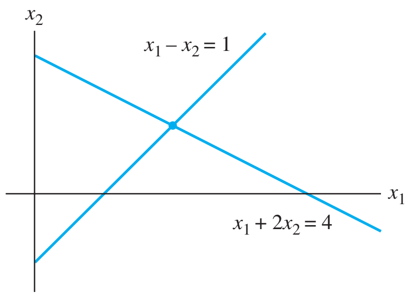
1.1 EXERCISES

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

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

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

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



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

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

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

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

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

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

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

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

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

Solve the systems in Exercises 11–14.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

23. a. Every elementary row operation is reversible.

b. A 5×6 matrix has six rows.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

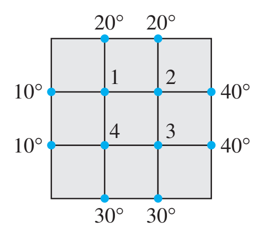
30. $\begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & -5 & 10 \end{bmatrix}, \begin{bmatrix} 1 & 3 & -4 \\ 0 & -2 & 6 \\ 0 & 1 & -2 \end{bmatrix}$

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

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

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

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



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

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

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

SOLUTIONS TO PRACTICE PROBLEMS

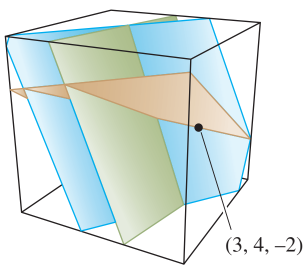
1. a. For "hand computation," the best choice is to interchange equations 3 and 4. Another possibility is to multiply equation 3 by $1/5$. Or, replace equation 4 by its sum with $-1/5$ times row 3. (In any case, do not use the x_2 in equation 2 to eliminate the $4x_2$ in equation 1. Wait until a triangular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

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

2. The system corresponding to the augmented matrix is

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

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



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

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

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

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

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

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

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

1.2 ROW REDUCTION AND ECHELON FORMS

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

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

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

THEOREM 2

Existence and Uniqueness Theorem

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

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

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

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

USING ROW REDUCTION TO SOLVE A LINEAR SYSTEM

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

PRACTICE PROBLEMS

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

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

2. Find the general solution of the system

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

1.2 EXERCISES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

WEB

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

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

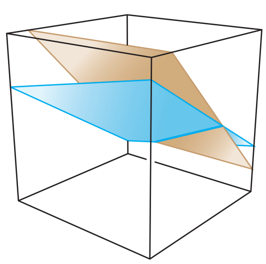
34. **[M]** In a wind tunnel experiment, the force on a projectile due to air resistance was measured at different velocities:

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

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

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

SOLUTIONS TO PRACTICE PROBLEMS



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

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

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

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

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

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

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

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

2. Row reduce the system's augmented matrix:

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

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

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

1.3 VECTOR EQUATIONS

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

Vectors in $\mathbb{R}^2$

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

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

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

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

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

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

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

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

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

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

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

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

Linear Combinations in Applications

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

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

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

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

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

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

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

SOLUTION

a. Compute

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

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

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

PRACTICE PROBLEMS

1. Prove that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for any $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$.
2. For what value(s) of h will $\mathbf{y}$ be in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if

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

1.3 EXERCISES

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

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

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

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

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

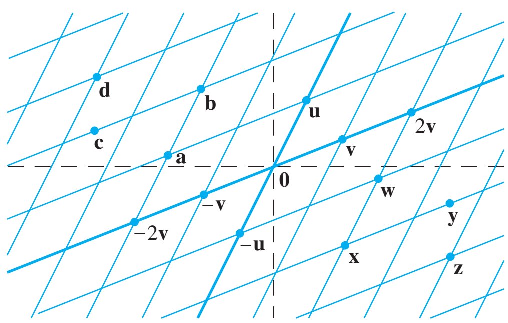
4. $\mathbf{u}$ and $\mathbf{v}$ as in Exercise 2

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

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

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

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



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

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

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

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

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

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

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

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

In Exercises 13 and 14, determine if $\mathbf{b}$ is a linear combination of the vectors formed from the columns of the matrix A.

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

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

15. Let $\mathbf{a}_1=\begin{bmatrix}1\\3\\-1\end{bmatrix}$, $\mathbf{a}_2=\begin{bmatrix}-5\\-8\\2\end{bmatrix}$, and $\mathbf{b}=\begin{bmatrix}3\\-5\\h\end{bmatrix}$. For what value(s) of h is $\mathbf{b}$ in the plane spanned by $\mathbf{a}_1$ and $\mathbf{a}_2$?

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

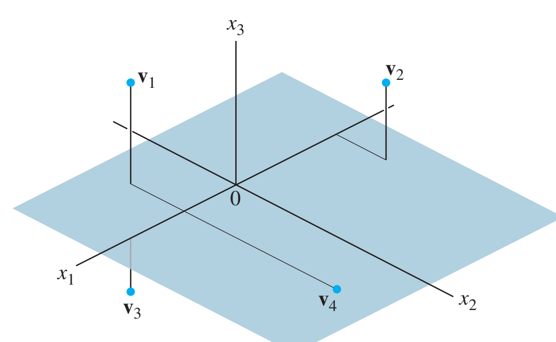
$$m = m_1 + \cdots + m_k$$

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

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

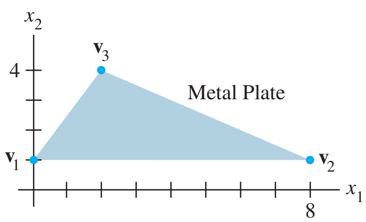
Compute the center of gravity of the system consisting of the following point masses (see the figure):

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



30. Let $\mathbf{v}$ be the center of mass of a system of point masses located at $\mathbf{v}_1, \ldots, \mathbf{v}_k$ as in Exercise 29. Is $\mathbf{v}$ in Span $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$? Explain.

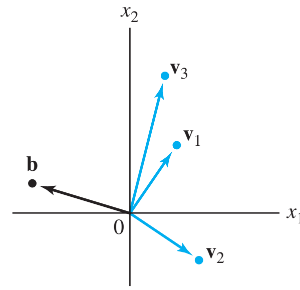
31. A thin triangular plate of uniform density and thickness has vertices at $\mathbf{v}_1 = (0, 1)$, $\mathbf{v}_2 = (8, 1)$, and $\mathbf{v}_3 = (2, 4)$, as in the figure below, and the mass of the plate is 3 g.



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

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

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



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

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

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

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

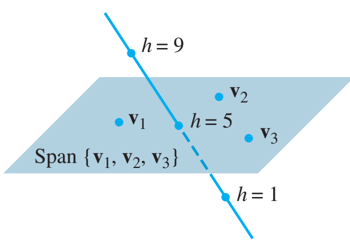
a. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$

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

SOLUTIONS TO PRACTICE PROBLEMS

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

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



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

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

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

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

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

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

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

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

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

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

PRACTICE PROBLEMS

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

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

1.4 EXERCISES

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

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

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

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

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

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

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

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

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

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

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

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

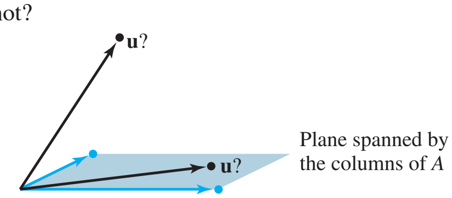
10. $\begin{aligned} 4x_1 - x_2 &= 8 \\ 5x_1 + 3x_2 &= 2 \\ 3x_1 - x_2 &= 1 \end{aligned}$

Given A and $\mathbf{b}$ in Exercises 11 and 12, write the augmented matrix for the linear system that corresponds to the matrix equation $A\mathbf{x} = \mathbf{b}$. Then solve the system and write the solution as a vector.

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

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

13. Let $\mathbf{u} = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$ and $A = \begin{bmatrix} 3 & -5 \\ -2 & 6 \\ 1 & 1 \end{bmatrix}$. Is $\mathbf{u}$ in the plane in $\mathbb{R}^3$ spanned by the columns of A? (See the figure.) Why or why not?



Where is $\mathbf{u}$?

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

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

16. Repeat the requests from Exercise 15 with

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

SG Mastering Linear Algebra Concepts: Span 1–18

WEB

SOLUTIONS TO PRACTICE PROBLEMS

1. The matrix equation

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

is equivalent to the vector equation

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

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

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

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

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

$$= \begin{bmatrix} 3 \\ 11 \end{bmatrix} + \begin{bmatrix} 19 \\ -4 \end{bmatrix} = \begin{bmatrix} 22 \\ 7 \end{bmatrix}$$

THEOREM 6

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

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

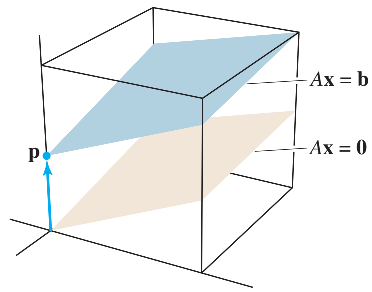


FIGURE 6 Parallel solution sets of $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{0}$.

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

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

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

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

PRACTICE PROBLEMS

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

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

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

1.5 EXERCISES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

17. Follow the method of Example 3 to describe the solutions of the following system in parametric vector form. Also, give a geometric description of the solution set and compare it to that in Exercise 5.

$$\begin{aligned} 2x_1 + 2x_2 + 4x_3 &= 8 \\ -4x_1 - 4x_2 - 8x_3 &= -16 \\ -3x_2 - 3x_3 &= 12 \end{aligned}$$

18. As in Exercise 17, describe the solutions of the following system in parametric vector form, and provide a geometric comparison with the solution set in Exercise 6.

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

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

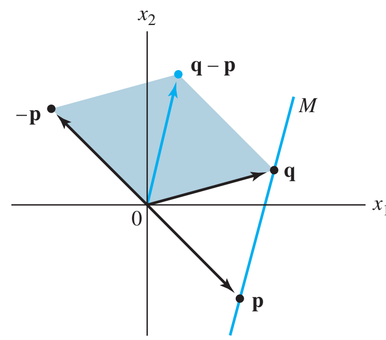
19. $\mathbf{a} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$

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

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

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

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



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

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

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

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

25. Prove Theorem 6:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

SOLUTIONS TO PRACTICE PROBLEMS

1. Row reduce the augmented matrix:

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

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

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

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

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

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

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

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

1.6 APPLICATIONS OF LINEAR SYSTEMS

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

A Homogeneous System in Economics

WEB

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

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

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

The following example shows how to find the equilibrium prices.

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

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

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

Also, the total flow into the network $(500 + 300 + 100 + 400)$ equals the total flow out of the network $(300 + x_3 + 600)$, which simplifies to $x_3 = 400$. Combine this equation with a rearrangement of the first four equations to obtain the following system of equations:

$$\begin{aligned} x_1 + x_2 \qquad\qquad\qquad &= 800 \\ x_2 - x_3 + x_4 \qquad &= 300 \\ x_4 + x_5 &= 500 \\ x_1 \qquad\qquad\qquad + x_5 &= 600 \\ x_3 \qquad\qquad &= 400 \end{aligned}$$

Row reduction of the associated augmented matrix leads to

$$\begin{aligned} x_1 \qquad\qquad + x_5 &= 600 \\ x_2 \qquad - x_5 &= 200 \\ x_3 \qquad\quad &= 400 \\ x_4 + x_5 &= 500 \end{aligned}$$

The general flow pattern for the network is described by

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ is free} \end{cases}$$

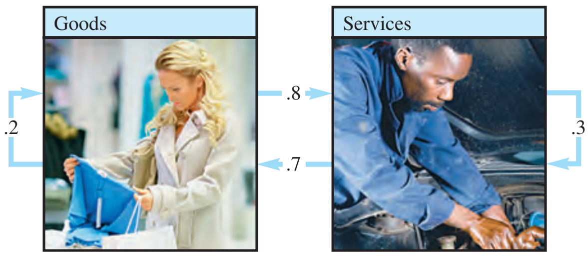
A negative flow in a network branch corresponds to flow in the direction opposite to that shown on the model. Since the streets in this problem are one-way, none of the variables here can be negative. This fact leads to certain limitations on the possible values of the variables. For instance, $x_5 \leq 500$ because x_4 cannot be negative. Other constraints on the variables are considered in Practice Problem 2. ■

PRACTICE PROBLEMS

1. Suppose an economy has three sectors: Agriculture, Mining, and Manufacturing. Agriculture sells 5% of its output to Mining and 30% to Manufacturing, and retains the rest. Mining sells 20% of its output to Agriculture and 70% to Manufacturing, and retains the rest. Manufacturing sells 20% of its output to Agriculture and 30% to Mining, and retains the rest. Determine the exchange table for this economy, where the columns describe how the output of each sector is exchanged among the three sectors.
2. Consider the network flow studied in Example 2. Determine the possible range of values of x_1 and x_2. [*Hint:* The example showed that $x_5 \leq 500$. What does this imply about x_1 and x_2? Also, use the fact that $x_5 \geq 0$.]

1.6 EXERCISES

1. Suppose an economy has only two sectors: Goods and Services. Each year, Goods sells 80% of its output to Services and keeps the rest, while Services sells 70% of its output to Goods and retains the rest. Find equilibrium prices for the annual outputs of the Goods and Services sectors that make each sector's income match its expenditures.



2. Find another set of equilibrium prices for the economy in Example 1. Suppose the same economy used Japanese yen instead of dollars to measure the values of the various sectors' outputs. Would this change the problem in any way? Discuss.

3. Consider an economy with three sectors: Fuels and Power, Manufacturing, and Services. Fuels and Power sells 80% of its output to Manufacturing, 10% to Services, and retains the rest. Manufacturing sells 10% of its output to Fuels and Power, 80% to Services, and retains the rest. Services sells 20% to Fuels and Power, 40% to Manufacturing, and retains the rest.
   a. Construct the exchange table for this economy.
   b. Develop a system of equations that leads to prices at which each sector's income matches its expenses. Then write the augmented matrix that can be row reduced to find these prices.
   c. [**M**] Find a set of equilibrium prices when the price for the Services output is 100 units.

4. Suppose an economy has four sectors: Mining, Lumber, Energy, and Transportation. Mining sells 10% of its output to Lumber, 60% to Energy, and retains the rest. Lumber sells 15% of its output to Mining, 50% to Energy, 20% to Transportation, and retains the rest. Energy sells 20% of its output to Mining, 15% to Lumber, 20% to Transportation, and retains the rest. Transportation sells 20% of its output to Mining, 10% to Lumber, 50% to Energy, and retains the rest.
   a. Construct the exchange table for this economy.
   b. [**M**] Find a set of equilibrium prices for the economy.

