The McGraw·Hill Companies

Mc Graw Hill Higher Education

INTRODUCTION TO LINEAR ALGEBRA WITH APPLICATIONS

Published by McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020.

Some ancillaries, including electronic and print components, may not be available to customers outside the United States.

This book is printed on acid-free paper.

1 2 3 4 5 6 7 8 9 0 DOC/DOC 0 9 8

ISBN 978–0–07–353235–6
MHID 0–07–353235–5

Editorial Director: *Stewart K. Mattson*
Senior Sponsoring Editor: *Elizabeth Covello*
Director of Development: *Kristine Tibbetts*
Developmental Editor: *Michelle Driscoll*
Director of Marketing: *Ryan Blankenship*
Marketing Coordinator: *Sabina Navsariwala*
Project Manager: *Joyce Watters*
Senior Production Supervisor: *Sherry L. Kane*
Senior Media Project Manager: *Tammy Juran*
Designer: *Laurie B. Janssen*
Cover Designer: *Ron Bissell*
(USE) Cover Image: *Royalty-Free/CORBIS*
Senior Photo Research Coordinator: *John C. Leland*
Supplement Producer: *Melissa M. Leick*
Compositor: *Laserwords Private Limited*
Typeface: *10.25/12 Times*
Printer: *R. R. Donnelly Crawfordsville, IN*

Library of Congress Cataloging-in-Publication Data

DeFranza, James, 1950–
Introduction to linear algebra / James DeFranza, Daniel Gagliardi. — 1st ed.
p. cm.
Includes index.
ISBN 978–0–07–353235–6 — ISBN 0–07–353235–5 (hard copy : alk. paper)
1. Algebras, Linear — Textbooks. 2. Algebras, Linear — Problems, exercises, etc. I. Gagliardi, Daniel. II. Title.

QA184.2.D44 2009
515′.5 — dc22

2008026020

To The Student

You are probably taking this course early in your undergraduate studies after two or three semesters of calculus, and most likely in your second year. Like calculus, linear algebra is a subject with elegant theory and many diverse applications. However, in this course you will be exposed to abstraction at a much higher level. To help with this transition, some colleges and universities offer a *Bridge Course to Higher Mathematics*. If you have not already taken such a course, this may likely be the first mathematics course where you will be expected to read and understand proofs of theorems, provide proofs of results as part of the exercise sets, and apply the concepts presented. All this is in the context of a specific body of knowledge. If you approach this task with an open mind and a willingness to read the text, some parts perhaps more than once, it will be an exciting and rewarding experience. Whether you are taking this course as part of a mathematics major or because linear algebra is applied in your specific area of study, a clear understanding of the theory is essential for applying the concepts of linear algebra to mathematics or other fields of science. The solved examples and exercises in the text are designed to prepare you for the types of problems you can expect to see in this course and other more advanced courses in mathematics. The organization of the material is based on our philosophy that *each topic should be fully developed before readers move onto the next*. The image of a tree on the front cover of the text is a metaphor for this learning strategy. It is particularly applicable to the study of mathematics. The trunk of the tree represents the material that forms the basis for everything that comes afterward. In our text, this material is contained in Chaps. 1 through 4. All other branches of the tree, representing more advanced topics and applications, extend from the foundational material of the trunk or from the ancillary material of the intervening branches. We have specifically designed our text so that you can read it and learn the concepts of linear algebra in a sequential and thorough manner. If you remain committed to learning this beautiful subject, the rewards will be significant in other courses you may take, and in your professional career. Good luck!

Jim DeFranza
jdefranza@stlawu.edu

Dan Gagliardi
gagliardid@canton.edu

DEFINITION 2 **Echelon Form** An $m \times n$ matrix is in **row echelon form** if

1. Every row with all 0 entries is below every row with nonzero entries.
2. If rows $1, 2, \ldots, k$ are the rows with nonzero entries and if the leading nonzero entry (**pivot**) in row i occurs in column c_i, for $1, 2, \ldots, k$, then $c_1 < c_2 < \cdots < c_k$.

The matrix is in **reduced row echelon form** if, in addition,

3. The first nonzero entry of each row is a 1.
4. Each column that contains a pivot has all other entries 0.

The process of transforming a matrix to reduced row echelon form is called *Gauss-Jordan elimination*.

EXAMPLE 3 Solve the linear system by transforming the augmented matrix to reduced row echelon form.

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

Solution The augmented matrix of the linear system is

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

To transform the matrix into reduced row echelon form, we first use the leading 1 in row 1 as a pivot to eliminate the terms in column 1 of rows 2, 3, and 4. To do this, we use the three row operations

$$\begin{aligned} -2R_1 + R_2 &\to R_2 \\ R_1 + R_3 &\to R_3 \\ -R_1 + R_4 &\to R_4 \end{aligned}$$

in succession, transforming the matrix

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

For the second step we use the leftmost 1 in row 2 as the pivot and eliminate the term in column 2 above the pivot, and the two terms below the pivot. The required row operations are

$$\begin{aligned} R_2 + R_1 &\to R_1 \\ -R_2 + R_3 &\to R_3 \\ -2R_2 + R_4 &\to R_4 \end{aligned}$$

then

$$\begin{array}{lll} a_{11} = -2 & a_{12} = 1 & a_{13} = 4 \\ a_{21} = 5 & a_{22} = 7 & a_{23} = 11 \\ a_{31} = 2 & a_{32} = 3 & a_{33} = 22 \end{array}$$

A **vector** is an $n \times 1$ matrix. The entries of a vector are called its **components**. For a given matrix A, it is convenient to refer to its *row vectors* and its *column vectors*. For example, let

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

Then the column vectors of A are

$$\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \qquad \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

while the row vectors of A, written vertically, are

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \qquad \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$$

Two $m \times n$ matrices A and B are **equal** if they have the same number of rows and columns and their corresponding entries are equal. Thus, $A = B$ if and only if $a_{ij} = b_{ij}$, for $1 \le i \le m$ and $1 \le j \le n$. Addition and scalar multiplication of matrices are also defined componentwise.

DEFINITION 1 **Addition and Scalar Multiplication** If A and B are two $m \times n$ matrices, then the **sum** of the matrices $A + B$ is the $m \times n$ matrix with the ij term given by $a_{ij} + b_{ij}$. The **scalar product** of the matrix A with the real number c, denoted by cA, is the $m \times n$ matrix with the ij term given by ca_{ij}.

EXAMPLE 1 Perform the operations on the matrices

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

a. $A + B$ **b.** $2A - 3B$

Solution

a. We add the two matrices by adding their corresponding entries, so that

$$\begin{aligned} A + B &= \begin{bmatrix} 2 & 0 & 1 \\ 4 & 3 & -1 \\ -3 & 6 & 5 \end{bmatrix} + \begin{bmatrix} -2 & 3 & -1 \\ 3 & 5 & 6 \\ 4 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2+(-2) & 0+3 & 1+(-1) \\ 4+3 & 3+5 & -1+6 \\ -3+4 & 6+2 & 5+1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 3 & 0 \\ 7 & 8 & 5 \\ 1 & 8 & 6 \end{bmatrix} \end{aligned}$$

b. To evaluate this expression, we first multiply each entry of the matrix A by 2 and each entry of the matrix B by -3. Then we add the resulting matrices. This gives

$$\begin{aligned} 2A + (-3B) &= 2\begin{bmatrix} 2 & 0 & 1 \\ 4 & 3 & -1 \\ -3 & 6 & 5 \end{bmatrix} + (-3)\begin{bmatrix} -2 & 3 & -1 \\ 3 & 5 & 6 \\ 4 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 0 & 2 \\ 8 & 6 & -2 \\ -6 & 12 & 10 \end{bmatrix} + \begin{bmatrix} 6 & -9 & 3 \\ -9 & -15 & -18 \\ -12 & -6 & -3 \end{bmatrix} \\ &= \begin{bmatrix} 10 & -9 & 5 \\ -1 & -9 & -20 \\ -18 & 6 & 7 \end{bmatrix} \end{aligned}$$

In Example 1(a) reversing the order of the addition of the matrices gives the same result. That is, $A + B = B + A$. This is so because addition of real numbers is commutative. This result holds in general, giving us that matrix addition is also a commutative operation. Some other familiar properties that hold for real numbers also hold for matrices and scalars. These properties are given in Theorem 4.

THEOREM 4 **Properties of Matrix Addition and Scalar Multiplication** Let A, B, and C be $m \times n$ matrices and c and d be real numbers.

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $c(A + B) = cA + cB$
4. $(c + d)A = cA + dA$
5. $c(dA) = (cd)A$
6. The $m \times n$ matrix with all zero entries, denoted by $\mathbf{0}$, is such that $A + \mathbf{0} = \mathbf{0} + A = A$.
7. For any matrix A, the matrix $-A$, whose components are the negative of each component of A, is such that $A + (-A) = (-A) + A = \mathbf{0}$.

Observe that the dot product of two vectors is a scalar. For example,

$$\begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} -5 \\ 1 \\ 4 \end{bmatrix} = (2)(-5) + (-3)(1) + (-1)(4) = -17$$

Now to motivate the concept and need for matrix multiplication we first introduce the operation of multiplying a vector by a matrix. As an illustration let

$$B = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

The **product** of B and $\mathbf{v}$, denoted by $B\mathbf{v}$, is a vector, in this case with two components. The first component of $B\mathbf{v}$ is the dot product of the first row vector of B with $\mathbf{v}$, while the second component is the dot product of the second row vector of B with $\mathbf{v}$, so that

$$B\mathbf{v} = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (1)(1) + (-1)(3) \\ (-2)(1) + (1)(3) \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Using this operation, the matrix B *transforms* the vector $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ to the vector $B\mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. If $A = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}$ is another matrix, then the product of A and $B\mathbf{v}$ is given by

$$A(B\mathbf{v}) = \begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

The question then arises, is there a single matrix which can be used to transform the original vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ to $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$? To answer this question, let

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

The product of B and $\mathbf{v}$ is

$$B\mathbf{v} = \begin{bmatrix} b_{11}x + b_{12}y \\ b_{21}x + b_{22}y \end{bmatrix}$$

Now, the product of A and $B\mathbf{v}$ is

$$\begin{aligned} A(B\mathbf{v}) &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11}x + b_{12}y \\ b_{21}x + b_{22}y \end{bmatrix} \\ &= \begin{bmatrix} a_{11}(b_{11}x + b_{12}y) + a_{12}(b_{21}x + b_{22}y) \\ a_{21}(b_{11}x + b_{12}y) + a_{22}(b_{21}x + b_{22}y) \end{bmatrix} \\ &= \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21})x + (a_{11}b_{12} + a_{12}b_{22})y \\ (a_{21}b_{11} + a_{22}b_{21})x + (a_{21}b_{12} + a_{22}b_{22})y \end{bmatrix} \\ &= \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \end{aligned}$$

37. Suppose that A is an $n \times n$ matrix. Show that if for each vector $\mathbf{x}$ in $\mathbb{R}^n$, $A\mathbf{x} = \mathbf{0}$, then A is the zero matrix.

38. For each positive integer n, let

$$A_n = \begin{bmatrix} 1-n & -n \\ n & 1+n \end{bmatrix}$$

Show that $A_n A_m = A_{n+m}$.

39. Find all 2×2 matrices that satisfy $AA^t = \mathbf{0}$.

40. Suppose that A and B are symmetric matrices. Show that if $AB = BA$, then AB is symmetric.

41. If A is an $m \times n$ matrix, show that AA^t and A^tA are both defined and are both symmetric.

42. An $n \times n$ matrix A is called *idempotent* provided that $A^2 = AA = A$. Suppose that A and B are $n \times n$ idempotent matrices. Show that if $AB = BA$, then the matrix AB is idempotent.

43. An $n \times n$ matrix A is *skew-symmetric* provided $A^t = -A$. Show that if a matrix is skew-symmetric, then the diagonal entries are 0.

44. The trace of an $n \times n$ matrix A is the sum of the diagonal terms, denoted $\mathbf{tr}(A)$.

a. If A and B are $n \times n$ matrices, show that $\mathbf{tr}(A + B) = \mathbf{tr}(A) + \mathbf{tr}(B)$.

b. If A is an $n \times n$ matrix and c is a scalar, show that $\mathbf{tr}(cA) = c\,\mathbf{tr}(A)$.

1.4 ▶ The Inverse of a Square Matrix

In the real number system the number 1 is the multiplicative identity. That is, for any real number a,

$$a \cdot 1 = 1 \cdot a = a$$

We also know that for every number x with $x \neq 0$, there exists the number $\frac{1}{x}$, also written x^{-1}, such that

$$x \cdot \frac{1}{x} = 1$$

We seek a similar relationship for square matrices. For an $n \times n$ matrix A, we can check that the $n \times n$ matrix

$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

is the **multiplicative identity**. That is, if A is any $n \times n$ matrix, then

$$AI = IA = A$$

This special matrix is called the **identity matrix**. For example, the 2×2, 3×3, and 4×4 identity matrices are, respectively,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

DEFINITION 1 **Inverse of a Square Matrix** Let A be an $n \times n$ matrix. If there exists an $n \times n$ matrix B such that

$$AB = I = BA$$

then the matrix B is a (multiplicative) **inverse** of the matrix A.

EXAMPLE 1 Find an inverse of the matrix

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

Solution In order for a 2×2 matrix $B = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$ to be an inverse of A, matrix B must satisfy

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 & x_2 + x_4 \\ x_1 + 2x_3 & x_2 + 2x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This matrix equation is equivalent to the linear system

$$\begin{cases} x_1 + x_3 = 1 \\ x_2 + x_4 = 0 \\ x_1 + 2x_3 = 0 \\ x_2 + 2x_4 = 1 \end{cases}$$

The augmented matrix and the reduced row echelon form are given by

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

Thus, the solution is $x_1 = 2, x_2 = -1, x_3 = -1, x_4 = 1$, and an inverse matrix is

$$B = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

The reader should verify that $AB = BA = I$.

Theorem 7 establishes the uniqueness, when it exists, of the multiplicative inverse.

THEOREM 7 The inverse of a matrix, if it exists, is unique.

Proof Assume that the square matrix A has an inverse and that B and C are both inverse matrices of A. That is, $AB = BA = I$ and $AC = CA = I$. We show that $B = C$. Indeed,

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

To illustrate the use of the formula, let

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

then

$$A^{-1} = \frac{1}{6 - (-1)} \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{7} & \frac{1}{7} \\ -\frac{1}{7} & \frac{2}{7} \end{bmatrix}$$

For an example which underscores the necessity of the condition that $ad - bc \neq 0$, we consider the matrix

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

Observe that in this case $ad - bc = 1 - 1 = 0$. Now, the matrix A is invertible if there is a $B = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$ such that

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

This matrix equation yields the inconsistent system

$$\begin{cases} x_1 + x_3 = 1 \\ x_2 + x_4 = 0 \\ x_1 + x_3 = 0 \\ x_2 + x_4 = 1 \end{cases}$$

Hence, A is not invertible.

To find the inverse of larger square matrices, we extend the method of augmented matrices. Let A be an $n \times n$ matrix. Let B be another $n \times n$ matrix, and let $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_n$ denote the n column vectors of B. Since $A\mathbf{B}_1, A\mathbf{B}_2, \ldots, A\mathbf{B}_n$ are the column vectors of AB, in order for B to be the inverse of A, we must have

$$A\mathbf{B}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad A\mathbf{B}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \ldots \qquad A\mathbf{B}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

That is, the *matrix equations*

$$A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad A\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \qquad \ldots \qquad A\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

must all have unique solutions. But all n linear systems can be solved simultaneously by row-reducing the $n \times 2n$ augmented matrix

$$\left[\begin{array}{cccc|cccc} a_{11} & a_{12} & \ldots & a_{1n} & 1 & 0 & \ldots & 0 \\ a_{21} & a_{22} & \ldots & a_{2n} & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 1 \end{array}\right]$$

On the left is the matrix A, and on the right is the matrix I. Then A **will have an inverse if and only if it is row equivalent to the identity matrix.** In this case, each of the linear systems can be solved. If the matrix A does not have an inverse, then the row-reduced matrix on the left will have a row of zeros, indicating at least one of the linear systems does not have a solution.

Example 2 illustrates the procedure.

EXAMPLE 2 Find the inverse of the matrix

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

Solution To find the inverse of this matrix, place the identity on the right to form the 3×6 matrix

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

Now use row operations to reduce the matrix on the left to the identity, while applying the same operations to the matrix on the right. The final result is

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

So the inverse matrix is

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

The reader should check that $AA^{-1} = A^{-1}A = I$.

EXAMPLE 3 Use the method of Example 2 to determine whether the matrix

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

is invertible.

Solution Following the procedure described above, we start with the matrix

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

Some examples of upper triangular matrices are

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & -1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

and some examples of lower triangular matrices are

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{bmatrix}$$

THEOREM 13 If A is an $n \times n$ triangular matrix, then the determinant of A is the product of the terms on the diagonal. That is,

$$\det(A) = a_{11} \cdot a_{22} \cdots a_{nn}$$

Proof We present the proof for an upper triangular matrix. The proof for a lower triangular matrix is identical. The proof is by induction on n. If $n = 2$, then $\det(A) = a_{11}a_{22} - 0$ and hence is the product of the diagonal terms.

Assume that the result holds for an $n \times n$ triangular matrix. We need to show that the same is true for an $(n+1) \times (n+1)$ triangular matrix A. To this end let

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & a_{1,n+1} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} & a_{2,n+1} \\ 0 & 0 & a_{33} & \cdots & a_{3n} & a_{3,n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} & a_{n,n+1} \\ 0 & 0 & 0 & \cdots & 0 & a_{n+1,n+1} \end{bmatrix}$$

Using the cofactor expansion along row $n + 1$, we have

$$\det(A) = (-1)^{(n+1)+(n+1)} a_{n+1,n+1} \begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}$$

Since the determinant on the right is $n \times n$ and upper triangular, by the inductive hypothesis

$$\begin{aligned} \det(A) &= (-1)^{2n+2}(a_{n+1,n+1})(a_{11}a_{22}\cdots a_{nn}) \\ &= a_{11}a_{22}\cdots a_{nn}a_{n+1,n+1} \end{aligned}$$

Properties of Determinants

Determinants for large matrices can be time-consuming to compute, so any properties of determinants that reduce the number of computations are useful. Theorem 14 shows how row operations affect the determinant.

EXAMPLE 6 Use Cramer's rule to solve the linear system.

$$\begin{cases} 2x + 3y = 2 \\ -5x + 7y = 3 \end{cases}$$

Solution The determinant of the coefficient matrix is given by

$$\begin{vmatrix} 2 & 3 \\ -5 & 7 \end{vmatrix} = 14 - (-15) = 29$$

and since the determinant is not zero, the system has a unique solution. The solution is given by

$$x = \frac{\begin{vmatrix} 2 & 3 \\ 3 & 7 \end{vmatrix}}{29} = \frac{14 - 9}{29} = \frac{5}{29} \qquad \text{and} \qquad y = \frac{\begin{vmatrix} 2 & 2 \\ -5 & 3 \end{vmatrix}}{29} = \frac{6 - (-10)}{29} = \frac{16}{29}$$

THEOREM 18 **Cramer's Rule** Let A be an $n \times n$ invertible matrix, and let $\mathbf{b}$ be a column vector with n components. Let A_i be the matrix obtained by replacing the ith column of A with $\mathbf{b}$. If $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is the unique solution to the linear system $A\mathbf{x} = \mathbf{b}$, then

$$x_i = \frac{\det(A_i)}{\det(A)} \qquad \text{for} \qquad i = 1, 2, \ldots, n$$

Proof Let I_i be the matrix obtained by replacing the ith column of the identity matrix with $\mathbf{x}$. Then the linear system is equivalent to the matrix equation

$$AI_i = A_i \qquad \text{so} \qquad \det(AI_i) = \det(A_i)$$

By Theorem 15, part 1, we have

$$\det(A)\det(I_i) = \det(AI_i) = \det(A_i)$$

Since A is invertible, $\det(A) \neq 0$ and hence

$$\det(I_i) = \frac{\det(A_i)}{\det(A)}$$

Expanding along the ith row to find the determinant of I_i gives

$$\det(I_i) = x_i \det(I) = x_i$$

where I is the $(n-1) \times (n-1)$ identity. Therefore,

$$x_i = \frac{\det(A_i)}{\det(A)}$$

39. a. Find the equation of the parabola in the form

$$Cy^2 + Dx + Ey + F = 0$$

that passes through the points $(-2, -2)$, $(3, 2)$, and $(4, -3)$.

b. Sketch the graph of the parabola.

40. a. Find the equation of the circle in the form

$$A(x^2 + y^2) + Dx + Ey + F = 0$$

that passes through the points $(-3, -3)$, $(-1, 2)$, and $(3, 0)$.

b. Sketch the graph of the circle.

41. a. Find the equation of the hyperbola in the form

$$Ax^2 + Cy^2 + Dx + Ey + F = 0$$

that passes through the points $(0, -4)$, $(0, 4)$, $(1, -2)$, and $(2, 3)$.

b. Sketch the graph of the hyperbola.

42. a. Find the equation of the ellipse in the form

$$Ax^2 + Cy^2 + Dx + Ey + F = 0$$

that passes through the points ([illegible], 2), $(-1, 3)$, ([illegible], -1), and $(4, 2)$.

b. Sketch the graph of the ellipse.

43. a. Find the equation of the ellipse in the form

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

that passes through the points $(-1, 0)$, $(0, 1)$, $(1, 0)$, $(2, 2)$, and $(3, 1)$.

b. Sketch the graph of the ellipse.

In Exercises 44–51, use Cramer's rule to solve the linear system.

