. Given

$$A = \begin{bmatrix} I & O & O \\ O & I & O \\ O & B & I \end{bmatrix}$$

where all of the submatrices are $n \times n$, determine the block form of A^{-1}.

12. Let A and B be 10×10 matrices that are partitioned into submatrices as follows

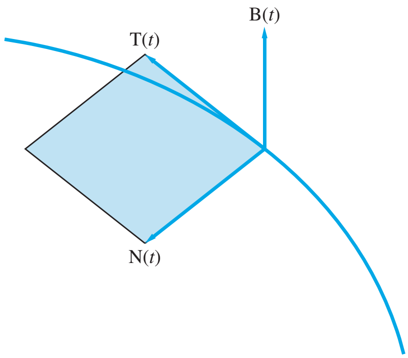
$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

(a) If A_{11} is a 6×5 matrix, and B_{11} is a $k \times r$ matrix, what conditions, if any, must k and r satisfy in order to make the block multiplication of A times B possible?

(b) Assuming that the block multiplication is possible, how would the $(2, 2)$ block of the product be determined?

CHAPTER

2



Determinants

With each square matrix, it is possible to associate a real number called the determinant of the matrix. The value of this number will tell us whether the matrix is singular.

In Section 2.1, the definition of the determinant of a matrix is given. In Section 2.2, we study properties of determinants and derive an elimination method for evaluating determinants. The elimination method is generally the simplest method to use for evaluating the determinant of an $n \times n$ matrix when $n > 3$. In Section 2.3, we see how determinants can be applied to solving $n \times n$ linear systems and how they can be used to calculate the inverse of a matrix. Two applications of determinants are presented in Section 2.3. Additional applications will also be presented later in Chapters 3 and 6.

2.1 The Determinant of a Matrix

With each $n \times n$ matrix A it is possible to associate a scalar, $\det(A)$, whose value will tell us whether the matrix is nonsingular. Before proceeding to the general definition, let us consider the following cases.

Case 1. 1 × 1 Matrices If $A = (a)$ is a 1×1 matrix, then A will have a multiplicative inverse if and only if $a \neq 0$. Thus, if we define

$$\det(A) = a$$

then A will be nonsingular if and only if $\det(A) \neq 0$.

Case 2. 2 × 2 Matrices Let

$$A = \begin{Bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Bmatrix}$$

By Theorem 1.5.2, A will be nonsingular if and only if it is row equivalent to I. Then, if $a_{11} \neq 0$, we can test whether A is row equivalent to I by performing the following operations:

APPENDIX

MATLAB

MATLAB is an interactive program for matrix computations. The original version of MATLAB, short for *matrix laboratory*, was developed by Cleve Moler from the Linpack and Eispack software libraries. Over the years MATLAB has undergone a series of expansions and revisions. Today it is the leading software for scientific computations. The MATLAB software is distributed by the MathWorks, Inc. of Natick, Massachusetts.

In addition to widespread use in industrial and engineering settings, MATLAB has become a standard instructional tool for undergraduate linear algebra courses. A Student Edition of MATLAB is available at a price affordable to undergraduates.

Another highly recommended resource for teaching linear algebra with MATLAB is *ATLAST Computer Exercises for Linear Algebra, 2nd ed.* (see [12]). This manual contains MATLAB-based exercises and projects for linear algebra and a collection of MATLAB utilities (M-files) that help students to visualize linear algebra concepts. The M-files are available for download from the ATLAST Web page:

www.umassd.edu/SpecialPrograms/ATLAST

The MATLAB Desktop Display

At start-up, MATLAB will display a desktop with three windows. The window on the right is the command window, in which MATLAB commands are entered and executed. The window on the top left displays either the Current Directory Browser or the Workspace Browser, depending on which button has been toggled.

The Workspace Browser allows you to view and make changes to the contents of the workspace. It is also possible to plot a data set using the Workspace window. Just highlight the data set to be plotted and then select the type of plot desired. MATLAB will display the graph in a new figure window. The Current Directory Browser allows you to view MATLAB and other files and to perform file operations such as opening and editing or searching for files.

The lower window on the left displays the Command History. It allows you view a log of all the commands that have been entered in the command window. To repeat a previous command, just click on the command to highlight it and then double-click to execute it. You can also recall and edit commands directly from the command window by using the arrow keys. From the command window, you can use the up arrow to recall previous commands. The commands can then be edited using the left and right arrow keys. Press the Enter key of your computer to execute the edited command.

Any of the MATLAB windows can be closed by clicking on the $\times$ in the upper-right corner of the window. To detach a window from the MATLAB desktop, click on the arrow that is next to the $\times$ in the upper right corner of the window.

Basic Data Elements

The basic elements that MATLAB uses are matrices. Once the matrices have been entered or generated, the user can quickly perform sophisticated computations with a minimal amount of programming.

Entering matrices in MATLAB is easy. To enter the matrix

$$\begin{Bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{Bmatrix}$$

type

$$A = [1 \quad 2 \quad 3 \quad 4; \quad 5 \quad 6 \quad 7 \quad 8; \quad 9 \quad 10 \quad 11 \quad 12; \quad 13 \quad 14 \quad 15 \quad 16]$$

or the matrix could be entered one row at a time:

$$\begin{array}{rrrrr} A = & [\,1 & 2 & 3 & 4 \\ & 5 & 6 & 7 & 8 \\ & 9 & 10 & 11 & 12 \\ & 13 & 14 & 15 & 16\,] \end{array}$$

Once a matrix has been entered, you can edit it in two ways. From the command window, you can redefine any entry with a MATLAB command. For example, the command $A(1, 3) = 5$ will change the third entry in the first row of A to 5. You can also edit the entries of a matrix from the Workspace Browser. To change the $(1, 3)$ entry of A with the Workspace Browser, we first locate A in the Name column of the browser and then click on the array icon to the left of A to open an array display of the matrix. To change the $(1, 3)$ entry to a 5, click on the corresponding cell of the array and enter 5.

Row vectors of equally spaced points can be generated using MATLAB's : operation. The command $\mathbf{x} = 2:6$ generates a row vector with integer entries going from 2 to 6.

$$\begin{array}{l} \mathbf{x} = \\ \qquad 2 \quad 3 \quad 4 \quad 5 \quad 6 \end{array}$$

It is not necessary to use integers or to have a step size of 1. For example, the command $\mathbf{x} = 1.2:0.2:2$ will produce

$$\begin{array}{l} \mathbf{x} = \\ \qquad 1.2000 \quad 1.4000 \quad 1.6000 \quad 1.8000 \quad 2.0000 \end{array}$$

Submatrices

To refer to a submatrix of the matrix A entered earlier, use the : to specify the rows and columns. For example, the submatrix consisting of the entries in the second two rows of columns 2 through 4 is given by $A(2:3, 2:4)$. Thus, the statement

$$C = A(2:3, 2:4)$$

generates

$$C =$$
$$\begin{matrix} 6 & 7 & 8 \\ 10 & 11 & 12 \end{matrix}$$

If the colon is used by itself for one of the arguments, either all the rows or all the columns of the matrix will be included. For example, $A(:, 2:3)$ represents the submatrix of A consisting of all the elements in the second and third columns, and $A(4, :)$ denotes the fourth row vector of A. We can generate a submatrix using nonadjacent rows or columns by using vector arguments to specify which rows and columns are to be included. For example, to generate a matrix whose entries are those which appear only in the first and third rows and second and fourth columns of A, set

$$E = A([1, 3], [2, 4])$$

The result will be

$$E =$$
$$\begin{matrix} 2 & 4 \\ 10 & 12 \end{matrix}$$

Generating Matrices

We can also generate by matrices using built-in MATLAB functions. For example, the command

$$B = \texttt{\textbf{rand}}(4)$$

will generate a 4×4 matrix whose entries are random numbers between 0 and 1. Other functions that can be used to generate matrices are **`eye`**, **`zeros`**, **`ones`**, **`magic`**, **`hilb`**, **`pascal`**, **`toeplitz`**, **`compan`**, and **`vander`**. To build triangular or diagonal matrices, we can use the MATLAB functions **`triu`**, **`tril`**, and **`diag`**.

The matrix building commands can be used to generate blocks of partitioned matrices. For example, the MATLAB command

$$E = [\ \texttt{\textbf{eye}}(2),\ \texttt{\textbf{ones}}(2, 3);\ \texttt{\textbf{zeros}}(2),\ [1:3;\quad 3:-1:1]\ ]$$

will generate the matrix

$$E =$$
$$\begin{matrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 2 & 1 \end{matrix}$$

Matrix Arithmetic

Addition and Multiplication of Matrices

Matrix arithmetic in MATLAB is straightforward. We can multiply our original matrix A times B simply by typing $A * B$. The sum and difference of A and B are given by $A + B$ and $A - B$, respectively. The transpose of the real matrix A is given by A'. For a matrix C with complex entries, the $'$ operation corresponds to conjugate transpose. Thus, C^H is given as C' in MATLAB.

Backslash or Matrix Left Division

If W is an $n \times n$ matrix and $\mathbf{b}$ represents a vector in R^n, the solution of the system $W\mathbf{x} = \mathbf{b}$ can be computed using MATLAB's backslash operator by setting

$$\mathbf{x} = W\backslash\mathbf{b}$$

For example, if we set

$$W = [1 \quad 1 \quad 1 \quad 1; \quad 1 \quad 2 \quad 3 \quad 4; \quad 3 \quad 4 \quad 6 \quad 2; \quad 2 \quad 7 \quad 10 \quad 5]$$

and $\mathbf{b} = [3; \quad 5; \quad 5; \quad 8]$, then the command

$$\mathbf{x} = W\backslash\mathbf{b}$$

will yield

$$\mathbf{x} = \begin{array}{r} 1.0000 \\ 3.0000 \\ -2.0000 \\ 1.0000 \end{array}$$

In the case that the $n \times n$ coefficient matrix is singular or has numerical rank less than n, the backslash operator will still compute a solution, but MATLAB will issue a warning. For example our original 4×4 matrix A is singular and the command

$$\mathbf{x} = A\backslash\mathbf{b}$$

yields

```
Warning: Matrix is close to singular or badly scaled.
Results may be inaccurate. RCOND = 1.387779e-018.
```

$$\mathbf{x} = \begin{array}{r} 1.0e+015* \\ \\ 2.2518 \\ -3.0024 \\ -0.7506 \\ 1.5012 \end{array}$$

The $1.0e+015$ indicates the exponent for each of the entries of $\mathbf{x}$. Thus each of the four entries listed is multiplied by 10^{15}. The value of RCOND is an estimate of the reciprocal of the condition number of the coefficient matrix. Even if the matrix were nonsingular, with a condition number on the order of 10^{18}, one could expect to lose as much as 18 digits of accuracy in the decimal representation of the computed solution. Since the computer keeps track of only 16 decimal digits, this means that the computed solution may not have any digits of accuracy.

If the coefficient matrix for a linear system has more rows than columns, then MATLAB assumes that a least squares solution of the system is desired. If we set

$$C = A(:, 1:2)$$

then C is a 4×2 matrix and the command

$$\mathbf{x} = C\backslash\mathbf{b}$$

will compute the least squares solution

$$\mathbf{x} = \begin{matrix} -2.2500 \\ 2.6250 \end{matrix}$$

If we now set

$$C = A(:, 1:3)$$

then C will be a 4×3 matrix with rank equal to 2. Although the least squares problem will not have a unique solution, MATLAB will still compute a solution and return a warning that the matrix is rank deficient. In this case, the command

$$\mathbf{x} = C\backslash\mathbf{b}$$

yields

```
Warning: Rank deficient, rank = 2, tol = 1.7852e-014.
```

$$\mathbf{x} = \begin{matrix} -0.9375 \\ 0 \\ 1.3125 \end{matrix}$$

Exponentiation

Powers of matrices are easily generated. The matrix A^5 is computed in MATLAB by typing $A\hat{}5$. We can also perform operations elementwise by preceding the operand by a period. For example, if $V = [1 \quad 2; \quad 3 \quad 4]$, then $V\hat{}2$ results in

$$\text{ans} = \begin{matrix} 7 & 10 \\ 15 & 22 \end{matrix}$$

while $V.\hat{}2$ will give

$$\text{ans} = \begin{matrix} 1 & 4 \\ 9 & 16 \end{matrix}$$

MATLAB Functions

To compute the eigenvalues of a square matrix A, we need only type **eig**(A). The eigenvectors and eigenvalues can be computed by setting

$$[X \quad D] = \mathbf{eig}(A)$$

Similarly, we can compute the determinant, inverse, condition number, norm, and rank of a matrix with simple one-word commands. Matrix factorizations such as the LU, QR, Cholesky, Schur decomposition, and singular value decomposition can be computed with a single command. For example, the command

$$[Q \quad R] = \mathbf{qr}(A)$$

will produce an orthogonal (or unitary) matrix Q and an upper triangular matrix R, with the same dimensions as A, such that $A = QR$.