5. An economy has four sectors: Agriculture, Manufacturing, Services, and Transportation. Agriculture sells 20% of its output to Manufacturing, 30% to Services, 30% to Transportation, and retains the rest. Manufacturing sells 35% of its output to Agriculture, 35% to Services, 20% to Transportation, and retains the rest. Services sells 10% of its output to Agriculture, 20% to Manufacturing, 20% to Transportation, and retains the rest. Transportation sells 20% of its output to Agriculture, 30% to Manufacturing, 20% to Services, and retains the rest.
   a. Construct the exchange table for this economy.
   b. [**M**] Find a set of equilibrium prices for the economy if the value of Transportation is \$10.00 per unit.
   c. The Services sector launches a successful "eat farm fresh" campaign, and increases its share of the output from the Agricultural sector to 40%, whereas the share of Agricultural production going to Manufacturing falls to 10%. Construct the exchange table for this new economy.
   d. [**M**] Find a set of equilibrium prices for this new economy if the value of Transportation is still \$10.00 per unit. What effect has the "eat farm fresh" campaign had on the equilibrium prices for the sectors in this economy?

Balance the chemical equations in Exercises 6–11 using the vector equation approach discussed in this section.

6. Aluminum oxide and carbon react to create elemental aluminum and carbon dioxide:

   $Al_2O_3 + C \rightarrow Al + CO_2$

   [For each compound, construct a vector that lists the numbers of atoms of aluminum, oxygen, and carbon.]

7. Alka-Seltzer contains sodium bicarbonate ($NaHCO_3$) and citric acid ($H_3C_6H_5O_7$). When a tablet is dissolved in water, the following reaction produces sodium citrate, water, and carbon dioxide (gas):

   $NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + H_2O + CO_2$

8. Limestone, $CaCO_3$, neutralizes the acid, H_3O, in acid rain by the following unbalanced equation:

   $H_3O + CaCO_3 \rightarrow H_2O + Ca + CO_2$

9. Boron sulfide reacts violently with water to form boric acid and hydrogen sulfide gas (the smell of rotten eggs). The unbalanced equation is

   $B_2S_3 + H_2O \rightarrow H_3BO_3 + H_2S$

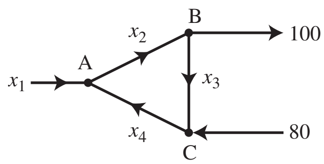
10. [**M**] If possible, use exact arithmetic or a rational format for calculations in balancing the following chemical reaction:

    $PbN_6 + CrMn_2O_8 \rightarrow Pb_3O_4 + Cr_2O_3 + MnO_2 + NO$

11. [**M**] The chemical reaction below can be used in some industrial processes, such as the production of arsene (AsH_3). Use exact arithmetic or a rational format for calculations to balance this equation.

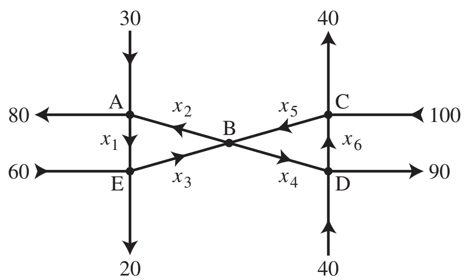
    $$MnS + As_2Cr_{10}O_{35} + H_2SO_4 \rightarrow HMnO_4 + AsH_3 + CrS_3O_{12} + H_2O$$

12. Find the general flow pattern of the network shown in the figure. Assuming that the flows are all nonnegative, what is the smallest possible value for x_4?



13. a. Find the general flow pattern of the network shown in the figure.

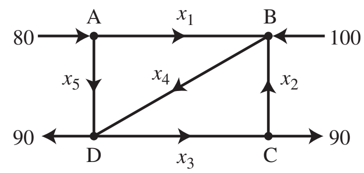
b. Assuming that the flow must be in the directions indicated, find the minimum flows in the branches denoted by x_2, x_3, x_4, and x_5.



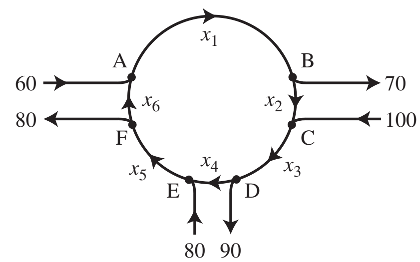
14. a. Find the general traffic pattern of the freeway network shown in the figure. (Flow rates are in cars/minute.)

b. Describe the general traffic pattern when the road whose flow is x_5 is closed.

c. When $x_5 = 0$, what is the minimum value of x_4?



15. Intersections in England are often constructed as one-way "roundabouts," such as the one shown in the figure. Assume that traffic must travel in the directions shown. Find the general solution of the network flow. Find the smallest possible value for x_6.



SOLUTIONS TO PRACTICE PROBLEMS

1. Write the percentages as decimals. Since all output must be taken into account, each column must sum to 1. This fact helps to fill in any missing entries.

Distribution of Output from:			
Agriculture	**Mining**	**Manufacturing**	**Purchased by:**
.65	.20	.20	Agriculture
.05	.10	.30	Mining
.30	.70	.50	Manufacturing

2. Since $x_5 \leq 500$, the equations D and A for x_1 and x_2 imply that $x_1 \geq 100$ and $x_2 \leq 700$. The fact that $x_5 \geq 0$ implies that $x_1 \leq 600$ and $x_2 \geq 200$. So, $100 \leq x_1 \leq 600$, and $200 \leq x_2 \leq 700$.

1.7 LINEAR INDEPENDENCE

The homogeneous equations in Section 1.5 can be studied from a different perspective by writing them as vector equations. In this way, the focus shifts from the unknown solutions of $A\mathbf{x} = \mathbf{0}$ to the vectors that appear in the vector equations.

SG Mastering: Linear Independence 1–31

In general, you should read a section thoroughly *several* times to absorb an important concept such as linear independence. The notes in the *Study Guide* for this section will help you learn to form mental images of key ideas in linear algebra. For instance, the following proof is worth reading carefully because it shows how the definition of linear independence can be *used*.

PROOF OF THEOREM 7 (Characterization of Linearly Dependent Sets)
If some $\mathbf{v}_j$ in S equals a linear combination of the other vectors, then $\mathbf{v}_j$ can be subtracted from both sides of the equation, producing a linear dependence relation with a nonzero weight (-1) on $\mathbf{v}_j$. [For instance, if $\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, then $\mathbf{0} = (-1)\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + 0\mathbf{v}_4 + \cdots + 0\mathbf{v}_p$.] Thus S is linearly dependent.

Conversely, suppose S is linearly dependent. If $\mathbf{v}_1$ is zero, then it is a (trivial) linear combination of the other vectors in S. Otherwise, $\mathbf{v}_1 \neq \mathbf{0}$, and there exist weights $c_1, \ldots, c_p$, not all zero, such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

Let j be the largest subscript for which $c_j \neq 0$. If $j = 1$, then $c_1\mathbf{v}_1 = \mathbf{0}$, which is impossible because $\mathbf{v}_1 \neq \mathbf{0}$. So $j > 1$, and

$$c_1\mathbf{v}_1 + \cdots + c_j\mathbf{v}_j + 0\mathbf{v}_{j+1} + \cdots + 0\mathbf{v}_p = \mathbf{0}$$
$$c_j\mathbf{v}_j = -c_1\mathbf{v}_1 - \cdots - c_{j-1}\mathbf{v}_{j-1}$$
$$\mathbf{v}_j = \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1} \quad \blacksquare$$

PRACTICE PROBLEMS

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -6 \\ 1 \\ 7 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 2 \end{bmatrix}$, and $\mathbf{z} = \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix}$.

1. Are the sets $\{\mathbf{u}, \mathbf{v}\}$, $\{\mathbf{u}, \mathbf{w}\}$, $\{\mathbf{u}, \mathbf{z}\}$, $\{\mathbf{v}, \mathbf{w}\}$, $\{\mathbf{v}, \mathbf{z}\}$, and $\{\mathbf{w}, \mathbf{z}\}$ each linearly independent? Why or why not?
2. Does the answer to Problem 1 imply that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ is linearly independent?
3. To determine if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ is linearly dependent, is it wise to check if, say, $\mathbf{w}$ is a linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{z}$?
4. Is $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$ linearly dependent?

1.7 EXERCISES

In Exercises 1–4, determine if the vectors are linearly independent. Justify each answer.

1. $\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 2 \\ -6 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ -8 \end{bmatrix}$

2. $\begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -8 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix}$

4. $\begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -9 \end{bmatrix}$

In Exercises 5–8, determine if the columns of the matrix form a linearly independent set. Justify each answer.

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

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

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

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

In Exercises 9 and 10, (a) for what values of h is $\mathbf{v}_3$ in Span $\{\mathbf{v}_1, \mathbf{v}_2\}$, and (b) for what values of h is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ linearly *dependent*? Justify each answer.

9. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ -6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 5 \\ -7 \\ h \end{bmatrix}$

10. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -5 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 9 \\ 15 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -5 \\ h \end{bmatrix}$

In Exercises 11–14, find the value(s) of h for which the vectors are linearly *dependent*. Justify each answer.

11. $\begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ -6 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ h \end{bmatrix}$

12. $\begin{bmatrix} 3 \\ -6 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 4 \\ -3 \end{bmatrix}, \begin{bmatrix} 9 \\ h \\ 3 \end{bmatrix}$

13. $\begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \begin{bmatrix} -2 \\ -9 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ h \\ -9 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ h \end{bmatrix}$

Determine by inspection whether the vectors in Exercises 15–20 are linearly *independent*. Justify each answer.

15. $\begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 7 \end{bmatrix}$

16. $\begin{bmatrix} 2 \\ -4 \\ 8 \end{bmatrix}, \begin{bmatrix} -3 \\ 6 \\ -12 \end{bmatrix}$

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

18. $\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \end{bmatrix}$

19. $\begin{bmatrix} -8 \\ 12 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$

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

In Exercises 21 and 22, mark each statement True or False. Justify each answer on the basis of a careful reading of the text.

21. a. The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution.

b. If S is a linearly dependent set, then each vector is a linear combination of the other vectors in S.

c. The columns of any 4×5 matrix are linearly dependent.

d. If $\mathbf{x}$ and $\mathbf{y}$ are linearly independent, and if $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ is linearly dependent, then $\mathbf{z}$ is in Span $\{\mathbf{x}, \mathbf{y}\}$.

22. a. If $\mathbf{u}$ and $\mathbf{v}$ are linearly independent, and if $\mathbf{w}$ is in Span $\{\mathbf{u}, \mathbf{v}\}$, then $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

b. If three vectors in $\mathbb{R}^3$ lie in the same plane in $\mathbb{R}^3$, then they are linearly dependent.

c. If a set contains fewer vectors than there are entries in the vectors, then the set is linearly independent.

d. If a set in $\mathbb{R}^n$ is linearly dependent, then the set contains more than n vectors.

In Exercises 23–26, describe the possible echelon forms of the matrix. Use the notation of Example 1 in Section 1.2.

23. A is a 2×2 matrix with linearly dependent columns.

24. A is a 3×3 matrix with linearly independent columns.

25. A is a 4×2 matrix, $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$, and $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$.

26. A is a 4×3 matrix, $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$, such that $\{\mathbf{a}_1, \mathbf{a}_2\}$ is linearly independent and $\mathbf{a}_3$ is not in Span $\{\mathbf{a}_1, \mathbf{a}_2\}$.

27. How many pivot columns must a 6×4 matrix have if its columns are linearly independent? Why?

28. How many pivot columns must a 4×6 matrix have if its columns span $\mathbb{R}^4$? Why?

29. Construct 3×2 matrices A and B such that $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, but $B\mathbf{x} = \mathbf{0}$ has only the trivial solution.

30. a. Fill in the blank in the following statement: "If A is an $m \times n$ matrix, then the columns of A are linearly independent if and only if A has _____ pivot columns."

b. Explain why the statement in (a) is true.

Exercises 31 and 32 should be solved *without performing row operations*. [*Hint:* Write $A\mathbf{x} = \mathbf{0}$ as a vector equation.]

31. Given $A = \begin{bmatrix} 2 & 3 & 5 \\ -5 & 1 & -4 \\ -3 & -1 & -4 \\ 1 & 0 & 1 \end{bmatrix}$, observe that the third column is the sum of the first two columns. Find a nontrivial solution of $A\mathbf{x} = \mathbf{0}$.

32. Given $A = \begin{bmatrix} 4 & 3 & -5 \\ -2 & -2 & 4 \\ -2 & -3 & 7 \end{bmatrix}$, observe that the first column minus three times the second column equals the third column. Find a nontrivial solution of $A\mathbf{x} = \mathbf{0}$.

Each statement in Exercises 33–38 is either true (in all cases) or false (for at least one example). If false, construct a specific example to show that the statement is not always true. Such an example is called a *counterexample* to the statement. If a statement is true, give a justification. (One specific example cannot explain why a statement is always true. You will have to do more work here than in Exercises 21 and 22.)

33. If $\mathbf{v}_1, \dots, \mathbf{v}_4$ are in $\mathbb{R}^4$ and $\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly dependent.

34. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^4$ and $\mathbf{v}_2$ is not a scalar multiple of $\mathbf{v}_1$, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.

35. If $\mathbf{v}_1, \dots, \mathbf{v}_5$ are in $\mathbb{R}^5$ and $\mathbf{v}_3 = \mathbf{0}$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ is linearly dependent.

36. If $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are in $\mathbb{R}^3$ and $\mathbf{v}_3$ is *not* a linear combination of $\mathbf{v}_1, \mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

37. If $\mathbf{v}_1, \dots, \mathbf{v}_4$ are in $\mathbb{R}^4$ and $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is also linearly dependent.

38. If $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is a linearly independent set of vectors in $\mathbb{R}^4$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also linearly independent. [*Hint:* Think about $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + 0 \cdot \mathbf{v}_4 = \mathbf{0}$.]

39. Suppose A is an $m \times n$ matrix with the property that for all $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has at most one solution. Use the

definition of linear independence to explain why the columns of A must be linearly independent.

40. Suppose an $m \times n$ matrix A has n pivot columns. Explain why for each $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ has at most one solution. [*Hint:* Explain why $A\mathbf{x} = \mathbf{b}$ cannot have infinitely many solutions.]

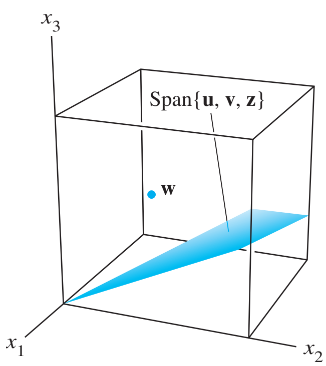
[**M**] In Exercises 41 and 42, use as many columns of A as possible to construct a matrix B with the property that the equation $B\mathbf{x} = \mathbf{0}$ has only the trivial solution. Solve $B\mathbf{x} = \mathbf{0}$ to verify your work.

41. $A = \begin{bmatrix} 3 & -4 & 10 & 7 & -4 \\ -5 & -3 & -7 & -11 & 15 \\ 4 & 3 & 5 & 2 & 1 \\ 8 & -7 & 23 & 4 & 15 \end{bmatrix}$

42. $A = \begin{bmatrix} 12 & 10 & -6 & 8 & 4 & -14 \\ -7 & -6 & 4 & -5 & -7 & 9 \\ 9 & 9 & -9 & 9 & 9 & -18 \\ -4 & -3 & -1 & 0 & -8 & 1 \\ 8 & 7 & -5 & 6 & 1 & -11 \end{bmatrix}$

43. [**M**] With A and B as in Exercise 41, select a column $\mathbf{v}$ of A that was not used in the construction of B and determine if $\mathbf{v}$ is in the set spanned by the columns of B. (Describe your calculations.)