44. $\begin{cases} 2x + 3y = 4 \\ 2x + 2y = 4 \end{cases}$

45. $\begin{cases} 5x - 5y = 7 \\ 2x - 3y = 6 \end{cases}$

46. $\begin{cases} 2x + 5y = 4 \\ 4x + y = 3 \end{cases}$

47. $\begin{cases} -9x - 4y = 3 \\ -7x + 5y = -10 \end{cases}$

48. $\begin{cases} -10x - 7y = -12 \\ 12x + 11y = 5 \end{cases}$

49. $\begin{cases} -x - 3y = 4 \\ -8x + 4y = 3 \end{cases}$

50. $\begin{cases} 2x + y - 4z = -8 \\ -4y + z = 3 \\ 4x - z = -8 \end{cases}$

51. $\begin{cases} 2x + 3y + 2z = -2 \\ -x - 3y - 8z = -2 \\ -3x + 2y - 7z = 2 \end{cases}$

52. An $n \times n$ matrix is *skew-symmetric* provided $A^t = -A$. Show that if A is skew-symmetric and n is an odd positive integer, then A is not invertible.

53. If A is a 3×3 matrix, show that $\det(A) = \det(A^t)$.

54. If A is an $n \times n$ upper triangular matrix, show that $\det(A) = \det(A^t)$.

1.7 ▶ Elementary Matrices and *LU* Factorization

In Sec. 1.2 we saw how the linear system $A\mathbf{x} = \mathbf{b}$ can be solved by using Gaussian elimination on the corresponding augmented matrix. Recall that the idea there was to use row operations to transform the coefficient matrix to row echelon form. The upper triangular form of the resulting matrix made it easy to find the solution by using back substitution. (See Example 1 of Sec. 1.2.) In a similar manner, if an augmented matrix is reduced to lower triangular form, then *forward substitution* can be used to find the solution of the corresponding linear system. For example, starting from the

Using forward substitution, we solve the system $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$, obtaining $y_1 = 1$, $y_2 = 5$, and $y_3 = 10$. Next we solve the linear system $U\mathbf{x} = \mathbf{y}$. That is,

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 10 \end{bmatrix}$$

Using back substitution, we obtain $x_3 = 2$, $x_2 = 1$, and $x_1 = 1$.

The following steps summarize the procedure for solving the linear system $A\mathbf{x} = \mathbf{b}$ when A admits an LU factorization.

1. Use Theorem 23 to write the linear system $A\mathbf{x} = \mathbf{b}$ as $L(U\mathbf{x}) = \mathbf{b}$.
2. Define the vector $\mathbf{y}$ by means of the equation $U\mathbf{x} = \mathbf{y}$.
3. Use forward substitution to solve the system $L\mathbf{y} = \mathbf{b}$ for $\mathbf{y}$.
4. Use back substitution to solve the system $U\mathbf{x} = \mathbf{y}$ for $\mathbf{x}$. Note that $\mathbf{x}$ is the solution to the original linear system.

PLU Factorization

We have seen that a matrix A has an LU factorization provided that it can be row-reduced without interchanging rows. We conclude this section by noting that when row interchanges are required to reduce A, a factorization is still possible. In this case the matrix A can be factored as $A = PLU$, where P is a *permutation matrix*, that is, a matrix that results from interchanging rows of the identity matrix. As an illustration, let

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

The matrix A can be reduced to

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

by means of the row operations $\mathcal{R}_1\colon\ R_1 \leftrightarrow R_3$, $\mathcal{R}_2\colon\ -R_1 + R_2 \longrightarrow R_2$, and $\mathcal{R}_3\colon\ -R_2 + R_3 \longrightarrow R_3$. The corresponding elementary matrices are given by

$$E_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Observe that the elementary matrix E_1 is a permutation matrix while E_2 and E_3 are lower triangular. Hence,

$$\begin{aligned} A &= E_1^{-1}\left(E_2^{-1}E_3^{-1}\right)U \\ &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 3 \\ 0 & 0 & -5 \end{bmatrix} \\ &= PLU \end{aligned}$$

Fact Summary

1. A row operation on a matrix A can be performed by multiplying A by an elementary matrix.
2. An elementary matrix is invertible, and the inverse is an elementary matrix.
3. An $n \times n$ matrix A is invertible if and only if it is the product of elementary matrices.
4. An $m \times n$ matrix A has an LU factorization if it can be reduced to an upper triangular matrix with no row interchanges.
5. If $A = LU$, then L is invertible.
6. An LU factorization of A provides an efficient method for solving $A\mathbf{x} = \mathbf{b}$.

Exercise Set 1.7

In Exercises 1–4:

a. Find the 3×3 elementary matrix E that performs the row operation.

b. Compute EA, where

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

1. $2R_1 + R_2 \longrightarrow R_2$

2. $R_1 \leftrightarrow R_2$

3. $-3R_2 + R_3 \longrightarrow R_3$

4. $-R_1 + R_3 \longrightarrow R_3$

In Exercises 5–10:

a. Find the elementary matrices required to reduce A to the identity.

b. Write A as the product of elementary matrices.

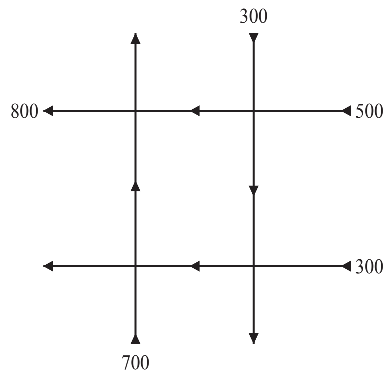
5. $A = \begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix}$

6. $A = \begin{bmatrix} -2 & 5 \\ 2 & 5 \end{bmatrix}$

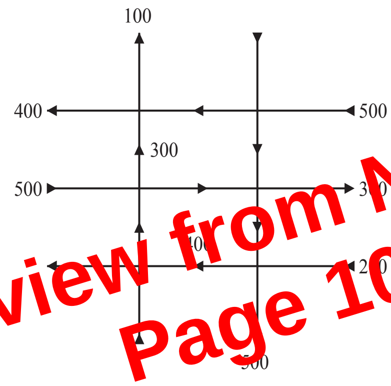
7. $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 5 & 3 \\ 1 & 2 & 0 \end{bmatrix}$

8. $A = \begin{bmatrix} -1 & 1 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 1 \end{bmatrix}$

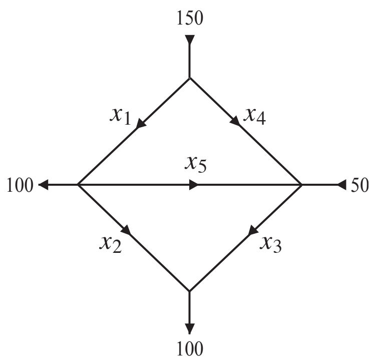
9. $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix}$



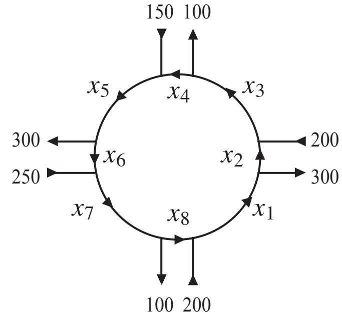
6. Find the traffic flow pattern for the network in the figure. Flow rates are in cars per hour. Give one specific solution.



7. Find the traffic flow pattern for the network in the figure. Flow rates are in cars per half-hour. What is the current status of the road labeled x_5?



8. Find the traffic flow pattern for the network in the figure. Flow rates are in cars per half-hour. What is the smallest possible value for x_8?



9. The table lists the number of milligrams of vitamin A, vitamin B, vitamin C, and niacin contained in 1 g of four different foods. A dietician wants to prepare a meal that provides 250 mg of vitamin A, 300 mg of vitamin B, 400 mg of vitamin C, and 70 mg of niacin. Determine how many grams of each food must be included, and describe any limitations on the quantities of each food that can be used.

	Group 1	Group 2	Group 3	Group 4
Vitamin A	20	30	40	10
Vitamin B	40	20	35	20
Vitamin C	50	40	10	30
Niacin	5	5	10	5

10. The table lists the amounts of sodium, potassium, carbohydrates, and fiber in a single serving of three food groups. Also listed are the daily recommended amounts based on a 2000-calorie diet. Is it possible to prepare a diet using the three food groups alone that meets the recommended amounts?

7. a. Explain why the matrix

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

is invertible.

b. Determine the maximum number of 1's that can be added to A such that the resulting matrix is invertible.

8. Show that if A is invertible, then A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$.

9. A matrix A is *skew-symmetric* provided $A^t = -A$.

a. Let A be an $n \times n$ matrix and define

$$B = A + A^t \qquad \text{and} \qquad C = A - A^t$$

Show that B is symmetric and C is skew-symmetric.

b. Show that every $n \times n$ matrix can be written as the sum of a symmetric and a skew-symmetric matrix.

10. Suppose $\mathbf{u}$ and $\mathbf{v}$ are solutions to the linear system $A\mathbf{x} = \mathbf{b}$. Show that if scalars α and β satisfy $\alpha + \beta = 1$, then $\alpha\mathbf{u} + \beta\mathbf{v}$ is also a solution to the linear system $A\mathbf{x} = \mathbf{b}$.

Chapter 1: Chapter Test

In Exercises 1–45, determine whether the statement is true or false.

1. A 2×2 linear system has one solution, no solutions, or infinitely many solutions.

2. A 3×3 linear system has no solutions, one solution, two solutions, three solutions, or infinitely many solutions.

3. If A and B are $n \times n$ matrices with no zero entries, then $AB \neq \mathbf{0}$.

4. Homogeneous linear systems always have at least one solution.

5. If A is an $n \times n$ matrix, then $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution if and only if the matrix A has an inverse.

6. If A and B are $n \times n$ matrices and $A\mathbf{x} = B\mathbf{x}$ for every $n \times 1$ matrix $\mathbf{x}$, then $A = B$.

7. If A, B, and C are invertible $n \times n$ matrices, then $(ABC)^{-1} = A^{-1}B^{-1}C^{-1}$.

8. If A is an invertible $n \times n$ matrix, then the linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

9. If A and B are $n \times n$ invertible matrices and $AB = BA$, then A commutes with B^{-1}.

10. If A and B commute, then $A^2B = BA^2$.

11. The matrix

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

does not have an inverse.

12. Interchanging two rows of a matrix changes the sign of its determinant.

13. Multiplying a row of a matrix by a nonzero constant results in the determinant being multiplied by the same nonzero constant.

14. If two rows of a matrix are equal, then the determinant of the matrix is 0.

15. Performing the operation $aR_i + R_j \to R_j$ on a matrix multiplies the determinant by the constant a.

16. If $A = \begin{bmatrix} 1 & 2 \\ 4 & 6 \end{bmatrix}$, then $A^2 - 7A = 2I$.

17. If A and B are invertible matrices, then $A + B$ is an invertible matrix.

18. If A and B are invertible matrices, then AB is an invertible matrix.

19. If A is an $n \times n$ matrix and A does not have an inverse, then the linear system $A\mathbf{x} = \mathbf{b}$ is inconsistent.

20. The linear system

$$\begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

is inconsistent.

21. The inverse of the matrix

$$\begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \quad \text{is} \quad \begin{bmatrix} 1 & 1 \\ -3 & 2 \end{bmatrix}$$

22. The matrix

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

does not have an inverse.

23. If the $n \times n$ matrix A is idempotent and invertible, then $A = I$.

24. If A and B commute, then A^t and B^t commute.

25. If A is an $n \times n$ matrix and $\det(A) = 3$, then $\det(A^t A) = 9$.

In Exercises 26–32, use the linear system

$$\begin{cases} 2x + 2y = 3 \\ x - y = 1 \end{cases}$$

26. The coefficient matrix is

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

27. The coefficient matrix A has determinant

$$\det(A) = 0$$

28. The linear system has a unique solution.

29. The only solution to the linear system is $x = -7/4$ and $y = -5/4$.

30. The inverse of the coefficient matrix A is

$$A^{-1} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} \end{bmatrix}$$

31. The linear system is equivalent to the matrix equation

$$\begin{bmatrix} 2 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

32. The solution to the system is given by the matrix equation

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

In Exercises 33–36, use the linear system

$$\begin{cases} x_1 + 2x_2 - 3x_3 = 1 \\ 2x_1 + 5x_2 - 8x_3 = 4 \\ -2x_1 - 4x_2 + 6x_3 = -2 \end{cases}$$

33. The determinant of the coefficient matrix is

$$\begin{vmatrix} 5 & -8 \\ -4 & 6 \end{vmatrix} + \begin{vmatrix} 2 & -8 \\ -2 & 6 \end{vmatrix} + \begin{vmatrix} 2 & 5 \\ -2 & -4 \end{vmatrix}$$

34. The determinant of the coefficient matrix is 0.

35. A solution to the linear system is $x_1 = -4$, $x_2 = 0$, and $x_3 = -1$.

36. The linear system has infinitely many solutions, and the general solution is given by x_3 is free, $x_2 = 2 + 2x_3$, and $x_1 = -3 - x_3$.

In Exercises 37–41, use the matrix

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

37. After the operation $R_1 \longleftrightarrow R_2$ is performed, the matrix becomes

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

38. After the operation $-2R_1 + R_3 \longrightarrow R_3$ is performed on the matrix found in Exercise 37, the matrix becomes

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

39. The matrix A is row equivalent to

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

3. Additive identity: The vector $\mathbf{0}$ satisfies $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$.

4. Additive inverse: For every vector $\mathbf{u}$, the vector $-\mathbf{u}$ satisfies $\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$.

5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

7. $c(d\mathbf{u}) = (cd)\mathbf{u}$

8. $(1)\mathbf{u} = \mathbf{u}$

By the associative property, the vector sum $\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_n$ can be computed unambiguously, without the need for parentheses. This will be important in Sec. 2.2.

EXAMPLE 3 Let

$$\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

Verify that the associative property holds for these three vectors. Also verify that for any scalars c and d, $c(d\mathbf{u}) = (cd)\mathbf{u}$.

Solution To verify the associative property, we have

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \left(\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) + \begin{bmatrix} 4 \\ -3 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -1+3 \end{bmatrix} + \begin{bmatrix} 4 \\ -3 \end{bmatrix}$$

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

and

$$\mathbf{u} + (\mathbf{v} + \mathbf{w}) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ -3 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix} + \begin{bmatrix} 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \end{bmatrix}$$

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

For the second verification, we have

$$c(d\mathbf{u}) = c\left(d\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = c\begin{bmatrix} d \\ -d \end{bmatrix} = \begin{bmatrix} cd \\ -cd \end{bmatrix} = cd\begin{bmatrix} 1 \\ -1 \end{bmatrix} = (cd)\mathbf{u}$$

The properties given in Theorem 1 can be used to establish other useful properties of vectors in $\mathbb{R}^n$. For example, if $\mathbf{u} \in \mathbb{R}^n$ and c is a scalar, then

$$0\mathbf{u} = 0\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0} \qquad \text{and} \qquad c\mathbf{0} = \mathbf{0}$$

29. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

$\mathbf{v}_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

30. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

31. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$

$\mathbf{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$

32. $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 8 \end{bmatrix}$

$\mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$

In Exercises 33–39, verify the indicated vector property of Theorem 1 for vectors in $\mathbb{R}^n$.

33. Property 2.

34. Property 3.

35. Property 4.

36. Property 5.

37. Property 6.

38. Property 7.

39. Property 8.

40. Prove that the zero vector in $\mathbb{R}^n$ is unique.

2.2 ▶ Linear Combinations

In three-dimensional Euclidean space $\mathbb{R}^3$ the *coordinate vectors* that define the three axes are the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Every vector in $\mathbb{R}^3$ can then be obtained from these three coordinate vectors, for example, the vector

$$\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Geometrically, the vector $\mathbf{v}$ is obtained by adding scalar multiples of the coordinate vectors, as shown in Fig. 1. The vectors $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ are not unique in this respect. For example, the vector $\mathbf{v}$ can also be written as a combination of the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

8. $\mathbf{v} = \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} -2 \\ -1 \\ -1 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} 3 \\ -1 \\ -3 \end{bmatrix}$

9. $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$

10. $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 5 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 10 \\ 3 \end{bmatrix}$

11. $\mathbf{v} = \begin{bmatrix} 3 \\ -17 \\ 17 \\ 7 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -3 \\ 4 \\ 1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 6 \\ -1 \\ 2 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$

12. $\mathbf{v} = \begin{bmatrix} 6 \\ 3 \\ 3 \\ 7 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ -3 \\ 1 \end{bmatrix}$

In Exercises 13–16, find all the ways that $\mathbf{v}$ can be written as a linear combination of the given vectors.

13. $\mathbf{v} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

14. $\mathbf{v} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$

15. $\mathbf{v} = \begin{bmatrix} 0 \\ -1 \\ -3 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} -2 \\ -1 \\ 2 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} -2 \\ -3 \\ -1 \end{bmatrix}$

$\mathbf{v}_4 = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$

16. $\mathbf{v} = \begin{bmatrix} -3 \\ -3 \\ 1 \end{bmatrix}$ $\mathbf{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

$\mathbf{v}_2 = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$ $\mathbf{v}_3 = \begin{bmatrix} 0 \\ -1 \\ -2 \end{bmatrix}$

$\mathbf{v}_4 = \begin{bmatrix} -3 \\ -1 \\ -2 \end{bmatrix}$

In Exercises 17–20, determine if the matrix M is a linear combination of the matrices M_1, M_2, and M_3.

17. $M = \begin{bmatrix} -2 & 4 \\ 4 & 0 \end{bmatrix}$

$M_1 = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$ $M_2 = \begin{bmatrix} -2 & 3 \\ 1 & 4 \end{bmatrix}$

$M_3 = \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix}$

$\mathbf{v}_k, c\mathbf{v}_k$, where c is a nonzero scalar. Show that $S_1 = S_2$.

36. Let S_1 be the set of all linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$, and S_2 be the set of all linear combinations of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_1 + \mathbf{v}_2$. Show that $S_1 = S_2$.

37. Suppose that $A\mathbf{x} = \mathbf{b}$ is a 3×3 linear system that is consistent. If there is a scalar c such that $\mathbf{A}_3 = c\mathbf{A}_1$, then show that the linear system has infinitely many solutions.

38. Suppose that $A\mathbf{x} = \mathbf{b}$ is a 3×3 linear system that is consistent. If $\mathbf{A}_3 = \mathbf{A}_1 + \mathbf{A}_2$, then show that the linear system has infinitely many solutions.

39. The equation

$$2y'' - 3y' + y = 0$$

is an example of a *differential equation*. Show that $y = f(x) = e^x$ and $y = g(x) = e^{\frac{1}{2}x}$ are solutions to the equation. Then show that any linear combination of $f(x)$ and $g(x)$ is another solution to the differential equation.

2.3 ▶ Linear Independence

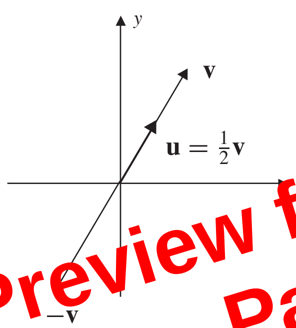


Figure 1

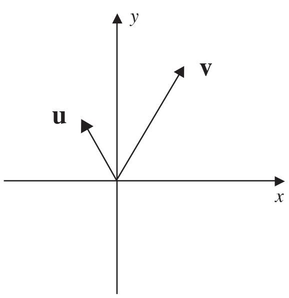


Figure 2

In Sec. 2.2 we saw that given a set S of vectors in $\mathbb{R}^n$, it is not always possible to express every vector in $\mathbb{R}^n$ as a linear combination of vectors from S. At the other extreme, there are infinitely many different subsets S such that the collection of all linear combinations of vectors from S is $\mathbb{R}^n$. For example, the collection of all linear combinations of the set of coordinate vectors $S = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is $\mathbb{R}^n$, but so is the collection of linear combinations of $T = \{\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{e}_1 + \mathbf{e}_2\}$. In this way S and T both *generate* $\mathbb{R}^n$. To characterize those minimal sets S that generate $\mathbb{R}^n$, we require the concept of *linear independence*. As motivation let two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$ lie on the same line, as shown in Fig. 1. Thus, there is a nonzero scalar c such that

$$\mathbf{u} = c\mathbf{v}$$

This condition can also be written as

$$\mathbf{u} - c\mathbf{v} = \mathbf{0}$$

In this case we say that the vectors $\mathbf{u}$ and $\mathbf{v}$ are *linearly dependent*. Evidently we have that two vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent provided that the zero vector is a nontrivial (not both scalars 0) linear combination of the vectors. On the other hand, the vectors shown in Fig. 2 are not linearly dependent. This concept is generalized to sets of vectors in $\mathbb{R}^n$.

DEFINITION 1 **Linearly Indpendent and Linearly Dependent** The set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ in $\mathbb{R}^n$ is **linearly independent** provided that the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$$

is the trivial solution $c_1 = c_2 = \cdots = c_m = 0$. If the above linear combination has a nontrivial solution, then the set S is called **linearly dependent**.

c. Show that the matrix

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

cannot be written as a linear combination of M_1, M_2, and M_3.

In Exercises 25 and 26, for the given matrix A determine if the linear system $A\mathbf{x} = \mathbf{b}$ has a unique solution.

25. $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 3 \\ 2 & 1 & 2 \end{bmatrix}$

26. $A = \begin{bmatrix} 3 & 2 & 4 \\ 1 & -1 & 4 \\ 0 & 2 & -4 \end{bmatrix}$

In Exercises 27–30, determine whether the set of polynomials is linearly independent or linearly dependent. A set of polynomials $S = \{p_1(x), p_2(x), \ldots, p_n(x)\}$ is linearly independent provided

$$c_1 p_1(x) + c_2 p_2(x) + \cdots + c_n p_n(x) = 0$$

for all x implies that

$$c_1 = c_2 = \cdots = c_n = 0$$

27. $p_1(x) = 1$ $p_2(x) = -2 + 4x^2$
$p_3(x) = 2x$ $p_4(x) = -12x + 8x^3$

28. $p_1(x) = 1$ $p_2(x) = x$
$p_3(x) = 5 + 2x - x^2$

29. $p_1(x) = 2$ $p_2(x) = x$ $p_3(x) = x^2$
$p_4(x) = 3x - 1$

30. $p_1(x) = x^3 - 2x^2 + 1$ $p_2(x) = 5x$
$p_3(x) = x^2 - 4$ $p_4(x) = x^3 + 2x$

In Exercises 31–34, show that the set of functions is linearly independent on the interval $[0, 1]$. A set of functions $S = \{f_1(x), f_2(x), \ldots, f_n(x)\}$ is linearly independent on the interval $[a, b]$ provided

$$c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) = 0$$

for all $x \in [a, b]$ implies that

$$c_1 = c_2 = \cdots = c_n = 0$$

31. $f_1(x) = \cos \pi x$ $f_2(x) = \sin \pi x$

32. $f_1(x) = e^x$ $f_2(x) = e^{-x}$
$f_3(x) = e^{2x}$

33. $f_1(x) = x$ $f_2(x) = x^2$ $f_3(x) = e^x$

34. $f_1(x) = x$ $f_2(x) = e^x$
$f_3(x) = \sin \pi x$

35. Verify that two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are linearly dependent if and only if one is a scalar multiple of the other.

36. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and

$$\mathbf{w_1} = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 \qquad \mathbf{w_2} = \mathbf{v}_2 + \mathbf{v}_3$$

and

$$\mathbf{w_3} = \mathbf{v}_3$$

Show that $T = \{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\}$ is linearly independent.

37. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and

$$\mathbf{w_1} = \mathbf{v}_1 + \mathbf{v}_2 \qquad \mathbf{w_2} = \mathbf{v}_2 - \mathbf{v}_3$$

and

$$\mathbf{w_3} = \mathbf{v}_2 + \mathbf{v}_3$$

Show that $T = \{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\}$ is linearly independent.

38. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and

$$\mathbf{w_1} = \mathbf{v}_2 \qquad \mathbf{w_2} = \mathbf{v}_1 + \mathbf{v}_3$$

and

$$\mathbf{w_3} = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$$

Determine whether the set $T = \{\mathbf{w_1}, \mathbf{w_2}, \mathbf{w_3}\}$ is linearly independent or linearly dependent.

39. Suppose that the set $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent. Show that if $\mathbf{v}_3$ cannot be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

40. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$.

a. Write $\mathbf{v}_1$ as a linear combination of the vectors in S in three different ways.

EXAMPLE 5 Let $V = \{(a, b) \mid a, b \in \mathbb{R}\}$. Let $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$. Define

$$(v_1, v_2) \oplus (w_1, w_2) = (v_1 + w_1 + 1, v_2 + w_2 + 1) \qquad \text{and}$$
$$c \odot (v_1, v_2) = (cv_1 + c - 1, cv_2 + c - 1)$$

Verify that V is a vector space.

Solution First observe that since the result of addition or scalar multiplication is an ordered pair, V is closed under addition and scalar multiplication. Since addition of real numbers is commutative and associative, axioms 2 and 3 hold for the $\oplus$ defined here. Now an element $\mathbf{w} \in V$ is the additive identity provided that for all $\mathbf{v} \in V$

$$\mathbf{v} \oplus \mathbf{w} = \mathbf{v} \qquad \text{or} \qquad (v_1 + w_1 + 1, v_2 + w_2 + 1) = (v_1, v_2)$$

Equating components gives

$$\begin{array}{lcl} v_1 + w_1 + 1 = v_1 & \text{and} & v_2 + w_2 + 1 = v_2 \qquad \text{so} \\ w_1 \quad = -1 & \text{and} & w_2 \quad = -1 \end{array}$$

This establishes the existence of an additive identity. Specifically, $\mathbf{0} = (-1, -1)$, so axiom 4 holds.

To show that each element $\mathbf{v}$ in V has an additive inverse, we must find a vector $\mathbf{w}$ such that

$$\mathbf{v} \oplus \mathbf{w} = \mathbf{0} = (-1, -1)$$

Since $\mathbf{v} \oplus \mathbf{w} = (v_1 + w_1 + 1, v_2 + w_2 + 1)$, this last equation requires that

$$\begin{array}{lcl} v_1 + w_1 + 1 = -1 & \text{and} & v_2 + w_2 + 1 = -1 \qquad \text{so that} \\ w_1 \quad = -v_1 - 2 & \text{and} & w_2 \quad = -v_2 - 2 \end{array}$$

Thus, for any element $\mathbf{v} = (v_1, v_2)$ in V, we have $-\mathbf{v} = (-v_1 - 2, -v_2 - 2)$. The remaining axioms all follow from the similar properties of the real numbers.

A **polynomial of degree** n is an expression of the form

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

where $a_0, \ldots, a_n$ are real numbers and $a_n \neq 0$. The degree of the **zero polynomial** is undefined since it can be written as $p(x) = 0x^n$ for any positive integer n. Polynomials comprise one of the most basic sets of functions and have many applications in mathematics.

EXAMPLE 6 **Vector Space of Polynomials** Let n be a fixed positive integer. Denote by $\mathcal{P}_n$ the set of all polynomials of degree n or less. Define addition by adding like terms. That is, if

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} + a_nx^n$$

29. Let V be the set of all real-valued functions defined on $\mathbb{R}$ with the standard operations that satisfy $f(0) = 1$. Determine whether V is a vector space.

30. Let V be the set of all real-valued functions defined on $\mathbb{R}$.

Define $f \oplus g$ by

$$(f \oplus g)(x) = f(x) + g(x)$$

and define $c \odot f$ by

$$(c \odot f)(x) = f(x + c)$$

Determine whether V is a vector space.

31. Let $f(x) = x^3$ defined on $\mathbb{R}$ and let

$$V = \{f(x + t) \mid t \in \mathbb{R}\}$$

Define

$$f(x + t_1) \oplus f(x + t_2) = f(x + t_1 + t_2)$$

$$c \odot f(x + t) = f(x + ct)$$

a. Determine the additive identity and additive inverses.

b. Show that V is a vector space.

3.2 ▶ Subspaces

Many interesting examples of vector spaces are subsets of a given vector space V that are vector spaces in their own right. For example, the xy plane in $\mathbb{R}^3$ given by

$$\left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \middle|\; x, y \in \mathbb{R} \right\}$$

is a subset of $\mathbb{R}^3$. It is also a vector space with the same standard componentwise operations defined on $\mathbb{R}^3$. Another example of a *subspace* of a vector space is given in Example 9 of Sec. 3.1. The determination as to whether a subset of a vector space is itself a vector space is simplified since many of the required properties are *inherited* from the parent space.

DEFINITION 1 **Subspace** A **subspace** W of a vector space V is a nonempty subset that is itself a vector space with respect to the inherited operations of vector addition and scalar multiplication on V.

The first requirement for a subset $W \subseteq V$ to be a subspace is that W be closed under the operations of V. For example, let V be the vector space $\mathbb{R}^2$ with the standard definitions of addition and scalar multiplication. Let $W \subseteq \mathbb{R}^2$ be the subset defined by

$$W = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \middle|\; a \in \mathbb{R} \right\}$$

Observe that the sum of any two vectors in W is another vector in W, since

$$\begin{bmatrix} a \\ 0 \end{bmatrix} \oplus \begin{bmatrix} b \\ 0 \end{bmatrix} = \begin{bmatrix} a + b \\ 0 \end{bmatrix}$$

In this way we say that W is closed under addition. The subset W is also closed under scalar multiplication since for any real number c,

$$c \odot \begin{bmatrix} a \\ 0 \end{bmatrix} = \begin{bmatrix} ca \\ 0 \end{bmatrix}$$

which is again in W.

On the other hand, the subset

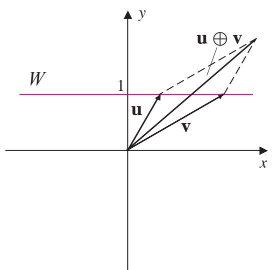
$$W = \left\{ \begin{bmatrix} a \\ 1 \end{bmatrix} \middle| a \in \mathbb{R} \right\}$$

is not closed under addition, since

$$\begin{bmatrix} a \\ 1 \end{bmatrix} \oplus \begin{bmatrix} b \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ 2 \end{bmatrix}$$

which is not in W. See Fig. 1. The subset W is also not closed under scalar multiplication since

$$c \odot \begin{bmatrix} a \\ 1 \end{bmatrix} = \begin{bmatrix} ca \\ c \end{bmatrix}$$



W is not a subspace of *V*

Figure 1

which is not in W for all values of $c \neq 1$.

Now let us suppose that a nonempty subset W is closed under both of the operations on V. To determine whether W is a subspace, we must show that each of the remaining vector space axioms hold. Fortunately, our task is simplified as most of these properties are inherited from the vector space V. For example, to show that the commutative property holds in W, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in W. Since $\mathbf{u}$ and $\mathbf{v}$ are also in V, then

$$\mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u}$$

Similarly, any three vectors in W satisfy the associative property, as this property is also inherited from V. To show that W contains the zero vector, let $\mathbf{w}$ be any vector in W. Since W is closed under scalar multiplication, $0 \odot \mathbf{w} \in W$. Now, by Theorem 2 of Sec. 3.1, we have $0 \odot \mathbf{w} = \mathbf{0}$. Thus, $\mathbf{0} \in W$. Similarly, for any $\mathbf{w} \in W$,

$$(-1) \odot \mathbf{w} = -\mathbf{w}$$

is also in W. All the other vector space properties, axioms 7 through 10, are inherited from V. This shows that W is a subspace of V. Conversely, if W is a subspace of V, then it is necessarily closed under addition and scalar multiplication. This proves Theorem 3.

THEOREM 3 Let W be a nonempty subset of the vector space V. Then W is a subspace of V if and only if W is closed under addition and scalar multiplication.

By Theorem 3, the first of the examples above with

$$W = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} \middle| a \in \mathbb{R} \right\}$$

We now consider what happens when subspaces are combined. In particular, let W_1 and W_2 be subspaces of a vector space V. Then the intersection $W_1 \cap W_2$ is also a subspace of V. To show this, let $\mathbf{u}$ and $\mathbf{v}$ be elements of $W_1 \cap W_2$ and let c be a scalar. Since W_1 and W_2 are both subspaces, then by Theorem 4, $\mathbf{u} \oplus (c \odot \mathbf{v})$ is in W_1 and is in W_2, and hence is in the intersection. Applying Theorem 4 again, we have that $W_1 \cap W_2$ is a subspace.

The extension to an arbitrary number of subspaces is stated in Theorem 5.

THEOREM 5 The intersection of any collection of subspaces of a vector space is a subspace of the vector space.

Example 6 shows that the union of two subspaces need not be a subspace.

EXAMPLE 6 Let W_1 and W_2 be the subspaces of $\mathbb{R}^2$ with the standard operations given by

$$W_1 = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \middle| \; x \in \mathbb{R} \right\} \quad \text{and} \quad W_2 = \left\{ \begin{bmatrix} 0 \\ y \end{bmatrix} \middle| \; y \in \mathbb{R} \right\}$$

Show that $W_1 \cup W_2$ is not a subspace.

Solution The subspaces W_1 and W_2 consist of all vectors that lie on the x axis and the y axis, respectively. Their union is the collection of all vectors that lie on either axis and is given by

$$W_1 \cup W_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \middle| \; x = 0 \text{ or } y = 0 \right\}$$

This set is not closed under addition since

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which is not in $W_1 \cup W_2$, as shown in Fig. 2.

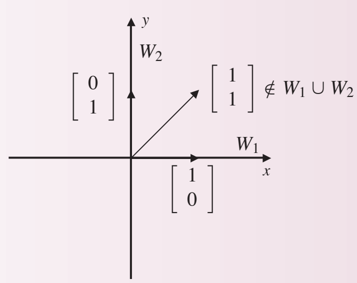


Figure 2

spans $\mathbb{R}^3$. These vectors are also linearly independent. To see this, observe that the matrix

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

whose column vectors are the vectors of S, is row equivalent to the 3×3 identity matrix, as seen in the solution to Example 9. [Another way of showing that S is linearly independent is to observe that $\det(A) = 1 \neq 0$.] Consequently, by Theorem 7 of Sec. 2.3, we have that every vector in $\mathbb{R}^3$ can be written in only one way as a linear combination of the vectors of S.

On the other hand, the span of the set of vectors

$$S' = \{\mathbf{v}'_1,\ \mathbf{v}'_2,\ \mathbf{v}'_3\} = \left\{ \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \\ 5 \end{bmatrix} \right\}$$

of Example 10 is a plane passing through the origin. Hence, not every vector in $\mathbb{R}^3$ can be written as a linear combination of the vectors in S'. As we expect, these vectors are linearly dependent since

$$\det\left(\begin{bmatrix} -1 & 4 & -6 \\ 2 & 1 & 3 \\ 1 & -3 & 5 \end{bmatrix} \right) = 0$$

In particular, $\mathbf{v}'_3 = 2\mathbf{v}'_1 - \mathbf{v}'_2$. The vectors $\mathbf{v}'_1$ and $\mathbf{v}'_2$ are linearly independent vectors which span the plane shown in Fig. 4, but not $\mathbb{R}^3$.

To pursue these notions a bit further, there are many sets of vectors which span $\mathbb{R}^3$. For example, the set

$$B = \{\mathbf{e}_1,\ \mathbf{e}_2,\ \mathbf{e}_3, \mathbf{v}\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$

spans $\mathbb{R}^3$, but by Theorem 3 of Sec. 2.3 must necessarily be linearly dependent. The ideal case, in terms of minimizing the number of vectors, is illustrated in Example 9 where the three linearly independent vectors of S span $\mathbb{R}^3$. In Sec. 3.3 we will see that S is a basis for $\mathbb{R}^3$, and that every basis for $\mathbb{R}^3$ consists of exactly three linearly independent vectors.

EXAMPLE 11 Show that the set of matrices

$$S = \left\{ \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

does not span $M_{2\times 2}$. Describe $\mathbf{span}(S)$.

Hence, the linear system has the unique solution $c_1 = 2a - c$, $c_2 = -a + c$, and $c_3 = -4a + b + 2c$, for all a, b, and c. Therefore, $\mathbf{span}(S) = \mathcal{P}_2$.

The Null Space and Column Space of a Matrix

Two special subspaces associated with every matrix A are the *null space* and *column space* of the matrix.

DEFINITION 4 **Null Space and Column Space** Let A be an $m \times n$ matrix.

1. The **null space** of A, denoted by $N(A)$, is the set of all vectors in $\mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{0}$.
2. The **column space** of A, denoted by $\mathbf{col}(A)$, is the set of all linear combinations of the column vectors of A.

Observe that $N(A)$ is a subset of $\mathbb{R}^n$ and $\mathbf{col}(A)$ is a subset of $\mathbb{R}^m$. Moreover, by Proposition 1, $\mathbf{col}(A)$ is a subspace of $\mathbb{R}^m$. Using this terminology, we give a restatement of Theorem 2 of Sec. 2.2.

THEOREM 6 Let A be an $m \times n$ matrix. The linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.

EXAMPLE 13 Let

$$A = \begin{bmatrix} 1 & -1 & -2 \\ -1 & 2 & 3 \\ 2 & -2 & -2 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

a. Determine whether $\mathbf{b}$ is in $\mathbf{col}(A)$.

b. Find $N(A)$.

Solution

a. By Theorem 6, the vector $\mathbf{b}$ is in $\mathbf{col}(A)$ if and only if there is a vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$. The corresponding augmented matrix is given by

$$\left[\begin{array}{rrr|r} 1 & -1 & -2 & 3 \\ -1 & 2 & 3 & 1 \\ 2 & -2 & -2 & -2 \end{array}\right] \quad \text{which reduces to} \quad \left[\begin{array}{rrr|r} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & -4 \end{array}\right]$$

This leads to the concept of a basis for an abstract vector space. As a first step, we generalize the concept of linear independence to abstract vector spaces introduced in Sec. 3.1.

DEFINITION 1 **Linear Independence and Linear Dependence** The set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ in a vector space V is called **linearly independent** provided that the only solution to the equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$$

is the trivial solution $c_1 = c_2 = \cdots = c_m = 0$. If the equation has a nontrivial solution, then the set S is called **linearly dependent**.

EXAMPLE 1 Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} -3 \\ 4 \\ 7 \end{bmatrix}$$

and let $W = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

a. Show that $\mathbf{v}_3$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

b. Show that $\text{span}\{\mathbf{v}_1, \mathbf{v}_2\} = W$.

c. Show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent.

Solution **a.** To solve the vector equation

$$c_1 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \\ 7 \end{bmatrix}$$

we row-reduce the corresponding augmented matrix for the linear system to obtain

$$\left[\begin{array}{rr|r} 1 & 0 & -3 \\ 0 & 2 & 4 \\ -1 & 2 & 7 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & 0 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

The solution to the vector equation above is $c_1 = -3$ and $c_2 = 2$, therefore

$$\mathbf{v}_3 = -3\mathbf{v}_1 + 2\mathbf{v}_2$$

Notice that the vector $\mathbf{v}_3$ lies in the plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, as shown in Fig. 1.

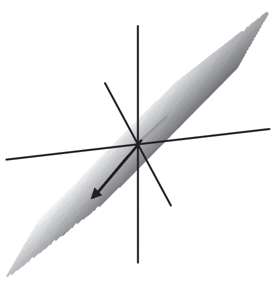


Figure 1

b. From part (a) an element of $W = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \mid c_1, c_2, c_3 \in \mathbb{R}\}$ can be written in the form

$$\begin{aligned} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(-3\mathbf{v}_1 + 2\mathbf{v}_2) \\ &= (c_1 - 3c_3)\mathbf{v}_1 + (c_2 + 2c_3)\mathbf{v}_2 \end{aligned}$$

Solution By using the methods presented in Chap. 2 it can be shown that the the column vectors of the matrix A are linearly independent. Since the column vectors of B consist of a set of five vectors in $\mathbb{R}^4$, by Theorem 3 of Sec. 2.3, the vectors are linearly dependent. In addition, the first four column vectors of B are the same as the linearly independent column vectors of A, hence by Theorem 5 of Sec. 2.3 the fifth column vector of B must be a linear combination of the other four vectors. Finally by Theorem 8, we know that $\mathbf{col}(A) = \mathbf{col}(B)$.

As a consequence of Theorem 8, a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ such that $V = \mathbf{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is minimal, in the sense of the number of spanning vectors, when they are linearly independent. We also saw in Chap. 2 that when a vector in $\mathbb{R}^n$ can be written as a linear combination of vectors from a linearly independent set, then the representation is unique. The same result holds for abstract vector spaces.

THEOREM 9 If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is a linearly independent set of vectors in a vector space V, then every vector in $\mathbf{span}(B)$ can be written uniquely as a linear combination of vectors from B.

Motivated by these ideas, we now define what we mean by a basis of a vector space.

DEFINITION 2 **Basis for a Vector Space** A subset B of a vector space V is a **basis** for V provided that

1. B is a linearly independent set of vectors in V
2. $\mathbf{span}(B) = V$

As an example, the set of coordinate vectors

$$S = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$$

spans $\mathbb{R}^n$ and is linearly independent, so that S is a basis for $\mathbb{R}^n$. This particular basis is called the **standard basis** for $\mathbb{R}^n$. In Example 3 we give a basis for $\mathbb{R}^3$, which is not the standard basis.

EXAMPLE 3 Show that the set

$$B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$$

is a basis for $\mathbb{R}^3$.

Dimension

We have already seen in Theorem 3 of Sec. 2.3 that any set of m vectors from $\mathbb{R}^n$, with $m > n$, must necessarily be linearly dependent. Hence, any basis of $\mathbb{R}^n$ contains at most n vectors. It can also be shown that any linearly independent set of m vectors, with $m < n$, does not span $\mathbb{R}^n$. For example, as we have already seen, two linearly independent vectors in $\mathbb{R}^3$ span a plane. Hence, any basis of $\mathbb{R}^n$ must contain exactly n vectors. The number n, an invariant of $\mathbb{R}^n$, is called the *dimension* of $\mathbb{R}^n$. Theorem 11 shows that this holds for abstract vector spaces.

THEOREM 11 If a vector space V has a basis with n vectors, then every basis has n vectors.

Proof Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V, and let $T = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\}$ be a subset of V with $m > n$. We claim that T is linearly dependent. To establish this result, observe that since B is a basis, then every vector in T can be written as a linear combination of the vectors from B. That is,

$$\begin{aligned}
\mathbf{u}_1 &= \lambda_{11}\mathbf{v}_1 + \lambda_{12}\mathbf{v}_2 + \cdots + \lambda_{1n}\mathbf{v}_n \\
\mathbf{u}_2 &= \lambda_{21}\mathbf{v}_1 + \lambda_{22}\mathbf{v}_2 + \cdots + \lambda_{2n}\mathbf{v}_n \\
&\vdots \\
\mathbf{u}_m &= \lambda_{m1}\mathbf{v}_1 + \lambda_{m2}\mathbf{v}_2 + \cdots + \lambda_{mn}\mathbf{v}_n
\end{aligned}$$

Now consider the equation

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = \mathbf{0}$$

Using the equations above, we can write this last equation in terms of the basis vectors. After collecting like terms, we obtain

$$\begin{aligned}
&(c_1\lambda_{11} + c_2\lambda_{21} + \cdots + c_m\lambda_{m1})\mathbf{v}_1 \\
&\quad + (c_1\lambda_{12} + c_2\lambda_{22} + \cdots + c_m\lambda_{m2})\mathbf{v}_2 \\
&\quad \vdots \\
&\quad + (c_1\lambda_{1n} + c_2\lambda_{2n} + \cdots + c_m\lambda_{mn})\mathbf{v}_n = \mathbf{0}
\end{aligned}$$

Since B is a basis, it is linearly independent, hence

$$\begin{aligned}
c_1\lambda_{11} + c_2\lambda_{21} + \cdots + c_m\lambda_{m1} &= 0 \\
c_1\lambda_{12} + c_2\lambda_{22} + \cdots + c_m\lambda_{m2} &= 0 \\
&\vdots \\
c_1\lambda_{1n} + c_2\lambda_{2n} + \cdots + c_m\lambda_{mn} &= 0
\end{aligned}$$

This last linear system is not square with n equations in the m variables $c_1, \ldots, c_m$. Since $m > n$, by Theorem 3 of Sec. 2.3 the linear system has a nontrivial solution, and hence T is linearly dependent.