Programming Features

MATLAB has all the flow control structures that you would expect in a high-level language, including **for** loops, **while** loops, and **if** statements. This allows the user to write his or her own MATLAB programs and to create additional MATLAB functions. Note that MATLAB prints out automatically the result of each command, unless the command line ends in a semicolon. *When using loops, we recommend ending each command with a semicolon to avoid printing all the results of the intermediate computations.*

M-files

It is possible to extend MATLAB by adding your own programs. MATLAB programs are all given the extension **.m** and are referred to as *M-files*. There are two basic types of M-files.

Script Files

Script files are files that contain a series of MATLAB commands. All the variables used in these commands are global, and consequently the values of these variables in your MATLAB session will change every time you run the script file. For example, if you wanted to determine the nullity of a matrix, you could create a script file **nullity.m** containing the following commands:

$$\mathbf{[m,n]} = \mathbf{size}(A);$$
$$\mathbf{nuldim} = \mathbf{n} - \mathbf{rank}(A)$$

Entering the command **nullity** would cause the two lines of code in the script file to be executed. The disadvantage of determining the nullity this way is that the matrix must be named A. Additionally, if you have been using the variables m and n, the values of these variables will be reassigned when you run the script file. An alternative would be to create a *function file*.

Function Files

Function files begin with a function declaration statement of the form

function [**oarg1,...,oargj**] = **fname**(**inarg1,...,inargk**)

All the variables used in the function M-file are local. When you call a function file, only the values of the output variables will change in your MATLAB session. For example, we could create a function file **nullity.m** to compute the nullity of a matrix as follows:

```
function k = nullity(A)
% The command nullity(A) computes the dimension
% of the nullspace of A.
[m,n] = size(A);
k = n - rank(A);
```

The lines beginning with **%** are comments that are not executed. These lines will be displayed whenever you type **help nullity** in a MATLAB session. Once the function is saved, it can be used in a MATLAB session in the same way that we use built-in MATLAB functions. For example, if we set

$$B = [1\ 2\ 3;\ 4\ 5\ 6;\ 7\ 8\ 9];$$

and then enter the command

n = **nullity**(B)

MATLAB will return the answer: **n** = 1.

The MATLAB Path

The M-files that you develop should be kept in a directory that can be added to the *MATLAB path*—the list of directories that MATLAB searches for M-files. To have your directories automatically appended to the MATLAB path at the start of a MATLAB session, create an M-file **startup.m** that includes commands to be executed at start-up. To append a directory to the MATLAB path, include a line in the **startup** file of the form

addpath dirlocation

For example, if you are working on a PC and the linear algebra files that you created are in drive **c** in a subdirectory **linalg** of the MATLAB directory, then, if you add the line

addpath c:\MATLAB\linalg

to the MATLAB start-up file, MATLAB will automatically preappend the **linalg** directory to its search path at start-up. On Windows platforms, the **startup.m** file should be placed in the **tools\local** subdirectory of your root MATLAB directory.

It is also possible to use files that are not in a directory on the MATLAB path. Simply use the Current Directory Browser to navigate to the directory containing the M-files. Double-click on the directory to set it as the current directory for the MATLAB session. MATLAB automatically looks in the current directory when it searches for M-files.

Relational and Logical Operators

MATLAB has six relational operators that are used for comparisons of scalars or elementwise comparisons of arrays. These operators are:

Relational Operators	
$<$	less than
$<=$	less than or equal
$>$	greater than
$>=$	greater than or equal
$==$	equal
$\sim=$	not equal

Given two $m \times n$ matrices A and B, the command

$$C = A < B$$

will generate an $m \times n$ matrix consisting of zeros and ones. The (i, j) entry will be equal to 1 if and only if $a_{ij} < b_{ij}$. For example, suppose that

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

The command $A >= 0$ will generate

$$\texttt{ans} = \begin{matrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

There are three logical operators in MATLAB:

Logical Operators	
&	AND
|	OR
$\sim$	NOT

These logical operators regard any nonzero scalar as corresponding to TRUE and 0 as corresponding to FALSE. The operator & corresponds to the logical AND. If a and b are scalars, the expression $a \,\&\, b$ will equal 1 if a and b are both nonzero (TRUE) and 0 otherwise. The operator | corresponds to the logical OR. The expression $a|b$ will have the value 0 if both a and b are 0; otherwise it will be equal to 1. The operator $\sim$ corresponds to the logical NOT. For a scalar a, it takes on the value 1 (TRUE) if $a = 0$ (FALSE) and the value 0 (FALSE) if $a \neq 0$ (TRUE).

For matrices, these operators are applied elementwise. Thus, if A and B are both $m \times n$ matrices, then $A \,\&\, B$ is a matrix of zeros and ones whose ijth entry is $a(i, j) \,\&\, b(i, j)$. For example, if

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{bmatrix}$$

then

$$A\&B = \begin{Bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{Bmatrix}, \quad A|B = \begin{Bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{Bmatrix}, \quad \sim A = \begin{Bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{Bmatrix}$$

The relational and logical operators are often used in **`if`** statements.

Columnwise Array Operators

MATLAB has a number of functions that, when applied to either a row or column vector $\mathbf{x}$, return a single number. For example, the command **`max`**($\mathbf{x}$) will compute the maximum entry of $\mathbf{x}$, and the command **`sum`**($\mathbf{x}$) will return the value of the sum of the entries of $\mathbf{x}$. Other functions of this form are **`min`**, **`prod`**, **`mean`**, **`all`**, and **`any`**. When used with a matrix argument, these functions are applied to each column vector and the results are returned as a row vector. For example, if

$$A = \begin{Bmatrix} -3 & 2 & 5 & 4 \\ 1 & 3 & 8 & 0 \\ -6 & 3 & 1 & 3 \end{Bmatrix}$$

then

$$\begin{aligned} \mathbf{min}(A) &= (-6, 2, 1, 0) \\ \mathbf{max}(A) &= (1, 3, 8, 4) \\ \mathbf{sum}(A) &= (-8, 8, 14, 7) \\ \mathbf{prod}(A) &= (18, 18, 40, 0) \end{aligned}$$

Graphics

If $\mathbf{x}$ and $\mathbf{y}$ are vectors of the same length, the command **`plot`**($\mathbf{x}, \mathbf{y}$) will produce a plot of all the (x_i, y_i) pairs, and each point will be connected to the next by a line segment. If the x-coordinates are taken close enough together, the graph should resemble a smooth curve. The command **`plot`**($\mathbf{x}, \mathbf{y}$, 'x') will plot the ordered pairs with x's, but will not connect the points.

For example, to plot the function $f(x) = \dfrac{\sin x}{x+1}$ on the interval $[0, 10]$, set

$$\mathbf{x} = 0 : 0.2 : 10 \qquad \text{and} \qquad \mathbf{y} = \sin(\mathbf{x})./(\mathbf{x} + 1)$$

The command **`plot`**($\mathbf{x}, \mathbf{y}$) will generate the graph of the function. To compare the graph to that of $\sin x$, we could set $\mathbf{z} = \sin(\mathbf{x})$ and use the command **`plot`**($\mathbf{x}, \mathbf{y}, \mathbf{x}, \mathbf{z}$) to plot both curves at the same time. We can include additional arguments in the command to specify the format of each plot. For example the command

$$\mathbf{plot}(\mathbf{x}, \mathbf{y}, \text{'}\mathbf{c}\text{'}, \mathbf{x}, \mathbf{z}, \text{'}--\text{'})$$

will plot the first function using a light blue (cyan) color and the second function using dashed lines. See Figure A.1.

It is also possible to do more sophisticated types of plots in MATLAB, including polar coordinates, three-dimensional surfaces, and contour plots.

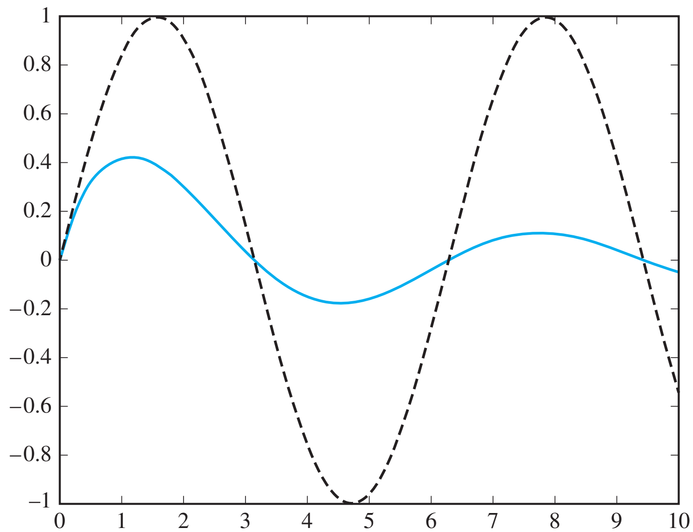


Figure A.1.

Symbolic Toolbox

In addition to doing numeric computations, it is possible to do symbolic calculations with MATLAB's symbolic toolbox. The symbolic toolbox allows us to manipulate symbolic expressions. It can be used to solve equations, differentiate and integrate functions, and perform symbolic matrix operations.

MATLAB's **`sym`** command can be used to turn any MATLAB data structure into a symbolic object. For example, the command **`sym`**('**`t`**') will turn the string '**`t`**' into a symbolic variable **`t`**, and the command **`sym(hilb(3))`** will produce the symbolic version of the 3×3 Hilbert matrix written in the form

$$\begin{matrix} [\ 1, & \frac{1}{2}, & \frac{1}{3}\ ] \\ [\ \frac{1}{2}, & \frac{1}{3}, & \frac{1}{4}\ ] \\ [\ \frac{1}{3}, & \frac{1}{4}, & \frac{1}{5}\ ] \end{matrix}$$

We can create a number of symbolic variables at once with the **`syms`** command. For example, the command

```
syms a b c
```

creates three symbolic variables **`a`**, **`b`**, and **`c`**. If we then set

$A =$ **`[a, b, c; b, c, a; c, a, b]`**

the result will be the symbolic matrix

```
A =
[ a, b, c ]
[ b, c, a ]
[ c, a, b ]
```

The MATLAB command **subs** can be used to substitute an expression or a value for a symbolic variable. For example, the command **subs(A,c,3)** will substitute 3 for each occurrence of **c** in the symbolic matrix A. Multiple substitutions are also possible: The command

```
subs(A,[a,b,c],[a-1,b+1,3])
```

will substitute **a** -1, **b** $+1$, and 3 for **a, b**, and **c**, respectively, in the matrix A.

The standard matrix operations $*$, $\hat{}$, $+$, $-$, and $'$ all work for symbolic matrices and also for combinations of symbolic and numeric matrices. If an operation involves two matrices and one of them is symbolic, the result will be a symbolic matrix. For example, the command

```
sym(hilb(3)) + eye(3)
```

will produce the symbolic matrix

$$
\begin{matrix}
[\ 2, & \frac{1}{2}, & \frac{1}{3}\ ] \\
[\ \frac{1}{2}, & \frac{4}{3}, & \frac{1}{4}\ ] \\
[\ \frac{1}{3}, & \frac{1}{4}, & \frac{6}{5}\ ]
\end{matrix}
$$

Standard MATLAB matrix commands such as

```
det, eig, inv, null, trace, sum, prod, poly
```

all work for symbolic matrices; however, others such as

```
rref, orth, rank, norm
```

do not. Likewise, none of the standard matrix factorizations are possible for symbolic matrices.