44. [**M**] Repeat Exercise 43 with the matrices A and B from Exercise 42. Then give an explanation for what you discover, assuming that B was constructed as specified.



SOLUTIONS TO PRACTICE PROBLEMS

1. Yes. In each case, neither vector is a multiple of the other. Thus each set is linearly independent.
2. No. The observation in Practice Problem 1, by itself, says nothing about the linear independence of $\{\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}\}$.
3. No. When testing for linear independence, it is usually a poor idea to check if one selected vector is a linear combination of the others. It may happen that the selected vector is not a linear combination of the others and yet the whole set of vectors is linearly dependent. In this practice problem, $\mathbf{w}$ is not a linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{z}$.
4. Yes, by Theorem 8. There are more vectors (four) than entries (three) in them.

1.8 INTRODUCTION TO LINEAR TRANSFORMATIONS

The difference between a matrix equation $A\mathbf{x} = \mathbf{b}$ and the associated vector equation $x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$ is merely a matter of notation. However, a matrix equation $A\mathbf{x} = \mathbf{b}$ can arise in linear algebra (and in applications such as computer graphics and signal processing) in a way that is not directly connected with linear combinations of vectors. This happens when we think of the matrix A as an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$.

For instance, the equations

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 5 \\ 8 \end{bmatrix}}_{\mathbf{b}} \quad \text{and} \quad \underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 3 \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} 0 \\ 0 \end{bmatrix}}_{\mathbf{0}}$$

say that multiplication by A transforms $\mathbf{x}$ into $\mathbf{b}$ and transforms $\mathbf{u}$ into the zero vector. See Fig. 1.

2. If $\mathbf{x}$ and $\mathbf{y}$ are production vectors, then the total cost vector associated with the combined production $\mathbf{x}+\mathbf{y}$ is precisely the sum of the cost vectors $T(\mathbf{x})$ and $T(\mathbf{y})$. ■

PRACTICE PROBLEMS

1. Suppose $T:\mathbb{R}^5 \to \mathbb{R}^2$ and $T(\mathbf{x}) = A\mathbf{x}$ for some matrix A and for each $\mathbf{x}$ in $\mathbb{R}^5$. How many rows and columns does A have?
2. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Give a geometric description of the transformation $\mathbf{x} \mapsto A\mathbf{x}$.
3. The line segment from $\mathbf{0}$ to a vector $\mathbf{u}$ is the set of points of the form $t\mathbf{u}$, where $0 \le t \le 1$. Show that a linear transformation T maps this segment into the segment between $\mathbf{0}$ and $T(\mathbf{u})$.

1.8 EXERCISES

1. Let $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, and define $T:\mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$. Find the images under T of $\mathbf{u} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

2. Let $A = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 3 \\ 6 \\ -9 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Define $T:\mathbb{R}^3 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$. Find $T(\mathbf{u})$ and $T(\mathbf{v})$.

In Exercises 3–6, with T defined by $T(\mathbf{x}) = A\mathbf{x}$, find a vector $\mathbf{x}$ whose image under T is $\mathbf{b}$, and determine whether $\mathbf{x}$ is unique.

3. $A = \begin{bmatrix} 1 & 0 & -3 \\ -3 & 1 & 6 \\ 2 & -2 & -1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 1 & -3 \\ 2 & -5 & 6 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -6 \\ -4 \\ -5 \end{bmatrix}$

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

6. $A = \begin{bmatrix} 1 & -3 & 2 \\ 3 & -8 & 8 \\ 0 & 1 & 2 \\ 1 & 0 & 8 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 3 \\ 10 \end{bmatrix}$

7. Let A be a 6×5 matrix. What must a and b be in order to define $T:\mathbb{R}^a \to \mathbb{R}^b$ by $T(\mathbf{x}) = A\mathbf{x}$?

8. How many rows and columns must a matrix A have in order to define a mapping from $\mathbb{R}^5$ into $\mathbb{R}^7$ by the rule $T(\mathbf{x}) = A\mathbf{x}$?

For Exercises 9 and 10, find all $\mathbf{x}$ in $\mathbb{R}^4$ that are mapped into the zero vector by the transformation $\mathbf{x} \mapsto A\mathbf{x}$ for the given matrix A.

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

10. $A = \begin{bmatrix} 3 & 2 & 10 & -6 \\ 1 & 0 & 2 & -4 \\ 0 & 1 & 2 & 3 \\ 1 & 4 & 10 & 8 \end{bmatrix}$

11. Let $\mathbf{b} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, and let A be the matrix in Exercise 9. Is $\mathbf{b}$ in the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$? Why or why not?

12. Let $\mathbf{b} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 4 \end{bmatrix}$, and let A be the matrix in Exercise 10. Is $\mathbf{b}$ in the range of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$? Why or why not?

In Exercises 13–16, use a rectangular coordinate system to plot $\mathbf{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, and their images under the given transformation T. (Make a separate and reasonably large sketch for each exercise.) Describe geometrically what T does to each vector $\mathbf{x}$ in $\mathbb{R}^2$.

13. $T(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

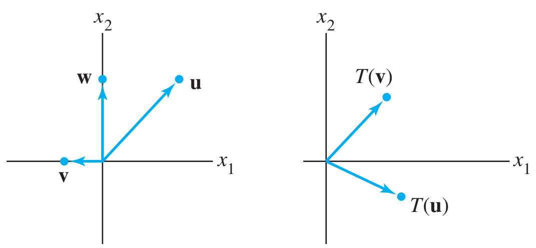
14. $T(\mathbf{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

15. $T(\mathbf{x}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

16. $T(\mathbf{x}) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

17. Let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ into $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ and maps $\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ into $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$. Use the fact that T is linear to find the images under T of $2\mathbf{u}$, $3\mathbf{v}$, and $2\mathbf{u} + 3\mathbf{v}$.

18. The figure shows vectors **u**, **v**, and **w**, along with the images $T(\mathbf{u})$ and $T(\mathbf{v})$ under the action of a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$. Copy this figure carefully, and draw the image $T(\mathbf{w})$ as accurately as possible. [*Hint:* First, write **w** as a linear combination of **u** and **v**.]



19. Let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y}_1 = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$, and $\mathbf{y}_2 = \begin{bmatrix} -1 \\ 6 \end{bmatrix}$, and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps $\mathbf{e}_1$ into $\mathbf{y}_1$ and maps $\mathbf{e}_2$ into $\mathbf{y}_2$. Find the images of $\begin{bmatrix} 5 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

20. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, and let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation that maps **x** into $x_1\mathbf{v}_1 + x_2\mathbf{v}_2$. Find a matrix A such that $T(\mathbf{x})$ is $A\mathbf{x}$ for each **x**.

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

21. a. A linear transformation is a special type of function.

b. If A is a 3×5 matrix and T is a transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then the domain of T is $\mathbb{R}^3$.

c. If A is an $m \times n$ matrix, then the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is $\mathbb{R}^m$.

d. Every linear transformation is a matrix transformation.

e. A transformation T is linear if and only if

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

for all $\mathbf{v}_1$ and $\mathbf{v}_2$ in the domain of T and for all scalars c_1 and c_2.

22. a. The range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the set of all linear combinations of the columns of A.

b. Every matrix transformation is a linear transformation.

c. If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation and if **c** is in $\mathbb{R}^m$, then a uniqueness question is "Is **c** in the range of T?"

d. A linear transformation preserves the operations of vector addition and scalar multiplication.

e. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ always maps the origin of $\mathbb{R}^n$ to the origin of $\mathbb{R}^m$.

23. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = mx + b$.

a. Show that f is a linear transformation when $b = 0$.

b. Find a property of a linear transformation that is violated when $b \neq 0$.

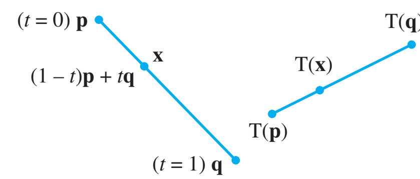
c. Why is f called a linear function?

24. An *affine transformation* $T : \mathbb{R}^n \to \mathbb{R}^m$ has the form $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, with A an $m \times n$ matrix and **b** in $\mathbb{R}^m$. Show that T is *not* a linear transformation when $\mathbf{b} \neq \mathbf{0}$. (Affine transformations are important in computer graphics.)

25. Given $\mathbf{v} \neq \mathbf{0}$ and **p** in $\mathbb{R}^n$, the line through **p** in the direction of **v** has the parametric equation $\mathbf{x} = \mathbf{p} + t\mathbf{v}$. Show that a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ maps this line onto another line or onto a single point (a *degenerate line*).

26. a. Show that the line through vectors **p** and **q** in $\mathbb{R}^n$ may be written in the parametric form $\mathbf{x} = (1 - t)\mathbf{p} + t\mathbf{q}$. (Refer to the figure with Exercises 21 and 22 in Section 1.5.)

b. The line segment from **p** to **q** is the set of points of the form $(1 - t)\mathbf{p} + t\mathbf{q}$ for $0 \le t \le 1$ (as shown in the figure below). Show that a linear transformation T maps this line segment onto a line segment or onto a single point.



27. Let **u** and **v** be linearly independent vectors in $\mathbb{R}^3$, and let P be the plane through **u**, **v**, and **0**. The parametric equation of P is $\mathbf{x} = s\mathbf{u} + t\mathbf{v}$ (with s, t in $\mathbb{R}$). Show that a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ maps P onto a plane through **0**, or onto a line through **0**, or onto just the origin in $\mathbb{R}^3$. What must be true about $T(\mathbf{u})$ and $T(\mathbf{v})$ in order for the image of the plane P to be a plane?

28. Let **u** and **v** be vectors in $\mathbb{R}^n$. It can be shown that the set P of all points in the parallelogram determined by **u** and **v** has the form $a\mathbf{u} + b\mathbf{v}$, for $0 \le a \le 1, 0 \le b \le 1$. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Explain why the image of a point in P under the transformation T lies in the parallelogram determined by $T(\mathbf{u})$ and $T(\mathbf{v})$.

29. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that reflects each point through the x_2-axis. Make two sketches similar to Fig. 6 that illustrate properties (i) and (ii) of a linear transformation.

30. Suppose vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ span $\mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Suppose $T(\mathbf{v}_i) = \mathbf{0}$ for $i = 1, \ldots, p$. Show that T is the zero transformation. That is, show that if **x** is any vector in $\mathbb{R}^n$, then $T(\mathbf{x}) = \mathbf{0}$.

31. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly dependent set in $\mathbb{R}^n$. Explain why the set $\{T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)\}$ is linearly dependent.

In Exercises 32–36, column vectors are written as rows, such as $\mathbf{x} = (x_1, x_2)$, and $T(\mathbf{x})$ is written as $T(x_1, x_2)$.

32. Show that the transformation T defined by $T(x_1, x_2) = (x_1 - 2|x_2|, x_1 - 4x_2)$ is not linear.

33. Show that the transformation T defined by $T(x_1, x_2) = (x_1 - 2x_2, x_1 - 3, 2x_1 - 5x_2)$ is not linear.

34. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the transformation that reflects each vector $\mathbf{x} = (x_1, x_2, x_3)$ through the plane $x_3 = 0$ onto $T(\mathbf{x}) = (x_1, x_2, -x_3)$. Show that T is a linear transformation. [See Example 4 for ideas.]

35. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the transformation that projects each vector $\mathbf{x} = (x_1, x_2, x_3)$ onto the plane $x_2 = 0$, so $T(\mathbf{x}) = (x_1, 0, x_3)$. Show that T is a linear transformation.

36. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Suppose $\{\mathbf{u}, \mathbf{v}\}$ is a linearly independent set, but $\{T(\mathbf{u}), T(\mathbf{v})\}$ is a linearly dependent set. Show that $T(\mathbf{x}) = \mathbf{0}$ has a nontrivial solution. [*Hint:* Use the fact that $c_1T(\mathbf{u}) + c_2T(\mathbf{v}) = \mathbf{0}$ for some weights c_1 and c_2, not both zero.]

[M] In Exercises 37 and 38, the given matrix determines a linear transformation T. Find all $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{0}$.

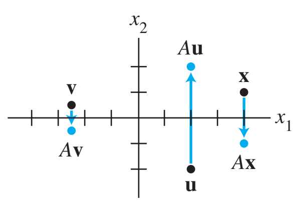
37. $\begin{bmatrix} 2 & 3 & 5 & -5 \\ -7 & 7 & 0 & 0 \\ -3 & 4 & 1 & 3 \\ -9 & 3 & -6 & -4 \end{bmatrix}$ **38.** $\begin{bmatrix} 3 & 4 & -7 & 0 \\ 5 & -8 & 7 & 4 \\ 6 & -8 & 6 & 4 \\ 9 & -7 & -2 & 0 \end{bmatrix}$

39. **[M]** Let $\mathbf{b} = \begin{bmatrix} 8 \\ 7 \\ 5 \\ -3 \end{bmatrix}$ and let A be the matrix in Exercise 37. Is $\mathbf{b}$ in the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$? If so, find an $\mathbf{x}$ whose image under the transformation is $\mathbf{b}$.

40. **[M]** Let $\mathbf{b} = \begin{bmatrix} 4 \\ -4 \\ -4 \\ -7 \end{bmatrix}$ and let A be the matrix in Exercise 38. Is $\mathbf{b}$ in the range of the transformation $\mathbf{x} \mapsto A\mathbf{x}$? If so, find an $\mathbf{x}$ whose image under the transformation is $\mathbf{b}$.

SG Mastering: Linear Transformations 1–34

SOLUTIONS TO PRACTICE PROBLEMS

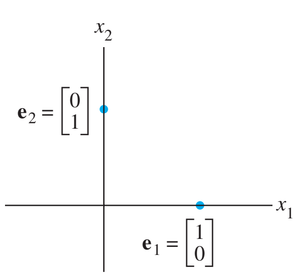


The transformation $\mathbf{x} \mapsto A\mathbf{x}$.

1. A must have five columns for $A\mathbf{x}$ to be defined. A must have two rows for the codomain of T to be $\mathbb{R}^2$.
2. Plot some random points (vectors) on graph paper to see what happens. A point such as $(4, 1)$ maps into $(4, -1)$. The transformation $\mathbf{x} \mapsto A\mathbf{x}$ reflects points through the x-axis (or x_1-axis).
3. Let $\mathbf{x} = t\mathbf{u}$ for some t such that $0 \le t \le 1$. Since T is linear, $T(t\mathbf{u}) = t\,T(\mathbf{u})$, which is a point on the line segment between $\mathbf{0}$ and $T(\mathbf{u})$.