The details of these observations are made clearer by considering a specific example. Let

$$S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} \right\}$$

We begin by considering the equation

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + c_4 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + c_5 \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

To solve this system, we reduce the corresponding augmented matrix to reduced echelon form. That is,

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

In the general solution, the variables c_1, c_2, and c_3 are the dependent variables corresponding to the leading ones in the reduced matrix, while c_4 and c_5 are free. Thus, the solution is given by

$$S = \{(-s, -s - t, -t, s, t) \mid s, t \in \mathbb{R}\}$$

Now to find a basis for $\mathbf{span}(S)$, we substitute these values into the original vector equation to obtain

$$-s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + (-s - t) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-t) \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + s \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

We claim that *each of the vectors corresponding to a free variable is a linear combination of the others*. To establish the claim in this case, let $s = 1$ and $t = 0$. The above vector equation now becomes

$$-\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

that is,

$$-\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_4 = \mathbf{0}$$

Thus, $\mathbf{v}_4$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Also, to see that $\mathbf{v}_5$ is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, we let $s = 0$ and $t = 1$.

In light of Theorem 8 we eliminate $\mathbf{v}_4$ and $\mathbf{v}_5$ from S to obtain $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Observe that S' is linearly independent since each of these vectors corresponds to a column with a leading 1. Thus, the equation

$$c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

has only the trivial solution.

23. Let

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

be the standard ordered basis for $\mathbb{R}^2$ and let

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right\}$$

be a second ordered basis.

a. Find $[I]_S^B$

b. Find the coordinates of

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix}$$

relative to the ordered basis B.

c. Draw the rectangle in the plane with vertices $(1, 2)$, $(1, 4)$, $(4, 1)$, and $(4, 4)$.

d. Draw the polygon in the plane with vertices given by the coordinates found in part (b).

24. Fix a real number θ and define the transition matrix from the standard ordered basis S on $\mathbb{R}^2$ to a second ordered basis B by

$$[I]_S^B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

a. If $[\mathbf{v}]_S = \begin{bmatrix} x \\ y \end{bmatrix}$, then find $[\mathbf{v}]_B$.

b. Draw the rectangle in the plane with vertices

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

c. Let $\theta = \frac{\pi}{2}$. Draw the rectangle in the plane with vertices the coordinates of the vectors, given in part (b), relative to the ordered basis B.

25. Suppose that $B_1 = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $B_2 = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are ordered bases for a vector space V such that $\mathbf{u}_1 = -\mathbf{v}_1 + 2\mathbf{v}_2$, $\mathbf{u}_2 = -\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3$, and $\mathbf{u}_3 = -\mathbf{v}_2 + \mathbf{v}_3$.

a. Find the transition matrix $[I]_{B_1}^{B_2}$

b. Find $[2\mathbf{u}_1 - 3\mathbf{u}_2 + \mathbf{u}_3]_{B_2}$

3.5 Application: Differential Equations

Differential equations arise naturally in virtually every branch of science and technology. They are used extensively by scientists and engineers to solve problems concerning growth, motion, vibrations, forces, or any problem involving the rates of change of variable quantities. Not surprisingly, mathematicians have devoted a great deal of effort to developing methods for solving differential equations. As it turns out, linear algebra is highly useful to these efforts. However, linear algebra also makes it possible to attain a deeper understanding of the theoretical foundations of these equations and their solutions. In this section and in Sec. 5.3 we give a brief introduction to the connection between linear algebra and differential equations.

As a first step, let y be a function of a single variable x. An equation that involves $x, y, y', y'', \ldots, y^{(n)}$, where n is a fixed positive integer, is called an **ordinary differential equation of order** n. We will henceforth drop the qualifier *ordinary* since none of the equations we investigate will involve partial derivatives. Also, for obvious reasons we will narrow the scope of our discussion and consider only equations of a certain type.

second derivatives $y' = re^{rx}$ and $y'' = r^2e^{rx}$, we see that $y = e^{rx}$ is a solution of the second-order equation if and only if

$$r^2e^{rx} + are^{rx} + be^{rx} = 0$$

that is,

$$e^{rx}(r^2 + ar + b) = 0$$

Since $e^{rx} > 0$ for every choice of r and x, we know e^{rx} is a solution of $y'' + ay' + by = 0$ if and only if

$$r^2 + ar + b = 0$$

This equation is called the **auxiliary equation**. As this equation is quadratic there are three possibilities for the roots r_1 and r_2. This in turn yields three possible variations for the solution of the differential equation. The auxiliary equation can have two distinct real roots, one real root, or two distinct complex roots. These cases are considered in order.

Case 1 The roots r_1 and r_2 are real and distinct.
In this case there are two solutions, given by

$$y_1(x) = e^{r_1x} \qquad \text{and} \qquad y_2(x) = e^{r_2x}$$

EXAMPLE 1 Find two distinct solutions to the differential equation $y'' - 3y' + 2y = 0$.

Solution Let $y = e^{rx}$. Since the auxiliary equation $r^2 - 3r + 2 = (r - 1)(r - 2) = 0$ has the distinct real roots $r_1 = 1$ and $r_2 = 2$, two distinct solutions for the differential equation are

$$y_1(x) = e^x \qquad \text{and} \qquad y_2(x) = e^{2x}$$

Case 2 There is one repeated root r. Although the auxiliary equation has only one root, there are still two distinct solutions, given by

$$y_1(x) = e^{rx} \qquad \text{and} \qquad y_2(x) = xe^{rx}$$

EXAMPLE 2 Find two distinct solutions to the differential equation $y'' - 2y' + y = 0$.

Solution Let $y = e^{rx}$. Since the auxiliary equation $r^2 - 2r + 1 = (r - 1)^2 = 0$ has the repeated root $r = 1$, two distinct solutions of the differential equation are

$$y_1(x) = e^x \qquad \text{and} \qquad y_2(x) = xe^x$$

Confirming Pages

b. Assume there exists a particular solution to the nonhomogeneous equation of the form

$$y_p(x) = A\cos 2x + B\sin 2x$$

Substitute $y_p(x)$ into the differential equation to find conditions on the coefficients A and B.

c. Verify that $y_c(x) + y_p(x)$ is a solution to the differential equation.

9. Let w be the weight of an object attached to a spring, g the constant acceleration due to gravity of 32 ft/s^2, k the spring constant, and d the distance in feet that the spring is stretched by the weight. Then the mass of the object is $m = \frac{w}{g}$ and $k = \frac{w}{d}$. Suppose that a 2-lb weight stretches a spring by 6-in. Find the equation of the motion of the weight if the object is pulled down by 3-in and then released. Notice that this system is *undamped*; that is, the damping coefficient is 0.

10. Suppose an 8-lb object is attached to a spring with a spring constant of 4 lb/ft and that the damping force on the system is twice the velocity. Find the equation of the motion if the object is pulled down 1-ft and given an upward velocity of 2 ft/s.

Review Exercises for Chapter 3

1. Determine for which values of k the vectors

$$\begin{bmatrix} 1 \\ -2 \\ 0 \\ 2 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \\ 4 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 3 \\ 4 \\ k \end{bmatrix}$$

form a basis for $\mathbb{R}^4$.

2. For which values of a, b, c, d, e, and f are the vectors

$$\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} b \\ c \\ 0 \end{bmatrix} \quad \begin{bmatrix} d \\ e \\ f \end{bmatrix}$$

a basis for $\mathbb{R}^3$?

3. Let

$$S = \left\{ \begin{bmatrix} a-b & a \\ b+c & a-c \end{bmatrix} \middle| \, a, b, c \in \mathbb{R} \right\}$$

a. Show that S is a subspace of $M_{2\times 2}$.

b. Is $\begin{bmatrix} 5 & 3 \\ -2 & 3 \end{bmatrix}$ in S?

c. Find a basis B for S.

d. Give a 2×2 matrix that is not in S.

4. Let $S = \{p(x) = a + bx + cx^2 \mid a + b + c = 0\}$.

a. Show that S is a subspace of $\mathcal{P}_2$.

b. Find a basis for S. Specify the dimension of S.

5. Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for a vector space V.

a. Determine whether the set $T = \{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3\}$ is a basis for V.

b. Determine whether the set $W = \{-\mathbf{v}_2 + \mathbf{v}_3, 3\mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_1 - \mathbf{v}_2 + 2\mathbf{v}_3\}$ is a basis for V.

6. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{v}_3 = \begin{bmatrix} 4 \\ -7 \\ 3 \\ 3 \end{bmatrix}$$

a. Explain why the set S is not a basis for $\mathbb{R}^4$.

b. Show that $\mathbf{v}_3$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

c. Find the dimension of the span of the set S.

d. Find a basis B for $\mathbb{R}^4$ that contains the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

24. If V is a vector space of dimension n and H is a subspace of dimension n, then $H = V$.

25. If B_1 and B_2 are bases for the vector space V, then the transition matrix from B_1 to B_2 is the inverse of the transition matrix from B_2 to B_1.

In Exercises 26–29, use the bases of $\mathbb{R}^2$

$$B_1 = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \end{bmatrix} \right\}$$

and

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\}$$

26. The coordinates of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, relative to B_1, are $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

27. The coordinates of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ relative to B_2 are $\frac{1}{4}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

28. The transition matrix from B_1 to B_2 is

$$[I]_{B_1}^{B_2} = \frac{1}{2}\begin{bmatrix} -1 & 3 \\ 1 & -1 \end{bmatrix}$$

29. The transition matrix from B_2 to B_1 is

$$[I]_{B_2}^{B_1} = \begin{bmatrix} 1 & 3 \\ 1 & 1 \end{bmatrix}$$

In Exercises 30–35, use the bases of $\mathcal{P}_3$,

$$B_1 = \{1, x, x^2, x^3\}$$

and

$$B_2 = \{x, x^2, 1, x^3\}$$

30. $[x^3 + 2x^2 - x]_{B_1} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$

31. $[x^3 + 2x^2 - x]_{B_1} = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}$

32. $[x^3 + 2x^2 - x]_{B_2} = \begin{bmatrix} 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}$

33. $[x^3 + 2x^2 - x]_{B_2} = \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$

34. $[(1 + x)^2 - 3(x^2 + x - 1) + x^3]_{B_2}$

$$= \begin{bmatrix} 4 \\ -1 \\ -2 \\ 1 \end{bmatrix}$$

35. The transition matrix from B_1 to B_2 is

$$[I]_{B_1}^{B_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$T\left(\begin{bmatrix} 7 \\ -1 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 3 & 2 \end{bmatrix}\begin{bmatrix} 7 \\ -1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Later in this chapter, in Sec. 4.4, we show that every linear transformation between finite dimensional vector spaces can be represented by a matrix. In Examples 1 and 2, we have discussed some of the algebraic properties of linear transformations. In Example 3 we consider the action of a linear transformation from a geometric perspective.

EXAMPLE 3 Define a linear transformation $T\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

a. Discuss the action of T on a vector in $\mathbb{R}^3$, and give a geometric interpretation of the equation

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) + T\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right)$$

b. Find the image of the set

$$S_1 = \left\{ t\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \,\middle|\, t \in \mathbb{R} \right\}$$

c. Find the image of the set

$$S_2 = \left\{ \begin{bmatrix} x \\ y \\ 3 \end{bmatrix} \,\middle|\, x, y \in \mathbb{R} \right\}$$

d. Describe the set

$$S_3 = \left\{ \begin{bmatrix} x \\ 0 \\ z \end{bmatrix} \,\middle|\, x, z \in \mathbb{R} \right\}$$

and find its image.

Solution **a.** The linear transformation T gives the *projection*, or shadow, of a vector in 3-space to its image in the xy plane. Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Operations with Linear Transformations

Linear transformations can be combined by using a natural addition and scalar multiplication to produce new linear transformations. For example, let $S, T: \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + y \\ -x \end{bmatrix} \qquad \text{and} \qquad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - y \\ x + 3y \end{bmatrix}$$

We then define

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}) \qquad \text{and} \qquad (cS)(\mathbf{v}) = c(S(\mathbf{v}))$$

To illustrate this definition, let $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$; then

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v}) = \begin{bmatrix} 2 + (-1) \\ -2 \end{bmatrix} + \begin{bmatrix} 2(2) - (-1) \\ 2 + 3(-1) \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$$

For scalar multiplication let $c = 3$. Then

$$(3T)(\mathbf{v}) = 3T(\mathbf{v}) = 3\begin{bmatrix} 5 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -3 \end{bmatrix}$$

In Theorem 1 we show that these operations on linear transformations produce linear transformations.

THEOREM 1 Let V and W be vector spaces and let $S, T: V \to W$ be linear transformations. The function $S + T$ defined by

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$$

is a linear transformation from V into W. If c is any scalar, the function cS defined by

$$(cS)(\mathbf{v}) = cS(\mathbf{v})$$

is a linear transformation from V into W.

Proof Let $\mathbf{u}, \mathbf{v} \in V$ and let d be any scalar. Then

$$\begin{aligned}
(S + T)(d\mathbf{u} + \mathbf{v}) &= S(d\mathbf{u} + \mathbf{v}) + T(d\mathbf{u} + \mathbf{v}) \\
&= S(d\mathbf{u}) + S(\mathbf{v}) + T(d\mathbf{u}) + T(\mathbf{v}) \\
&= dS(\mathbf{u}) + S(\mathbf{v}) + dT(\mathbf{u}) + T(\mathbf{v}) \\
&= d(S(\mathbf{u}) + T(\mathbf{u})) + S(\mathbf{v}) + T(\mathbf{v}) \\
&= d(S + T)(\mathbf{u}) + (S + T)(\mathbf{v})
\end{aligned}$$

Fact Summary

Let V, W, and Z be vector spaces and S and T functions from V into W.

1. The function T is a linear transformation provided that for all $\mathbf{u}, \mathbf{v}$ in V and all scalars c, $T(c\mathbf{u} + \mathbf{v}) = cT(\mathbf{u}) + T(\mathbf{v})$.
2. If A is an $m \times n$ matrix and T is defined by $T(\mathbf{x}) = A\mathbf{x}$, then T is a linear transformation from $\mathbb{R}^n$ into $\mathbb{R}^m$.
3. If T is a linear transformation, then the zero vector in V is mapped to the zero vector in W, that is, $T(\mathbf{0}) = \mathbf{0}$.
4. If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an ordered basis for V and $W = \mathbb{R}^n$, then the coordinate mapping $T(\mathbf{v}) = [\mathbf{v}]_B$ is a linear transformation.
5. If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors in V and T is a linear transformation, then
$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n)$$
for all scalars $c_1, \ldots, c_n$.
6. If S and T are linear transformations and c is a scalar, then $S + T$ and cT are linear transformations.
7. If $T: V \longrightarrow W$ is a linear transformation and $L : W \longrightarrow Z$ is a linear transformation, then $L \circ T: V \longrightarrow Z$ is a linear transformation.

Exercise Set 4.1

In Exercises 1–6, determine whether the function $T: \mathbb{R}^2 \to \mathbb{R}^2$ is a linear transformation.

1. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}$

2. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + y \\ x - y + 2 \end{bmatrix}$

3. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y^2 \end{bmatrix}$

4. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - y \\ x + 3y \end{bmatrix}$

5. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 0 \end{bmatrix}$

6. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} \frac{x+y}{2} \\ \frac{x+y}{2} \end{bmatrix}$

In Exercises 7–16, determine whether the function is a linear transformation between vector spaces.

7. $T: \mathbb{R} \to \mathbb{R}$, $T(x) = x^2$

8. $T: \mathbb{R} \to \mathbb{R}$, $T(x) = -2x$

9. $T: \mathbb{R}^2 \to \mathbb{R}$, $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x^2 + y^2$

10. $T: \mathbb{R}^3 \to \mathbb{R}^2$,
$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$$

11. $T: \mathbb{R}^3 \to \mathbb{R}^3$,
$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + y - z \\ 2xy \\ x + z + 1 \end{bmatrix}$$

b. Let $\mathbf{w} = \begin{bmatrix} 7 \\ -6 \\ -9 \end{bmatrix}$. Determine whether there is a vector $\mathbf{v}$ in $\mathbb{R}^3$ such that $T(\mathbf{v}) = \mathbf{w}$.

34. Define $T\colon \mathcal{P}_2 \to \mathcal{P}_2$ by

$$T(p(x)) = p'(x) - p(0)$$

a. Find all vectors that are mapped to $\mathbf{0}$.

b. Find two polynomials $p(x)$ and $q(x)$ such that $T(p(x)) = T(q(x)) = 6x - 3$.

c. Is T a linear operator?

35. Suppose $T_1\colon V \to \mathbb{R}$ and $T_2\colon V \to \mathbb{R}$ are linear transformations. Define $T\colon V \to \mathbb{R}^2$ by

$$T(\mathbf{v}) = \begin{bmatrix} T_1(\mathbf{v}) \\ T_2(\mathbf{v}) \end{bmatrix}$$

Show that T is a linear transformation.

36. Define $T\colon M_{n\times n} \to \mathbb{R}$ by $T(A) = \operatorname{tr}(A)$. Show that T is a linear transformation.

37. Suppose that B is a fixed $n \times n$ matrix. Define $T\colon M_{n\times n} \to M_{n\times n}$ by $T(A) = AB - BA$. Show that T is a linear operator.

38. Define $T\colon \mathbb{R} \to \mathbb{R}$ by $T(x) = mx + b$. Determine when T is a linear operator.

39. Define $T\colon C^{(0)}[0, 1] \to \mathbb{R}$ by

$$T(f) = \int_0^1 f(x)\,dx$$

for each function f in $C^{(0)}[0, 1]$.

a. Show that T is a linear operator.

b. Find $T(2x^2 - x + 3)$.

40. Suppose that $T\colon V \to W$ is a linear transformation and $T(\mathbf{u}) = \mathbf{w}$. If $T(\mathbf{v}) = \mathbf{0}$, then find $T(\mathbf{u} + \mathbf{v})$.

41. Suppose that $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and $\{\mathbf{v}, \mathbf{w}\}$ is a linearly independent subset of $\mathbb{R}^n$. If $\{T(\mathbf{v}), T(\mathbf{w})\}$ is linearly dependent, show that $T(\mathbf{u}) = \mathbf{0}$ has a nontrivial solution.

42. Suppose that $T\colon V \to V$ is a linear operator and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. Show that $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is linearly dependent.

43. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly independent subset of $\mathbb{R}^3$. Find a linear operator $T\colon \mathbb{R}^3 \to \mathbb{R}^3$, such that $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ is linearly dependent.

44. Suppose that $T_1\colon V \to V$ and $T_2\colon V \to V$ are linear operators and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V. If $T_1(\mathbf{v}_i) = T_2(\mathbf{v}_i)$, for each $i = 1, 2, \ldots, n$, show that $T_1(\mathbf{v}) = T_2(\mathbf{v})$ for all $\mathbf{v}$ in V.

45. Verify that $\pounds(U, V)$ is a vector space.

4.2 ▶ The Null Space and Range

In Sec. 3.2, we defined the null space of an $m \times n$ matrix to be the subspace of $\mathbb{R}^n$ of all vectors $\mathbf{x}$ with $A\mathbf{x} = \mathbf{0}$. We also defined the column space of A as the subspace of $\mathbb{R}^m$ of all linear combinations of the column vectors of A. In this section we extend these ideas to linear transformations.

DEFINITION 1 **Null Space and Range** Let V and W be vector spaces. For a linear transformation $T\colon V \longrightarrow W$ the **null space** of T, denoted by $N(T)$, is defined by

$$N(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$$

The **range** of T, denoted by $R(T)$, is defined by

$$R(T) = \{T(\mathbf{v}) \mid \mathbf{v} \in V\}$$

The null space of a linear transformation is then the set of all vectors in V that are mapped to the zero vector, with the range being the set of all images of the mapping, as shown in Fig. 1.

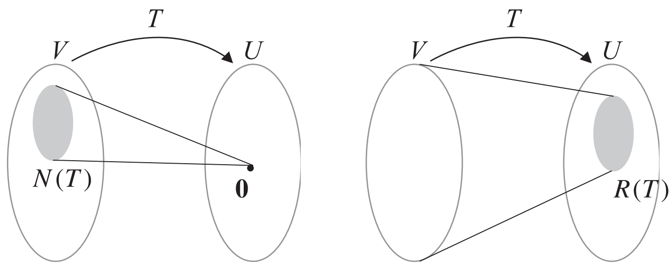


Figure 1

In Theorem 3 we see that the null space and the range of a linear transformation are both subspaces.

THEOREM 3 Let V and W be vector spaces and $T: V \longrightarrow W$ a linear transformation.

1. The null space of T is a subspace of V.
2. The range of T is a subspace of W.

Proof (1) Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be in $N(T)$, so that $T(\mathbf{v}_1) = \mathbf{0}$ and $T(\mathbf{v}_2) = \mathbf{0}$. If c is a scalar, then using the linearity of T, we have

$$T(c\mathbf{v}_1 + \mathbf{v}_2) = cT(\mathbf{v}_1) + T(\mathbf{v}_2) = c\mathbf{0} + \mathbf{0} = \mathbf{0}$$

Thus, $c\mathbf{v}_1 + \mathbf{v}_2$ is in $N(T)$, and by Theorem 4 of Sec. 3.2, $N(T)$ is a subspace of V.
(2) Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be in $R(T)$. Then there are vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in V such that $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. Then for any scalar c,

$$T(c\mathbf{v}_1 + \mathbf{v}_2) = cT(\mathbf{v}_1) + T(\mathbf{v}_2) = c\mathbf{w}_1 + \mathbf{w}_2$$

so that $c\mathbf{w}_1 + \mathbf{w}_2$ is in $R(T)$ and hence $R(T)$ is a subspace of W.

EXAMPLE 1 Define the linear transformation $T: \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \left(\begin{bmatrix} a+b \\ b-c \\ a+d \end{bmatrix}\right)$$

a. Find a basis for the null space of T and its dimension.

Applying A^{-1} to $\mathbf{w}$, we obtain

$$\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 + v_2 \\ -v_1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = T^{-1}(\mathbf{w})$$

THEOREM 11 If V and W are vector spaces of dimension n, then V and W are isomorphic.