Help Facility

MATLAB includes a HELP facility that provides help on all MATLAB features. To access MATLAB's help browser, click on the help button in the toolbar (this is the button with the ? symbol) or type **helpbrowser** in the command window. You can also access HELP by selecting it from the View menu. The help facility gives information on getting started with MATLAB and on using and customizing the desktop. It lists and describes all the MATLAB functions, operations, and commands.

You can also obtain help information on any of the MATLAB commands directly from the command window. Simply enter **`help`** followed by the name of the command. For example, the MATLAB command **`eig`** is used to compute eigenvalues. For information on how to use this command, you could either find the command using the help browser or simply type **`help eig`** in the command window.

From the command window, you also can obtain help on any MATLAB operator. Simply type **`help`** followed by the name of the operator. To do this, you need to know the name that MATLAB gives to the operator. You can obtain a complete list of all operator names by entering **`help`** followed by any operator symbol. For example, to obtain help on the backslash operation, first type **`help`** \. MATLAB will respond by displaying the list of all operator names. The backslash operator is listed as **`mldivide`** (short for "matrix left divide"). To find out how the operator works, simply type **`help mldivide`**.

Conclusions

MATLAB is a powerful tool for matrix computations that is also user friendly. The fundamentals can be mastered easily, and consequently students are able to begin numerical experiments with only a minimal amount of preparation. Indeed, the material in this appendix, together with the MATLAB help facility, should be enough to get you started.

The MATLAB exercises at the end of each chapter are designed to enhance understanding of linear algebra. The exercises do not assume familiarity with MATLAB. Often specific commands are given to guide the reader through the more complicated MATLAB constructions. Consequently, you should be able to work through all the exercises without resorting to additional MATLAB books or manuals.

Although this appendix summarizes the features of MATLAB that are relevant to an undergraduate course in linear algebra, many other advanced capabilities have not been discussed. References [18] and [26] describe MATLAB in greater detail.

BIBLIOGRAPHY

A Linear Algebra and Matrix Theory

[1] Brualdi, Richard A., and Herbert J. Ryser, *Combinatorial Matrix Theory*. New York: Cambridge University Press, 1991.

[2] Carlson, David, Charles R. Johnson, David C. Lay, and A. Duane Porter, *Linear Algebra Gems: Assets for Undergraduate Mathematics*. Washington, DC: MAA, 2001.

[3] Carlson, David, Charles R. Johnson, David C. Lay, A. Duane Porter, Ann Watkins, and William Watkins, eds., *Resources for Teaching Linear Algebra*. Washington, DC: MAA, 1997.

[4] Gantmacher, Felix R., *The Theory of Matrices*, 2 vols. New York: Chelsea Publishing Co., 1960.

[5] Hill, David R., and David E. Zitarelli, *Linear Algebra Labs with MATLAB*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 2004.

[6] Hogben, Leslie, ed., *Handbook of Linear Algebra*, 2nd ed. Boca Raton, FL: Chapman and Hall/CRC Press, 2013.

[7] Horn, Roger A., and Charles R. Johnson, *Matrix Analysis*, 2nd ed. New York: Cambridge University Press, 2012.

[8] Horn, Roger A., and Charles R. Johnson, *Topics in Matrix Analysis*. New York: Cambridge University Press, 1991.

[9] Keith, Sandra, *Visualizing Linear Algebra Using Maple*. Upper Saddle River, NJ: Prentice Hall, 2001.

[10] Kleinfeld, Erwin, and Margaret Kleinfeld, *Understanding Linear Algebra Using MATLAB*. Upper Saddle River, NJ: Prentice Hall, 2001.

[11] Lancaster, Peter, and M. Tismenetsky, *The Theory of Matrices with Applications*, 2nd ed. New York: Academic Press, 1985.

[12] Leon, Steven J., Eugene Herman, and Richard Faulkenberry, *ATLAST Computer Exercises for Linear Algebra*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2003.

[13] Ortega, James M., *Matrix Theory: A Second Course*. New York: Plenum Press, 1987.

B Applied and Numerical Linear Algebra

[14] Anderson, E., Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson, *LAPACK Users' Guide*, 3rd ed. Philadelphia: SIAM, 1999.

[15] Bellman, Richard, *Introduction to Matrix Analysis*, 2nd ed. New York: McGraw-Hill Book Co., 1970.

[16] Björck, Åke, *Numerical Methods for Least Squares Problems*. Philadelphia: SIAM, 1996.

[17] Chan, Raymond H., Chen Grief, and Dianne P. O'Leary, *Milestones in Matrix Computation The Selected Works of Gene H. Golub With Commentaries*. Oxford: Oxford University Press, 2007.

[18] Coleman, Thomas F., and Charles Van Loan, *Handbook for Matrix Computations*. Philadelphia: SIAM, 1988.

[19] Conte, Samuel, D., and Carl De Boor, *Elementary Numerical Analysis: An Algorithmic Approach*, 3rd ed. New York: McGraw-Hill Book Co., 1980.

[20] Dahlquist, G., and Å. Björck, *Numerical Methods in Scientific Computing*, Vol. 1. Philadelphia: SIAM, 2008.

[21] Datta, Biswa Nath, *Numerical Linear Algebra and Applications*, 2nd ed. Philadelphia: SIAM 2010.

[22] Demmel, James W., *Applied Numerical Linear Algebra*. Philadelphia: SIAM, 1997.

[23] Fletcher, Trevor J., *Linear Algebra Through Its Applications*. New York: Van Nostrand Reinhold, 1972.

[24] Golub, Gene H., and Charles F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD: Johns Hopkins University Press, 2013.

[25] Greenbaum, Anne, *Iterative Methods for Solving Linear Systems*. Philadelphia: SIAM, 1997.

[26] Higham, Desmond J., and Nicholas J. Higham, *MATLAB Guide*. Philadelphia: SIAM, 2000.

[27] O'Leary, Dianne P., *Scientific Computing with Case Studies*. Philadelphia: SIAM, 2009.

[28] Parlett, Beresford N., *The Symmetric Eigenvalue Problem*. Philadelphia: SIAM, 1997. (Reprint of Prentice-Hall 1980 edition)

[29] Saad, Yousef, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia: SIAM, 2003.

[30] Stewart, G. W., *Matrix Algorithms, Volume I: Basic Decompositions*. Philadelphia: SIAM, 1998.

[31] Stewart, G. W., *Matrix Algorithms, Volume II: Eigensystems*. Philadelphia: SIAM, 2001.

[32] Strang, Gilbert, *Essays in Linear Algebra*. Wellesley, MA: Wellesley-Cambridge Press, 2012.

[33] Trefethen, Loyd N., *Numerical Linear Algebra*. Philadelphia: SIAM, 1997.

[34] Watkins, David S., *Fundamentals of Matrix Computation*, 2nd ed. New York: John Wiley & Sons, 2002.

[35] Watkins, David S., *The Matrix Eigenvalue Problem GR and Krylov Subspace Methods*, Philadelphia: SIAM, 2007

[36] Wilkinson, J. H., *The Algebraic Eigenvalue Problem*. New York: Oxford University Press, 1965.

[37] Wilkinson, J. H., and C. Reinsch, *Handbook for Automatic Computation*, Vol. II: *Linear Algebra*. New York: Springer-Verlag, 1971.

C Books of Related Interest

[38] Chiang, Alpha C., *Fundamental Methods of Mathematical Economics*. New York: McGraw-Hill Book Co., 1967.
[39] Courant, R., and D. Hilbert, *Methods of Mathematical Physics*, Vol. I. New York: Wiley-Interscience, 1953.
[40] Edwards, Allen L., *Multiple Regression and the Analysis of Variance and Covariance*. New York: W. H. Freeman and Co., 1985.
[41] Gander, Walter, and Jiří Hřebíček, *Solving Problems in Scientific Computing Using Maple and MATLAB*, 4th ed. Berlin: Springer-Verlag, 2004.
[42] Higham, Nicholas J., *Accuracy and Stability of Numerical Algorithms*, 2nd ed. Philadelphia: SIAM, 2002.
[43] Rivlin, Theodore J., *The Chebyshev Polynomials*. New York: Wiley-Interscience, 1974.
[44] Van Loan, Charles, *Computational Frameworks for the Fast Fourier Transform*. Philadelphia: SIAM, 1992.

References [5], [12], [18], and [26] contain information on MATLAB. Reference [12] may be used as a companion volume to this book. (See the Preface for more details and for information on how to obtain the ATLAST collection of M-files for linear algebra.) Extended bibliographies are included in the following references: [4], [7], [21], [24], [28], [29], and [36].

Answers to Selected Exercises

Chapter 1

1.1 1. (a) $(11, 3)$; (b) $(4, 1, 3)$; (c) $(-2, 0, 3, 1)$;
(d) $(-2, 3, 0, 3, 1)$

2. (a) $\begin{bmatrix} 1 & -3 \\ 0 & 2 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$;

(c) $\begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 3 & 1 & -2 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

3. (a) One solution. The two lines intersect at the point $(3, 1)$.
(b) No solution. The lines are parallel.
(c) Infinitely many solutions. Both equations represent the same line.
(d) No solution. Each pair of lines intersect in a point; however, there is no point that is on all three lines.

4. (a) $\left[\begin{array}{rr|r} 1 & 1 & 4 \\ 1 & -1 & 2 \end{array}\right]$; (c) $\left[\begin{array}{rr|r} 2 & -1 & 3 \\ -4 & 2 & -6 \end{array}\right]$;

(d) $\left[\begin{array}{rr|r} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -1 & 3 & 3 \end{array}\right]$

6. (a) $(1, -2)$; (b) $(3, 2)$; (c) $(\frac{1}{2}, \frac{2}{3})$;
(d) $(1, 1, 2)$; (e) $(-3, 1, 2)$;
(f) $(-1, 1, 1)$; (g) $(1, 1, -1)$;
(h) $(4, -3, 1, 2)$

7. (a) $(2, -1)$; (b) $(-2, 3)$

8. (a) $(-1, 2, 1)$; (b) $(3, 1, -2)$

1.2 1. Row echelon form: (a), (c), (d), (g), and (h); reduced row echelon form: (c), (d), and (g)

2. (a) Inconsistent;
(c) consistent, infinitely many solutions;
(d) consistent $(4, 5, 2)$; (e) inconsistent;
(f) consistent, $(5, 3, 2)$

3. (b) $\emptyset$;
(c) $\{(2 + 3\alpha, \alpha, -2) \mid \alpha \text{ real}\}$;
(d) $\{(5 - 2\alpha - \beta, \alpha, 4 - 3\beta, \beta) \mid \alpha, \beta \text{ real}\}$;
(e) $\{(3 - 5\alpha + 2\beta, \alpha, \beta, 6) \mid \alpha, \beta \text{ real}\}$;
(f) $\{(\alpha, 2, -1) \mid \alpha \text{ real}\}$

4. (a) x_1, x_2, x_3 are lead variables.
(c) x_1, x_3 are lead variables and x_2 is a free variable.
(e) x_1, x_4 are lead variables and x_2, x_3 are free variables.