1.9 THE MATRIX OF A LINEAR TRANSFORMATION

Whenever a linear transformation T arises geometrically or is described in words, we usually want a "formula" for $T(\mathbf{x})$. The discussion that follows shows that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is actually a matrix transformation $\mathbf{x} \mapsto A\mathbf{x}$ and that important properties of T are intimately related to familiar properties of A. The key to finding A is to observe that T is completely determined by what it does to the columns of the $n \times n$ identity matrix I_n.



EXAMPLE 1 The columns of $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^3$ such that

$$T(\mathbf{e}_1) = \begin{bmatrix} 5 \\ -7 \\ 2 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = \begin{bmatrix} -3 \\ 8 \\ 0 \end{bmatrix}$$

With no additional information, find a formula for the image of an arbitrary $\mathbf{x}$ in $\mathbb{R}^2$.

THEOREM 12

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let A be the standard matrix for T. Then:

a. T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if the columns of A span $\mathbb{R}^m$;

b. T is one-to-one if and only if the columns of A are linearly independent.

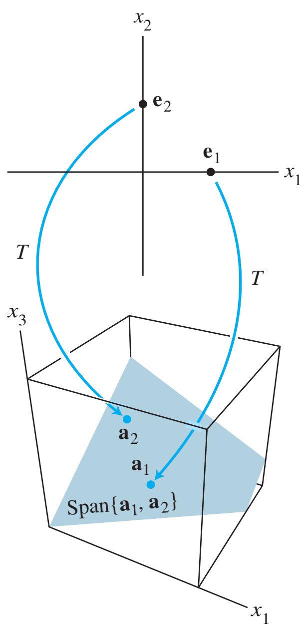
PROOF

a. By Theorem 4 in Section 1.4, the columns of A span $\mathbb{R}^m$ if and only if for each $\mathbf{b}$ in $\mathbb{R}^m$ the equation $A\mathbf{x} = \mathbf{b}$ is consistent—in other words, if and only if for every $\mathbf{b}$, the equation $T(\mathbf{x}) = \mathbf{b}$ has at least one solution. This is true if and only if T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$.

b. The equations $T(\mathbf{x}) = \mathbf{0}$ and $A\mathbf{x} = \mathbf{0}$ are the same except for notation. So, by Theorem 11, T is one-to-one if and only if $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. This happens if and only if the columns of A are linearly independent, as was already noted in the boxed statement (3) in Section 1.7. ■

Statement (a) in Theorem 12 is equivalent to the statement "T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if every vector in $\mathbb{R}^m$ is a linear combination of the columns of A." See Theorem 4 in Section 1.4.

In the next example and in some exercises that follow, column vectors are written in rows, such as $\mathbf{x} = (x_1, x_2)$, and $T(\mathbf{x})$ is written as $T(x_1, x_2)$ instead of the more formal $T((x_1, x_2))$.



The transformation T is not onto $\mathbb{R}^3$.

EXAMPLE 5 Let $T(x_1, x_2) = (3x_1 + x_2, 5x_1 + 7x_2, x_1 + 3x_2)$. Show that T is a one-to-one linear transformation. Does T map $\mathbb{R}^2$ onto $\mathbb{R}^3$?

SOLUTION When $\mathbf{x}$ and $T(\mathbf{x})$ are written as column vectors, you can determine the standard matrix of T by inspection, visualizing the row–vector computation of each entry in $A\mathbf{x}$.

$$T(\mathbf{x}) = \begin{bmatrix} 3x_1 + x_2 \\ 5x_1 + 7x_2 \\ x_1 + 3x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 5 & 7 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4}$$

So T is indeed a linear transformation, with its standard matrix A shown in (4). The columns of A are linearly independent because they are not multiples. By Theorem 12(b), T is one-to-one. To decide if T is onto $\mathbb{R}^3$, examine the span of the columns of A. Since A is 3×2, the columns of A span $\mathbb{R}^3$ if and only if A has 3 pivot positions, by Theorem 4. This is impossible, since A has only 2 columns. So the columns of A do not span $\mathbb{R}^3$, and the associated linear transformation is not onto $\mathbb{R}^3$. ■

PRACTICE PROBLEM

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation that first performs a horizontal shear that maps $\mathbf{e}_2$ into $\mathbf{e}_2 - .5\mathbf{e}_1$ (but leaves $\mathbf{e}_1$ unchanged) and then reflects the result through the x_2-axis. Assuming that T is linear, find its standard matrix. [*Hint:* Determine the final location of the images of $\mathbf{e}_1$ and $\mathbf{e}_2$.]

1.9 EXERCISES

In Exercises 1–10, assume that T is a linear transformation. Find the standard matrix of T.

1. $T: \mathbb{R}^2 \to \mathbb{R}^4$, $T(\mathbf{e}_1) = (3, 1, 3, 1)$, and $T(\mathbf{e}_2) = (-5, 2, 0, 0)$, where $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$.

2. $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(\mathbf{e}_1) = (1, 4)$, $T(\mathbf{e}_2) = (-2, 9)$, and $T(\mathbf{e}_3) = (3, -8)$, where $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are the columns of the 3×3 identity matrix.

3. $T: \mathbb{R}^2 \to \mathbb{R}^2$ is a vertical shear transformation that maps $\mathbf{e}_1$ into $\mathbf{e}_1 - 3\mathbf{e}_2$, but leaves $\mathbf{e}_2$ unchanged.

4. $T: \mathbb{R}^2 \to \mathbb{R}^2$ is a horizontal shear transformation that leaves $\mathbf{e}_1$ unchanged and maps $\mathbf{e}_2$ into $\mathbf{e}_2 + 2\mathbf{e}_1$.

5. $T: \mathbb{R}^2 \to \mathbb{R}^2$ rotates points (about the origin) through $\pi/2$ radians (counterclockwise).

6. $T: \mathbb{R}^2 \to \mathbb{R}^2$ rotates points (about the origin) through $-3\pi/2$ radians (clockwise).

7. $T: \mathbb{R}^2 \to \mathbb{R}^2$ first rotates points through $-3\pi/4$ radians (clockwise) and then reflects points through the horizontal x_1-axis. [*Hint:* $T(\mathbf{e}_1) = (-1/\sqrt{2}, 1/\sqrt{2})$.]

8. $T: \mathbb{R}^2 \to \mathbb{R}^2$ first performs a horizontal shear that transforms $\mathbf{e}_2$ into $\mathbf{e}_2 + 2\mathbf{e}_1$ (leaving $\mathbf{e}_1$ unchanged) and then reflects points through the line $x_2 = -x_1$.

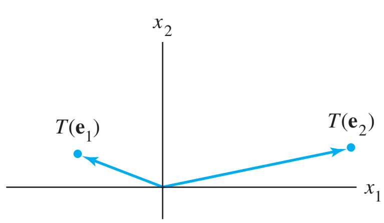
9. $T: \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the horizontal x_1-axis and then rotates points $-\pi/2$ radians.

10. $T: \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the horizontal x_1-axis and then reflects points through the line $x_2 = x_1$.

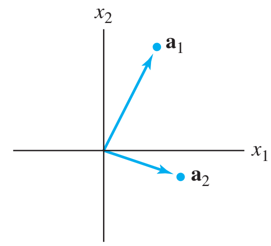
11. A linear transformation $T: \mathbb{R}^2 \to \mathbb{R}^2$ first reflects points through the x_1-axis and then reflects points through the x_2-axis. Show that T can also be described as a linear transformation that rotates points about the origin. What is the angle of that rotation?

12. Show that the transformation in Exercise 10 is merely a rotation about the origin. What is the angle of the rotation?

13. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation such that $T(\mathbf{e}_1)$ and $T(\mathbf{e}_2)$ are the vectors shown in the figure. Using the figure, sketch the vector $T(2, 1)$.



14. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation with standard matrix $A = [\,\mathbf{a}_1 \;\; \mathbf{a}_2\,]$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are shown in the figure at the top of column 2. Using the figure, draw the image of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ under the transformation T.



In Exercises 15 and 16, fill in the missing entries of the matrix, assuming that the equation holds for all values of the variables.

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

16. $$\begin{bmatrix} ? & ? \\ ? & ? \\ ? & ? \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3x_1 - 2x_2 \\ x_1 + 4x_2 \\ x_2 \end{bmatrix}$$

In Exercises 17–20, show that T is a linear transformation by finding a matrix that implements the mapping. Note that $x_1, x_2, \ldots$ are not vectors but are entries in vectors.

17. $T(x_1, x_2, x_3, x_4) = (x_1 + 2x_2, 0, 2x_2 + x_4, x_2 - x_4)$

18. $T(x_1, x_2) = (x_1 + 4x_2, 0, x_1 - 3x_2, x_1)$

19. $T(x_1, x_2, x_3) = (x_1 - 5x_2 + 4x_3, x_2 - 6x_3)$

20. $T(x_1, x_2, x_3, x_4) = 3x_1 + 4x_3 - 2x_4$ (Notice: $T: \mathbb{R}^4 \to \mathbb{R}$)

21. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that $T(x_1, x_2) = (x_1 + x_2, 4x_1 + 5x_2)$. Find $\mathbf{x}$ such that $T(\mathbf{x}) = (3, 8)$.

22. Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation with $T(x_1, x_2) = (2x_1 - x_2, -3x_1 + x_2, 2x_1 - 3x_2)$. Find $\mathbf{x}$ such that $T(\mathbf{x}) = (0, -1, -4)$.

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

23. a. A linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is completely determined by its effect on the columns of the $n \times n$ identity matrix.

    b. If $T: \mathbb{R}^2 \to \mathbb{R}^2$ rotates vectors about the origin through an angle φ, then T is a linear transformation.

    c. When two linear transformations are performed one after another, the combined effect may not always be a linear transformation.

    d. A mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ is onto $\mathbb{R}^m$ if every vector $\mathbf{x}$ in $\mathbb{R}^n$ maps onto some vector in $\mathbb{R}^m$.

    e. If A is a 3×2 matrix, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ cannot be one-to-one.

24. a. If A is a 4×3 matrix, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^3$ onto $\mathbb{R}^4$.

b. Every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation.

c. The columns of the standard matrix for a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ are the images of the columns of the $n \times n$ identity matrix under T.

d. A mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one if each vector in $\mathbb{R}^n$ maps onto a unique vector in $\mathbb{R}^m$.

e. The standard matrix of a horizontal shear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ has the form $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$, where a and d are ± 1.

In Exercises 25–28, determine if the specified linear transformation is (a) one-to-one and (b) onto. Justify each answer.

25. The transformation in Exercise 17

26. The transformation in Exercise 2

27. The transformation in Exercise 19

28. The transformation in Exercise 14

In Exercises 29 and 30, describe the possible echelon forms of the standard matrix for a linear transformation T. Use the notation of Example 1 in Section 1.2.

29. $T : \mathbb{R}^3 \to \mathbb{R}^4$ is one-to-one.

30. $T : \mathbb{R}^4 \to \mathbb{R}^3$ is onto.

31. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, with A its standard matrix. Complete the following statement to make it true: "T is one-to-one if and only if A has ____ pivot columns." Explain why the statement is true. [*Hint:* Look in the exercises for Section 1.7.]

32. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, with A its standard matrix. Complete the following statement to make it true: "T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$ if and only if A has ____ pivot columns." Find some theorems that explain why the statement is true.

33. Verify the uniqueness of A in Theorem 10. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation such that $T(\mathbf{x}) = B\mathbf{x}$ for some $m \times n$ matrix B. Show that if A is the standard matrix for T, then $A = B$. [*Hint:* Show that A and B have the same columns.]

34. Let $S : \mathbb{R}^p \to \mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations. Show that the mapping $\mathbf{x} \mapsto T(S(\mathbf{x}))$ is a linear transformation (from $\mathbb{R}^p$ to $\mathbb{R}^m$). [*Hint:* Compute $T(S(c\mathbf{u} + d\mathbf{v}))$ for $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^p$ and scalars c and d. Justify each step of the computation, and explain why this computation gives the desired conclusion.]

35. If a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$, can you give a relation between m and n? If T is one-to-one, what can you say about m and n?

36. Why is the question "Is the linear transformation T onto?" an existence question?

[**M**] In Exercises 37–40, let T be the linear transformation whose standard matrix is given. In Exercises 37 and 38, decide if T is a one-to-one mapping. In Exercises 39 and 40, decide if T maps $\mathbb{R}^5$ onto $\mathbb{R}^5$. Justify your answers.

37. $\begin{bmatrix} -5 & 6 & -5 & -6 \\ 8 & 3 & -3 & 8 \\ 2 & 9 & 5 & -12 \\ -3 & 2 & 7 & -12 \end{bmatrix}$

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

39. $\begin{bmatrix} 4 & -7 & 3 & 7 & 5 \\ 6 & -8 & 5 & 12 & -8 \\ -7 & 10 & -8 & -9 & 14 \\ 3 & -5 & 4 & 2 & -6 \\ -5 & 6 & -6 & -7 & 3 \end{bmatrix}$

40. $\begin{bmatrix} 9 & 43 & 5 & 6 & -1 \\ 14 & 15 & -7 & -5 & 4 \\ -8 & -6 & 12 & -5 & -9 \\ -5 & -6 & -4 & 9 & 8 \\ 13 & 14 & 15 & 3 & 11 \end{bmatrix}$

SOLUTION TO PRACTICE PROBLEM

WEB

Follow what happens to $\mathbf{e}_1$ and $\mathbf{e}_2$. See Fig. 5. First, $\mathbf{e}_1$ is unaffected by the shear and then is reflected into $-\mathbf{e}_1$. So $T(\mathbf{e}_1) = -\mathbf{e}_1$. Second, $\mathbf{e}_2$ goes to $\mathbf{e}_2 - .5\mathbf{e}_1$ by the shear

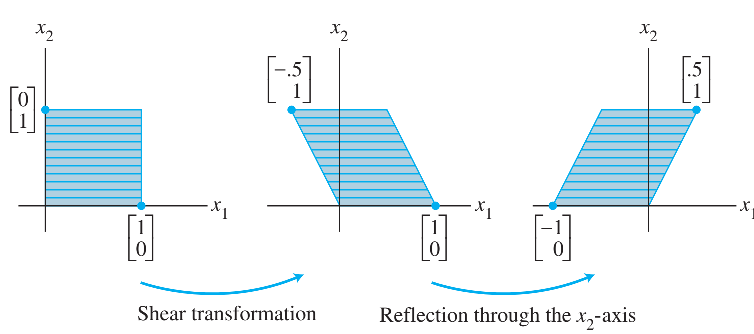


FIGURE 5 The composition of two transformations.

transformation. Since reflection through the x_2-axis changes $\mathbf{e}_1$ into $-\mathbf{e}_1$ and leaves $\mathbf{e}_2$ unchanged, the vector $\mathbf{e}_2 - .5\mathbf{e}_1$ goes to $\mathbf{e}_2 + .5\mathbf{e}_1$. So $T(\mathbf{e}_2) = \mathbf{e}_2 + .5\mathbf{e}_1$. Thus the standard matrix of T is

$$\begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} -\mathbf{e}_1 & \mathbf{e}_2 + .5\mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} -1 & .5 \\ 0 & 1 \end{bmatrix}$$

1.10 LINEAR MODELS IN BUSINESS, SCIENCE, AND ENGINEERING

The mathematical models in this section are all *linear*; that is, each describes a problem by means of a linear equation, usually in vector or matrix form. The first model concerns nutrition but actually is representative of a general technique in linear programming problems. The second model comes from electrical engineering. The third model introduces the concept of a *linear difference equation*, a powerful mathematical tool for studying dynamic processes in a wide variety of fields such as engineering, ecology, economics, telecommunications, and the management sciences. Linear models are important because natural phenomena are often linear or nearly linear when the variables involved are held within reasonable bounds. Also, linear models are more easily adapted for computer calculation than are complex nonlinear models.