Proof By Theorem 10, there are isomorphisms $T_1\colon V \longrightarrow \mathbb{R}^n$ and $T_2\colon W \longrightarrow \mathbb{R}^n$, as shown in Fig. 1. Let $\phi = T_2^{-1} \circ T_1\colon V \longrightarrow W$. To show that ϕ is linear, we first note that T_2^{-1} is linear by Proposition 3. Next by Theorem 2 of Sec. 4.1, the composition $T_2^{-1} \circ T_1$ is linear. Finally, by Theorem 4 of Sec. A.2, the mapping ϕ is one-to-one and onto and is therefore a vector space isomorphism.

$$\begin{tikzcd} V \arrow[r, "T_1"] \arrow[d, "\phi"'] & \mathbb{R}^n \\ W \arrow[ur, "T_2"'] & \end{tikzcd}$$

$$\phi = T_2^{-1} \circ T_1\colon V \longrightarrow W$$

Figure 1

EXAMPLE 4 Find an explicit isomorphism from $\mathcal{P}_2$ onto the vector space of 2×2 symmetric matrices $S_{2\times 2}$.

Solution To use the method given in the proof of Theorem 11, first let

$$B_1 = \{1, x, x^2\} \qquad \text{and} \qquad B_2 = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$$

be ordered bases for $\mathcal{P}_2$ and $S_{2\times 2}$, respectively. Let T_1 and T_2 be the respective coordinate maps from $\mathcal{P}_2$ and $S_{2\times 2}$ into $\mathbb{R}^3$. Then

$$T_1(ax^2 + bx + c) = \begin{bmatrix} c \\ b \\ a \end{bmatrix} \qquad \text{and} \qquad T_2\left(\begin{bmatrix} a & b \\ b & c \end{bmatrix} \right) = \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

22. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -2 & 3 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

23. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -2 & 0 & 1 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

24. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

In Exercises 25–28, determine whether the matrix mapping $T: V \to V$ is an isomorphism.

25. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

26. $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -3 & 1 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$

27. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 & -1 & -1 \\ 2 & 0 & 2 \\ 1 & 1 & -3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

28. $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 3 & 0 \\ -1 & -2 & -3 \\ 0 & -1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$

29. Show that $T: M_{n\times n} \to M_{n\times n}$ defined by

$$T(A) = A^t$$

is an isomorphism.

30. Show that $T: \mathcal{P}_3 \to \mathcal{P}_3$ defined by

$$T(p(x)) = p'''(x) + p''(x) + p'(x) + p(x)$$

is an isomorphism.

31. Let A be an $n \times n$ invertible matrix. Show that $T: M_{n\times n} \to M_{n\times n}$ defined by

$$T(B) = ABA^{-1}$$

is an isomorphism.

32. Find an isomorphism from $M_{2\times 2}$ onto $\mathbb{R}^4$.

33. Find an isomorphism from $\mathbb{R}^4$ onto $\mathcal{P}_3$.

34. Find an isomorphism from $M_{2\times 2}$ onto $\mathcal{P}_3$.

35. Let

$$V = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \middle| \, x + 2y - z = 0 \right\}$$

Find an isomorphism from V onto $\mathbb{R}^2$.

36. Let

$$V = \left\{ \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \middle| \, a, b, c \in \mathbb{R} \right\}$$

Find an isomorphism from $\mathcal{P}_2$ onto V.

37. Suppose $T: \mathbb{R}^3 \to \mathbb{R}^3$ is an isomorphism. Show that T takes lines through the origin to lines through the origin and planes through the origin to planes through the origin.

4.4 ▶ Matrix Representation of a Linear Transformation

Matrices have played an important role in our study of linear algebra. In this section we establish the connection between matrices and linear transformations. To illustrate the idea, recall from Sec. 4.1 that given any $m \times n$ matrix A, we can define a linear transformation $T: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by

$$T(\mathbf{v}) = A\mathbf{v}$$

In Example 8 of Sec. 4.1, we showed how a linear transformation $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ is completely determined by the images of the coordinate vectors $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$ of $\mathbb{R}^3$. The key was to recognize that a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ can be written as

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3 \qquad \text{so that} \qquad T(\mathbf{v}) = v_1T(\mathbf{e}_1) + v_2T(\mathbf{e}_2) + v_3T(\mathbf{e}_3)$$

In that example, T was defined so that

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, T(\mathbf{e}_2) = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \qquad T(\mathbf{e}_3) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Now let A be the 2×3 matrix whose column vectors are $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$. Then

$$T(\mathbf{v}) = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \mathbf{v} = A\mathbf{v}$$

That is, the linear transformation T is given by a matrix product. In general, if $T\colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation, then it is possible to write

$$T(\mathbf{v}) = A\mathbf{v}$$

where A is the $m \times n$ matrix whose jth column vector is $T(\mathbf{e}_j)$ for $j = 1, 2, \ldots, n$. The matrix A is called the **matrix representation of T relative to the standard bases** of $\mathbb{R}^n$ and $\mathbb{R}^m$.

In this section we show that *every linear transformation between finite dimensional vector spaces can be written as a matrix multiplication*. Specifically, let V and W be finite dimensional vector spaces with fixed ordered bases B and B', respectively. If $T\colon V \longrightarrow W$ is a linear transformation, then there exists a matrix A such that

$$[T(\mathbf{v})]_{B'} = A[\mathbf{v}]_B$$

In the case for which $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, and B and B' are, respectively, the standard bases, the last equation is equivalent to

$$T(\mathbf{v}) = A\mathbf{v}$$

as above. We now present the details.

Let V and W be vector spaces with ordered bases $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $B' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, respectively, and let $T\colon V \longrightarrow W$ be a linear transformation. Now let $\mathbf{v}$ be any vector in V and let

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

be the coordinate vector of $\mathbf{v}$ relative to the basis B. Thus,

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

Applying T to both sides of this last equation gives

$$\begin{aligned} T(\mathbf{v}) &= T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n) \end{aligned}$$

Note that for each $i = 1, 2, \ldots, n$ the vector $T(\mathbf{v}_i)$ is in W. Thus, there are unique scalars a_{ij} with $1 \le i \le m$ and $1 \le j \le n$ such that

$$B' = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$
$$\mathbf{v} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

6. $T: \mathbb{R}^3 \to \mathbb{R}^3$,

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x - z \\ -x + y + z \\ 2z \end{bmatrix}$$
$$B = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}$$
$$B' = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$
$$\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

7. $T: \mathbb{R}^2 \to \mathbb{R}^2$,

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x \\ x + y \end{bmatrix}$$
$$B = \left\{ \begin{bmatrix} -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$
$$B' = \left\{ \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \end{bmatrix} \right\}$$
$$\mathbf{v} = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$$

8. $T: \mathbb{R}^3 \to \mathbb{R}^3$,

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + z \\ 2y - x \\ y + z \end{bmatrix}$$
$$B = \left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$
$$B' = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} \right\}$$
$$\mathbf{v} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$$

9. $T: \mathcal{P}_2 \to \mathcal{P}_2$,

$$T(ax^2 + bx + c) = ax^2 + bx + c$$
$$B = \{1, 1 - x, (1 - x)^2\}$$
$$B' = \{1, x, x^2\}$$
$$\mathbf{v} = x^2 - 3x + 3$$

10. $T: \mathcal{P}_2 \to \mathcal{P}_2$,

$$T(p(x)) = p'(x) + p(x)$$
$$B = \{1 - x - x^2, 1, 1 + x^2\}$$
$$B' = \{-1 + x, -1 + x + x^2, x\}$$
$$\mathbf{v} = 1 - x$$

11. Let

$$H = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

and let T be the linear operator on all 2×2 matrices with trace 0, defined by

$$T(A) = AH - HA$$
$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$
$$B' = B$$
$$\mathbf{v} = \begin{bmatrix} 2 & 1 \\ 3 & -2 \end{bmatrix}$$

12. $T: M_{2\times2} \to M_{2\times2}$,

$$T(A) = 2A^t + A$$

B and B' the standard basis on $M_{2\times2}$

$$\mathbf{v} = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$$

13. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear operator defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + 2y \\ x - y \end{bmatrix}$$

Let B be the standard ordered basis for $\mathbb{R}^2$ and B' the ordered basis for $\mathbb{R}^2$ defined by

$$B' = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right\}$$

a. Find $[T]_B$.

b. Find $[T]_{B'}$.

to show directly that the matrices $[T]_{B_1}$ and $[T]_{B_2}$ are similar.

16. Let $T\colon \mathcal{P}_2 \longrightarrow \mathcal{P}_2$ be the linear operator defined by $T(p(x)) = xp'(x) + p''(x)$. Find the matrix representation $[T]_{B_1}$ relative to the basis $B_1 = \{1, x, x^2\}$ and the matrix representation $[T]_{B_2}$ relative to $B_2 = \{1, x, 1 + x^2\}$. Find the transition matrix $P = [I]_{B_2}^{B_1}$, and use Theorem 15 to show directly that the matrices $[T]_{B_1}$ and $[T]_{B_2}$ are similar.

17. Show that if A and B are similar matrices and B and C are similar matrices, then A and C are similar matrices.

18. Show that if A and B are similar matrices, then $\det(A) = \det(B)$.

19. Show that if A and B are similar matrices, then $\mathbf{tr}(A) = \mathbf{tr}(B)$.

20. Show that if A and B are similar matrices, then A^t and B^t are similar matrices.

21. Show that if A and B are similar matrices, then A^n and B^n are similar matrices for each positive integer n.

22. Show that if A and B are similar matrices and λ is any scalar, then $\det(A - \lambda I) = \det(B - \lambda I)$.

4.6 ▶ Application: Computer Graphics

The rapid development of increasingly more powerful computers has led to the explosive growth of digital media. Computer-generated visual content is ubiquitous, found in almost every arena from advertising and entertainment to science and medicine. The branch of computer science known as *computer graphics* is devoted to the study of the generation and manipulation of digital images. Computer graphics are based on displaying two- or three-dimensional objects in two-dimensional space. Images displayed on a computer screen are stored in memory using data items called **pixels**, which is short for picture elements. A single picture can be comprised of millions of pixels, which collectively determine the image. Each pixel contains information on how to color the corresponding point on a computer screen, as shown in Fig. 1. If an image contains curves or lines, the pixels which describe the object may be connected by a mathematical formula. The saddle shown in Fig. 1 is an example.

Figure 1

Reflection

The **reflection** of a geometric object through a line produces the *mirror image* of the object across the line. The linear operator that reflects a vector through the x axis is given by

$$R_x\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ -y \end{bmatrix}$$

A reflection through the y axis is given by

$$R_y\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x \\ y \end{bmatrix}$$

and a reflection through the line $y = x$ is given by

$$R_{y=x}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix}$$

The matrix representations, relative to the standard basis B, for each of these are given by

$$[R_x]_B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad [R_y]_B = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \qquad [R_{y=x}]_B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

EXAMPLE 3 Perform the following reflections on the triangle T of Fig. 4.

a. Reflection through the x axis.

b. Reflection through the y axis.

c. Reflection through the line $y = x$.

Solution

a. The vertices of the triangle in Fig. 4 are given by

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Applying the matrix $[R_x]_B$ to the vertices of the original triangle, we obtain

$$\mathbf{v}_1' = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \qquad \mathbf{v}_2' = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \qquad \mathbf{v}_3' = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

The image of the triangle is shown in Fig. 9(a).

b. Applying the matrix $[R_y]_B$ to the vertices of the original triangle, we obtain

$$\mathbf{v}_1' = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2' = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \qquad \mathbf{v}_3' = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

The image of the triangle with this reflection is shown in Fig. 9(b).

c. Finally, applying the matrix $[R_{x=y}]_B$ to the vertices of the original triangle, we obtain

$$\mathbf{v}_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \mathbf{v}_2' = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{v}_3' = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

The image of the triangle is shown in Fig. 9(c).

given by

$$S_\theta\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$$

The matrix of S_θ relative to the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2\}$ for $\mathbb{R}^2$ is given by

$$[S_\theta]_B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

When using homogeneous coordinates, we apply the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

EXAMPLE 7 Find the image of the triangle of Fig. 4 under a translation by the vector $\mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, followed by a rotation of 30°, or $\pi/6$ rad, in the counterclockwise direction.

Solution The matrix for the combined operation is given by

$$\begin{bmatrix} \cos\frac{\pi}{6} & -\sin\frac{\pi}{6} & 0 \\ \sin\frac{\pi}{6} & \cos\frac{\pi}{6} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} + \frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} & \frac{1}{2} - \frac{\sqrt{3}}{2} \\ 0 & 0 & 1 \end{bmatrix}$$

The vertices of the triangle in homogeneous coordinates are given by

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

After applying the above matrix to each of these vectors, we obtain

$$\mathbf{v}'_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} \qquad \mathbf{v}'_2 = \begin{bmatrix} \frac{3\sqrt{3}}{2} \\ \frac{3}{2} \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}'_3 = \begin{bmatrix} \sqrt{3} - 1 \\ \sqrt{3} + 1 \\ 1 \end{bmatrix}$$

The resulting triangle is shown in Fig. 13.

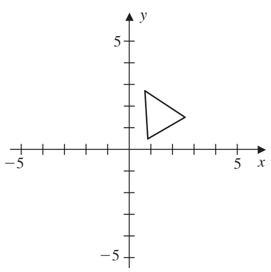


Figure 13

Projection

Rendering a picture of a three-dimensional object on a flat computer screen requires projecting points in 3-space to points in 2-space. We discuss only one of many methods to project points in $\mathbb{R}^3$ to points in $\mathbb{R}^2$ that preserve the natural appearance of an object.

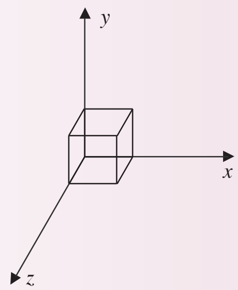


Figure 15

b. Find a 3×3 matrix that will rotate the (projected) vertices of the cube by 30° and another that will translate the cube by the vector $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Solution

a. We can arbitrarily set $z_d = -1$. Then

$$\tan\psi = \tan 30^\circ \approx 0.577 = \frac{y_d}{x_d} \quad \text{and} \quad (\tan\phi)^2 = (\tan 26.6^\circ)^2 \approx (0.5)^2 = x_d^2 + y_d^2$$

so that

$$y_d = 0.577 x_d \quad \text{and} \quad x_d^2 + y_d^2 = \frac{1}{4}$$

Solving the last two equations gives $x_d \approx 0.433$ and $y_d \approx 0.25$, so that the direction vector is

$$\mathbf{v}_d = \begin{bmatrix} 0.433 \\ 0.25 \\ -1 \end{bmatrix}$$

Using the formulas for a projected point given above, we can project each vertex of the cube into $\mathbb{R}^2$. Connecting the images by line segments gives the picture shown in Fig. 16. The projected points are given in Table 1.

Table 1

Vertex	Projected Point
(0,0,1)	(0.433, 0.25)
(1,0,1)	(1.433, 0.25)
(1,0,0)	(1, 0)
(0,0,0)	(0, 0)
(0,1,1)	(0.433, 1.25)
(1,1,1)	(1.433, 1.25)
(1,1,0)	(1, 1)
(0,1,0)	(0, 1)

Alternatively, the set

$$V_\lambda = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} = \lambda\mathbf{v}\} = \{\mathbf{v} \in \mathbb{R}^n \mid (A - \lambda I)\mathbf{v} = 0\} = N(A - \lambda I)$$

Since V_λ is the null space of the matrix $A - \lambda I$, by Theorem 3 of Sec. 4.2 it is a subspace of $\mathbb{R}^n$.

EXAMPLE 2 Find the eigenvalues and corresponding eigenvectors of

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Give a description of the eigenspace corresponding to each eigenvalue.

Solution By Theorem 1 to find the eigenvalues, we solve the characteristic equation

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 2-\lambda & -12 \\ 1 & -5-\lambda \end{vmatrix} \\ &= (2-\lambda)(-5-\lambda) - (1)(-12) \\ &= \lambda^2 + 3\lambda + 2 \\ &= (\lambda+1)(\lambda+2) = 0 \end{aligned}$$

Thus the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = -2$. To find the eigenvectors, we need to find all nonzero vectors in the null spaces of $A - \lambda_1 I$ and $A - \lambda_2 I$. First, for $\lambda_1 = -1$,

$$A - \lambda_1 I = A + I = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -12 \\ 1 & -4 \end{bmatrix}$$

The null space of $A + I$ is found by row-reducing the augmented matrix

$$\left[\begin{array}{cc|c} 3 & -12 & 0 \\ 1 & -4 & 0 \end{array}\right] \qquad \text{to} \qquad \left[\begin{array}{cc|c} 1 & -4 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

The solution set for this linear system is given by $S = \left\{ \begin{bmatrix} 4t \\ t \end{bmatrix} \middle| \, t \in \mathbb{R} \right\}$. Choosing $t = 1$, we obtain the eigenvector $\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$. Hence, the eigenspace corresponding to $\lambda_1 = -1$ is

$$V_{\lambda_1} = \left\{ t \begin{bmatrix} 4 \\ 1 \end{bmatrix} \middle| \, t \text{ is any real number} \right\}$$

For $\lambda_2 = -2$,

$$A - \lambda_2 I = \begin{bmatrix} 4 & -12 \\ 1 & -3 \end{bmatrix}$$

In a similar way we find that the vector $\mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ is an eigenvector corresponding to $\lambda_2 = -2$. The corresponding eigenspace is

THEOREM 3 Let A be an $n \times n$ matrix, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be distinct eigenvalues with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Then the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is linearly independent.

Proof The proof is by contradiction. Assume that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct eigenvalues of A with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, and assume that the set of eigenvectors is linearly dependent. Then by Theorem 5 of Sec. 2.3, at least one of the vectors can be written as a linear combination of the others. Moreover, the eigenvectors can be reordered so that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$, with $m < n$, are linearly independent, but $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{m+1}$ are linearly dependent with $\mathbf{v}_{m+1}$ a nontrivial linear combination of the first m vectors. Therefore, there are scalars $c_1, \ldots, c_m$, not all 0, such that

$$\mathbf{v}_{m+1} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$$

This is the statement that will result in a contradiction. We multiply the last equation by A to obtain

$$\begin{aligned} A\mathbf{v}_{m+1} &= A(c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m) \\ &= c_1A(\mathbf{v}_1) + \cdots + c_mA(\mathbf{v}_m) \end{aligned}$$

Further, since $\mathbf{v}_i$ is an eigenvector corresponding to the eigenvalue λ_i, then $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$, and after substitution in the previous equation, we have

$$\lambda_{m+1}\mathbf{v}_{m+1} = c_1\lambda_1\mathbf{v}_1 + \cdots + c_m\lambda_m\mathbf{v}_m$$

Now multiplying both sides of $\mathbf{v}_{m+1} = c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m$ by λ_{m+1}, we also have

$$\lambda_{m+1}\mathbf{v}_{m+1} = c_1\lambda_{m+1}\mathbf{v}_1 + \cdots + c_m\lambda_{m+1}\mathbf{v}_m$$

By equating the last two expressions for $\lambda_{m+1}\mathbf{v}_{m+1}$ we obtain

$$c_1\lambda_1\mathbf{v}_1 + \cdots + c_m\lambda_m\mathbf{v}_m = c_1\lambda_{m+1}\mathbf{v}_1 + \cdots + c_m\lambda_{m+1}\mathbf{v}_m$$

or equivalently,

$$c_1(\lambda_1 - \lambda_{m+1})\mathbf{v}_1 + \cdots + c_m(\lambda_m - \lambda_{m+1})\mathbf{v}_m = \mathbf{0}$$

Since the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly independent, the only solution to the previous equation is the trivial solution, that is,

$$c_1(\lambda_1 - \lambda_{m+1}) = 0 \qquad c_2(\lambda_2 - \lambda_{m+1}) = 0 \qquad \ldots \qquad c_m(\lambda_m - \lambda_{m+1}) = 0$$

Since all the eigenvalues are distinct, we have

$$\lambda_1 - \lambda_{m+1} \neq 0 \qquad \lambda_2 - \lambda_{m+1} \neq 0 \qquad \ldots \qquad \lambda_m - \lambda_{m+1} \neq 0$$

and consequently

$$c_1 = 0 \qquad c_2 = 0 \qquad \ldots \qquad c_m = 0$$

This contradicts the assumption that the nonzero vector $\mathbf{v}_{m+1}$ is a nontrivial linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$.

COROLLARY 1 If A is an $n \times n$ matrix with n distinct eigenvalues, then A is diagonalizable.

EXAMPLE 4 Show that every 2×2 real symmetric matrix is diagonalizable.

Solution Recall that the matrix A is symmetric if and only if $A = A^t$. Every 2×2 symmetric matrix has the form

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

See Example 5 of Sec. 1.3. The eigenvalues are found by solving the characteristic equation

$$\det(A - \lambda I) = \begin{vmatrix} a-\lambda & b \\ b & d-\lambda \end{vmatrix} = \lambda^2 - (a+d)\lambda + ad - b^2 = 0$$

By the quadratic formula, the eigenvalues are

$$\lambda = \frac{a + d \pm \sqrt{(a-d)^2 + 4b^2}}{2}$$

Since the discriminant $(a-d)^2 + 4b^2 \geq 0$, the characteristic equation has either one or two real roots. If $(a-d)^2 + 4b^2 = 0$, then $(a-d)^2 = 0$ and $b^2 = 0$, which holds if and only if $a = d$ and $b = 0$. Hence, the matrix A is diagonal. If $(a-d)^2 + 4b^2 > 0$, then A has two distinct eigenvalues; so by Corollary 1, the matrix A is diagonalizable.

By Theorem 2, if A is diagonalizable, then A is similar to a diagonal matrix whose eigenvalues are the same as the eigenvalues of A. In Theorem 4 we show that the same can be said about any two similar matrices.

THEOREM 4 Let A and B be similar $n \times n$ matrices. Then A and B have the same eigenvalues.

Proof Since A and B are similar matrices, there is an invertible matrix P such that $B = P^{-1}AP$. Now

$$\begin{aligned} \det(B - \lambda I) &= \det(P^{-1}AP - \lambda I) \\ &= \det(P^{-1}(AP - P(\lambda I))) \\ &= \det(P^{-1}(AP - \lambda I P)) \\ &= \det(P^{-1}(A - \lambda I)P) \end{aligned}$$

EXAMPLE 6 Define the linear operator $T: \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 - x_2 + 2x_3 \\ 2x_1 + 2x_3 \\ x_1 + 3x_2 \end{bmatrix}$$

Show that T is diagonalizable.