5. (a) $(5, 1)$; (b) inconsistent; (c) $(0, 0)$;
(d) $\left\{\left(\dfrac{5-\alpha}{4}, \dfrac{1+7\alpha}{8}, \alpha\right) \middle| \alpha \text{ real}\right\}$;
(e) $\{(8 - 2\alpha, \alpha - 5, \alpha)\}$;
(f) inconsistent;
(g) inconsistent; (h) inconsistent;
(i) $(0, \frac{3}{2}, 1)$;
(j) $\{(2 - 6\alpha, 4 + \alpha, 3 - \alpha, \alpha)\}$;
(k) $\{(\frac{15}{4} - \frac{5}{8}\alpha - \beta, -\frac{1}{4} - \frac{1}{8}\alpha, \alpha, \beta)\}$;

6. (a) $(0, -1)$;
(b) $\{(\frac{3}{4} - \frac{5}{8}\alpha, -\frac{1}{4} - \frac{1}{8}\alpha, \alpha, 3) \mid \alpha \text{ is real}\}$;
(d) $\{\alpha(-\frac{4}{3}, 0, \frac{1}{3}, 1)\}$

8. $a \neq -2$

9. $\beta = 2$

10. (a) $a = 5, b = 4$; (b) $a = 5, b \neq 4$

11. (a) $(-2, 2)$; (b) $(-7, 4)$

12. (a) $(-3, 2, 1)$; (b) $(2, -2, 1)$

15. $x_1 = 280, x_2 = 230, x_3 = 350, x_4 = 590$

19. $x_1 = 2,\ x_2 = 3,\ x_3 = 12,\ x_4 = 6$

20. 6 moles N_2, 18 moles H_2, 21 moles O_2

21. All three should be equal, i.e., $x_1 = x_2 = x_3$.

22. (a) $(5, 3, -2)$; (b) $(2, 4, 2)$;
(c) $(2, 0, -2, -2, 0, 2)$

1.3 1. (a) $\begin{bmatrix} 6 & 2 & 8 \\ -4 & 0 & 2 \\ 2 & 4 & 4 \end{bmatrix}$;

(b) $\begin{bmatrix} 4 & 1 & 6 \\ -5 & 1 & 2 \\ 3 & -2 & 3 \end{bmatrix}$;

(c) $\begin{bmatrix} 3 & 2 & 2 \\ 5 & -3 & -1 \\ -4 & 16 & 1 \end{bmatrix}$;

(d) $\begin{bmatrix} 3 & 5 & -4 \\ 2 & -3 & 16 \\ 2 & -1 & 1 \end{bmatrix}$;

(f) $\begin{bmatrix} 5 & 5 & 8 \\ -10 & -1 & -9 \\ 15 & 4 & 6 \end{bmatrix}$;

(h) $\begin{bmatrix} 5 & -10 & 15 \\ 5 & -1 & 4 \\ 8 & -9 & 6 \end{bmatrix}$

2. (a) $\begin{bmatrix} 15 & 19 \\ 4 & 0 \end{bmatrix}$; (c) $\begin{bmatrix} 19 & 21 \\ 17 & 21 \\ 8 & 10 \end{bmatrix}$;

(f) $\begin{bmatrix} 6 & 4 & 8 & 10 \\ -3 & -2 & -4 & -5 \\ 9 & 6 & 12 & 15 \end{bmatrix}$

(b) and (e) are not possible.

3. (a) 3×3; (b) 1×2

4. (a) $\begin{bmatrix} 3 & 2 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ 3 & -2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$;

(c) $\begin{bmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \\ 3 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix}$

9. (a) $\mathbf{b} = 2\mathbf{a}_1 + \mathbf{a}_2$

10. (a) inconsistent; (b) consistent; (c) inconsistent

13. $\mathbf{b} = (8, -7, -1, 7)^T$

14. $\mathbf{w} = (\frac{1}{2}, \frac{1}{3}, \frac{1}{6})^T$, $\mathbf{r} = (\frac{43}{120}, \frac{45}{120}, \frac{32}{120})^T$

18. $b = a_{22} - \dfrac{a_{12}a_{21}}{a_{11}}$

1.4 7. $A = A^2 = A^3 = A^n$

8. $A^{2n} = I$, $A^{2n+1} = A$

13. (a) $\begin{bmatrix} 1 & -2 \\ -3 & 7 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & -\frac{3}{2} \\ -1 & 2 \end{bmatrix}$

31. 4500 married, 5500 single

32. (b) 0 walks of length 2 from V_2 to V_3 and 3 walks of length 2 from V_2 to V_5;

(c) 6 walks of length 3 from V_2 to V_3 and 2 walks of length 3 from V_2 to V_5

33. (a) $A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$;

(c) 5 walks of length 3 from V_2 to V_4 and 7 walks of length 3 or less

1.5 1. (a) type I;

(b) not an elementary matrix;

(c) type III; (d) type II

3. (a) $\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$;

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

4. (a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$;

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

5. (a) $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$;

(b) $F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

6. (a) $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$;

(b) $E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$;

(c) $E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

8. (a) $\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{bmatrix}$

9. (b) (i) $(0, -1, 1)^T$, (ii) $(-4, -2, 5)^T$, (iii) $(0, 3, -2)^T$

10. (a) $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$;

(c) $\begin{bmatrix} -4 & 3 \\ \frac{3}{2} & -1 \end{bmatrix}$; (d) $\begin{bmatrix} \frac{1}{3} & 0 \\ -1 & \frac{1}{3} \end{bmatrix}$;

(f) $\begin{bmatrix} 3 & 0 & -5 \\ 0 & \frac{1}{3} & 0 \\ -1 & 0 & 2 \end{bmatrix}$;

(g) $\begin{bmatrix} 2 & -3 & 3 \\ -\frac{3}{5} & \frac{6}{5} & -1 \\ -\frac{2}{5} & -\frac{1}{5} & 0 \end{bmatrix}$;

(h) $\begin{bmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ -2 & -1 & -1 \\ \frac{3}{2} & 1 & \frac{1}{2} \end{bmatrix}$

11. (a) $\begin{bmatrix} -1 & 0 \\ 4 & 2 \end{bmatrix}$; (b) $\begin{bmatrix} -8 & 5 \\ -14 & 9 \end{bmatrix}$

12. (a) $\begin{bmatrix} 20 & -5 \\ -34 & 7 \end{bmatrix}$; (c) $\begin{bmatrix} 0 & -2 \\ -2 & 2 \end{bmatrix}$

1.6 1. (b) $\begin{bmatrix} I \\ A^{-1} \end{bmatrix}$; (c) $\begin{bmatrix} A^TA & A^T \\ A & I \end{bmatrix}$;

(d) $AA^T + I$; (e) $\begin{bmatrix} I & A^{-1} \\ A & I \end{bmatrix}$

3. (a) $A\mathbf{b}_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $A\mathbf{b}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 & 1 \end{bmatrix} B = \begin{bmatrix} 3 & 4 \end{bmatrix}$,

$\begin{bmatrix} 2 & -1 \end{bmatrix} B = \begin{bmatrix} 3 & -1 \end{bmatrix}$;

(c) $AB = \begin{bmatrix} 3 & 4 \\ 3 & -1 \end{bmatrix}$

4. (a) $\left[\begin{array}{cc|cc} 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \\ \hline 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \end{array}\right]$;

(c) $\left[\begin{array}{cc|cc} 2 & 2 & 2 & 2 \\ 2 & 4 & 2 & 2 \\ \hline 3 & 1 & 1 & 1 \\ 3 & 2 & 1 & 2 \end{array}\right]$;

(d) $\left[\begin{array}{cc|cc} 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \hline 3 & 2 & 1 & 2 \\ 3 & 1 & 1 & 1 \end{array}\right]$

5. (b) $\left[\begin{array}{ccc|c} 0 & 2 & 0 & -2 \\ 8 & 5 & 8 & -5 \\ 3 & 2 & 3 & -2 \\ \hline 5 & 3 & 5 & -3 \end{array}\right]$;

(d) $\left[\begin{array}{cc} 3 & -3 \\ 2 & -2 \\ 1 & -1 \\ \hline 5 & -5 \\ 4 & -4 \end{array}\right]$

13. $A^2 = \begin{bmatrix} B & O \\ O & B \end{bmatrix}$, $A^4 = \begin{bmatrix} B^2 & O \\ O & B^2 \end{bmatrix}$

14. (a) $\begin{bmatrix} O & I \\ I & O \end{bmatrix}$; (b) $\begin{bmatrix} I & O \\ -B & I \end{bmatrix}$

CHAPTER TEST A

1. False 2. True 3. True 4. True 5. False
6. False 7. False 8. False 9. False 10. True
11. True 12. True 13. True 14. False
15. True

Chapter 2

2.1 1. (a) $\det(M_{21}) = -8$, $\det(M_{22}) = -2$, $\det(M_{23}) = 5$;

(b) $A_{21} = 8$, $A_{22} = -2$, $A_{23} = -5$

2. (a) and (c) are nonsingular.

3. (a) 1; (b) 4; (c) 0; (d) 58; (e) -39; (f) 0; (g) 8; (h) 20

4. (a) 2; (b) -4; (c) 0; (d) 0

5. $-x^3 + ax^2 + bx + c$

6. $\lambda = 6$ or -1

2.2 1. (a) -24; (b) 30; (c) -1

2. (a) 10; (b) 20

3. (a), (e), and (f) are singular while (b), (c), and (d) are nonsingular.

4. $c = 5$ or -3

7. (a) 20; (b) 108; (c) 160; (d) $\frac{5}{4}$

9. (a) -6; (c) 6; (e) 1

13. $\det(A) = u_{11}u_{22}u_{33}$

2.3 1. (a) $\det(A) = -7$, $\operatorname{adj} A = \begin{bmatrix} -1 & -2 \\ -3 & 1 \end{bmatrix}$,

$A^{-1} = \begin{bmatrix} \frac{1}{7} & \frac{2}{7} \\ \frac{3}{7} & -\frac{1}{7} \end{bmatrix}$;

(c) $\det(A) = 3$, $\operatorname{adj} A = \begin{bmatrix} -3 & 5 & 2 \\ 0 & 1 & 1 \\ 6 & -8 & -5 \end{bmatrix}$,

$A^{-1} = \frac{1}{3} \operatorname{adj} A$

2. (a) $(\frac{5}{7}, \frac{8}{7})$; (b) $(\frac{11}{5}, -\frac{4}{5})$;

(c) $(4, -2, 2)$; (d) $(2, -1, 2)$;

(e) $(-\frac{2}{3}, \frac{2}{3}, \frac{1}{3}, 0)$

3. $-\frac{3}{4}$

4. $(\frac{1}{2}, -\frac{3}{4}, 1)^T$

5. (a) $\det(A) = 0$, so A is singular.

(b) $\operatorname{adj} A = \begin{bmatrix} -1 & 2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix}$ and

$$A \operatorname{adj} A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

9. (a) $\det(\operatorname{adj}(A)) = 8$ and $\det(A) = 2$;

(b) $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & -1 & 1 \\ 0 & -6 & 2 & -2 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

14. Do Your Homework.

CHAPTER TEST A

1. True 2. False 3. False 4. True 5. False
6. True 7. True 8. True 9. False 10. True

Chapter 3

3.1

1. (a) $\|\mathbf{x}_1\| = 10$, $\|\mathbf{x}_2\| = \sqrt{17}$;
   (b) $\|\mathbf{x}_3\| = 13 < \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$
2. (a) $\|\mathbf{x}_1\| = \sqrt{5}$, $\|\mathbf{x}_2\| = 3\sqrt{5}$;
   (b) $\|\mathbf{x}_3\| = 4\sqrt{5} = \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$
7. If $\mathbf{x} + \mathbf{y} = \mathbf{x}$ for all $\mathbf{x}$ in the vector space, then $\mathbf{0} = \mathbf{0} + \mathbf{y} = \mathbf{y}$.
8. If $\mathbf{x} + \mathbf{y} = \mathbf{x} + \mathbf{z}$, then $-\mathbf{x} + (\mathbf{x} + \mathbf{y}) = -\mathbf{x} + (\mathbf{x} + \mathbf{z})$ and the conclusion follows using axioms 1, 2, 3, and 4.
11. V is not a vector space. Axiom 6 does not hold.

3.2

1. (a) and (c) are subspaces; (b), (d), and (e) are not.
2. (b) and (c) are subspaces; (a) and (d) are not.
3. (a), (c), (e), and (f) are subspaces; (b), (d), and (g) are not.
4. (a) $\{(0, 0)^T\}$;
   (b) $\operatorname{Span}((-2, 1, 0, 0)^T, (3, 0, 1, 0)^T)$;
   (c) $\operatorname{Span}((1, 1, 1)^T)$;
   (d) $\operatorname{Span}((-5, 0, -3, 1)^T, (-1, 1, 0, 0)^T)$
5. Only the set in part (c) is a subspace of P_4.
6. (a), (b), and (d) are subspaces.
11. (a), (c), and (e) are spanning sets.
12. (a) and (b) are spanning sets.
19. (b) and (c)

3.3

1. (a) and (e) are linearly independent; (b), (c), and (d) are linearly dependent.
2. (a) and (e) are linearly independent; (b), (c), and (d) are not.
3. (a) and (b) are all of 3-space;
   (c) a plane through $(0, 0, 0)$;
   (d) a line through $(0, 0, 0)$;
   (e) a plane through $(0, 0, 0)$
4. (a) linearly independent;
   (b) linearly independent;
   (c) linearly dependent
8. (a) and (b) are linearly dependent while (c) and (d) are linearly independent.
11. When α is an odd multiple of $\pi/2$. If the graph of $y = \cos x$ is shifted to the left or right by an odd multiple of $\pi/2$, we obtain the graph of either $\sin x$ or $-\sin x$.