As you read about each model, pay attention to how its linearity reflects some property of the system being modeled.

Constructing a Nutritious Weight-Loss Diet

WEB

The formula for the Cambridge Diet, a popular diet in the 1980s, was based on years of research. A team of scientists headed by Dr. Alan H. Howard developed this diet at Cambridge University after more than eight years of clinical work with obese patients.[1] The very low-calorie powdered formula diet combines a precise balance of carbohydrate, high-quality protein, and fat, together with vitamins, minerals, trace elements, and electrolytes. Millions of persons have used the diet to achieve rapid and substantial weight loss.

To achieve the desired amounts and proportions of nutrients, Dr. Howard had to incorporate a large variety of foodstuffs in the diet. Each foodstuff supplied several of the required ingredients, but not in the correct proportions. For instance, nonfat milk was a major source of protein but contained too much calcium. So soy flour was used for part of the protein because soy flour contains little calcium. However, soy flour contains proportionally too much fat, so whey was added since it supplies less fat in relation to calcium. Unfortunately, whey contains too much carbohydrate. . . .

The following example illustrates the problem on a small scale. Listed in Table 1 are three of the ingredients in the diet, together with the amounts of certain nutrients supplied by 100 grams (g) of each ingredient.[2]

EXAMPLE 1 If possible, find some combination of nonfat milk, soy flour, and whey to provide the exact amounts of protein, carbohydrate, and fat supplied by the diet in one day (Table 1).

[1] The first announcement of this rapid weight-loss regimen was given in the *International Journal of Obesity* (1978) **2**, 321–332.

[2] Ingredients in the diet as of 1984; nutrient data for ingredients adapted from USDA Agricultural Handbooks No. 8-1 and 8-6, 1976.

The vectors in (6) and (7) account for all of the population in 2001.[3] Thus

$$\begin{bmatrix} r_1 \\ s_1 \end{bmatrix} = r_0 \begin{bmatrix} .95 \\ .05 \end{bmatrix} + s_0 \begin{bmatrix} .03 \\ .97 \end{bmatrix} = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} r_0 \\ s_0 \end{bmatrix}$$

That is,

$$\mathbf{x}_1 = M\mathbf{x}_0 \tag{8}$$

where M is the **migration matrix** determined by the following table:

From: City	From: Suburbs	To:
.95	.03	City
.05	.97	Suburbs

Equation (8) describes how the population changes from 2000 to 2001. If the migration percentages remain constant, then the change from 2001 to 2002 is given by

$$\mathbf{x}_2 = M\mathbf{x}_1$$

and similarly for 2002 to 2003 and subsequent years. In general,

$$\mathbf{x}_{k+1} = M\mathbf{x}_k \quad \text{for } k = 0, 1, 2, \ldots \tag{9}$$

The sequence of vectors $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ describes the population of the city/suburban region over a period of years.

EXAMPLE 3 Compute the population of the region just described for the years 2001 and 2002, given that the population in 2000 was 600,000 in the city and 400,000 in the suburbs.

SOLUTION The initial population in 2000 is $\mathbf{x}_0 = \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix}$. For 2001,

$$\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 600{,}000 \\ 400{,}000 \end{bmatrix} = \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix}$$

For 2002,

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix} \begin{bmatrix} 582{,}000 \\ 418{,}000 \end{bmatrix} = \begin{bmatrix} 565{,}440 \\ 434{,}560 \end{bmatrix}$$

■

The model for population movement in (9) is *linear* because the correspondence $\mathbf{x}_k \mapsto \mathbf{x}_{k+1}$ is a linear transformation. The linearity depends on two facts: the number of people who chose to move from one area to another is *proportional* to the number of people in that area, as shown in (6) and (7), and the cumulative effect of these choices is found by *adding* the movement of people from the different areas.

PRACTICE PROBLEM

Find a matrix A and vectors $\mathbf{x}$ and $\mathbf{b}$ such that the problem in Example 1 amounts to solving the equation $A\mathbf{x} = \mathbf{b}$.

[3] For simplicity, we ignore other influences on the population such as births, deaths, and migration into and out of the city/suburban region.

1.10 EXERCISES

1. The container of a breakfast cereal usually lists the number of calories and the amounts of protein, carbohydrate, and fat contained in one serving of the cereal. The amounts for two common cereals are given below. Suppose a mixture of these two cereals is to be prepared that contains exactly 295 calories, 9 g of protein, 48 g of carbohydrate, and 8 g of fat.
   a. Set up a vector equation for this problem. Include a statement of what the variables in your equation represent.
   b. Write an equivalent matrix equation, and then determine if the desired mixture of the two cereals can be prepared.

Nutrition Information per Serving

Nutrient	**General Mills Cheerios®**	**Quaker® 100% Natural Cereal**
Calories	110	130
Protein (g)	4	3
Carbohydrate (g)	20	18
Fat (g)	2	5

2. One serving of Shredded Wheat supplies 160 calories, 5 g of protein, 6 g of fiber, and 1 g of fat. One serving of Crispix® supplies 110 calories, 2 g of protein, .1 g of fiber, and .4 g of fat.
   a. Set up a matrix B and a vector $\mathbf{u}$ such that $B\mathbf{u}$ gives the amounts of calories, protein, fiber, and fat contained in a mixture of three servings of Shredded Wheat and two servings of Crispix.
   b. **[M]** Suppose that you want a cereal with more fiber than Crispix but fewer calories than Shredded Wheat. Is it possible for a mixture of the two cereals to supply 130 calories, 3.20 g of protein, 2.46 g of fiber, and .64 g of fat? If so, what is the mixture?

3. After taking a nutrition class, a big Annie's® Mac and Cheese fan decides to improve the levels of protein and fiber in her favorite lunch by adding broccoli and canned chicken. The nutritional information for the foods referred to in this exercise are given in the table below.

Nutrition Information per Serving

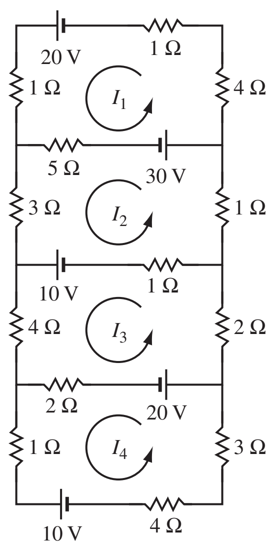
Nutrient	**Mac and Cheese**	**Broccoli**	**Chicken**	**Shells**
Calories	270	51	70	260
Protein (g)	10	5.4	15	9
Fiber (g)	2	5.2	0	5

   a. **[M]** If she wants to limit her lunch to 400 calories but get 30 g of protein and 10 g of fiber, what proportions of servings of Mac and Cheese, broccoli, and chicken should she use?
   b. **[M]** She found that there was too much broccoli in the proportions from part (a), so she decided to switch from classical Mac and Cheese to Annie's® Whole Wheat Shells and White Cheddar. What proportions of servings of each food should she use to meet the same goals as in part (a)?

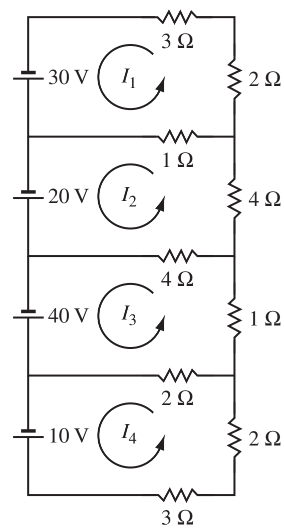
4. The Cambridge Diet supplies .8 g of calcium per day, in addition to the nutrients listed in the Table 1 for Example 1. The amounts of calcium per unit (100 g) supplied by the three ingredients in the Cambridge Diet are as follows: 1.26 g from nonfat milk, .19 g from soy flour, and .8 g from whey. Another ingredient in the diet mixture is isolated soy protein, which provides the following nutrients in each unit: 80 g of protein, 0 g of carbohydrate, 3.4 g of fat, and .18 g of calcium.
   a. Set up a matrix equation whose solution determines the amounts of nonfat milk, soy flour, whey, and isolated soy protein necessary to supply the precise amounts of protein, carbohydrate, fat, and calcium in the Cambridge Diet. State what the variables in the equation represent.
   b. **[M]** Solve the equation in (a) and discuss your answer.

In Exercises 5–8, write a matrix equation that determines the loop currents. **[M]** If MATLAB or another matrix program is available, solve the system for the loop currents.

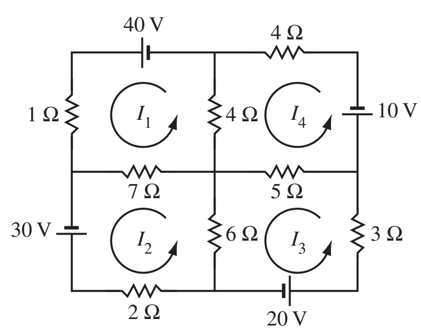
5.



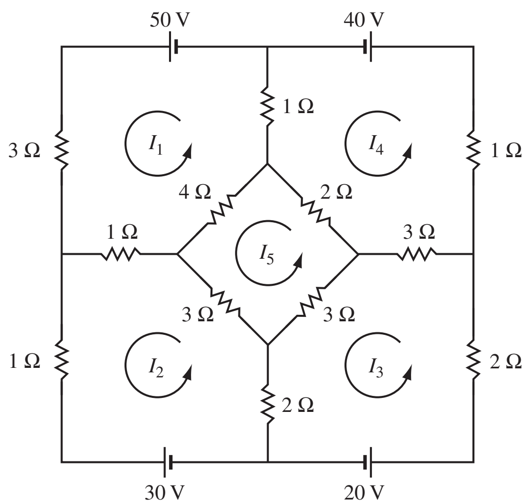
6.



7.



8.



9. In a certain region, about 7% of a city's population moves to the surrounding suburbs each year, and about 5% of the suburban population moves into the city. In 2010, there were 800,000 residents in the city and 500,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2010. Then estimate the populations in the city and in the suburbs two years later, in 2012. (Ignore other factors that might influence the population sizes.)

10. In a certain region, about 6% of a city's population moves to the surrounding suburbs each year, and about 4% of the suburban population moves into the city. In 2010, there were 10,000,000 residents in the city and 800,000 in the suburbs. Set up a difference equation that describes this situation, where $\mathbf{x}_0$ is the initial population in 2010. Then estimate the populations in the city and in the suburbs two years later, in 2012.

11. In 1994, the population of California was 31,524,000, and the population living in the United States but *outside* California was 228,680,000. During the year, it is estimated that 516,100 persons moved from California to elsewhere in the United States, while 381,262 persons moved to California from elsewhere in the United States.[4]

a. Set up the migration matrix for this situation, using five decimal places for the migration rates into and out of California. Let your work show how you produced the migration matrix.

b. [**M**] Compute the projected populations in the year 2000 for California and elsewhere in the United States, assuming that the migration rates did not change during the 6-year period. (These calculations do not take into account births, deaths, or the substantial migration of persons into California and elsewhere in the United States from other countries.)

[4] Migration data supplied by the Demographic Research Unit of the California State Department of Finance.

12. [**M**] Budget® Rent A Car in Wichita, Kansas has a fleet of about 500 cars, at three locations. A car rented at one location may be returned to any of the three locations. The various fractions of cars returned to the three locations are shown in the matrix below. Suppose that on Monday there are 295 cars at the airport (or rented from there), 55 cars at the east side office, and 150 cars at the west side office. What will be the approximate distribution of cars on Wednesday?

Cars Rented From:

Airport	East	West	Returned To:
.97	.05	.10	Airport
.00	.90	.05	East
.03	.05	.85	West

13. [**M**] Let M and $\mathbf{x}_0$ be as in Example 3.

a. Compute the population vectors $\mathbf{x}_k$ for $k = 1, \dots, 20$. Discuss what you find.

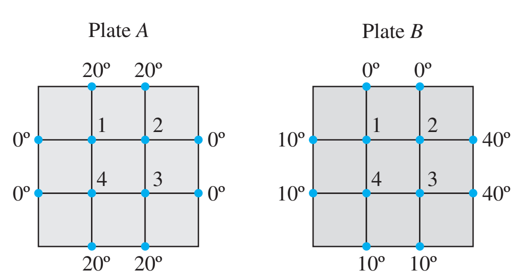
b. Repeat part (a) with an initial population of 350,000 in the city and 650,000 in the suburbs. What do you find?

14. [**M**] Study how changes in boundary temperatures on a steel plate affect the temperatures at interior points on the plate.

a. Begin by estimating the temperatures T_1, T_2, T_3, T_4 at each of the sets of four points on the steel plate shown in the figure. In each case, the value of T_k is approximated by the average of the temperatures at the four closest points. See Exercises 33 and 34 in Section 1.1, where the values (in degrees) turn out to be (20, 27.5, 30, 22.5). How is this list of values related to your results for the points in set (a) and set (b)?

b. Without making any computations, guess the interior temperatures in (a) when the boundary temperatures are all multipled by 3. Check your guess.

c. Finally, make a general conjecture about the correspondence from the list of eight boundary temperatures to the list of four interior temperatures.



(a) (b)

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

CHAPTER 1 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. (If true, cite appropriate facts or theorems. If false, explain why or give a counterexample that shows why the statement is not true in every case.

a. Every matrix is row equivalent to a unique matrix in echelon form.

b. Any system of n linear equations in n variables has at most n solutions.

c. If a system of linear equations has two different solutions, it must have infinitely many solutions.

d. If a system of linear equations has no free variables, then it has a unique solution.

e. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ is transformed into $[\,C \quad \mathbf{d}\,]$ by elementary row operations, then the equations $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have exactly the same solution sets.

f. If a system $A\mathbf{x} = \mathbf{b}$ has more than one solution, then so does the system $A\mathbf{x} = \mathbf{0}$.