Solution Let $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the standard basis for $\mathbb{R}^3$. Then the matrix for T relative to B is

$$[T]_B = \begin{bmatrix} 3 & -1 & 2 \\ 2 & 0 & 2 \\ 1 & 3 & 0 \end{bmatrix}$$

Observe that the eigenvalues of $[T]_B$ are $\lambda_1 = -2$, $\lambda_2 = 4$, and $\lambda_3 = 1$ with corresponding eigenvectors, respectively,

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} -5 \\ 4 \\ 7 \end{bmatrix}$$

Now let $B' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and

$$P = \begin{bmatrix} 1 & 1 & -5 \\ 1 & 1 & 4 \\ -2 & 1 & 7 \end{bmatrix}$$

Then

$$[T]_{B'} = \frac{1}{9}\begin{bmatrix} -1 & 4 & -3 \\ 5 & 1 & 3 \\ -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & -1 & 2 \\ 2 & 0 & 2 \\ 1 & 3 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 & -5 \\ 1 & 1 & 4 \\ -2 & 1 & 7 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Fact Summary

Let A be an $n \times n$ matrix.

1. If A is diagonalizable, then $A = PDP^{-1}$ or equivalently $D = P^{-1}AP$. The matrix D is a diagonal matrix with diagonal entries the eigenvalues of A. The matrix P is invertible whose column vectors are the corresponding eigenvectors.
2. If A is diagonalizable, then the diagonalizing matrix P is not unique. If the columns of P are permuted, then the diagonal entries of D are permuted in the same way.
3. The matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
4. If A has n distinct eigenvalues, then A is diagonalizable.

b. Find the matrix B for T relative to the basis $\{x, x-1, x^2\}$.

c. Show the eigenvalues of A and B are the same.

d. Explain why T is not diagonalizable.

36. Define a vector space $V = \textbf{span}\{\sin x, \cos x\}$ and a linear operator $T\colon V \to V$ by $T(f(x)) = f'(x)$. Show that T is diagonalizable.

37. Define a linear operator $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + 2x_2 + 2x_3 \\ -x_1 + 2x_2 + x_3 \\ x_1 - x_2 \end{bmatrix}$$

Show that T is not diagonalizable.

38. Define a linear operator $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 4x_1 + 2x_2 + 4x_3 \\ 4x_1 + 2x_2 + 4x_3 \\ 4x_3 \end{bmatrix}$$

Show that T is diagonalizable.

39. Let T be a linear operator on a finite dimensional vector space, A the matrix for T relative to a basis B_1, and B the matrix for T relative to a basis B_2. Show that A is diagonalizable if and only if B is diagonalizable.

5.3 ▶ Application: Systems of Linear Differential Equations

In Sec. 3.5 we considered only a single differential equation where the solution involved a single function. However, in many modeling applications, an equation that involves the derivatives of only one function is not sufficient. It is more likely that the rate of change of a variable quantity will be linked to other functions outside itself. This is the fundamental idea behind the notion of a *dynamical system*. One of the most familiar examples of this is the *predator-prey* model. For example, suppose we wish to create a model to predict the number of foxes and rabbits in some habitat. The growth rate of the foxes is dependent on not only the number of foxes but also the number of rabbits in their territory. Likewise, the growth rate of the rabbit population in part is dependent on their current number, but is obviously mitigated by the number of foxes in their midst. The mathematical model required to describe this relationship is a *system of differential equations* of the form

$$\begin{cases} y_1'(t) &= f(t, y_1, y_2) \\ y_2'(t) &= g(t, y_1, y_2) \end{cases}$$

In this section we consider systems of linear differential equations. Problems such as predator-prey problems involve systems of *nonlinear* differential equations.

Uncoupled Systems

At the beginning of Sec. 3.5 we saw that the differential equation given by

$$y' = ay$$

has the solution $y(t) = Ce^{at}$, where $C = y(0)$. An extension of this to two dimensions is the **system of differential equations**

$$\begin{cases} y_1' &= ay_1 \\ y_2' &= by_2 \end{cases}$$

In Example 2 we describe the solution for a system when the eigenvalues have the same sign.

EXAMPLE 2 Find the general solution to the system of differential equations

$$\begin{cases} y_1' = y_1 + 3y_2 \\ y_2' = \phantom{y_1 + {}} 2y_2 \end{cases}$$

Solution The system of differential equations is given in matrix form by

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

The eigenvalues of A are $\lambda_1 = 1$ and $\lambda_2 = 2$ with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

The matrix that diagonalizes A is then

$$P = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \quad \text{with} \quad P^{-1} = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix}$$

The uncoupled system is given by

$$\mathbf{w}' = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \mathbf{w}$$

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

with general solution

$$\mathbf{w}(t) = \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \mathbf{w}(0)$$

Hence, the solution to the original system is given by

$$\mathbf{y}(t) = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \mathbf{y}(0)$$

$$= \begin{bmatrix} e^t & -3e^t + 3e^{2t} \\ 0 & e^{2t} \end{bmatrix} \mathbf{y}(0)$$

The general solution can also be written in the form

$$y_1(t) = \left[y_1(0) - 3y_2(0)\right] e^t + 3y_2(0)e^{2t} \qquad \text{and} \qquad y_2(t) = y_2(0)e^{2t}$$

Hence, the solution to the original system is given by

$$\begin{aligned}
\mathbf{y}(t) &= \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} e^{-\frac{3}{20}t} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} \mathbf{y}(0) \\
&= \frac{1}{3} \begin{bmatrix} 2e^{-\frac{3}{20}t} + 1 & -e^{-\frac{3}{20}t} + 1 \\ -2e^{-\frac{3}{20}t} + 2 & e^{-\frac{3}{20}t} + 2 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \end{bmatrix} \\
&= \frac{8}{3} \begin{bmatrix} 2e^{-\frac{3}{20}t} + 1 \\ -2e^{-\frac{3}{20}t} + 2 \end{bmatrix}
\end{aligned}$$

c. The solution to the system in equation form is given by

$$y_1(t) = \frac{8}{3}\left(2e^{-\frac{3}{20}t} + 1\right) \qquad \text{and} \qquad y_2(t) = \frac{8}{3}\left(-2e^{-\frac{3}{20}t} + 2\right)$$

To find the amount of salt in each tank as t goes to infinity, we compute the limits

$$\lim_{t\to\infty} \frac{8}{3}\left(2e^{-\frac{3}{20}t} + 1\right) = \frac{8}{3}(0+1) = \frac{8}{3}$$

and

$$\lim_{t\to\infty} \frac{8}{3}\left(-2e^{-\frac{3}{20}t} + 2\right) = \frac{8}{3}(0+2) = \frac{16}{3}$$

These values make sense intuitively as we expect that the 8 lb of salt should eventually be thoroughly mixed, and divided proportionally between the two tanks in a ratio of 1 : 2.

Exercise Set 5.3

In Exercises 1–6, find the general solution to the system of differential equations.

1. $\begin{cases} y_1' = -y_1 + y_2 \\ y_2' = -2y_2 \end{cases}$

2. $\begin{cases} y_1' = -y_1 + 2y_2 \\ y_2' = y_1 \end{cases}$

3. $\begin{cases} y_1' = y_1 - 3y_2 \\ y_2' = -3y_1 + y_2 \end{cases}$

4. $\begin{cases} y_1' = y_1 - y_2 \\ y_2' = -y_1 + y_2 \end{cases}$

5. $\begin{cases} y_1' = -4y_1 - 3y_2 - 3y_3 \\ y_2' = 2y_1 + 3y_2 + 2y_3 \\ y_3' = 4y_1 + 2y_2 + 3y_3 \end{cases}$

6. $\begin{cases} y_1' = -3y_1 - 4y_2 - 4y_3 \\ y_2' = 7y_1 + 11y_2 + 13y_3 \\ y_3' = -5y_1 - 8y_2 - 10y_3 \end{cases}$

In Exercises 7 and 8, solve the initial-value problem.

7. $\begin{cases} y_1' = -y_1 \\ y_2' = 2y_1 + y_2 \end{cases} \qquad y_1(0) = 1 \qquad y_2(0) = -1$

For reasons that will soon be clear, we scale $\mathbf{v}_1$ (by the reciprocal of the sum of its components) so that it becomes a probability vector. Observe that this new vector

$$\widehat{\mathbf{v}}_1 = \begin{bmatrix} \frac{5}{8} \\ \frac{3}{8} \end{bmatrix}$$

is also an eigenvector since it is in the eigenspace V_{λ_1}. Since the 2×2 transition matrix has two distinct eigenvalues, by Corollary 1 of Sec. 5.2, T is diagonalizable and, by Theorem 2 of Sec. 5.2, can be written as

$$\begin{aligned} T &= PDP^{-1} \\ &= \begin{bmatrix} \frac{5}{8} & -1 \\ \frac{3}{8} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{2}{10} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -\frac{3}{8} & \frac{5}{8} \end{bmatrix} \end{aligned}$$

By Exercise 27 of Sec. 5.2, the powers of T are given by

$$T^n = PD^nP^{-1} = P \begin{bmatrix} 1^n & 0 \\ 0 & \left(\frac{2}{10}\right)^n \end{bmatrix} P^{-1}$$

As mentioned above, this gives us an easier way to compute the state vector for large values of n. Another benefit from this representation is that the matrix D^n approaches

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

as n gets large. This suggests that the eigenvector corresponding to $\lambda = 1$ is useful in determining the limiting proportion of sunny days to cloudy days far into the future.

Steady-State Vector

Given an initial state vector $\mathbf{v}$, of interest is the long-run behavior of this vector in a Markov chain, that is, the tendency of the vector $T^n\mathbf{v}$ for large n. If for any initial state vector $\mathbf{v}$ there is some vector $\mathbf{s}$ such that $T^n\mathbf{v}$ approaches $\mathbf{s}$, then $\mathbf{s}$ is called a **steady-state vector** for the Markov process.

In our weather model we saw that the transition matrix T has an eigenvalue $\lambda = 1$ and a corresponding probability eigenvector given by

$$\widehat{\mathbf{v}}_1 = \begin{bmatrix} \frac{5}{8} \\ \frac{3}{8} \end{bmatrix} = \begin{bmatrix} 0.625 \\ 0.375 \end{bmatrix}$$

We claim that this vector is a steady-state vector for the weather model. As verification, let $\mathbf{u}$ be an initial probability vector, say, $\mathbf{u} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$. We then compute

$$T^{10}\mathbf{u} = \begin{bmatrix} 0.6249999954 \\ 0.3750000046 \end{bmatrix} \quad \text{and} \quad T^{20}\mathbf{u} = \begin{bmatrix} 0.6250000002 \\ 0.3750000002 \end{bmatrix}$$

which suggests that $T^n\mathbf{u}$ approaches $\widehat{\mathbf{v}}_1$. That this is in fact the case is stated in Theorem 6. Before doing so, we note that a **regular** transition matrix T is a transition matrix such that for some n, all the entries of T^n are positive.

THEOREM 6 If a Markov chain has a regular stochastic transition matrix T, then there is a unique probability vector $\mathbf{s}$ with $T\mathbf{s} = \mathbf{s}$. Moreover, $\mathbf{s}$ is the steady-state vector for any initial probability vector.

EXAMPLE 1 A group insurance plan allows three different options for participants, plan A, B, or C. Suppose that the percentages of the total number of participants enrolled in each plan are 25 percent, 30 percent, and 45 percent, respectively. Also, from past experience assume that participants change plans as shown in the table.

	A	B	C
A	0.75	0.25	0.2
B	0.15	0.45	0.4
C	0.1	0.3	0.4

a. Find the percent of participants enrolled in each plan after 5 years.

b. Find the steady-state vector for the system.

Solution Let T be the matrix given by

$$T = \begin{bmatrix} 0.75 & 0.25 & 0.2 \\ 0.15 & 0.45 & 0.4 \\ 0.1 & 0.3 & 0.4 \end{bmatrix}$$

a. The number of participants enrolled in each plan after 5 years is approximated by the vector

$$T^5\mathbf{v} = \begin{bmatrix} 0.49776 & 0.46048 & 0.45608 \\ 0.28464 & 0.30432 & 0.30664 \\ 0.21760 & 0.23520 & 0.23728 \end{bmatrix} \begin{bmatrix} 0.25 \\ 0.30 \\ 0.45 \end{bmatrix} = \begin{bmatrix} 0.47 \\ 0.30 \\ 0.22 \end{bmatrix}$$

so approximately 47 percent will be enrolled in plan A, 30 percent in plan B, and 22 percent in plan C.

b. The steady-state vector for the system is the probability eigenvector corresponding to the eigenvalue $\lambda = 1$, that is,

$$\mathbf{s} = \begin{bmatrix} 0.48 \\ 0.30 \\ 0.22 \end{bmatrix}$$

a. Let D be the diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & \dots & \dots & \dots & \lambda_n \end{bmatrix}$$

and find e^D.

b. Suppose A is diagonalizable and $D = P^{-1}AP$. Show that $e^A = Pe^D P^{-1}$.

c. Use parts (a) and (b) to compute e^A for the matrix

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

Chapter 5: Chapter Test

In Exercises 1–40, determine whether the statement is true or false.

1. The matrix

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

diagonalizes the matrix

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

2. The matrix

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

is similar to the matrix

$$D = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

3. The matrix

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

is diagonalizable.

4. The eigenvalues of

$$A = \begin{bmatrix} -1 & 0 \\ -4 & -3 \end{bmatrix}$$

are $\lambda_1 = -3$ and $\lambda_2 = -1$.

5. The characteristic polynomial of

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

is $\lambda^3 + 2\lambda^2 + \lambda - 4$.

6. The eigenvectors of

$$A = \begin{bmatrix} -4 & 0 \\ 3 & -5 \end{bmatrix}$$

are $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

7. The matrix

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

has an eigenvalue $\lambda_1 = 1$ and V_{λ_1} has dimension 1.

8. If

$$A = \begin{vmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{vmatrix}$$

then $AA^t = I$.

9. If A is a 2×2 matrix with $\det(A) < 0$, then A has two real eigenvalues.

10. If A is a 2×2 matrix that has two distinct eigenvalues λ_1 and λ_2, then $\mathbf{tr}(A) = \lambda_1 + \lambda_2$.

11. If $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$, then the eigenvalues of A are $\lambda_1 = a + b$ and $\lambda_2 = b - a$.

12. For all integers k the matrix $A = \begin{bmatrix} 1 & k \\ 1 & 1 \end{bmatrix}$ has only one eigenvalue.

13. If A is a 2×2 invertible matrix, then A and A^{-1} have the same eigenvalues.

14. If A is similar to B, then $\mathbf{tr}(A) = \mathbf{tr}(B)$.

15. The matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is diagonalizable.

PROPOSITION 2 Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are orthogonal if and only if $\mathbf{u} \cdot \mathbf{v} = 0$. The zero vector is orthogonal to every vector in $\mathbb{R}^n$.

One consequence of Proposition 2 is that if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, then

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) &= \|\mathbf{u}\|^2 + 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{v}\|^2 \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \end{aligned}$$

This is a generalization of the Pythagorean theorem to $\mathbb{R}^n$.

Theorem 3 gives several useful properties of the norm in $\mathbb{R}^n$.

THEOREM 3 **Properties of the Norm in $\mathbb{R}^n$** Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$ and c a scalar.

1. $\|\mathbf{v}\| \geq 0$
2. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$
3. $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$
4. (Triangle inequality) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Proof Parts 1 and 2 follow immediately from Definition 1 and Theorem 1. Part 3 is established in Example 1. To establish part 4, we have

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) \\ &= (\mathbf{u} \cdot \mathbf{u}) + 2(\mathbf{u} \cdot \mathbf{v}) + (\mathbf{v} \cdot \mathbf{v}) \\ &= \|\mathbf{u}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2 \end{aligned}$$

Now, by the Cauchy-Schwartz inequality, $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$, so that

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &\leq \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

After taking square roots of both sides of this equation, we obtain

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

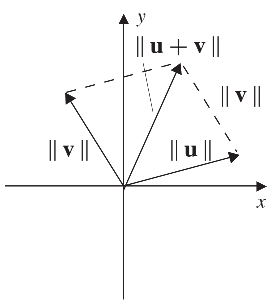


Figure 4

Geometrically, part 4 of Theorem 3 confirms our intuition that the shortest distance between two points is a straight line, as seen in Fig. 4.

PROPOSITION 3 Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^n$. Then $\|\mathbf{u}+\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$ if and only if the vectors have the same direction.

Proof First suppose that the vectors have the same direction. Then the angle between the vectors is 0, so that $\cos\theta = 1$ and $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|$. Therefore,

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})\\ &= \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2\\ &= \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2\\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2\end{aligned}$$

Taking square roots of both sides of the previous equation gives $\|\mathbf{u}+\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$.

Conversely, suppose that $\|\mathbf{u}+\mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$. After squaring both sides, we obtain

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2$$

However, we also have

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2\mathbf{u}\cdot\mathbf{v} + \|\mathbf{v}\|^2$$

Equating both expressions for $\|\mathbf{u}+\mathbf{v}\|^2$ gives

$$\|\mathbf{u}\|^2 + 2\,\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\mathbf{u}\cdot\mathbf{v} + \|\mathbf{v}\|^2$$

Simplifying the last equation, we obtain $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|$ and hence

$$\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = 1$$

Therefore, $\cos\theta = 1$, so that $\theta = 0$ and the vectors have the same direction.

Fact Summary

All vectors are in $\mathbb{R}^n$.

1. The length of a vector and the distance between two vectors are natural extensions of the same geometric notions in $\mathbb{R}^2$ and $\mathbb{R}^3$.
2. The dot product of a vector with itself gives the square of its length and is 0 only when the vector is the zero vector. The dot product of two vectors is commutative and distributes through vector addition.
3. By using the Cauchy-Schwartz inequality $|\mathbf{u}\cdot\mathbf{v}| \le \|\mathbf{u}\|\,\|\mathbf{v}\|$, the angle between vectors is defined by
$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$
4. Two vectors are orthogonal if and only if the dot product of the vectors is 0.

In Exercises 19–22, let

$$\mathbf{v_1} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \qquad \mathbf{v_2} = \begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix}$$

$$\mathbf{v_3} = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} \qquad \mathbf{v_4} = \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$\mathbf{v_5} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

19. Determine which of the vectors are orthogonal.

20. Determine which of the vectors are in the same direction.

21. Determine which of the vectors are in the opposite direction.

22. Determine which of the vectors are unit vectors.

In Exercises 23–28, find the projection of $\mathbf{u}$ onto $\mathbf{v}$ given by

$$\mathbf{w} = \frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v}$$

The vector $\mathbf{w}$ is called the *orthogonal projection* of $\mathbf{u}$ onto $\mathbf{v}$. Sketch the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

23. $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

24. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

25. $\mathbf{u} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$

26. $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

27. $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$

28. $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$

29. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and suppose $\mathbf{v} \cdot \mathbf{u}_i = 0$ for each $i = 1, \ldots, n$. Show that $\mathbf{v}$ is orthogonal to every vector in $\mathbf{span}(S)$.

30. Let $\mathbf{v}$ be a fixed vector in $\mathbb{R}^n$ and define $S = \{\mathbf{u} \mid \mathbf{u} \cdot \mathbf{v} = 0\}$. Show that S is a subspace of $\mathbb{R}^n$.

31. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a set of nonzero vectors which are pairwise orthogonal. That is, if $i \neq j$, then $\mathbf{v}_i \cdot \mathbf{v}_j = 0$. Show that S is linearly independent.

32. Let A be an $n \times n$ invertible matrix. Show that if $i \neq j$, then row vector i of A and column vector j of A^{-1} are orthogonal.

33. Show that for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$,

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

34. **a.** Find a vector that is orthogonal to every vector in the plane P: $x + 2y - z = 0$.

b. Find a matrix A such that the null space $N(A)$ is the plane $x + 2y - z = 0$.

35. Suppose that the column vectors of an $n \times n$ matrix A are pairwise orthogonal. Find $A^t A$.

36. Let A be an $n \times n$ matrix and $\mathbf{u}$ and $\mathbf{v}$ vectors in $\mathbb{R}^n$. Show that

$$\mathbf{u} \cdot (A\mathbf{v}) = (A^t\mathbf{u}) \cdot \mathbf{v}$$

37. Let A be an $n \times n$ matrix. Show that A is symmetric if and only if

$$(A\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (A\mathbf{v})$$

for all $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. *Hint*: See Exercise 36.

and linearly independent. Theorem 5 relates the notions of orthogonality and linear independence in an inner product space.

THEOREM 5 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal set of nonzero vectors in an inner product space V, then S is linearly independent.

Proof Since the set S is an orthogonal set of nonzero vectors,

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0 \qquad \text{for } i \neq j \qquad \text{and} \qquad \langle \mathbf{v}_i, \mathbf{v}_i \rangle = \|\mathbf{v}_i\|^2 \neq 0 \qquad \text{for all } i$$

Now suppose that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

The vectors are linearly independent if and only if the only solution to the previous equation is the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. Now let $\mathbf{v}_j$ be an element of S. Take the inner product on both sides of the previous equation with $\mathbf{v}_j$ so that

$$\langle \mathbf{v}_j, (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{j-1}\mathbf{v}_{j-1} + c_j\mathbf{v}_j + c_{j+1}\mathbf{v}_{j+1} + \cdots + c_n\mathbf{v}_n) \rangle = \langle \mathbf{v}_j, \mathbf{0} \rangle$$

By the linearity of the inner product and the fact that S is orthogonal, this equation reduces to

$$c_j \langle \mathbf{v}_j, \mathbf{v}_j \rangle = \langle \mathbf{v}_j, \mathbf{0} \rangle$$

Now, by Proposition 4 and the fact that $\|\mathbf{v}_j\| \neq 0$, we have

$$c_j \|\mathbf{v}_j\|^2 = 0 \qquad \text{so that} \qquad c_j = 0$$

Since this holds for each $j = 1, \ldots, n$, then $c_1 = c_2 = \cdots = c_n = 0$ and therefore S is linearly independent.