3.4

1. Only in parts (a) and (e) do they form a basis.
2. Only in part (a) do they form a basis.
3. (c) 2
4. 1
5. (c) 2;
   (d) a plane through $(0, 0, 0)$ in 3-space
6. (b) $\{(1, 1, 1)^T\}$, dimension 1;
   (c) $\{(1, 0, 1)^T, (0, 1, 1)^T\}$, dimension 2
7. basis $\{(1, 1, 0, 0)^T, (1, -1, 1, 0)^T, (0, 2, 0, 1)^T\}$
11. $\{x^2 + 2, x + 3\}$
12. (a) $\{E_{11}, E_{22}\}$; (c) $\{E_{11}, E_{21}, E_{22}\}$;
    (e) $\{E_{12}, E_{21}, E_{22}\}$;
    (f) $\{E_{11}, E_{22}, E_{21} + E_{12}\}$
13. 2
14. (a) 3; (b) 3; (c) 2; (d) 2
15. (a) $\{x, x^2\}$; (b) $\{x - 1, (x - 1)^2\}$;
    (c) $\{x(x - 1)\}$

3.5

1. (a) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$;
   (c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
2. (a) $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$; (b) $\begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$;
   (c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

3. (a) $\begin{bmatrix} \frac{5}{2} & \frac{7}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$; (b) $\begin{bmatrix} 11 & 14 \\ -4 & -5 \end{bmatrix}$;

(c) $\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$

4. $[\mathbf{x}]_E = (-1, 2)^T$, $[\mathbf{y}]_E = (5, -8)^T$, $[\mathbf{z}]_E = (-1, 5)^T$

5. (a) $\begin{bmatrix} 2 & 0 & -1 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$; (b) $(1, -4, 3)^T$;

(c) $(0, -1, 1)^T$; (d) $(2, 2, -1)^T$

6. (a) $\begin{bmatrix} 1 & -1 & -2 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 7 \\ 5 \\ -2 \end{bmatrix}$

7. $\mathbf{w}_1 = (5, 9)^T$ and $\mathbf{w}_2 = (1, 4)^T$

8. $\mathbf{u}_1 = (0, -1)^T$ and $\mathbf{u}_2 = (1, 5)^T$

9. (a) $\begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$; (b) $\begin{bmatrix} \frac{1}{4} & -\frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$

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

3.6

2. (a) 3; (b) 3; (c) 2

3. (a) $\mathbf{u}_2$, $\mathbf{u}_4$, $\mathbf{u}_5$ are the column vectors of U corresponding to the free variables. $\mathbf{u}_2 = 2\mathbf{u}_1$, $\mathbf{u}_4 = 5\mathbf{u}_1 - \mathbf{u}_3$, $\mathbf{u}_5 = -3\mathbf{u}_1 + 2\mathbf{u}_3$

4. (a) consistent; (b) inconsistent;

(e) consistent

5. (a) infinitely many solutions;

(c) unique solution

8. rank of $A = 3$; $\dim N(B) = 1$;

18. (b) $n - 1$

32. If $\mathbf{x}_j$ is a solution to $A\mathbf{x} = \mathbf{e}_j$ for $j = 1, \ldots, m$ and $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m)$, then $AX = I_m$.

CHAPTER TEST A

1. True 2. False 3. False 4. False 5. True
6. True 7. False 8. True 9. True 10. False
11. True 12. False 13. True 14. False
15. False

Chapter 4

4.1

1. (a) reflection about x_2 axis;

(b) reflection about the origin;

(c) reflection about the line $x_2 = x_1$;

(d) the length of the vector is halved;

(e) projection onto x_2 axis

4. $(7, 18)^T$

5. All except (c) are linear transformations from R^3 into R^2.

6. (b) and (c) are linear transformations from R^2 into R^3.

7. (a), (b), and (d) are linear transformations.

9. (a) and (c) are linear transformations from P_2 into P_3.

10. $L(e^x) = e^x - 1$ and $L(x^2) = x^3/3$.

11. (a) and (c) are linear transformations from $C[0, 1]$ into R^1.

17. (a) $\ker(L) = \{\mathbf{0}\}$, $L(R^3) = R^3$;

(c) $\ker(L) = \text{Span}(\mathbf{e}_2, \mathbf{e}_3)$, $L(R^3) = \text{Span}((1, 1, 1)^T)$

18. (a) $L(S) = \text{Span}(\mathbf{e}_2, \mathbf{e}_3)$;

(b) $L(S) = \text{Span}(\mathbf{e}_1, \mathbf{e}_2)$

19. (a) $\ker(L) = P_1$, $L(P_3) = \text{Span}(x^2, x)$;

(c) $\ker(L) = \text{Span}(x^2 - x)$, $L(P_3) = P_2$

23. The operator in part (a) is one-to-one and onto.

4.2

1. (a) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$; (c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;

(d) $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$; (e) $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

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

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

3. (a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$;

(c) $\begin{bmatrix} 0 & 0 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$

4. (a) $(0, 0, 0)^T$; (b) $(2, -1, -1)^T$;

(c) $(-15, 9, 6)^T$

5. (a) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$; (b) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;

(c) $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$; (d) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

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

7. (b) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix}$

8. (a) $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & -2 & -1 \end{bmatrix}$;
(b) (i) $7\mathbf{y}_1 + 6\mathbf{y}_2 - 8\mathbf{y}_3$, (ii) $3\mathbf{y}_1 + 3\mathbf{y}_2 - 3\mathbf{y}_3$, (iii) $\mathbf{y}_1 + 5\mathbf{y}_2 + 3\mathbf{y}_3$

9. (a) square; (b) (i) contraction by a factor $\frac{1}{2}$, (ii) clockwise rotation by 45°, (iii) translation 2 units to the right and 3 units down

10. (a) $\begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$;
(b) $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$; (d) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & \frac{1}{2} \\ 1 & 0 \end{bmatrix}$;

14. $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ -2 & 0 & 0 \end{bmatrix}$; (a) $\begin{bmatrix} \frac{1}{2} \\ -2 \end{bmatrix}$ (d) $\begin{bmatrix} 5 \\ -8 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$;

18. (a) $\begin{bmatrix} -1 & -3 & 1 \\ 0 & 2 & 0 \end{bmatrix}$; (c) $\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 3 \end{bmatrix}$

4.3

1. For the matrix A, see the answers to Exercise 1 of Section 4.2.
(a) $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; (b) $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$;
(c) $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$; (d) $B = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$;
(e) $B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$

2. (a) $\begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix}$; (b) $\begin{bmatrix} 1 & 0 \\ -4 & -1 \end{bmatrix}$

3. $B = A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$
(Note: in this case the matrices A and U commute; so $B = U^{-1}AU = U^{-1}UA = A$.)

4. $V = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & -2 \\ 1 & 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

5. (a) $\begin{bmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$; (b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$;
(c) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; (d) $a_1x + a_2 2^n(1 + x^2)$

6. (a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$; (b) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$;
(c) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

CHAPTER TEST A

1. False 2. True 3. True 4. False 5. False 6. True 7. True 8. True 9. True 10. False

Chapter 5

5.1

1. (a) 0°; (b) 90°

2. (a) $\sqrt{14}$ (scalar projection), $(2, 1, 3)^T$ (vector projection);
(b) $0, \mathbf{0}$; (c) $\frac{14\sqrt{13}}{13}^T$, $(\frac{42}{13}, \frac{28}{13})^T$;
(d) $\frac{8\sqrt{21}}{21}^T$, $(\frac{8}{21}, \frac{16}{21}, \frac{32}{21})^T$

3. (a) $\mathbf{p} = (3, 0)^T$, $\mathbf{x} - \mathbf{p} = (0, 4)^T$, $\mathbf{p}^T(\mathbf{x} - \mathbf{p}) = 3 \cdot 0 + 0 \cdot 4 = 0$;
(c) $\mathbf{p} = (3, 3, 3)^T$, $\mathbf{x} - \mathbf{p} = (-1, 1, 0)^T$, $\mathbf{p}^T(\mathbf{x} - \mathbf{p}) = -1 \cdot 3 + 1 \cdot 3 + 0 \cdot 3 = 0$

5. (1.8, 3.6)

6. (1.4, 3.8)

7. 0.4

8. (a) $2x + 4y + 3z = 0$; (c) $z - 4 = 0$

9. $\frac{5}{3}$

10. $\frac{8}{7}$

20. The correlation matrix with entries rounded to two decimal places is

$$\begin{bmatrix} 1.00 & -0.04 & 0.41 \\ -0.04 & 1.00 & 0.87 \\ 0.41 & 0.87 & 1.00 \end{bmatrix}$$

5.2 1. (a) $\{(3,4)^T\}$ basis for $R(A^T)$,
$\{(-4,3)^T\}$ basis for $N(A)$,
$\{(1,2)^T\}$ basis for $R(A)$,
$\{(-2,1)^T\}$ basis for $N(A^T)$;

(d) basis for $R(A^T)$:
$\{(1,0,0,0)^T,(0,1,0,0)^T(0,0,1,1)^T\}$,
basis for $N(A)$: $\{(0,0,-1,1)^T\}$,
basis for $R(A)$:
$\{(1,0,0,1)^T,(0,1,0,1)^T(0,0,1,1)^T\}$,
basis for $N(A^T)$: $\{(1,1,1,-1)^T\}$

2. (a) $\{(1,1,0)^T,(-1,0,1)^T\}$

3. (b) The orthogonal complement is spanned by $(-5,1,3)^T$.

4. $\{(-1,2,0,1)^T,(2,-3,1,0)^T\}$ is one basis for $S^\perp$.

6. (a) $\mathbf{N}=(8,-2,1)^T$; (b) $8x-2y+z=7$

10. $\dim N(A)=n-r$, $\dim N(A^T)=m-r$

5.3 1. (a) $(2,1)^T$; (c) $(1.6,0.6,1.2)^T$

2. (1a) $\mathbf{p}=(3,1,0)^T$, $\mathbf{r}=(0,0,2)^T$
(1c) $\mathbf{p}=(3.4,0.2,0.6,2.8)^T$,
$\mathbf{r}=(0.6,-0.2,0.4,-0.8)^T$

3. (a) $\{(1-2\alpha,\alpha)^T \mid \alpha \text{ real}\}$;
(b) $\{(2-2\alpha,1-\alpha,\alpha)^T \mid \alpha \text{ real}\}$

4. (a) $\mathbf{p}=(1,2,-1)^T$, $\mathbf{b}-\mathbf{p}=(2,0,2)^T$;
(b) $\mathbf{p}=(3,1,4)^T$, $\mathbf{p}-\mathbf{b}=(-5,-1,4)^T$

5. (a) $y=1.8+2.9x$

6. $0.55+1.65x+1.25x^2$

14. The least squares circle will have center $(0.58,-0.64)$ and radius 2.73 (answers rounded to two decimal places).

15. (a) $\mathbf{w}=(0.1995,\ 0.2599,\ 0.3412,\ 0.1995)^T$
(b) $\mathbf{r}=(0.2605,\ 0.2337,\ 0.2850,\ 0.2208)^T$

5.4 1. $\|\mathbf{x}\|_2=2$, $\|\mathbf{y}\|_2=6$, $\|\mathbf{x}+\mathbf{y}\|_2=2\sqrt{10}$

2. (a) $\theta=\frac{\pi}{4}$; $\mathbf{p}=(\frac{4}{3},\frac{1}{3},\frac{1}{3},0)^T$