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

h. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ can be transformed by elementary row operations into reduced echelon form, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent.

i. If matrices A and B are row equivalent, they have the same reduced echelon form.

j. The equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution if and only if there are no free variables.

k. If A is an $m \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$, then A has m pivot columns.

l. If an $m \times n$ matrix A has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ in $\mathbb{R}^m$.

m. If an $n \times n$ matrix A has n pivot positions, then the reduced echelon form of A is the $n \times n$ identity matrix.

n. If 3×3 matrices A and B each have three pivot positions, then A can be transformed into B by elementary row operations.

o. If A is an $m \times n$ matrix, if the equation $A\mathbf{x} = \mathbf{b}$ has at least two different solutions, and if the equation $A\mathbf{x} = \mathbf{c}$ is consistent, then the equation $A\mathbf{x} = \mathbf{c}$ has many solutions.

p. If A and B are row equivalent $m \times n$ matrices and if the columns of A span $\mathbb{R}^m$, then so do the columns of B.

q. If none of the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$ is a multiple of one of the other vectors, then S is linearly independent.

r. If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, then $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are not in $\mathbb{R}^2$.

s. In some cases, it is possible for four vectors to span $\mathbb{R}^5$.

t. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^m$, then $-\mathbf{u}$ is in Span$\{\mathbf{u}, \mathbf{v}\}$.

u. If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^2$, then $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

v. If $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, then $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$.

w. Suppose that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are in $\mathbb{R}^5$, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$, and $\mathbf{v}_3$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

x. A linear transformation is a function.

y. If A is a 6×5 matrix, the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ cannot map $\mathbb{R}^5$ onto $\mathbb{R}^6$.

z. If A is an $m \times n$ matrix with m pivot columns, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one mapping.

2. Let a and b represent real numbers. Describe the possible solution sets of the (linear) equation $ax = b$. [*Hint:* The number of solutions depends upon a and b.]

3. The solutions (x, y, z) of a single linear equation

$$ax + by + cz = d$$

form a plane in $\mathbb{R}^3$ when a, b, and c are not all zero. Construct sets of three linear equations whose graphs (a) intersect in a single line, (b) intersect in a single point, and (c) have no

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 36 & 51 & 13 \\ 52 & 34 & 74 \\ 0 & 7 & 1.1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 33 \\ 45 \\ 3 \end{bmatrix}$$

CHAPTER 1 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. (If true, cite appropriate facts or theorems. If false, explain why or give a counterexample that shows why the statement is not true in every case.

a. Every matrix is row equivalent to a unique matrix in echelon form.

b. Any system of n linear equations in n variables has at most n solutions.

c. If a system of linear equations has two different solutions, it must have infinitely many solutions.

d. If a system of linear equations has no free variables, then it has a unique solution.

e. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ is transformed into $[\,C \quad \mathbf{d}\,]$ by elementary row operations, then the equations $A\mathbf{x} = \mathbf{b}$ and $C\mathbf{x} = \mathbf{d}$ have exactly the same solution sets.

f. If a system $A\mathbf{x} = \mathbf{b}$ has more than one solution, then so does the system $A\mathbf{x} = \mathbf{0}$.

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

h. If an augmented matrix $[\,A \quad \mathbf{b}\,]$ can be transformed by elementary row operations into reduced echelon form, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent.

i. If matrices A and B are row equivalent, they have the same reduced echelon form.

j. The equation $A\mathbf{x} = \mathbf{0}$ has the trivial solution if and only if there are no free variables.

k. If A is an $m \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^m$, then A has m pivot columns.

l. If an $m \times n$ matrix A has a pivot position in every row, then the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ in $\mathbb{R}^m$.

m. If an $n \times n$ matrix A has n pivot positions, then the reduced echelon form of A is the $n \times n$ identity matrix.

n. If 3×3 matrices A and B each have three pivot positions, then A can be transformed into B by elementary row operations.

o. If A is an $m \times n$ matrix, if the equation $A\mathbf{x} = \mathbf{b}$ has at least two different solutions, and if the equation $A\mathbf{x} = \mathbf{c}$ is consistent, then the equation $A\mathbf{x} = \mathbf{c}$ has many solutions.

p. If A and B are row equivalent $m \times n$ matrices and if the columns of A span $\mathbb{R}^m$, then so do the columns of B.

q. If none of the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$ is a multiple of one of the other vectors, then S is linearly independent.

r. If $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly independent, then $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are not in $\mathbb{R}^2$.

s. In some cases, it is possible for four vectors to span $\mathbb{R}^5$.

t. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^m$, then $-\mathbf{u}$ is in Span$\{\mathbf{u}, \mathbf{v}\}$.

u. If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are nonzero vectors in $\mathbb{R}^2$, then $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

v. If $\mathbf{w}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, then $\mathbf{u}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$.

w. Suppose that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are in $\mathbb{R}^5$, $\mathbf{v}_2$ is not a multiple of $\mathbf{v}_1$, and $\mathbf{v}_3$ is not a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

x. A linear transformation is a function.

y. If A is a 6×5 matrix, the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ cannot map $\mathbb{R}^5$ onto $\mathbb{R}^6$.

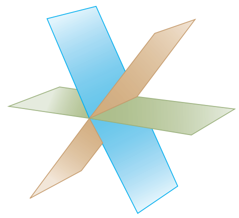
z. If A is an $m \times n$ matrix with m pivot columns, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is a one-to-one mapping.

2. Let a and b represent real numbers. Describe the possible solution sets of the (linear) equation $ax = b$. [*Hint:* The number of solutions depends upon a and b.]

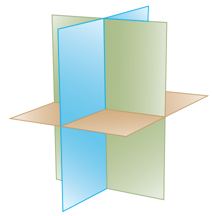
3. The solutions (x, y, z) of a single linear equation

$$ax + by + cz = d$$

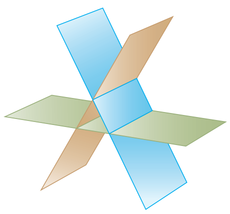
form a plane in $\mathbb{R}^3$ when a, b, and c are not all zero. Construct sets of three linear equations whose graphs (a) intersect in a single line, (b) intersect in a single point, and (c) have no points in common. Typical graphs are illustrated in the figure.



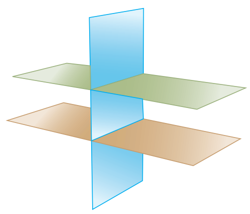
Three planes intersecting in a line
(a)



Three planes intersecting in a point
(b)



Three planes with no intersection
(c)



Three planes with no intersection
(c')

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

5. Determine h and k such that the solution set of the system (i) is empty, (ii) contains a unique solution, and (iii) contains infinitely many solutions.

a. $\begin{aligned} x_1 + 3x_2 &= k \\ 4x_1 + hx_2 &= 8 \end{aligned}$ b. $\begin{aligned} -2x_1 + hx_2 &= 1 \\ 6x_1 + kx_2 &= -2 \end{aligned}$

6. Consider the problem of determining whether the following system of equations is consistent:

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

a. Define appropriate vectors, and restate the problem in terms of linear combinations. Then solve that problem.

b. Define an appropriate matrix, and restate the problem using the phrase "columns of A."

c. Define an appropriate linear transformation T using the matrix in (b), and restate the problem in terms of T.

7. Consider the problem of determining whether the following system of equations is consistent for all b_1, b_2, b_3:

$$\begin{aligned} 2x_1 - 4x_2 - 2x_3 &= b_1 \\ -5x_1 + x_2 + x_3 &= b_2 \\ 7x_1 - 5x_2 - 3x_3 &= b_3 \end{aligned}$$

a. Define appropriate vectors, and restate the problem in terms of Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Then solve that problem.

b. Define an appropriate matrix, and restate the problem using the phrase "columns of A."

c. Define an appropriate linear transformation T using the matrix in (b), and restate the problem in terms of T.

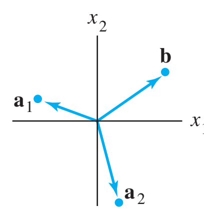
8. Describe the possible echelon forms of the matrix A. Use the notation of Example 1 in Section 1.2.

a. A is a 2×3 matrix whose columns span $\mathbb{R}^2$.

b. A is a 3×3 matrix whose columns span $\mathbb{R}^3$.

9. Write the vector $\begin{bmatrix} 5 \\ 6 \end{bmatrix}$ as the sum of two vectors, one on the line $\{(x, y) : y = 2x\}$ and one on the line $\{(x, y) : y = x/2\}$.

10. Let $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{b}$ be the vectors in $\mathbb{R}^2$ shown in the figure, and let $A = [\mathbf{a}_1 \ \ \mathbf{a}_2]$. Does the equation $A\mathbf{x} = \mathbf{b}$ have a solution? If so, is the solution unique? Explain.



11. Construct a 2×3 matrix A, not in echelon form, such that the solution of $A\mathbf{x} = \mathbf{0}$ is a line in $\mathbb{R}^3$.

12. Construct a 2×3 matrix A, not in echelon form, such that the solution of $A\mathbf{x} = \mathbf{0}$ is a plane in $\mathbb{R}^3$.

13. Write the *reduced* echelon form of a 3×3 matrix A such that the first two columns of A are pivot columns and $A\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

14. Determine the value(s) of a such that $\left\{ \begin{bmatrix} 1 \\ a \end{bmatrix}, \begin{bmatrix} a \\ a+2 \end{bmatrix} \right\}$ is linearly independent.

15. In (a) and (b), suppose the vectors are linearly independent. What can you say about the numbers $a, \ldots, f$? Justify your answers. [*Hint:* Use a theorem for (b).]

a. $\begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \end{bmatrix}$ b. $\begin{bmatrix} a \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} b \\ c \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} d \\ e \\ f \\ 1 \end{bmatrix}$

16. Use Theorem 7 in Section 1.7 to explain why the columns of the matrix A are linearly independent.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 10 \end{bmatrix}$$

17. Explain why a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{R}^5$ must be linearly independent when $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent and $\mathbf{v}_4$ is *not* in Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

18. Suppose $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set in $\mathbb{R}^n$. Show that $\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2\}$ is also linearly independent.

19. Suppose $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are distinct points on one line in $\mathbb{R}^3$. The line need not pass through the origin. Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.

20. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and suppose $T(\mathbf{u}) = \mathbf{v}$. Show that $T(-\mathbf{u}) = -\mathbf{v}$.

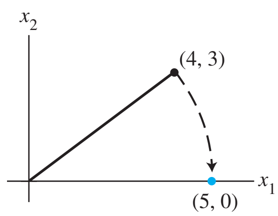
21. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation that reflects each vector through the plane $x_2 = 0$. That is, $T(x_1, x_2, x_3) = (x_1, -x_2, x_3)$. Find the standard matrix of T.

22. Let A be a 3×3 matrix with the property that the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^3$ onto $\mathbb{R}^3$. Explain why the transformation must be one-to-one.

23. A *Givens rotation* is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ used in computer programs to create a zero entry in a vector (usually a column of a matrix). The standard matrix of a Givens rotation in $\mathbb{R}^2$ has the form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \qquad a^2 + b^2 = 1$$

Find a and b such that $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ is rotated into $\begin{bmatrix} 5 \\ 0 \end{bmatrix}$.



A Givens rotation in $\mathbb{R}^2$.

WEB

24. The following equation describes a Givens rotation in $\mathbb{R}^3$. Find a and b.

$$\begin{bmatrix} a & 0 & -b \\ 0 & 1 & 0 \\ b & 0 & a \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2\sqrt{5} \\ 3 \\ 0 \end{bmatrix}, \qquad a^2 + b^2 = 1$$

25. A large apartment building is to be built using modular construction techniques. The arrangement of apartments on any particular floor is to be chosen from one of three basic floor plans. Plan A has 18 apartments on one floor, including 3 three-bedroom units, 7 two-bedroom units, and 8 one-bedroom units. Each floor of plan B includes 4 three-bedroom units, 4 two-bedroom units, and 8 one-bedroom units. Each floor of plan C includes 5 three-bedroom units, 3 two-bedroom units, and 9 one-bedroom units. Suppose the building contains a total of x_1 floors of plan A, x_2 floors of plan B, and x_3 floors of plan C.

a. What interpretation can be given to the vector $x_1 \begin{bmatrix} 3 \\ 7 \\ 8 \end{bmatrix}$?

b. Write a formal linear combination of vectors that expresses the total numbers of three-, two-, and one-bedroom apartments contained in the building.

c. [**M**] Is it possible to design the building with exactly 66 three-bedroom units, 74 two-bedroom units, and 136 one-bedroom units? If so, is there more than one way to do it? Explain your answer.

NUMERICAL NOTES

1. The fastest way to obtain AB on a computer depends on the way in which the computer stores matrices in its memory. The standard high-performance algorithms, such as in LAPACK, calculate AB by columns, as in our definition of the product. (A version of LAPACK written in C++ calculates AB by rows.)
2. The definition of AB lends itself well to parallel processing on a computer. The columns of B are assigned individually or in groups to different processors, which independently and hence simultaneously compute the corresponding columns of AB.

PRACTICE PROBLEMS

1. Since vectors in $\mathbb{R}^n$ may be regarded as $n \times 1$ matrices, the properties of transposes in Theorem 3 apply to vectors, too. Let

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

Compute $(A\mathbf{x})^T$, $\mathbf{x}^T A^T$, $\mathbf{x}\mathbf{x}^T$, and $\mathbf{x}^T\mathbf{x}$. Is $A^T\mathbf{x}^T$ defined?

2. Let A be a 4×4 matrix and let $\mathbf{x}$ be a vector in $\mathbb{R}^4$. What is the fastest way to compute $A^2\mathbf{x}$? Count the multiplications.

2.1 EXERCISES

In Exercises 1 and 2, compute each matrix sum or product if it is defined. If an expression is undefined, explain why. Let

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

$$C = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & 5 \\ -1 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

1. $-2A$, $B - 2A$, AC, CD
2. $A + 3B$, $2C - 3E$, DB, EC

In the rest of this exercise set and in those to follow, assume that each matrix expression is defined. That is, the sizes of the matrices (and vectors) involved "match" appropriately.

3. Let $A = \begin{bmatrix} 2 & -5 \\ 3 & -2 \end{bmatrix}$. Compute $3I_2 - A$ and $(3I_2)A$.
4. Compute $A - 5I_3$ and $(5I_3)A$, where

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

In Exercises 5 and 6, compute the product AB in two ways: (a) by the definition, where $A\mathbf{b}_1$ and $A\mathbf{b}_2$ are computed separately, and (b) by the row–column rule for computing AB.

5. $A = \begin{bmatrix} -1 & 3 \\ 2 & 4 \\ 5 & -3 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -2 \\ -2 & 3 \end{bmatrix}$
6. $A = \begin{bmatrix} 4 & -3 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 4 \\ 3 & -2 \end{bmatrix}$
7. If a matrix A is 5×3 and the product AB is 5×7, what is the size of B?
8. How many rows does B have if BC is a 5×4 matrix?
9. Let $A = \begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 9 \\ -3 & k \end{bmatrix}$. What value(s) of k, if any, will make $AB = BA$?
10. Let $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix}$. Verify that $AB = AC$ and yet $B \neq C$.
11. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$. Compute AD and DA. Explain how the columns or rows of A change when A is multiplied by D on the right or on the left. Find a 3×3 matrix B, not the identity matrix or the zero matrix, such that $AB = BA$.
12. Let $A = \begin{bmatrix} 3 & -6 \\ -2 & 4 \end{bmatrix}$. Construct a 2×2 matrix B such that AB is the zero matrix. Use two different nonzero columns for B.