COROLLARY 1 If V is an inner product space of dimension n, then any orthogonal set of n nonzero vectors is a basis for V.

The proof of this corollary is a direct result of Theorem 12 of Sec. 3.3. Theorem 6 provides us with an easy way to find the coordinates of a vector relative to an orthonormal basis. This property underscores the usefulness and desirability of orthonormal bases.

THEOREM 6 If $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an ordered orthonormal basis for an inner product space V and $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$, then the coordinates of $\mathbf{v}$ relative to B are given by $c_i = \langle \mathbf{v}_i, \mathbf{v} \rangle$ for each $i = 1, 2, \ldots, n$.

Proof Let $\mathbf{v}_i$ be a vector in B. Taking the inner product on both sides of

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_i\mathbf{v}_i + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_n\mathbf{v}_n$$

To justify this substitution, note that $\mathbf{w}_1 \cdot \mathbf{w}_2 = 0$; that is, multiplying $\mathbf{w}_2$ by a scalar does not change the fact that it is orthogonal to $\mathbf{w}_1$. To find $\mathbf{w}_3$, we use the computation

$$\mathbf{w}_3 = \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2}\mathbf{w}_2$$

$$= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \left(-\frac{1}{3}\right)\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{6}\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

As before we replace $\mathbf{w}_3$ with

$$\mathbf{w}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

An orthogonal basis for U is then given by

$$B' = \left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$$

Normalizing each of the vectors of B' produces the orthonormal basis

$$B'' = \left\{ \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}$$

Fact Summary

1. Every finite dimensional inner product space has an orthonormal basis.
2. The Gram-Schmidt process is an algorithm to construct an orthonormal basis from any basis of the vector space.

Exercise Set 6.3

In Exercises 1–8, use the standard inner product on $\mathbb{R}^n$.

a. Find $\text{proj}_{\mathbf{v}}\, \mathbf{u}$.

b. Find the vector $\mathbf{u} - \text{proj}_{\mathbf{v}}\mathbf{u}$ and verify this vector is orthogonal to $\mathbf{v}$.

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

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

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

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

5. $\mathbf{u} = \begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$

6. $\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$

7. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

8. $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

In Exercises 9–12, use the inner product on $\mathcal{P}_2$ defined by

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$$

a. Find $\mathbf{proj}_q p$.

b. Find the vector $p - \mathbf{proj}_q p$ and verify that this vector is orthogonal to q.

9. $p(x) = x^2 - x + 1, q(x) = 3x - 1$

10. $p(x) = x^2 - x + 1, q(x) = 2x - 1$

11. $p(x) = 2x^2 + 1, q(x) = x^2 - 1$

12. $p(x) = -4x + 1, q(x) = x$

In Exercises 13–16, use the standard inner product on $\mathbb{R}^n$. Use the basis B and the Gram-Schmidt process to find an orthonormal basis for $\mathbb{R}^n$.

13. $B = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}$

14. $B = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$

15. $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} \right\}$

16. $B = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

In Exercises 17 and 18, use the inner product on $\mathcal{P}_2$ defined by

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$$

Use the given basis B and the Gram-Schmidt process to find an orthonormal basis for $\mathcal{P}_2$.

17. $B = \{x - 1, x + 2, x^2\}$

18. $B = \{x^2 - x, x, 2x + 1\}$

In Exercises 19–22, use the standard inner product on $\mathbb{R}^n$ to find an orthonormal basis for the subspace $\mathbf{span}(W)$.

19. $W = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\}$

20. $W = \left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$

21. $W = \left\{ \begin{bmatrix} -1 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}$

22. $W = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ -1 \end{bmatrix} \right\}$

In Exercises 23 and 24, use the inner product on $\mathcal{P}_3$ defined by

$$\langle p, q \rangle = \int_0^1 p(x)q(x)\,dx$$

to find an orthonormal basis for the subspace $\mathbf{span}(W)$.

23. $W = \{x, 2x + 1\}$

24. $W = \{1, x + 2, x^3 - 1\}$

Solution **a.** Let

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

Notice that $\mathbf{w}_1$ and $\mathbf{w}_2$ are orthogonal and hence by Theorem 5 of Sec. 6.2 are linearly independent. Thus, $\{\mathbf{w}_1, \mathbf{w}_2\}$ is a basis for W.

b. Now by Proposition 5, the vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

is in $W^{\perp}$ if and only if $\mathbf{v} \cdot \mathbf{w}_1 = 0$ and $\mathbf{v} \cdot \mathbf{w}_2 = 0$. This requirement leads to the linear system

$$\begin{cases} x \quad\;\; - z - w = 0 \\ \quad\; y - z + w = 0 \end{cases}$$

The two-parameter solution set for this linear system is

$$S = \left\{ \begin{bmatrix} s + t \\ s - t \\ s \\ t \end{bmatrix} \middle|\; s, t \in \mathbb{R} \right\}$$

The solution to this system, in vector form, provides a description of the orthogonal complement of W and is given by

$$W^{\perp} = \textbf{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}$$

Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal, by Theorem 5 of Sec. 6.2 they are linearly independent and hence a basis for $W^{\perp}$.

c. Let B be the set of vectors $B = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{v}_1, \mathbf{v}_2\}$. Since B is an orthogonal set of four vectors in $\mathbb{R}^4$, then by Corollary 1 of Sec. 6.2, B is a basis for $\mathbb{R}^4$. Dividing each of these vectors by its length, we obtain the (ordered) orthonormal basis for $\mathbb{R}^4$ given by

22. $W = \textbf{span}\left\{\begin{bmatrix}1\\2\end{bmatrix}\right\}$ $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$

23. $W = \textbf{span}\left\{\begin{bmatrix}3\\-1\end{bmatrix}\right\}$ $\mathbf{v} = \begin{bmatrix}0\\1\end{bmatrix}$

24. $W = \textbf{span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}\right\}$

$\mathbf{v} = \begin{bmatrix}1\\1\\1\end{bmatrix}$

25. $W = \textbf{span}\left\{\begin{bmatrix}1\\1\\-1\end{bmatrix}, \begin{bmatrix}-1\\2\\4\end{bmatrix}\right\}$

$\mathbf{v} = \begin{bmatrix}2\\1\\1\end{bmatrix}$

26. Show that if V is an inner product space, then $V^{\perp} = \{\mathbf{0}\}$ and $\{\mathbf{0}\}^{\perp} = V$.

27. Show that if W_1 and W_2 are finite dimensional subspaces of an inner product space and $W_1 \subset W_2$, then $W_2^{\perp} \subset W_1^{\perp}$.

28. Let $V = C^{(0)}[-1, 1]$ with the inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$$

and $W = \{f \in V \mid f(-x) = f(x)\}$.

a. Show that W is a subspace of V.

b. Show $W^{\perp} = \{f \in V \mid f(-x) = -f(x)\}$.

c. Verify that $W \cap W^{\perp} = \{\mathbf{0}\}$.

d. Let $g(x) = \frac{1}{2}[f(x) + f(-x)]$ and $h(x) = \frac{1}{2}[f(x) - f(-x)]$. Verify $g(-x) = g(x)$ and $h(-x) = -h(x)$, so every f can be written as the sum of a function in W and a function in $W^{\perp}$.

29. Let $V = M_{2\times 2}$ with the inner product

$$\langle A, B\rangle = \mathbf{tr}(B^t A)$$

Let $W = \{A \in V \mid A \text{ is symmetric}\}$.

a. Show that

$$W^{\perp} = \{A \in V \mid A \text{ is skew symmetric}\}$$

b. Show that every A in V can be written as the sum of matrices from W and $W^{\perp}$.

30. In $\mathbb{R}^2$ with the standard inner product, the transformation that sends a vector to the orthogonal projection onto a subspace W is a linear transformation. Let $W = \textbf{span}\left\{\begin{bmatrix}2\\1\end{bmatrix}\right\}$.

a. Find the matrix representation P relative to the standard basis for the orthogonal projection of $\mathbb{R}^2$ onto W.

b. Let $\mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$. Find $\text{proj}_W\mathbf{v}$ and verify the result is the same by applying the matrix P found in part (a).

c. Show $P^2 = P$.

31. If W is a finite dimensional subspace of an inner product space, show that $(W^{\perp})^{\perp} = W$.

6.5 ▶ Application: Least Squares Approximation

There are many applications in mathematics and science where an exact solution to a problem cannot be found, but an approximate solution exists that is sufficient to satisfy the demands of the application. Consider the problem of finding the equation of a line going through the points $(1, 2)$, $(2, 1)$, and $(3, 3)$. Observe from Fig. 1 that this problem has no solution as the three points are noncollinear.

This leads to the problem of finding the line that is the *best fit* for these three points based on some criteria for measuring *goodness of fit*. There are different ways of solving this new problem. One way, which uses calculus, is based on the idea

1. Find the eigenvalues and corresponding eigenvectors of A.
2. Since A is diagonalizable, there are n linearly independent eigenvectors. If necessary, use the Gram-Schmidt process to find an orthonormal set of eigenvectors.
3. Form the orthogonal matrix P with column vectors determined in Step 2.
4. The matrix $P^{-1}AP = P^tAP = D$ is a diagonal matrix.

EXAMPLE 3 Let

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

Find an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix.

Solution The characteristic equation for A is given by

$$\det(A - \lambda I) = -\lambda^3 + 3\lambda + 2 = -(\lambda - 2)(\lambda + 1)^2 = 0$$

Thus, the eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 2$. The corresponding eigenspaces are

$$V_{\lambda_1} = \textbf{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\} \quad \text{and} \quad V_{\lambda_2} = \textbf{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Let B be the set of vectors

$$B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Since B is a linearly independent set of three vectors, by Theorem 2 of Sec. 5.2, A is diagonalizable. To find an orthogonal matrix P which diagonalizes A, we use the Gram-Schmidt process on B. This was done in Example 3 of Sec. 6.3, yielding the orthonormal basis

$$B' = \left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\}$$

Now let P be the matrix given by

$$P = \begin{bmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & 0 & \frac{\sqrt{6}}{3} \end{bmatrix}$$

Observe that P is an orthogonal matrix with $P^{-1} = P^t$. Morevover,

$$P^tAP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Since the eigenvalues have opposite sign, the conic section C is a hyperbola. To describe the hyperbola, we first diagonalize A. Using the unit eigenvectors, the orthogonal matrix that diagonalizes A is

$$P = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \qquad \text{with} \qquad \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix} = P^tAP$$

Making the substitution $\mathbf{x} = P\mathbf{x}'$ in the equation $\mathbf{x}^tA\mathbf{x} + \mathbf{b}^t\mathbf{x} + f = 0$ gives

$$[x' \;\; y']\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} + [-4 \;\; -8]\begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} + 14 = 0$$

After simplification of this equation we obtain

$$-2(x')^2 - 4\sqrt{5}x' + 3(y')^2 + 14 = 0$$

that is,

$$-2[(x')^2 + 2\sqrt{5}(x')] + 3(y')^2 + 14 = 0$$

After completing the square on x', we obtain

$$-2[(x')^2 + 2\sqrt{5}(x') + 5] + 3(y')^2 = -14 - 10$$

that is,

$$\frac{(x' + \sqrt{5})^2}{12} - \frac{(y')^2}{8} = 1$$

This last equation describes a hyperbola with x' as the major axis. An additional transformation translating the x' axis allows us to simplify the result even further. If we let

$$x'' = x' + \sqrt{5} \qquad \text{and} \qquad y'' = y'$$

then the equation now becomes

$$\frac{(x'')^2}{12} - \frac{(y'')^2}{8} = 1$$

The graph is shown in Fig. 3.

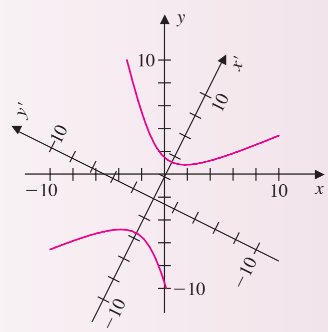


Figure 3

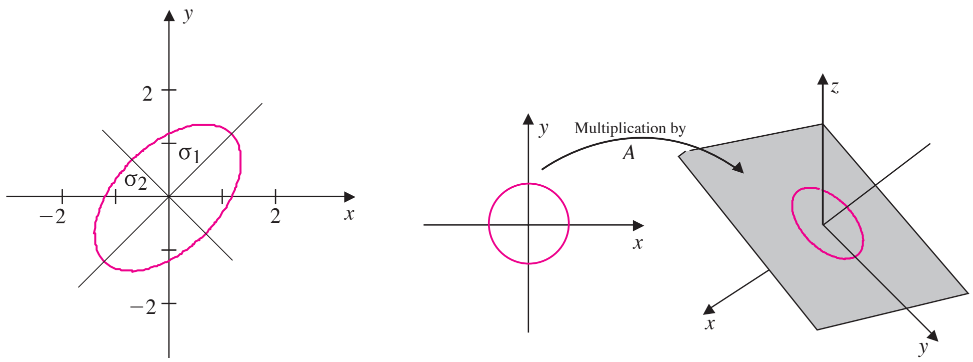


Figure 1

For certain matrices, some of the singular values may be zero. As an illustration, consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$. For this matrix, we have $\mathbf{col}(A) = \mathbf{span}\begin{bmatrix} 1 \\ 3 \end{bmatrix}$. The reduced row echelon form for A is the matrix $\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$, which has only one pivot column. Hence the rank of A is equal to 1. The eigenvalues of A^tA are $\lambda_1 = 50$ and $\lambda_2 = 0$ with corresponding unit eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

The singular values of A are given by $\sigma_1 = 5\sqrt{2}$ and $\sigma_2 = 0$. Now, multiplying $\mathbf{v}_1$ and $\mathbf{v}_2$ by A gives

$$A\mathbf{v}_1 = \begin{bmatrix} \sqrt{5} \\ 3\sqrt{5} \end{bmatrix} \qquad \text{and} \qquad A\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Observe that $A\mathbf{v}_1$ spans the one dimensional column space of A. In this case, the linear transformation $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined by $T(\mathbf{x}) = A\mathbf{x}$ maps the unit circle to the line segment

$$\left\{ t \begin{bmatrix} \sqrt{5} \\ 3\sqrt{5} \end{bmatrix} \middle| -1 \le t \le 1 \right\}$$

as shown in Fig. 2.

Singular Value Decomposition (SVD)

We now turn our attention to the problem of finding a singular value decomposition of an $m \times n$ matrix A.

THEOREM 18 **SVD** Let A be an $m \times n$ matrix of rank r, with r nonzero singular values $\sigma_1, \sigma_2, \ldots, \sigma_r$. Then there exists an $m \times n$ matrix Σ, an $m \times m$ orthogonal matrix

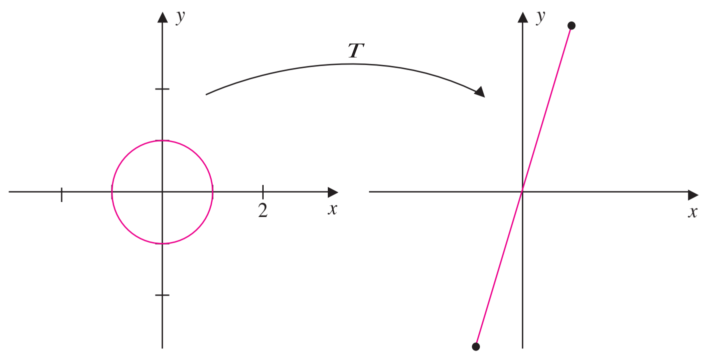


Figure 2

U, and an $n \times n$ orthogonal matrix V such that

$$A = U\Sigma V^t$$

Proof Since A^tA is an $n \times n$ symmetric matrix, by Theorem 14 of Sec. 6.6 there is an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of A^tA. Now by Theorem 17, $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is an orthogonal basis for $\mathbf{col}(A)$. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be the orthonormal basis for $\mathbf{col}(A)$, given by

$$\mathbf{u}_i = \frac{1}{\| A\mathbf{v}_i \|} A\mathbf{v}_i = \frac{1}{\sigma_i} A\mathbf{v}_i \qquad \text{for} \qquad i = 1, \ldots, r$$

Next, extend $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ to the orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ of $\mathbb{R}^m$. We can now define the orthogonal matrices V and U, using the vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$, respectively, as column vectors, so that

$$V = [\ \mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n\ ] \qquad \text{and} \qquad U = [\ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\ ]$$

Moreover, since $A\mathbf{v}_i = \sigma_i \mathbf{u}_i$, for $i = 1, \ldots, r$, then

$$AV = \begin{bmatrix} A\mathbf{v}_1 & \cdots & A\mathbf{v}_r & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \sigma_1\mathbf{u}_1 & \cdots & \sigma_r\mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

Now let Σ be the $m \times n$ matrix given by

$$\Sigma = \left[\begin{array}{cccc|ccc} \sigma_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \sigma_r & 0 & \ldots & 0 \\ \hline 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \ldots & \ldots & 0 & 0 & \ldots & 0 \end{array}\right]$$

Then

$$U\Sigma = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \end{bmatrix} \Sigma$$

$$= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \cdots & \sigma_r \mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}$$

$$= AV$$

Since V is orthogonal, then $V^t = V^{-1}$, and hence, $A = U\Sigma V^t$.

EXAMPLE 3 Find a singular value decomposition of the matrix

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

Solution A procedure for finding an SVD of A is included in the proof of Theorem 18. We present the solution as a sequence of steps.

Step 1. *Find the eigenvalues and corresponding orthonormal eigenvectors of A^tA and define the matrix V.*

The eigenvalues of the matrix

$$A^tA = \begin{bmatrix} 6 & -6 \\ -6 & 6 \end{bmatrix}$$

in decreasing order are given by $\lambda_1 = 12$ and $\lambda_2 = 0$. The corresponding orthonormal eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Since the column vectors of V are given by the orthonormal eigenvectors of A^tA, the matrix V is given by

$$V = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Step 2. *Find the singular values of A and define the matrix Σ.*

The singular values of A are the square roots of the eigenvalues of A^tA, so that

$$\sigma_1 = \sqrt{\lambda_1} = 2\sqrt{3} \quad \text{and} \quad \sigma_2 = \sqrt{\lambda_2} = 0$$

Since Σ has the same dimensions as A, then Σ is 3×2. In this case,

$$\Sigma = \begin{bmatrix} 2\sqrt{3} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

singular values) has the SVD $A = U\Sigma V^t$. That is,

$$
\begin{aligned}
A = U\Sigma V^t &= \begin{bmatrix} \sigma_1\mathbf{u}_1 & \cdots & \sigma_r\mathbf{u}_r & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} V^t \\
&= \sigma_1\mathbf{u}_1\mathbf{v}_1^t + \sigma_2\mathbf{u}_2\mathbf{v}_2^t + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^t \\
&= \sigma_1\left(\frac{1}{\sigma_1}A\mathbf{v}_1\right)\mathbf{v}_1^t + \sigma_2\left(\frac{1}{\sigma_2}A\mathbf{v}_2\right)\mathbf{v}_2^t + \cdots + \sigma_r\left(\frac{1}{\sigma_r}A\mathbf{v}_r\right)\mathbf{v}_r^t \\
&= (A\mathbf{v}_1)\mathbf{v}_1^t + (A\mathbf{v}_2)\mathbf{v}_2^t + \cdots + (A\mathbf{v}_r)\mathbf{v}_r^t
\end{aligned}
$$

Figure 4

Observe that each of the terms $A\mathbf{v}_i\mathbf{v}_i^t$ is a matrix of rank 1. Consequently, the sum of the first k terms of the last equation is a matrix of rank $k \le r$, which gives an approximation to the matrix A. This factorization of a matrix has application in many areas.

As an illustration of the utility of such an approximation, suppose that A is the 356×500 matrix, where each entry is a numeric value for a pixel, of the gray scale image of the surface of Mars shown in Fig. 4. A simple algorithm using the method above for approximating the image stored in the matrix A is given by the following:

1. Find the eigenvectors of the $n \times n$ symmetric matrix A^tA.
2. Compute $A\mathbf{v}_i$ for $i = 1, \ldots, k$, with $k \le r = \mathbf{rank}(A)$.
3. The matrix $(A\mathbf{v}_1)\mathbf{v}_1^t + (A\mathbf{v}_2)\mathbf{v}_2^t + \cdots + (A\mathbf{v}_k)\mathbf{v}_k^t$ is an approximation of the original image.

To transmit the kth approximation of the image and reproduce it back on earth requires the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of A^tA and the vectors $A\mathbf{v}_1, \ldots, A\mathbf{v}_k$.

The images in Fig. 5 are produced using matrices of ranks 1, 4, 10, 40, 80, and 100, respectively.

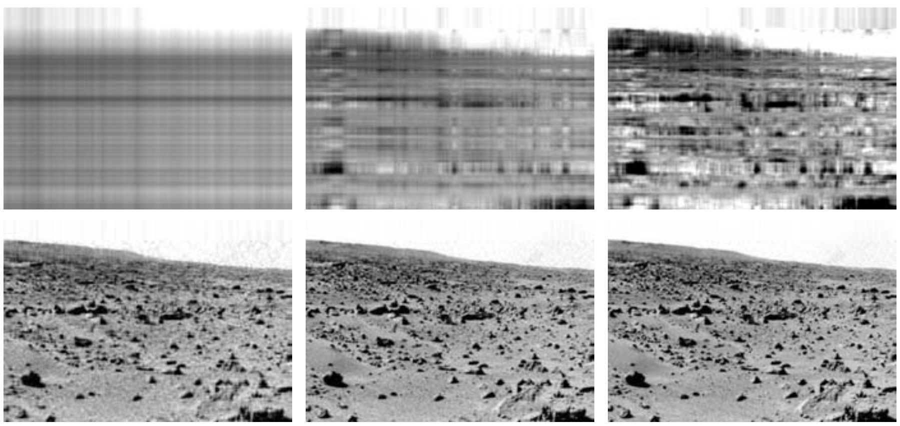


Figure 5

The storage requirements for each of the images are given in Table 1.

read as "Q implies P" or "P is necessary for Q." For example, let P be the statement *Mary lives in Iowa* and Q the statement that *Mary lives in the United States*. Then certainly $P \implies Q$ is a theorem since every resident of Iowa is a resident of the United States. But $Q \implies P$ is not a theorem since, for example, if Mary is a resident of California, then she is a resident of the United States but not a resident of Iowa. So the statement $Q \implies P$ is not always true given that Q is true. In terms of sets, if A is the set of residents of Iowa and B is the set of residents of the United States, then the statement P is *Mary is in* A and Q is *Mary is in* B. Then *Mary is in* A implies *Mary is in* B. It is also clear that if *Mary is in* $B\backslash A$, then *Mary is in* B does not imply that *Mary is in* A.