3. (b) $\|\mathbf{x}\|=1$, $\|\mathbf{y}\|=3$

4. (a) 0; (b) 5; (c) 7; (d) $\sqrt{74}$

7. (a) 1; (b) $\frac{1}{\pi}$

8. (a) $\frac{\pi}{6}$; (b) $\mathbf{p}=\frac{3}{2}x$

11. (a) $\frac{\sqrt{10}}{2}$; (b) $\frac{\sqrt{34}}{4}$

15. (a) $\|\mathbf{x}\|_1=7$, $\|\mathbf{x}\|_2=5$, $\|\mathbf{x}\|_\infty=4$;
(b) $\|\mathbf{x}\|_1=4$, $\|\mathbf{x}\|_2=\sqrt{6}$, $\|\mathbf{x}\|_\infty=2$;
(c) $\|\mathbf{x}\|_1=3$, $\|\mathbf{x}\|_2=\sqrt{3}$, $\|\mathbf{x}\|_\infty=1$

16. $\|\mathbf{x}-\mathbf{y}\|_1=5$, $\|\mathbf{x}-\mathbf{y}\|_2=3$, $\|\mathbf{x}-\mathbf{y}\|_\infty=2$

28. (a) not a norm; (b) norm; (c) norm

5.5 1. (a) and (d)

2. (b) $\mathbf{x}=-\frac{\sqrt{2}}{3}\mathbf{u}_1+\frac{5}{3}\mathbf{u}_2$,
$$\|\mathbf{x}\|=\left[\left(-\frac{\sqrt{2}}{3}\right)^2+\left(\frac{5}{3}\right)^2\right]^{1/2}=\sqrt{3}$$

3. $\mathbf{p}=(\frac{23}{18},\frac{41}{18},\frac{8}{9})^T$, $\mathbf{p}-\mathbf{x}=(\frac{5}{18},\frac{5}{18},-\frac{10}{9})^T$

4. (b) $c_1=y_1\cos\theta+y_2\sin\theta$,
$c_2=-y_1\sin\theta+y_2\cos\theta$

6. (a) 15; (b) $\|\mathbf{u}\|=3$, $\|\mathbf{v}\|=5\sqrt{2}$; (c) $\frac{\pi}{4}$

9. (b) (i) 0, (ii) $-\frac{\pi}{2}$, (iii) 0, (iv) $\frac{\pi}{8}$

21. (b) (i) $(2,-2)^T$, (ii) $(5,2)^T$, (iii) $(3,1)^T$

22. (a) $P=\begin{bmatrix}\frac{1}{2}&\frac{1}{2}&0&0\\ \frac{1}{2}&\frac{1}{2}&0&0\\ 0&0&\frac{1}{2}&\frac{1}{2}\\ 0&0&\frac{1}{2}&\frac{1}{2}\end{bmatrix}$;

23. (b) $Q=\begin{bmatrix}\frac{1}{2}&-\frac{1}{2}&0&0\\ -\frac{1}{2}&\frac{1}{2}&0&0\\ 0&0&\frac{1}{2}&-\frac{1}{2}\\ 0&0&-\frac{1}{2}&\frac{1}{2}\end{bmatrix}$

29. (b) $\|1\|=\sqrt{2}$, $\|x\|=\frac{\sqrt{6}}{3}$; (c) $l(x)=\frac{9}{7}x$

5.6 1. (a) $\left\{\left(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^T,\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)^T\right\}$;
(b) $\left\{\left(\frac{2}{\sqrt{5}},\frac{1}{\sqrt{5}}\right)^T,\left(-\frac{1}{\sqrt{5}},\frac{2}{\sqrt{5}}\right)^T\right\}$

2. (a) $\begin{bmatrix}-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}\end{bmatrix}\begin{bmatrix}\sqrt{2}&\sqrt{2}\\ 0&4\sqrt{2}\end{bmatrix}$;
(b) $\begin{bmatrix}\frac{2}{\sqrt{5}}&-\frac{1}{\sqrt{5}}\\ \frac{1}{\sqrt{5}}&\frac{2}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}\sqrt{5}&4\sqrt{5}\\ 0&3\sqrt{5}\end{bmatrix}$

3. $\{(\frac{1}{3},\frac{2}{3},-\frac{2}{3})^T,\ (\frac{2}{3},\frac{1}{3},\frac{2}{3})^T,\ (-\frac{2}{3},\frac{2}{3},\frac{1}{3})^T\}$

4. $u_1(x)=\frac{1}{\sqrt{2}}$, $u_2(x)=\frac{\sqrt{6}}{2}x$,
$u_3(x)=\frac{3\sqrt{10}}{4}\left(x^2-\frac{1}{3}\right)$

5. (a) $\left\{\frac{1}{3}(2,1,2)^T,\ \frac{\sqrt{2}}{6}(-1,4,-1)^T\right\}$;

(b) $Q = \begin{bmatrix} \frac{2}{3} & \frac{-\sqrt{2}}{6} \\ \frac{1}{3} & \frac{2\sqrt{2}}{3} \\ \frac{2}{3} & \frac{-\sqrt{2}}{6} \end{bmatrix}; \quad R = \begin{bmatrix} 3 & \frac{5}{3} \\ 0 & \frac{\sqrt{2}}{3} \end{bmatrix};$

(c) $\mathbf{x} = \begin{bmatrix} 9 \\ -3 \end{bmatrix}$

6. (b) $\begin{bmatrix} \frac{3}{5} & -\frac{4}{5\sqrt{2}} \\ \frac{4}{5} & \frac{3}{5\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 0 & 2\sqrt{2} \end{bmatrix};$

(c) $(2.1, 5.5)^T$

7. $\left\{\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)^T, \left(\frac{\sqrt{2}}{3}, \frac{\sqrt{2}}{3}, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{6}\right)^T\right\}$

8. $\left\{\begin{bmatrix} \frac{4}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{1}{5} \end{bmatrix}, \begin{bmatrix} \frac{1}{5} \\ -\frac{2}{5} \\ -\frac{2}{5} \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}\right\}$

5.7 1. (a) $T_4 = 8x^4 - 8x^2 + 1$, $T_5 = 16x^5 - 20x^3 + 5x$;

(b) $H_4 = 16x^4 - 48x^2 + 12$,
$H_5 = 32x^5 - 160x^3 + 120x$

2. $p_1(x) = x$, $p_2(x) = x^2 - \frac{4}{\pi} + 1$

4. $p(x) = (\sinh 1)P_0(x) + \frac{3}{e}P_1(x) + 5\left(\sinh 1 - \frac{3}{e}\right)P_2(x)$
$p(x) \approx 0.9963 + 1.1036x + 0.5367x^2$

6. (a) $U_0 = 1$, $U_1 = 2xU_2 = 4x^2 - 1$

11. $p(x) = (x-2)(x-3) + (x-1)(x-3) + 2(x-1)(x-2)$

13. $1 \cdot f\left(-\frac{1}{\sqrt{3}}\right) + 1 \cdot f\left(\frac{1}{\sqrt{3}}\right)$

14. (a) degree 3 or less; (b) the formula gives the exact answer for the first integral. The approximate value for the second integral is 1.5, while the exact answer is $\dfrac{\pi}{2}$.

CHAPTER TEST A

1. False 2. False 3. False 4. False 5. True
6. False 7. True 8. True 9. True 10. False

Chapter 6

6.1 1. (a) $\lambda_1 = 5$, the eigenspace is spanned by $(1,1)^T$, $\lambda_2 = -1$, the eigenspace is spanned by $(1,-2)^T$;

(b) $\lambda_1 = 3$, the eigenspace is spanned by $(4,3)^T$, $\lambda_2 = 2$, the eigenspace is spanned by $(1,1)^T$;

(c) $\lambda_1 = \lambda_2 = 2$, the eigenspace is spanned by $(1,1)^T$,

(d) $\lambda_1 = 3 + 4i$, the eigenspace is spanned by $(2i,1)^T$, $\lambda_2 = 3 - 4i$, the eigenspace is spanned by $(-2i,1)^T$;

(e) $\lambda_1 = 2 + i$, the eigenspace is spanned by $(1, 1+i)^T$, $\lambda_2 = 2 - i$, the eigenspace is spanned by $(1, 1-i)^T$;

(f) $\lambda_1 = \lambda_2 = \lambda_3 = 0$, the eigenspace is spanned by $(1,0,0)^T$;

(g) $\lambda_1 = 2$, the eigenspace is spanned by $(1,1,0)^T$, $\lambda_2 = 1$, the eigenspace is spanned by $(1,0,0)^T$, $(0,1,-1)^T$;

(h) $\lambda_1 = 1$, the eigenspace is spanned by $(1,0,0)^T$, $\lambda_2 = 4$, the eigenspace is spanned by $(1,1,1)^T$,
$\lambda_3 = -2$, the eigenspace is spanned by $(-1,-1,5)^T$;

(i) $\lambda_1 = 2$, the eigenspace is spanned by $(7,3,1)^T$, $\lambda_2 = 1$, the eigenspace is spanned by $(3,2,1)^T$, $\lambda_3 = 0$, the eigenspace is spanned by $(1,1,1)^T$;

(j) $\lambda_1 = \lambda_2 = \lambda_3 = -1$, the eigenspace is spanned by $(1,0,1)^T$;

(k) $\lambda_1 = \lambda_2 = 2$, the eigenspace is spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$, $\lambda_3 = 3$, the eigenspace is spanned by $\mathbf{e}_3$, $\lambda_4 = 4$, the eigenspace is spanned by $\mathbf{e}_4$;

(l) $\lambda_1 = 3$, the eigenspace is spanned by $(1,2,0,0)^T$, $\lambda_2 = 1$, the eigenspace is spanned by $(0,1,0,0)^T$,
$\lambda_3 = \lambda_4 = 2$, the eigenspace is spanned by $(0,0,1,0)^T$

10. β is an eigenvalue of B if and only if $\beta = \lambda - \alpha$ for some eigenvalue λ of A.

14. $\lambda_1 = 6$, $\lambda_2 = 2$;

24. $\lambda_1 \mathbf{x}^T\mathbf{y} = (A\mathbf{x})^T\mathbf{y} = \mathbf{x}^T A^T \mathbf{y} = \lambda_2 \mathbf{x}^T\mathbf{y}$

6.2 1. (a) $\begin{bmatrix} c_1e^{2t} + c_2e^{3t} \\ c_1e^{2t} + 2c_2e^{3t} \end{bmatrix};$

(b) $\begin{bmatrix} -c_1e^{-2t} - 4c_2e^{t} \\ c_1e^{-2t} + c_2e^{t} \end{bmatrix};$

(c) $\begin{bmatrix} 2c_1 + c_2e^{5t} \\ c_1 - 2c_2e^{5t} \end{bmatrix};$

(d) $\begin{bmatrix} -c_1e^t \sin t + c_2e^t \cos t \\ c_1e^t \cos t + c_2e^t \sin t \end{bmatrix}$;

(e) $\begin{bmatrix} -c_1e^{3t} \sin 2t + c_2e^{3t} \cos 2t \\ c_1e^{3t} \cos 2t + c_2e^{3t} \sin 2t \end{bmatrix}$;

(f) $\begin{bmatrix} -c_1 + c_2e^{5t} + c_3e^t \\ -3c_1 + 8c_2e^{5t} \\ c_1 + 4c_2e^{5t} \end{bmatrix}$

2. (a) $\begin{bmatrix} e^{-3t} + 2e^t \\ -e^{-3t} + 2e^t \end{bmatrix}$;

(b) $\begin{bmatrix} e^t \cos 2t + 2e^t \sin 2t \\ e^t \sin 2t - 2e^t \cos 2t \end{bmatrix}$;

(c) $\begin{bmatrix} -6e^t + 2e^{-t} + 6 \\ -3e^t + e^{-t} + 4 \\ -e^t + e^{-t} + 2 \end{bmatrix}$;

(d) $\begin{bmatrix} -2 - 3e^t + 6e^{2t} \\ 1 + 3e^t - 3e^{2t} \\ 1 + 3e^{2t} \end{bmatrix}$

4. $y_1(t) = 15e^{-0.24t} + 25e^{-0.08t}$,
$y_2(t) = -30e^{-0.24t} + 50e^{-0.08t}$

5. (a) $\begin{bmatrix} -2c_1e^t - 2c_2e^{-t} + c_3e^{\sqrt{2}t} + c_4e^{-\sqrt{2}t} \\ c_1e^t + c_2e^{-t} - c_3e^{\sqrt{2}t} - c_4e^{-\sqrt{2}t} \end{bmatrix}$