13. Let $\mathbf{r}_1, \ldots, \mathbf{r}_p$ be vectors in $\mathbb{R}^n$, and let Q be an $m \times n$ matrix. Write the matrix $[\, Q\mathbf{r}_1 \ \cdots \ Q\mathbf{r}_p \,]$ as a *product* of two matrices (neither of which is an identity matrix).

14. Let U be the 3×2 cost matrix described in Example 6 in Section 1.8. The first column of U lists the costs per dollar of output for manufacturing product B, and the second column lists the costs per dollar of output for product C. (The costs are categorized as materials, labor, and overhead.) Let $\mathbf{q}_1$ be a vector in $\mathbb{R}^2$ that lists the output (measured in dollars) of products B and C manufactured during the first quarter of the year, and let $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{q}_4$ be the analogous vectors that list the amounts of products B and C manufactured in the second, third, and fourth quarters, respectively. Give an economic description of the data in the matrix UQ, where $Q = [\, \mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \mathbf{q}_3 \ \ \mathbf{q}_4 \,]$.

Exercises 15 and 16 concern arbitrary matrices A, B, and C for which the indicated sums and products are defined. Mark each statement True or False. Justify each answer.

15. a. If A and B are 2×2 matrices with columns $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{b}_1, \mathbf{b}_2$, respectively, then $AB = [\, \mathbf{a}_1\mathbf{b}_1 \ \ \mathbf{a}_2\mathbf{b}_2 \,]$.
    b. Each column of AB is a linear combination of the columns of B using weights from the corresponding column of A.
    c. $AB + AC = A(B + C)$
    d. $A^T + B^T = (A + B)^T$
    e. The transpose of a product of matrices equals the product of their transposes in the same order.

16. a. The first row of AB is the first row of A multiplied on the right by B.
    b. If A and B are 3×3 matrices and $B = [\, \mathbf{b}_1 \ \ \mathbf{b}_2 \ \ \mathbf{b}_3 \,]$, then $AB = [\, A\mathbf{b}_1 + A\mathbf{b}_2 + A\mathbf{b}_3 \,]$.
    c. If A is an $n \times n$ matrix, then $(A^2)^T = (A^T)^2$
    d. $(ABC)^T = C^T A^T B^T$
    e. The transpose of a sum of matrices equals the sum of their transposes.

17. If $A = \begin{bmatrix} 1 & -3 \\ -3 & 5 \end{bmatrix}$ and $AB = \begin{bmatrix} -3 & -11 \\ 1 & 17 \end{bmatrix}$, determine the first and second columns of B.

18. Suppose the third column of B is all zeros. What can be said about the third column of AB?

19. Suppose the third column of B is the sum of the first two columns. What can be said about the third column of AB? Why?

20. Suppose the first two columns, $\mathbf{b}_1$ and $\mathbf{b}_2$, of B are equal. What can be said about the columns of AB? Why?

21. Suppose the last column of AB is entirely zeros but B itself has no column of zeros. What can be said about the columns of A?

22. Show that if the columns of B are linearly dependent, then so are the columns of AB.

23. Suppose $CA = I_n$ (the $n \times n$ identity matrix). Show that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A cannot have more columns than rows.

24. Suppose A is a $3 \times n$ matrix whose columns span $\mathbb{R}^3$. Explain how to construct an $n \times 3$ matrix D such that $AD = I_3$.

25. Suppose A is an $m \times n$ matrix and there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. Prove that $m = n$ and $C = D$. [*Hint:* Think about the product CAD.]

26. Suppose $AD = I_m$ (the $m \times m$ identity matrix). Show that for any $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution. [*Hint:* Think about the equation $AD\mathbf{b} = \mathbf{b}$.] Explain why A cannot have more rows than columns.

In Exercises 27 and 28, view vectors in $\mathbb{R}^n$ as $n \times 1$ matrices. For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the matrix product $\mathbf{u}^T\mathbf{v}$ is a 1×1 matrix, called the **scalar product**, or **inner product**, of $\mathbf{u}$ and $\mathbf{v}$. It is usually written as a single real number without brackets. The matrix product $\mathbf{u}\mathbf{v}^T$ is an $n \times n$ matrix, called the **outer product** of $\mathbf{u}$ and $\mathbf{v}$. The products $\mathbf{u}^T\mathbf{v}$ and $\mathbf{u}\mathbf{v}^T$ will appear later in the text.

27. Let $\mathbf{u} = \begin{bmatrix} -3 \\ 2 \\ -5 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Compute $\mathbf{u}^T\mathbf{v}$, $\mathbf{v}^T\mathbf{u}$, $\mathbf{u}\mathbf{v}^T$, and $\mathbf{v}\mathbf{u}^T$.

28. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^n$, how are $\mathbf{u}^T\mathbf{v}$ and $\mathbf{v}^T\mathbf{u}$ related? How are $\mathbf{u}\mathbf{v}^T$ and $\mathbf{v}\mathbf{u}^T$ related?

29. Prove Theorem 2(b) and 2(c). Use the row–column rule. The (i, j)-entry in $A(B + C)$ can be written as
$$a_{i1}(b_{1j} + c_{1j}) + \cdots + a_{in}(b_{nj} + c_{nj})$$
or
$$\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$$

30. Prove Theorem 2(d). [*Hint:* The (i, j)-entry in $(rA)B$ is $(ra_{i1})b_{1j} + \cdots + (ra_{in})b_{nj}$.]

31. Show that $I_m A = A$ where A is an $m \times n$ matrix. Assume $I_m\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^m$.

32. Show that $AI_n = A$ when A is an $m \times n$ matrix. [*Hint:* Use the (column) definition of AI_n.]

33. Prove Theorem 3(d). [*Hint:* Consider the jth row of $(AB)^T$.]

34. Give a formula for $(AB\mathbf{x})^T$, where $\mathbf{x}$ is a vector and A and B are matrices of appropriate sizes.

35. [**M**] Read the documentation for your matrix program, and write the commands that will produce the following matrices (without keying in each entry of the matrix).
    a. A 4×5 matrix of zeros
    b. A 5×3 matrix of ones
    c. The 5×5 identity matrix
    d. A 4×4 diagonal matrix, with diagonal entries 3, 4, 2, 5

A useful way to test new ideas in matrix algebra, or to make conjectures, is to make calculations with matrices selected at random. Checking a property for a few matrices does not prove that the property holds in general, but it makes the property more believable. Also, if the property is actually false, making a few calculations may help to discover this.

36. [**M**] Write the command(s) that will create a 5×6 matrix with random entries. In what range of numbers do the entries lie? Tell how to create a 4×4 matrix with random integer entries between -9 and 9. [*Hint:* If x is a random number such that $0 < x < 1$, then $-9.5 < 19(x - .5) < 9.5$.]

37. [**M**] Construct random 4×4 matrices A and B to test whether $AB = BA$. The best way to do this is to compute $AB - BA$ and check whether this difference is the zero matrix. Then test $AB - BA$ for three more pairs of random 4×4 matrices. Report your conclusions.

38. [**M**] Construct a random 5×5 matrix A and test whether $(A + I)(A - I) = A^2 - I$. The best way to do this is to compute $(A + I)(A - I) - (A^2 - I)$ and verify that this difference is the zero matrix. Do this for three random matrices. Then test $(A + B)(A - B) = A^2 - B^2$ the same way for three pairs of random 4×4 matrices. Report your conclusions.

39. [**M**] Use at least three pairs of random 4×4 matrices A and B to test the equalities $(A + B)^T = A^T + B^T$ and $(AB)^T = B^T A^T$, as well as $(AB)^T = A^T B^T$. (See Exercise 37.) Report your conclusions. [*Note*: Most matrix programs use A' for A^T.]

40. [**M**] Let

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

Compute S^k for $k = 2, \ldots, 6$.

41. [**M**] Describe in words what happens when A^5, A^{10}, A^{20}, and A^{30} are computed for

$$A = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/3 & 1/6 \\ 1/4 & 1/6 & 7/12 \end{bmatrix}$$

SOLUTIONS TO PRACTICE PROBLEMS

1. $A\mathbf{x} = \begin{bmatrix} 1 & -3 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$. So $(A\mathbf{x})^T = \begin{bmatrix} -4 & 2 \end{bmatrix}$. Also,

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

The quantities $(A\mathbf{x})^T$ and $\mathbf{x}^T A^T$ are equal, by Theorem 3(d). Next,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T\mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 25 + 9 \end{bmatrix} = 34$$

A 1×1 matrix such as $\mathbf{x}^T\mathbf{x}$ is usually written without the brackets. Finally, $A^T\mathbf{x}^T$ is not defined, because $\mathbf{x}^T$ does not have two rows to match the two columns of A^T.

2. The fastest way to compute $A^2\mathbf{x}$ is to compute $A(A\mathbf{x})$. The product $A\mathbf{x}$ requires 16 multiplications, 4 for each entry, and $A(A\mathbf{x})$ requires 16 more. In contrast, the product A^2 requires 64 multiplications, 4 for each of the 16 entries in A^2. After that, $A^2\mathbf{x}$ takes 16 more multiplications, for a total of 80.

2.2 THE INVERSE OF A MATRIX

Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers. This section investigates the matrix analogue of the reciprocal, or multiplicative inverse, of a nonzero number.

NUMERICAL NOTE

WEB

In practical work, A^{-1} is seldom computed, unless the entries of A^{-1} are needed. Computing both A^{-1} and $A^{-1}\mathbf{b}$ takes about three times as many arithmetic operations as solving $A\mathbf{x} = \mathbf{b}$ by row reduction, and row reduction may be more accurate.

PRACTICE PROBLEMS

1. Use determinants to determine which of the following matrices are invertible.

a. $\begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix}$ b. $\begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix}$ c. $\begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix}$

2. Find the inverse of the matrix $A = \begin{bmatrix} 1 & -2 & -1 \\ -1 & 5 & 6 \\ 5 & -4 & 5 \end{bmatrix}$, if it exists.

2.2 EXERCISES

Find the inverses of the matrices in Exercises 1–4.

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

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

5. Use the inverse found in Exercise 1 to solve the system

$$\begin{aligned} 8x_1 + 6x_2 &= 2 \\ 5x_1 + 4x_2 &= -1 \end{aligned}$$

6. Use the inverse found in Exercise 3 to solve the system

$$\begin{aligned} 7x_1 + 3x_2 &= -9 \\ -6x_1 - 3x_2 &= 4 \end{aligned}$$

7. Let $A = \begin{bmatrix} 1 & 2 \\ 5 & 12 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$, and $\mathbf{b}_4 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.

a. Find A^{-1}, and use it to solve the four equations

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad A\mathbf{x} = \mathbf{b}_3, \quad A\mathbf{x} = \mathbf{b}_4$$

b. The four equations in part (a) can be solved by the *same* set of row operations, since the coefficient matrix is the same in each case. Solve the four equations in part (a) by row reducing the augmented matrix $[\,A \;\; \mathbf{b}_1 \;\; \mathbf{b}_2 \;\; \mathbf{b}_3 \;\; \mathbf{b}_4\,]$.

8. Suppose P is invertible and $A = PBP^{-1}$. Solve for B in terms of A.

In Exercises 9 and 10, mark each statement True or False. Justify each answer.

9. a. In order for a matrix B to be the inverse of A, the equations $AB = I$ and $BA = I$ must both be true.

b. If A and B are $n \times n$ and invertible, then $A^{-1}B^{-1}$ is the inverse of AB.

c. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $ab - cd \neq 0$, then A is invertible.

d. If A is an invertible $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ is consistent for *each* $\mathbf{b}$ in $\mathbb{R}^n$.

e. Each elementary matrix is invertible.

10. a. If A is invertible, then elementary row operations that reduce A to the identity I_n also reduce A^{-1} to I_n.

b. If A is invertible, then the inverse of A^{-1} is A itself.

c. A product of invertible $n \times n$ matrices is invertible, and the inverse of the product is the product of their inverses in the same order.

d. If A is an $n \times n$ matrix and $A\mathbf{x} = \mathbf{e}_j$ is consistent for every $j \in \{1, 2, \ldots, n\}$, then A is invertible. Note: $\mathbf{e}_1, \ldots, \mathbf{e}_n$ represent the columns of the identity matrix.

e. If A can be row reduced to the identity matrix, then A must be invertible.

11. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Show that the equation $AX = B$ has a unique solution $A^{-1}B$.

12. Use matrix algebra to show that if A is invertible and D satisfies $AD = I$, then $D = A^{-1}$.

13. Suppose $AB = AC$, where B and C are $n \times p$ matrices and A is invertible. Show that $B = C$. Is this true, in general, when A is not invertible?

14. Suppose $(B - C)D = 0$, where B and C are $m \times n$ matrices and D is invertible. Show that $B = C$.

15. Let A be an invertible $n \times n$ matrix, and let B be an $n \times p$ matrix. Explain why $A^{-1}B$ can be computed by row reduction:

If $[\, A \quad B \,] \sim \cdots \sim [\, I \quad X \,]$, then $X = A^{-1}B$.

If A is larger than 2×2, then row reduction of $[A \quad B]$ is much faster than computing both A^{-1} and $A^{-1}B$.

16. Suppose A and B are $n \times n$ matrices, B is invertible, and AB is invertible. Show that A is invertible. [*Hint:* Let $C = AB$, and solve this equation for A.]

17. Suppose A, B, and C are invertible $n \times n$ matrices. Show that ABC is also invertible by producing a matrix D such that $(ABC)D = I$ and $D(ABC) = I$.

18. Solve the equation $AB = BC$ for A, assuming that A, B, and C are square and B is invertible.

19. If A, B, and C are $n \times n$ invertible matrices, does the equation $C^{-1}(A + X)B^{-1} = I_n$ have a solution, X? If so, find it.

20. Suppose A, B, and X are $n \times n$ matrices with A, X, and $A - AX$ invertible, and suppose

$$(A - AX)^{-1} = X^{-1}B \tag{3}$$

a. Explain why B is invertible.

b. Solve equation (3) for X. If a matrix needs to be inverted, explain why that matrix is invertible.

21. Explain why the columns of an $n \times n$ matrix A are linearly independent when A is invertible.

22. Explain why the columns of an $n \times n$ matrix A span $\mathbb{R}^n$ when A is invertible. [*Hint:* Review Theorem 4 in Section 1.4.]

23. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Explain why A has n pivot columns and A is row equivalent to I_n. By Theorem 7, this shows that A must be invertible. (This exercise and Exercise 24 will be cited in Section 2.3.)

24. Suppose A is $n \times n$ and the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Explain why A must be invertible. [*Hint:* Is A row equivalent to I_n?]

Exercises 25 and 26 prove Theorem 4 for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

25. Show that if $ad - bc = 0$, then the equation $A\mathbf{x} = \mathbf{0}$ has more than one solution. Why does this imply that A is not invertible? [*Hint:* First, consider $a = b = 0$. Then, if a and b are not both zero, consider the vector $\mathbf{x} = \begin{bmatrix} -b \\ a \end{bmatrix}$.]

26. Show that if $ad - bc \neq 0$, the formula for A^{-1} works.

Exercises 27 and 28 prove special cases of the facts about elementary matrices stated in the box following Example 5. Here A is a 3×3 matrix and $I = I_3$. (A general proof would require slightly more notation.)