A statement that is equivalent to the theorem $P \implies Q$ is the **contrapositive** statement $\sim Q \implies \sim P$, that is, *not Q implies not P*. In the example above, if Mary is not a resident of the United States, then Mary is not a resident of Iowa. An equivalent formulation of the statement, in the terminology of sets, is that if Mary $\notin B$, then it implies Mary $\notin A$.

There are other statements in mathematics that require proof. **Lemmas** are preliminary results used to prove theorems, **propositions** are results not as important as theorems, and **corollaries** are special cases of a theorem. A statement that is not yet proven is called a **conjecture**. One of the most famous conjectures is the celebrated Riemann hypothesis. A single **counterexample** is enough to refute a false conjecture. For example, the statement *All lions have green eyes* is rendered invalid by the discovery of a single blue-eyed lion.

In this section we briefly introduce three main types of proof. A fourth type, called mathematical induction, is discussed in Sec. A.4.

Direct Argument

In a **direct argument**, a sequence of logical steps links the hypotheses P to the conclusion Q. Example 1 provides an illustration of this technique.

EXAMPLE 1 Prove that if p and q are odd integers, then $p+q$ is an even integer.

Solution To prove this statement with a direct argument, we assume that p and q are odd integers. Then there are integers m and n such that

$$p = 2m + 1 \qquad \text{and} \qquad q = 2n + 1$$

Adding p and q gives

$$\begin{aligned} p + q &= 2m + 1 + 2n + 1 \\ &= 2(m + n) + 2 \\ &= 2(m + n + 1) \end{aligned}$$

Since $p+q$ is a multiple of 2, it is an even integer.

Since for every natural number $n \geq 1$ it is also the case that $n + 1 \geq 2$, we have

$$(n+1)! \geq (n+1)2^{n-1} \geq 2 \cdot 2^{n-1} = 2^n$$

Consequently, the statement $n! \geq 2^{n-1}$ is true for every natural number n.

EXAMPLE 4 For any natural number n, find the sum of the odd natural numbers from 1 to $2n - 1$.

Solution The first five cases are given in Table 3.

Table 3

n	$2n-1$	$1+3+\cdots+(2n-1)$
1	1	1
2	3	$1+3=4$
3	5	$1+3+5=9$
4	7	$1+3+5+7=16$
5	9	$1+3+5+7+9=25$

The data in Table 3 suggest that for each $n \geq 1$,

$$1 + 3 + 5 + 7 + \cdots + (2n-1) = n^2$$

Starting with the case for $n = 1$, we see that the left-hand side is 1 and the expression on the right is $1^2 = 1$. Hence, the statement holds when $n = 1$. Next, we assume that $1 + 3 + 5 + \cdots + (2n-1) = n^2$. For the next case when the index is $n + 1$, we consider the sum

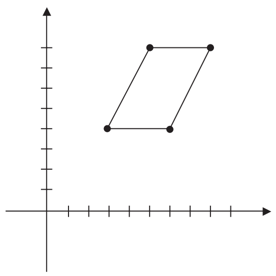
$$1 + 3 + 5 + \cdots + (2n-1) + [2(n+1) - 1] = 1 + 3 + 5 + \cdots + (2n-1) + (2n+1)$$

Using the inductive hypothesis, we get

$$\begin{aligned} \underbrace{1 + 3 + 5 + \cdots + (2n-1)}_{n^2} + [2(n+1) - 1] &= n^2 + (2n+1) \\ &= n^2 + 2n + 1 \\ &= (n+1)^2 \end{aligned}$$

Therefore, by induction the statement holds for all natural numbers.

d.



25. a. $[I]_{B_1}^{B_2} = \begin{bmatrix} -1 & -1 & 0 \\ 2 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$

b. $[2\mathbf{u}_1 - 2\mathbf{u}_2 + \mathbf{u}_3]_{B_2} = [I]_{B_1}^{B_2} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$

Section 3.5

1. a. $y_1 = e^{2x}, y_2 = e^{3x}$

b. $W[y_1, y_2](x) = \begin{vmatrix} e^{2x} & e^{3x} \\ 2e^{2x} & 3e^{3x} \end{vmatrix} = e^{5x} > 0$ for all x

c. $y(x) = C_1 e^{2x} + C_2 e^{3x}$

3. a. $y_1 = e^{-2x}, y_2 = xe^{-2x}$

b. $W[y_1, y_2](x) = \begin{vmatrix} e^{-2x} & xe^{-2x} \\ -2e^{-2x} & e^{-2x} - 2xe^{-2x} \end{vmatrix} =$
$e^{-4x} > 0$ for all x

c. $y(x) = C_1 e^{-2x} + C_2 xe^{-2x}$

5. $y(x) = e^x + 2xe^x$

7. a. $y_c(x) = C_1 e^{3x} + C_2 e^x$

b. $a = 1, b = 3, c = 4$

9. $y(x) = \frac{1}{4} \cos(8x)$

Review Exercises Chapter 3

1. $k \neq 69$

3. a. Since S is closed under vector addition and scalar multiplication, S is a subspace of $M_{2\times 2}$.

b. Yes, let $a = 3, b = -2, c = 0$.

c. $B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \right\}$

d. The matrix $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ is not in S.

5. a. The set T is a basis since it is a linearly independent set of three vectors in the three-dimensional vector space V.

b. The set W is not a basis for V since it is not linearly independent.

7. Since $\mathbf{v}_1$ can be written as

$$\mathbf{v}_1 = \left(\frac{-c_2}{c_1}\right)\mathbf{v}_2 + \left(\frac{-c_3}{c_1}\right)\mathbf{v}_3 + \cdots + \left(\frac{-c_n}{c_1}\right)\mathbf{v}_n$$

then

$$V = \mathbf{span}\{\mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$$

9. a. The set $B = \{\mathbf{u}, \mathbf{v}\}$ is a basis for $\mathbb{R}^2$ since it is linearly independent. To see this, consider

$$a\mathbf{u} + b\mathbf{v} = \mathbf{0}$$

Now take the dot product of both sides with first $\mathbf{u}$, then $\mathbf{v}$, to show that $a = b = 0$.

b. If $[\mathbf{w}]_B = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, then

$$\alpha = \frac{\begin{vmatrix} x & v_1 \\ y & v_2 \end{vmatrix}}{\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}} = \frac{xv_2 - yv_1}{u_1v_2 - v_1u_2}$$

and

$$\beta = \frac{\begin{vmatrix} u_1 & x \\ u_2 & y \end{vmatrix}}{\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}} = \frac{yu_1 - xu_2}{u_1v_2 - v_1u_2}.$$

Chapter Test: Chapter 3

1.	F	**2.**	T
3.	F	**4.**	F
5.	T	**6.**	F
7.	F	**8.**	F
9.	T	**10.**	T
11.	T	**12.**	T
13.	T	**14.**	T
15.	T	**16.**	F
17.	F	**18.**	F
19.	T	**20.**	T
21.	T	**22.**	T
23.	T	**24.**	T
25.	T	**26.**	F
27.	T	**28.**	T
29.	T	**30.**	F
31.	T	**32.**	F
33.	T	**34.**	F
35.	T		

e. $[T^n]_B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}^n$

9. Since $T^2 - T + I = \mathbf{0}, T - T^2 = I$. Then

$$\begin{aligned}(T \circ (I - T))(\mathbf{v}) &= T((I - T)(\mathbf{v})) = T(\mathbf{v} - T(\mathbf{v})) \\ &= T(\mathbf{v}) - T^2(\mathbf{v}) = I(\mathbf{v}) = \mathbf{v}\end{aligned}$$

Chapter Test: Chapter 4

1. F	**2.** F
3. T	**4.** T
5. T	**6.** F
7. F	**8.** F
9. T	**10.** F
11. T	**12.** F
13. T	**14.** T
15. T	**16.** T
17. T	**18.** T
19. T	**20.** T
21. F	**22.** F
23. T	**24.** F
25. T	**26.** F
27. F	**28.** T
29. F	**30.** T
31. F	**32.** F
33. T	**34.** F
35. F	**36.** T
37. T	**38.** T
39. F	**40.** T

Chapter 5

Section 5.1

1. $\lambda = 3$

3. $\lambda = 0$

5. $\lambda = 1$

7. a. $\lambda^2 + 5\lambda = 0$

b. $\lambda_1 = 0, \lambda_2 = -5$

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

d. $\begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix}$;

$\begin{bmatrix} -2 & 2 \\ 3 & -3 \end{bmatrix}\begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ -15 \end{bmatrix} = -5\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

9. a. $(\lambda - 1)^2 = 0$

b. $\lambda_1 = 1$

c. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

d. $\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

11. a. $(\lambda + 1)^2(\lambda - 1) = 0$

b. $\lambda_1 = -1, \lambda_2 = 1$

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

d. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

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

13. a. $(\lambda - 2)(\lambda - 1)^2 = 0$

b. $\lambda_1 = 2, \lambda_2 = 1$

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

d. $\begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & -1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

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

15. a. $(\lambda + 1)(\lambda - 2)(\lambda + 2)(\lambda - 4) = 0$

b. $\lambda_1 = -1, \lambda_2 = 2, \lambda_3 = -2, \lambda_4 = 4$

c. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$

$\mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

39. Since A and B are matrix representations for the same linear operator, they are similar. Let $A = Q^{-1}BQ$. The matrix A is diagonalizable if and only if $D = P^{-1}AP$ for some invertible matrix P and diagonal matrix D. Then

$$D = P^{-1}(Q^{-1}BQ)P = (QP)^{-1}B(QP)$$

so B is diagonalizable. The proof of the converse is identical.

Section 5.3

1.
$$y_1(t) = [y_1(0) + y_2(0)]e^{-t} - y_2(0)e^{-2t}$$
$$y_2(t) = y_2(0)e^{-2t}$$

3.
$$\begin{aligned} y_1(t) &= \frac{1}{2}[y_1(0) - y_2(0)]e^{4t} \\ &\quad + \frac{1}{2}[y_1(0) + y_2(0)]e^{-2t} \\ y_2(t) &= \frac{1}{2}[-y_1(0) + y_2(0)]e^{4t} \\ &\quad + \frac{1}{2}[y_1(0) + y_2(0)]e^{-2t} \end{aligned}$$

5.
$$\begin{aligned} y_1(t) &= [2y_1(0) + y_2(0) + y_3(0)]e^{-t} \\ &\quad + [-y_1(0) - y_2(0) - y_3(0)]e^{2t} \\ y_2(t) &= [-2y_1(0) - y_2(0) - 2y_3(0)]e^{-t} \\ &\quad + 2[y_1(0) + y_2(0) + y_3(0)]e^{2t} \\ y_3(t) &= [-2y_1(0) - y_2(0) - y_3(0)]e^{-t} \\ &\quad + [2y_1(0) + y_2(0) + 2y_3(0)]e^{t} \end{aligned}$$

7. $y_1(t) = e^{-t}, y_2(t) = -e^{-t}$

9. **a.** $y_1'(t) = -\frac{1}{60}y_1 + \frac{1}{120}y_2,$

$y_2'(t) = \frac{1}{60}y_1 - \frac{1}{120}y_2$

$y_1(0) = 12, \; y_2(0) = 0$

b. $y_1(t) = 4 + 8e^{-\frac{1}{40}t}, y_2(t) = 8 - 8e^{-\frac{1}{40}t}$

c. $\lim_{t\to\infty} y_1(t) = 4, \lim_{t\to\infty} y_2(t) = 8$
The 12 lb of salt will be evenly distributed in a ratio of 1:2 between the two tanks.

Section 5.4

1. **a.** $T = \begin{bmatrix} 0.85 & 0.08 \\ 0.15 & 0.92 \end{bmatrix}$

b. $T^{10}\begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \approx \begin{bmatrix} 0.37 \\ 0.63 \end{bmatrix}$

c. $\begin{bmatrix} 0.35 \\ 0.65 \end{bmatrix}$

3. $T = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.4 & 0.4 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{bmatrix}$

$T^3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \approx \begin{bmatrix} 0.36 \\ 0.35 \\ 0.29 \end{bmatrix}$

$T^{10}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \approx \begin{bmatrix} 0.33 \\ 0.33 \\ 0.33 \end{bmatrix}$

5. $T = \begin{bmatrix} 0.5 & 0 & 0 \\ 0.5 & 0.75 & 0 \\ 0 & 0.25 & 1 \end{bmatrix}$

The steady-state probability vector is $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, and hence the disease will not be eradicated.

7. **a.** $T = \begin{bmatrix} 0.33 & 0.25 & 0.17 & 0.25 \\ 0.25 & 0.33 & 0.25 & 0.17 \\ 0.17 & 0.25 & 0.33 & 0.25 \\ 0.25 & 0.17 & 0.25 & 0.33 \end{bmatrix}$

b. $T\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.5(0.16)^n + 0.25 \\ 0.25 \\ -0.5(0.16)^n + 0.25 \\ 0.25 \end{bmatrix}$

c. $\begin{bmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{bmatrix}$

9. Eigenvalues of T: $\lambda_1 = -q + p + 1, \lambda_2 = 1$, with corresponding eigenvectors $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} q/p \\ 1 \end{bmatrix}$.

The steady-state probability vector is

$$\frac{1}{1 + q/p}\begin{bmatrix} q/p \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{q}{p+q} \\ \frac{p}{p+q} \end{bmatrix}.$$

Review Exercises Chapter 5

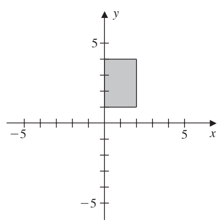
1. **a.** $\begin{bmatrix} a & b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a+b \\ a+b \end{bmatrix}$

$= (a+b)\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

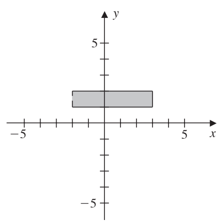
b. $\lambda_1 = a + b, \lambda_2 = a - b$

c. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

17.



19.



21. $(A \cap B) \cap C = \{5\} = A \cap (B \cap C)$

23. $A \cap (B \cup C) = \{1, 2, 5, 7\} = (A \cap B) \cup (A \cap C)$

25. $A \backslash (B \cup C) = \{3, 9, 11\} = (A \backslash B) \cap (A \backslash C)$

Section A.2

1. Since for each first coordinate there is a unique second coordinate, f is a function.

3. Since there is no x such that $f(x) = 14$, the function is not onto. The range of f is the set $\{-2, -1, 3, 9, 11\}$.

5. $f^{-1}(\{-2\}) = \{1, 4\}$

7. Since f is not one-to-one, f does not have an inverse.

9. $\{(1, -2), (2, -1), (3, 3), (4, 5), (5, 9), (6, 11)\}$

11. $f(A \cup B) = f((-3, 7)) = [0, 49)$

$f(A) \cup f(B) = [0, 25] \cup [0, 49) = [0, 49)$

13. $f(A \cap B) = f(\{0\}) = \{0\}$

$f(A) \cap f(B) = [0, 4] \cap [0, 4] = [0, 4]$

Therefore, $f(A \cap B) \subset f(A) \cap f(B)$, but $f(A \cap B) \neq f(A) \cap f(B)$.

15. $f^{-1}(x) = \frac{x-b}{a}$

17. If n is odd, then $f^{(n)}(x) = -x + c$. If n is even, then $f^{(n)}(x) = x$.

19. **a.** To show that f is one-to-one, we have

$$\begin{aligned} e^{2x_1 - 1} &= e^{2x_2 - 1} \\ \Leftrightarrow 2x_1 - 1 &= 2x_2 - 1 \\ \Leftrightarrow x_1 &= x_2 \end{aligned}$$

b. Since the exponential function is always positive, f is not onto $\mathbb{R}$.

c. Define $g : \mathbb{R} \to (0, \infty)$ by $g(x) = e^{2x-1}$.

d. $g^{-1}(x) = \frac{1}{2}(1 + \ln x)$.

21. **a.** To show that f is one-to-one, we have $2n_1 = 2n_2$ if and only if $n_1 = n_2$.

b. Since every image is an even number, the range of f is a proper subset of $\mathbb{N}$.

c. $f^{-1}(E) = \mathbb{N}$; $f^{-1}(O) = \phi$

23. **a.** $f(A) = \{2k + 1 \mid k \in \mathbb{Z}\}$

b. $f(B) = \{2k + 1 \mid k \in \mathbb{Z}\}$

c. $f^{-1}(\{0\}) = \{(m, n) \mid n = -2m\}$

d. $f^{-1}(E) = \{(m, n) \mid n \text{ is even}\}$

e. $f^{-1}(O) = \{(m, n) \mid n \text{ is odd}\}$

f. Since $f((1, -2)) = 0 = f((0, 0))$, then f is not one-to-one.

g. If $z \in \mathbb{Z}$, let $m = 0$ and $n = z$ so that $f(m, n) = z$.

Section A.3

1. If the side is x, then $h^2 = x^2 + x^2 = 2x^2$, so $h = \sqrt{2}x$.

3. If the side is x, then the height is $h = \frac{\sqrt{3}}{2}x$, so the area is $A = \frac{1}{2}x\frac{\sqrt{3}}{2}x = \frac{\sqrt{3}}{4}x^2$.

5. If a divides b, there is some k such that $ak = b$; and if b divides c, there is some ℓ such that $b\ell = c$. Then $c = b\ell = (ak)\ell = (k\ell)a$, so a divides c.

7. If n is odd, there is some k such that $n = 2k + 1$. Then $n^2 = (2k + 1)^2 = 2(2k^2 + k) + 1$, so n^2 is odd.

9. If $b = a + 1$, then $(a + b)^2 = (2a + 1)^2 = 2(2a^2 + 2a) + 1$, so $(a + b)^2$ is odd.

11. Let $m = 2$ and $n = 3$. Then $m^2 + n^2 = 13$, which is not divisible by 4.

13. Contrapositive: Suppose n is even, so there is some k such that $n = 2k$. Then $n^2 = 4k^2$, so n^2 is even.

15. Contrapositive: Suppose $p = q$. Then $\sqrt{pq} = \sqrt{p^2} = p = (p + q)/2$.

17. Contrapositive: Suppose $x > 0$. If $\epsilon = x/2 > 0$, then $x > \epsilon$.

19. Contradiction: Suppose $\sqrt[3]{2} = p/q$ such that p and q have no common factors. Then $2q^3 = p^3$, so p^3 is even and hence p is even. This gives that q is also even, which contradicts the assumption that p and q have no common factors.

21. If $7xy \leq 3x^2 + 2y^2$, then $3x^2 - 7xy + 2y^2 = (3x - y)(x - 2y) \geq 0$. There are two cases: either both factors are greater than or equal to 0, or both are less

Base case: $n = 1 : 2 = 1(2)$
Inductive hypothesis: Assume the summation formula holds for the natural number n.
Consider

$$\begin{aligned} &2 + 4 + 6 + \cdots + 2n + 2(n+1) \\ &\quad = n(n+1) + 2(n+1) \\ &\quad = (n+1)(n+2) \end{aligned}$$

13. Base case: $n = 5 : 32 = 2^5 > 25 = 5^2$
Inductive hypothesis: Assume $2^n > n^2$ holds for the natural number n.
Consider $2^{n+1} = 2(2^n) > 2n^2$. But since $2n^2 - (n+1)^2 = n^2 - 2n - 1 = (n-1)^2 - 2 > 0$, for all $n \geq 5$, we have $2^{n+1} > (n+1)^2$.

15. Base case: $n = 1 : 1^2 + 1 = 2$, which is divisible by 2.
Inductive hypothesis: Assume $n^2 + n$ is divisible by 2.
Consider $(n+1)^2 + (n+1) = n^2 + n + 2n + 2$. By the inductive hypothesis, $n^2 + n$ is divisible by 2, so since both terms on the right are divisible by 2, then $(n+1)^2 + (n+1)$ is divisible by 2. Alternatively, observe that $n^2 + n = n(n+1)$, which is the product of consecutive integers and is therefore even.

17. Base case: $n = 1 : 1 = \frac{r-1}{r-1}$
Inductive hypothesis: Assume the formula holds for the natural number n.
Consider

$$\begin{aligned} &1 + r + r^2 + \cdots + r^{n-1} + r^n \\ &\quad = \frac{r^n - 1}{r - 1} + r^n \\ &\quad = \frac{r^n - 1 + r^n(r-1)}{r-1} \\ &\quad = \frac{r^{n+1} - 1}{r - 1} \end{aligned}$$

19. Base case: $n = 2 : A \cap (B_1 \cup B_2) = (A \cap B_1) \cup (A \cap B_2)$, by Theorem 1 of Sec. A.1
Inductive hypothesis: Assume the formula holds for the natural number n.
Consider

$$\begin{aligned} &A \cap (B_1 \cup B_2 \cup \cdots \cup B_n \cup B_{n+1}) \\ &\quad = A \cap [(B_1 \cup B_2 \cup \cdots \cup B_n) \cup B_{n+1}] \\ &\quad = [A \cap (B_1 \cup B_2 \cup \cdots \cup B_n)] \cup (A \cap B_{n+1}) \\ &\quad = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n) \cup (A \cap B_{n+1}) \end{aligned}$$

21.

$$\begin{aligned} \binom{n}{r} &= \frac{n!}{r!(n-r)!} \\ &= \frac{n!}{(n-r)!(n-(n-r))!} \\ &= \binom{n}{n-r} \end{aligned}$$

23. By the binomial theorem,

$$2^n = (1+1)^n = \sum_{k=0}^{n} \binom{n}{k}$$

N