(b) $\begin{bmatrix} c_1e^{2t} + c_2e^{-2t} - c_3e^t - c_4e^{-t} \\ c_1e^{2t} - c_2e^{-2t} + c_3e^t - c_4e^{-t} \end{bmatrix}$

6. $y_1(t) = -e^{2t} + e^{-2t} + e^t$;
$y_2(t) = -e^{2t} - e^{-2t} + 2e^t$

8. $x_1(t) = \cos t + 3 \sin t + \frac{1}{\sqrt{3}} \sin \sqrt{3}t$,
$x_2(t) = \cos t + 3 \sin t - \frac{1}{\sqrt{3}} \sin \sqrt{3}t$

10. (a) $m_1x_1''(t) = -kx_1 + k(x_2 - x_1)$
$m_2x_2''(t) = -k(x_2 - x_1) + k(x_3 - x_2)$
$m_3x_3''(t) = -k(x_3 - x_2) - kx_3$

(b) $\begin{bmatrix} 0.1 \cos 2\sqrt{3}t + 0.9 \cos \sqrt{2}t \\ -0.2 \cos 2\sqrt{3}t + 1.2 \cos \sqrt{2}t \\ 0.1 \cos 2\sqrt{3}t + 0.9 \cos \sqrt{2}t \end{bmatrix}$

11. $p(\lambda) = (-1)^n(\lambda^n - a_{n-1}\lambda^{n-1} - \cdots - a_1\lambda - a_0)$

6.3

8. (b) $\alpha = 2$; (c) $\alpha = 3$ or $\alpha = -1$;
(d) $\alpha = 1$; (e) $\alpha = 0$; (g) all values of α

21. The transition matrix and steady-state vector for the Markov chain are

$$\begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}$$

In the long run we would expect 60 percent of the employees to be enrolled.

22. (a) $A = \begin{bmatrix} 0.70 & 0.20 & 0.10 \\ 0.20 & 0.70 & 0.10 \\ 0.10 & 0.10 & 0.80 \end{bmatrix}$

(c) The membership of all three groups will approach 100,000 as n gets large.

26. The transition matrix is

$$A = 0.85 \begin{bmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{3} & 0 & 0 & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{3} & 0 & 1 & \frac{1}{4} \end{bmatrix} + 0.15 \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

30. (b) $\begin{bmatrix} e & e \\ 0 & e \end{bmatrix}$

31. (a) $\begin{bmatrix} 3-2e & 1-e \\ -6+6e & -2+3e \end{bmatrix}$;

(c) $\begin{bmatrix} e & -1+e & -1+e \\ 1-e & 2-e & 1-e \\ -1+e & -1+e & e \end{bmatrix}$

32. (a) $\begin{bmatrix} e^{-t} \\ e^{-t} \end{bmatrix}$; (b) $\begin{bmatrix} -3e^t - e^{-t} \\ e^t + e^{-t} \end{bmatrix}$;

(c) $\begin{bmatrix} 3e^t - 2 \\ 2 - e^{-t} \\ e^{-t} \end{bmatrix}$

6.4

1. (a) $\|\mathbf{z}\| = 6$, $\|\mathbf{w}\| = 3$, $\langle \mathbf{z}, \mathbf{w} \rangle = -4 + 4i$, $\langle \mathbf{w}, \mathbf{z} \rangle = -4 - 4i$;

(b) $\|\mathbf{z}\| = 4$, $\|\mathbf{w}\| = 7$, $\langle \mathbf{z}, \mathbf{w} \rangle = -4 + 10i$, $\langle \mathbf{w}, \mathbf{z} \rangle = -4 - 10i$

2. (b) $\mathbf{z} = 4\mathbf{z}_1 + 2\sqrt{2}\mathbf{z}_2$

3. (a) $\mathbf{u}_1^H\mathbf{z} = 4 + 2i$, $\mathbf{z}^H\mathbf{u}_1 = 4 - 2i$, $\mathbf{u}_2^H\mathbf{z} = 6 - 5i$, $\mathbf{z}^H\mathbf{u}_2 = 6 + 5i$;

(b) $\|\mathbf{z}\| = 9$

4. (b) and (f) are Hermitian while (b), (c), (e), and (f) are normal.

14. (b) $\|U\mathbf{x}\|^2 = (U\mathbf{x})^HU\mathbf{x} = \mathbf{x}^HU^HU\mathbf{x} = \mathbf{x}^H\mathbf{x} = \|\mathbf{x}\|^2$

15. U is unitary, since $U^HU = (I - 2\mathbf{u}\mathbf{u}^H)^2 = I - 4\mathbf{u}\mathbf{u}^H + 4\mathbf{u}(\mathbf{u}^H\mathbf{u})\mathbf{u}^H = I$.

24. $\lambda_1 = 1, \lambda_2 = -1,$
$\mathbf{u}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$, $\mathbf{u}_2 = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T$,
$$A = 1\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} + (-1)\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

6.5 2. (a) $\sigma_1 = \sqrt{10}, \sigma_2 = 0$;
(b) $\sigma_1 = 3, \sigma_2 = 2$;
(c) $\sigma_1 = 4, \sigma_2 = 2$;
(d) $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$. The matrices U and V are not unique. The reader may check his or her answers by multiplying out $U\Sigma V^T$.

3. (b) rank of $A = 2$, $A' = \begin{bmatrix} 1.2 & -2.4 \\ -0.6 & 1.2 \end{bmatrix}$

4. The closest matrix of rank 2 is
$$\begin{bmatrix} -2 & 8 & 20 \\ 14 & 19 & 10 \\ 0 & 0 & 0 \end{bmatrix},$$
The closest matrix of rank 1 is
$$\begin{bmatrix} 6 & 12 & 12 \\ 8 & 16 & 16 \\ 0 & 0 & 0 \end{bmatrix}$$

5. (a) basis for $R(A^T)$:
$\{\mathbf{v}_1 = (\frac{2}{3}, \frac{2}{3}, \frac{1}{3})^T, \mathbf{v}_2 = (-\frac{2}{3}, \frac{1}{3}, \frac{2}{3})^T\}$;
basis for $N(A)$: $\{\mathbf{v}_3 = (\frac{1}{3}, -\frac{2}{3}, \frac{2}{3})^T\}$

6.6 1. (a) $\begin{bmatrix} 3 & -\frac{5}{2} \\ -\frac{5}{2} & 1 \end{bmatrix}$; (b) $\begin{bmatrix} 2 & \frac{1}{2} & -1 \\ \frac{1}{2} & 3 & \frac{3}{2} \\ -1 & \frac{3}{2} & 1 \end{bmatrix}$

3. (a) $Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, $\frac{(x')^2}{4} + \frac{(y')^2}{12} = 1$, ellipse;
(d) $Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$,
$\left(y' + \frac{\sqrt{2}}{2}\right)^2 = -\frac{\sqrt{2}}{2}(x' - \sqrt{2})$ or
$(y'')^2 = -\dfrac{\sqrt{2}}{2}x''$, parabola

6. (a) positive definite; (b) indefinite;
(d) negative definite; (e) indefinite

7. (a) minimum; (b) saddle point;
(c) saddle point; (f) local maximum

6.7 1. (a) $\det(A_1) = 2$, $\det(A_2) = 3$, positive definite;
(b) $\det(A_1) = 3$, $\det(A_2) = -10$, not positive definite;
(c) $\det(A_1) = 6$, $\det(A_2) = 14$, $\det(A_3) = -38$, not positive definite;
(d) $\det(A_1) = 4$, $\det(A_2) = 8$, $\det(A_3) = 13$, positive definite

2. $a_{11} = 3$, $a_{22}^{(1)} = 2$, $a_{33}^{(2)} = \frac{4}{3}$

4. (a) $\begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$;
(b) $\begin{bmatrix} 1 & 0 \\ -\frac{1}{3} & 1 \end{bmatrix}\begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -\frac{1}{3} \\ 0 & 1 \end{bmatrix}$;
(c) $\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ \frac{1}{4} & -1 & 1 \end{bmatrix}\begin{bmatrix} 16 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$;
(d) $\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{2}{3} & 1 & 1 \end{bmatrix}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{3} \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

5. (a) $\begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$;
(b) $\begin{bmatrix} 3 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix}$;
(c) $\begin{bmatrix} 4 & 0 & 0 \\ 2 & \sqrt{2} & 0 \\ 1 & -\sqrt{2} & 2 \end{bmatrix}\begin{bmatrix} 4 & 2 & 1 \\ 0 & \sqrt{2} & -\sqrt{2} \\ 0 & 0 & 2 \end{bmatrix}$;
(d) $\begin{bmatrix} 3 & 0 & 0 \\ 1 & \sqrt{3} & 0 \\ -2 & \sqrt{3} & \sqrt{2} \end{bmatrix}\begin{bmatrix} 3 & 1 & -2 \\ 0 & \sqrt{3} & \sqrt{3} \\ 0 & 0 & \sqrt{2} \end{bmatrix}$

6.8 1. (a) $\lambda_1 = 4$, $\lambda_2 = -1$, $\mathbf{x}_1 = (3, 2)^T$;
(b) $\lambda_1 = 8$, $\lambda_2 = 3$, $\mathbf{x}_1 = (1, 2)^T$;
(c) $\lambda_1 = 7$, $\lambda_2 = 2$, $\lambda_3 = 0$, $\mathbf{x}_1 = (1, 1, 1)^T$

2. (a) $\lambda_1 = 3, \lambda_2 = -1, \mathbf{x}_1 = (3, 1)^T$;
(b) $\lambda_1 = 2 = 2\exp(0)$,
$\lambda_2 = -2 = 2\exp(\pi i)$, $\mathbf{x}_1 = (1, 1)^T$;
(c) $\lambda_1 = 2 = 2\exp(0)$,
$\lambda_2 = -1 + \sqrt{3}i = 2\exp\left(\frac{2\pi i}{3}\right)$,
$\lambda_3 = -1 - \sqrt{3}i = 2\exp\left(\frac{4\pi i}{3}\right)$,
$\mathbf{x}_1 = (4, 2, 1)^T$

3. $x_1 = 70{,}000$, $x_2 = 56{,}000$, $x_3 = 44{,}000$

4. $x_1 = x_2 = x_3$

5. $(I - A)^{-1} = I + A + \cdots + A^{m-1}$

6. (a) $(I - A)^{-1} = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{bmatrix}$;

(b) $A^2 = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

$A^3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

7. (b) and (c) are reducible.

15. (d) $\mathbf{w} = (\frac{12}{29}, \frac{12}{29}, \frac{3}{29}, \frac{2}{29})^T$
$\approx (0.4138, 0.4138, 0.1034, 0.0690)^T$

CHAPTER TEST A

1. True 2. False 3. True 4. False 5. False
6. False 7. False 8. False 9. True 10. False
11. True 12. True 13. True 14. False
15. True

Chapter 7

7.1

1. (a) 0.231×10^4; (b) 0.326×10^2;
(c) 0.128×10^{-1}; (d) 0.824×10^5

2. (a) $\epsilon = -2$; $\delta \approx -8.7 \times 10^{-4}$;
(b) $\epsilon = 0.04$; $\delta \approx 1.2 \times 10^{-3}$;
(c) $\epsilon = 3.0 \times 10^{-5}$; $\delta \approx 2.3 \times 10^{-3}$;
(d) $\epsilon = -31$; $\delta \approx -3.8 \times 10^{-4}$

3. (a) $(1.0101)_2 \times 2^4$; (b) $(1.1000)_2 \times 2^{-2}$;
(c) $(1.0100)_2 \times 2^3$; (d) $-(1.1010)_2 \times 2^{-4}$

4. (a) 10,420, $\epsilon = -0.0018$, $\delta \approx -1.7 \times 10^{-7}$;
(b) 0, $\epsilon = -8$, $\delta = -1$;
(c) 1×10^{-4}, $\epsilon = 5 \times 10^{-5}$, $\delta = 1$;
(d) 82,190, $\epsilon = 25.7504$, $\delta \approx 3.1 \times 10^{-4}$

5. (a) 0.1043×10^6; (b) 0.1045×10^6;
(c) 0.1045×10^6

8. 23

9. (a) $(1.00111000000000000000000)_2 \times 2^3$ or 9.75

7.2

1. $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3 \end{bmatrix}$

2. (a) $(2, -1, 3)^T$; (b) $(1, -1, 3)^T$;
(c) $(1, 5, 1)^T$

3. (a) n^2 multiplications and $n(n-1)$ additions;
(b) n^3 multiplications and $n^2(n-1)$ additions;
(c) $(AB)\mathbf{x}$ requires $n^3 + n^2$ multiplications and $n^3 - n$ additions; $A(B\mathbf{x})$ requires $2n^2$ multiplications and $2n(n-1)$ additions.

4. (b) (i) 156 multiplications and 105 additions,
(ii) 47 multiplications and 24 additions,
(iii) 100 multiplications and 60 additions

8. $5n - 4$ multiplications/divisions, $3n - 3$ additions/subtractions

9. (a) $[(n-j)(n-j+1)]/2$ multiplications; $[(n-j-1)(n-j)]/2$ additions;
(c) It requires on the order of $\frac{2}{3}n^3$ additional multiplications/divisions to compute A^{-1} given the LU factorization.

7.3

1. (a) $(1, 1, -2)$;
(b) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & 8 \\ 0 & 0 & -23 \end{bmatrix}$

2. (a) $(1, 2, 2)$; (b) $(4, -3, 0)$;
(c) $(1, 1, 1)$

3. $P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $L = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{3} & 1 \end{bmatrix}$,

$U = \begin{bmatrix} 2 & 4 & -6 \\ 0 & 6 & 9 \\ 0 & 0 & 5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 6 \\ -\frac{1}{2} \\ 1 \end{bmatrix}$

4. $P = Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,

$PAQ = LU = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 0 & 2 \end{bmatrix}$,

$\mathbf{x} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

5. (a) $\hat{\mathbf{c}} = P\mathbf{c} = (-4, 6)^T$,
$\mathbf{y} = L^{-1}\hat{\mathbf{c}} = (-4, 8)^T$,
$\mathbf{z} = U^{-1}\mathbf{y} = (-3, 4)^T$
(b) $\mathbf{x} = Q\mathbf{z} = (4, -3)^T$

6. (b) $P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$,

$L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ \frac{1}{2} & \frac{2}{3} & 1 \end{bmatrix}$, $U = \begin{bmatrix} 8 & 6 & 2 \\ 0 & 6 & 3 \\ 0 & 0 & 2 \end{bmatrix}$

7. Error $\dfrac{-2000e}{0.6} \approx -3333e$. If $e = 0.001$, then $\delta = -\frac{2}{3}$.

8. $(1.667, 1.001)$

9. (5.002, 1.000)
10. (5.001, 1.001)

7.4

1. (a) $\|A\|_F = \sqrt{2}$, $\|A\|_\infty = 1$, $\|A\|_1 = 1$;
   (b) $\|A\|_F = 5$, $\|A\|_\infty = 5$, $\|A\|_1 = 6$;
   (c) $\|A\|_F = \|A\|_\infty = \|A\|_1 = 1$;
   (d) $\|A\|_F = 7$, $\|A\|_\infty = 6$, $\|A\|_1 = 10$;
   (e) $\|A\|_F = 9$, $\|A\|_\infty = 10$, $\|A\|_1 = 12$
2. 2
4. $\|I\|_1 = \|I\|_\infty = 1$, $\|I\|_F = \sqrt{n}$;
6. (a) 10; (b) $(-1, 1, -1)^T$
27. (a) Since for any vector $\mathbf{y}$ in $\mathbb{R}^n$ we have

$$\|\mathbf{y}\|_\infty \leq \|\mathbf{y}\|_2 \leq \sqrt{n}\,\|\mathbf{y}\|_\infty$$

it follows that

$$\begin{aligned} \|A\mathbf{x}\|_\infty &\leq \|A\mathbf{x}\|_2 \\ &\leq \|A\|_2\|\mathbf{x}\|_2 \leq \sqrt{n}\,\|A\|_2\|\mathbf{x}\|_\infty \end{aligned}$$

29. $\mathrm{cond}_\infty A = 400$
30. The solutions are $\begin{bmatrix} -0.48 \\ 0.8 \end{bmatrix}$ and $\begin{bmatrix} -2.902 \\ 2.0 \end{bmatrix}$
31. $\mathrm{cond}_\infty(A) = 28$
33. (a) $A_n^{-1} = \begin{bmatrix} 1-n & n \\ n & -n \end{bmatrix}$;
    (b) $\mathrm{cond}_\infty A_n = 4n$;
    (c) $\lim_{n\to\infty} \mathrm{cond}_\infty A_n = \infty$;
34. $\sigma_1 = 8$, $\sigma_2 = 8$, $\sigma_3 = 4$
35. (a) $\mathbf{r} = (-0.06, 0.02)^T$ and the relative residual is 0.012;
    (b) 20;
    (d) $\mathbf{x} = (1, 1)^T$, $\|\mathbf{x} - \mathbf{x}'\|_\infty = 0.12$;
36. $\mathrm{cond}_1(A) = 6$
37. 0.3
38. (a) $\|\mathbf{r}\|_\infty = 0.10$, $\mathrm{cond}_\infty(A) = 32$;
    (b) 0.64;
    (c) $\mathbf{x} = (12.50, 4.26, 2.14, 1.10)^T$, $\delta = 0.04$

7.5

1. (a) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$; (b) $\begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$;
   (c) $\begin{bmatrix} -\frac{4}{5} & \frac{3}{5} \\ -\frac{3}{5} & -\frac{4}{5} \end{bmatrix}$
2. (a) $\begin{bmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \\ \frac{4}{5} & 0 & -\frac{3}{5} \end{bmatrix}$;
   (b) $\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$;
   (c) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix}$;
   (d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$
3. $H = I - \frac{1}{\beta}\mathbf{v}\mathbf{v}^T$ for the given β and $\mathbf{v}$.
   (a) $\beta = 90$, $\mathbf{v} = (-10, 8, -4)^T$;
   (b) $\beta = 70$, $\mathbf{v} = (10, 6, 2)^T$;
   (c) $\beta = 15$, $\mathbf{v} = (-5, -3, 4)^T$
4. (a) $\beta = 90$, $\mathbf{v} = (0, 10, 4, 8)^T$;
   (b) $\beta = 15$, $\mathbf{v} = (0, 0, -5, -1, 2)^T$
6. (a) $H_2H_1A = R$, where $H_i = I - \dfrac{1}{\beta_i}\mathbf{v}_i\mathbf{v}_i^T$, $i = 1, 2$, and $\beta_1 = 12$, $\beta_2 = 45$.

$$\mathbf{v}_1 = \begin{bmatrix} -4 \\ 2 \\ -2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 9 \\ -3 \end{bmatrix},$$

$$R = \begin{bmatrix} 3 & \frac{19}{2} & \frac{9}{2} \\ 0 & -5 & -3 \\ 0 & 0 & 6 \end{bmatrix},$$

$$\mathbf{c} = H_2H_1\mathbf{b} = \begin{bmatrix} -\frac{5}{2} \\ -5 \\ 0 \end{bmatrix};$$

   (b) $\mathbf{x} = (-4, 1, 0)^T$
7. (a) $G = \begin{bmatrix} \dfrac{3}{5} & \dfrac{4}{5} \\ \dfrac{4}{5} & -\dfrac{3}{5} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$
8. It takes three multiplications, two additions, and one square root to determine H. It takes four multiplications/divisions, one addition, and one square root to determine G. The calculation of GA requires $4n$ multiplications

and $2n$ additions, while the calculation of HA requires $3n$ multiplications/divisions and $3n$ additions.

9. (a) $n-k+1$ multiplications/divisions, $2n-2k+1$ additions;
   (b) $n(n-k+1)$ multiplications/divisions, $n(2n-2k+1)$ additions
10. (a) $4(n-k)$ multiplications/divisions, $2(n-k)$ additions;
    (b) $4n(n-k)$ multiplications, $2n(n-k)$ additions
11. (a) rotation; (b) rotation;
    (c) Givens transformation;
    (d) Givens transformation

7.6

1. (a) $\mathbf{u}_1=\begin{bmatrix}1\\1\end{bmatrix}$; (b) $A_2=\begin{bmatrix}2&0\\0&0\end{bmatrix}$;
   (c) $\lambda_1=2,\ \lambda_2=0$; the eigenspace corresponding to λ_1 is spanned by $\mathbf{u}_1$.
2. (a) $\mathbf{v}_1=\begin{bmatrix}3\\5\\3\end{bmatrix}$, $\mathbf{u}_1=\begin{bmatrix}0.6\\1.0\\0.6\end{bmatrix}$,
   $\mathbf{v}_2=\begin{bmatrix}2.2\\4.2\\2.2\end{bmatrix}$, $\mathbf{u}_2=\begin{bmatrix}0.52\\1.00\\0.52\end{bmatrix}$,
   $\mathbf{v}_3=\begin{bmatrix}2.05\\4.05\\2.05\end{bmatrix}$;
   (b) $\lambda_1'=4.05$; (c) $\lambda_1=4,\ \delta=0.0125$
3. (b) A has no dominant eigenvalue.
4. $A_2=\begin{bmatrix}3&-1\\-1&1\end{bmatrix}$, $A_3=\begin{bmatrix}3.4&0.2\\0.2&0.6\end{bmatrix}$,
   $\lambda_1=2+\sqrt{2}\approx 3.414,\ \lambda_2=2-\sqrt{2}\approx 0.586$
5. (b) $H=I-\dfrac{1}{\beta}\mathbf{v}\mathbf{v}^T$, where $\beta=\frac{1}{3}$ and $\mathbf{v}=(-\frac{1}{3},-\frac{2}{3},\frac{1}{3})^T$;
   (c) $\lambda_2=3,\ \lambda_3=1,\ HAH=\begin{bmatrix}4&0&3\\0&5&-4\\0&2&-1\end{bmatrix}$

7.7

1. (a) $(\sqrt{2},0)^T$; (b) $(1-3\sqrt{2},3\sqrt{2},-\sqrt{2})^T$;
   (c) $(1,0)^T$; (d) $(1-\sqrt{2},\sqrt{2},-\sqrt{2})^T$
2. $x_i=\dfrac{d_ib_i+e_ib_{n+i}}{d_i^2+e_i^2},\ i=1,\ldots,n$
4. (a) $Q=\begin{bmatrix}\frac{1}{2}&-\frac{1}{6}\\ \frac{1}{2}&-\frac{1}{2}\\ \frac{1}{2}&\frac{5}{6}\\ \frac{1}{2}&-\frac{1}{6}\end{bmatrix}$, $R=\begin{bmatrix}2&12\\0&6\end{bmatrix}$
   (b) $\mathbf{x}=\begin{bmatrix}0&\frac{1}{3}\end{bmatrix}^T$
5. (a) $\sigma_1=\sqrt{2+\rho^2},\ \sigma_2=\rho$;
   (b) $\lambda_1'=2,\ \lambda_2'=0,\ \sigma_1'=\sqrt{2},\ \sigma_2'=0$
12. $A^+=\begin{bmatrix}\frac{1}{4}&\frac{1}{4}&0\\ \frac{1}{4}&\frac{1}{4}&0\end{bmatrix}$
13. (a) $A^+=\begin{bmatrix}\frac{1}{10}&-\frac{1}{10}\\ \frac{2}{10}&-\frac{2}{10}\end{bmatrix}$;
    (b) $A^+\mathbf{b}=\begin{bmatrix}1\\2\end{bmatrix}$;
    (c) $\left\{\mathbf{y}\ \middle|\ \mathbf{y}=\begin{bmatrix}1\\2\end{bmatrix}+\alpha\begin{bmatrix}-2\\1\end{bmatrix}\right\}$
15. $\|A_1-A_2\|_F=\rho$, $\|A_1^+-A_2^+\|_F=1/\rho$. As $\rho\to 0$, $\|A_1-A_2\|_F\to 0$ and $\|A_1^+-A_2^+\|_F\to\infty$.

CHAPTER TEST A

1. False 2. False 3. False 4. True 5. False
6. False 7. True 8. False 9. False 10. False

INDEX