27. Let A be a 3×3 matrix.

a. Use equation (2) from Section 2.1 to show that $\text{row}_i(A) = \text{row}_i(I) \cdot A$, for $i = 1, 2, 3$.

b. Show that if rows 1 and 2 of A are interchanged, then the result may be written as EA, where E is an elementary matrix formed by interchanging rows 1 and 2 of I.

c. Show that if row 3 of A is multiplied by 5, then the result may be written as EA, where E is formed by multiplying row 3 of I by 5.

28. Suppose row 2 of A is replaced by $\text{row}_2(A) - 3 \cdot \text{row}_1(A)$. Show that the result is EA, where E is formed from I by replacing $\text{row}_2(I)$ by $\text{row}_2(I) - 3 \cdot \text{row}_1(A)$.

Find the inverses of the matrices in Exercises 29–32, if they exist. Use the algorithm introduced in this section.

29. $\begin{bmatrix} 1 & -3 \\ 4 & -9 \end{bmatrix}$ **30.** $\begin{bmatrix} 3 & 6 \\ 4 & 7 \end{bmatrix}$

31. $\begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{bmatrix}$ **32.** $\begin{bmatrix} 1 & 2 & -1 \\ -4 & -7 & 3 \\ -2 & -6 & 4 \end{bmatrix}$

33. Use the algorithm from this section to find the inverses of

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

Let A be the corresponding $n \times n$ matrix, and let B be its inverse. Guess the form of B, and then show that $AB = I$.

34. Repeat the strategy of Exercise 33 to guess the inverse B of

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 2 & 2 & 0 & & 0 \\ 3 & 3 & 3 & & 0 \\ \vdots & & & \ddots & \vdots \\ n & n & n & \cdots & n \end{bmatrix}.$$

Show that $AB = I$.

35. Let $A = \begin{bmatrix} -1 & -7 & -3 \\ 2 & 15 & 6 \\ 1 & 3 & 2 \end{bmatrix}$. Find the third column of A^{-1} without computing the other columns.

36. [**M**] Let $A = \begin{bmatrix} -25 & -9 & -27 \\ 536 & 185 & 537 \\ 154 & 52 & 143 \end{bmatrix}$. Find the second and third columns of A^{-1} without computing the first column.

37. Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$. Construct a 2×3 matrix C (by trial and error) using only 1, -1, and 0 as entries, such that $CA = I_2$. Compute AC and note that $AC \neq I_3$.

38. Let $A = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{bmatrix}$. Construct a 4×2 matrix D using only 1 and 0 as entries, such that $AD = I_2$. Is it possible that $CA = I_4$ for some 4×2 matrix C? Why or why not?

39. [**M**] Let

$$D = \begin{bmatrix} .011 & .003 & .001 \\ .003 & .009 & .003 \\ .001 & .003 & .011 \end{bmatrix}$$

be a flexibility matrix, with flexibility measured in inches per pound. Suppose that forces of 40, 50, and 30 lb are applied at points 1, 2, and 3, respectively, in Fig. 1 of Example 3. Find the corresponding deflections.

40. [**M**] Compute the stiffness matrix D^{-1} for D in Exercise 39. List the forces needed to produce a deflection of .04 in. at point 3, with zero deflections at the other points.

41. [**M**] Let

$$D = \begin{bmatrix} .0130 & .0050 & .0020 & .0010 \\ .0050 & .0100 & .0040 & .0020 \\ .0020 & .0040 & .0100 & .0050 \\ .0010 & .0020 & .0050 & .0130 \end{bmatrix}$$

be a flexibility matrix for an elastic beam such as the one in Example 3, with four points at which force is applied. Units are centimeters per newton of force. Measurements at the four points show deflections of .07, .12, .16, and .12 cm. Determine the forces at the four points.

42. [**M**] With D as in Exercise 41, determine the forces that produce a deflection of .22 cm at the second point on the beam, with zero deflections at the other three points. How is the answer related to the entries in D^{-1}? [*Hint:* First answer the question when the deflection is 1 cm at the second point.]

SOLUTIONS TO PRACTICE PROBLEMS

1. a. $\det \begin{bmatrix} 3 & -9 \\ 2 & 6 \end{bmatrix} = 3 \cdot 6 - (-9) \cdot 2 = 18 + 18 = 36$. The determinant is nonzero, so the matrix is invertible.

b. $\det \begin{bmatrix} 4 & -9 \\ 0 & 5 \end{bmatrix} = 4 \cdot 5 - (-9) \cdot 0 = 20 \neq 0$. The matrix is invertible.

c. $\det \begin{bmatrix} 6 & -9 \\ -4 & 6 \end{bmatrix} = 6 \cdot 6 - (-9)(-4) = 36 - 36 = 0$. The matrix is not invertible.

2. $$[\,A \quad I\,] \sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ -1 & 5 & 6 & 0 & 1 & 0 \\ 5 & -4 & 5 & 0 & 0 & 1 \end{bmatrix}$$

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

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

So $[\,A \quad I\,]$ is row equivalent to a matrix of the form $[\,B \quad D\,]$, where B is square and has a row of zeros. Further row operations will not transform B into I, so we stop. A does not have an inverse.

2.3 CHARACTERIZATIONS OF INVERTIBLE MATRICES

This section provides a review of most of the concepts introduced in Chapter 1, in relation to systems of n linear equations in n unknowns and to *square* matrices. The main result is Theorem 8.

THEOREM 9

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation and let A be the standard matrix for T. Then T is invertible if and only if A is an invertible matrix. In that case, the linear transformation S given by $S(\mathbf{x}) = A^{-1}\mathbf{x}$ is the unique function satisfying equations (1) and (2).

PROOF Suppose that T is invertible. Then (2) shows that T is onto $\mathbb{R}^n$, for if $\mathbf{b}$ is in $\mathbb{R}^n$ and $\mathbf{x} = S(\mathbf{b})$, then $T(\mathbf{x}) = T(S(\mathbf{b})) = \mathbf{b}$, so each $\mathbf{b}$ is in the range of T. Thus A is invertible, by the Invertible Matrix Theorem, statement (i).

Conversely, suppose that A is invertible, and let $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Then, S is a linear transformation, and S obviously satisfies (1) and (2). For instance,

$$S(T(\mathbf{x})) = S(A\mathbf{x}) = A^{-1}(A\mathbf{x}) = \mathbf{x}$$

Thus T is invertible. The proof that S is unique is outlined in Exercise 38. ■

EXAMPLE 2 What can you say about a one-to-one linear transformation T from $\mathbb{R}^n$ into $\mathbb{R}^n$?

SOLUTION The columns of the standard matrix A of T are linearly independent (by Theorem 12 in Section 1.9). So A is invertible, by the Invertible Matrix Theorem, and T maps $\mathbb{R}^n$ *onto* $\mathbb{R}^n$. Also, T is invertible, by Theorem 9. ■

NUMERICAL NOTES

In practical work, you might occasionally encounter a "nearly singular" or **ill-conditioned** matrix—an invertible matrix that can become singular if some of its entries are changed ever so slightly. In this case, row reduction may produce fewer than n pivot positions, as a result of roundoff error. Also, roundoff error can sometimes make a singular matrix appear to be invertible.

WEB

Some matrix programs will compute a **condition number** for a square matrix. The larger the condition number, the closer the matrix is to being singular. The condition number of the identity matrix is 1. A singular matrix has an infinite condition number. In extreme cases, a matrix program may not be able to distinguish between a singular matrix and an ill-conditioned matrix.

Exercises 41–45 show that matrix computations can produce substantial error when a condition number is large.

PRACTICE PROBLEMS

1. Determine if $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 4 \\ 2 & 3 & 4 \end{bmatrix}$ is invertible.
2. Suppose that for a certain $n \times n$ matrix A, statement (g) of the Invertible Matrix Theorem is *not* true. What can you say about equations of the form $A\mathbf{x} = \mathbf{b}$?
3. Suppose that A and B are $n \times n$ matrices and the equation $AB\mathbf{x} = \mathbf{0}$ has a nontrivial solution. What can you say about the matrix AB?

2.3 EXERCISES

Unless otherwise specified, assume that all matrices in these exercises are $n \times n$. Determine which of the matrices in Exercises 1–10 are invertible. Use as few calculations as possible. Justify your answers.

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

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

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

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

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

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

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

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

9. [M] $\begin{bmatrix} 4 & 0 & -3 & -7 \\ -6 & 9 & 9 & 9 \\ 7 & -5 & 10 & 19 \\ -1 & 2 & 4 & -1 \end{bmatrix}$

10. [M] $\begin{bmatrix} 5 & 3 & 1 & 7 & 9 \\ 6 & 4 & 2 & 8 & -8 \\ 7 & 5 & 3 & 10 & 9 \\ 9 & 6 & 4 & -9 & -5 \\ 8 & 5 & 2 & 11 & 4 \end{bmatrix}$

In Exercises 11 and 12, the matrices are all $n \times n$. Each part of the exercises is an *implication* of the form "If ⟨ statement 1 ⟩, then ⟨ statement 2 ⟩." Mark an implication as True if the truth of ⟨ statement 2 ⟩ *always* follows whenever ⟨ statement 1 ⟩ happens to be true. An implication is False if there is an instance in which ⟨ statement 2 ⟩ is false but ⟨ statement 1 ⟩ is true. Justify each answer.

11. a. If the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, then A is row equivalent to the $n \times n$ identity matrix.

b. If the columns of A span $\mathbb{R}^n$, then the columns are linearly independent.

c. If A is an $n \times n$ matrix, then the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

d. If the equation $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, then A has fewer than n pivot positions.

e. If A^T is not invertible, then A is not invertible.

12. a. If there is an $n \times n$ matrix D such that $AD = I$, then $DA = I$.

b. If the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ into $\mathbb{R}^n$, then the row reduced echelon form of A is I.

c. If the columns of A are linearly independent, then the columns of A span $\mathbb{R}^n$.

d. If the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one.

e. If there is a $\mathbf{b}$ in $\mathbb{R}^n$ such that the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then the solution is unique.

13. An $m \times n$ **upper triangular matrix** is one whose entries *below* the main diagonal are 0's (as in Exercise 8). When is a square upper triangular matrix invertible? Justify your answer.

14. An $m \times n$ **lower triangular matrix** is one whose entries *above* the main diagonal are 0's (as in Exercise 3). When is a square lower triangular matrix invertible? Justify your answer.

15. Is it possible for a 4×4 matrix to be invertible when its columns do not span $\mathbb{R}^4$? Why or why not?

16. If an $n \times n$ matrix A is invertible, then the columns of A^T are linearly independent. Explain why.

17. Can a square matrix with two identical columns be invertible? Why or why not?

18. Can a square matrix with two identical rows be invertible? Why or why not?

19. If the columns of a 7×7 matrix D are linearly independent, what can be said about the solutions of $D\mathbf{x} = \mathbf{b}$? Why?

20. If A is a 5×5 matrix and the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^5$, is it possible that for some $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has more than one solution? Why or why not?

21. If the equation $C\mathbf{u} = \mathbf{v}$ has more than one solution for some $\mathbf{v}$ in $\mathbb{R}^n$, can the columns of the $n \times n$ matrix C span $\mathbb{R}^n$? Why or why not?

22. If $n \times n$ matrices E and F have the property that $EF = I$, then E and F commute. Explain why.

23. Assume that F is an $n \times n$ matrix. If the equation $F\mathbf{x} = \mathbf{y}$ is inconsistent for some $\mathbf{y}$ in $\mathbb{R}^n$, what can you say about the equation $F\mathbf{x} = \mathbf{0}$? Why?

24. If an $n \times n$ matrix G cannot be row reduced to I_n, what can you say about the columns of G? Why?

25. Verify the boxed statement preceding Example 1.

26. Explain why the columns of A^2 span $\mathbb{R}^n$ whenever the columns of an $n \times n$ matrix A are linearly independent.

27. Let A and B be $n \times n$ matrices. Show that if AB is invertible, so is A. You cannot use Theorem 6(b), because you cannot *assume* that A and B are invertible. [*Hint:* There is a matrix W such that $ABW = I$. Why?]

28. Let A and B be $n \times n$ matrices. Show that if AB is invertible, so is B.

29. If A is an $n \times n$ matrix and the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one, what else can you say about this transformation? Justify your answer.

30. If A is an $n \times n$ matrix and the equation $A\mathbf{x} = \mathbf{b}$ has more than one solution for some $\mathbf{b}$, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is not one-to-one. What else can you say about this transformation? Justify your answer.

31. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$. Without using Theorems 5 or 8, explain why each equation $A\mathbf{x} = \mathbf{b}$ has in fact exactly one solution.

32. Suppose A is an $n \times n$ matrix with the property that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. Without using the Invertible Matrix Theorem, explain directly why the equation $A\mathbf{x} = \mathbf{b}$ must have a solution for each $\mathbf{b}$ in $\mathbb{R}^n$.

In Exercises 33 and 34, T is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$. Show that T is invertible and find a formula for T^{-1}.

33. $T(x_1, x_2) = (-5x_1 + 9x_2, 4x_1 - 7x_2)$

34. $T(x_1, x_2) = (2x_1 - 8x_2, -2x_1 + 7x_2)$

35. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Explain why T is both one-to-one and onto $\mathbb{R}^n$. Use equations (1) and (2). Then give a second explanation using one or more theorems.

36. Suppose a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ has the property that $T(\mathbf{u}) = T(\mathbf{v})$ for some pair of distinct vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. Can T map $\mathbb{R}^n$ onto $\mathbb{R}^n$? Why or why not?

37. Suppose T and U are linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that $T(U(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Is it true that $U(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$? Why or why not?

38. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation, and let S and U be functions from $\mathbb{R}^n$ into $\mathbb{R}^n$ such that $S(T(\mathbf{x})) = \mathbf{x}$ and $U(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. Show that $U(\mathbf{v}) = S(\mathbf{v})$ for all $\mathbf{v}$ in $\mathbb{R}^n$. This will show that T has a unique inverse, as asserted in Theorem 9. [*Hint:* Given any $\mathbf{v}$ in $\mathbb{R}^n$, we can write $\mathbf{v} = T(\mathbf{x})$ for some $\mathbf{x}$. Why? Compute $S(\mathbf{v})$ and $U(\mathbf{v})$.]

39. Let T be a linear transformation that maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Show that T^{-1} exists and maps $\mathbb{R}^n$ onto $\mathbb{R}^n$. Is T^{-1} also one-to-one?

40. Suppose T and S satisfy the invertibility equations (1) and (2), where T is a linear transformation. Show directly that S is a linear transformation. [*Hint:* Given $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, let $\mathbf{x} = S(\mathbf{u})$, $\mathbf{y} = S(\mathbf{v})$. Then $T(\mathbf{x}) = \mathbf{u}$, $T(\mathbf{y}) = \mathbf{v}$. Why? Apply S to both sides of the equation $T(\mathbf{x}) + T(\mathbf{y}) = T(\mathbf{x} + \mathbf{y})$. Also, consider $T(c\mathbf{x}) = cT(\mathbf{x})$.]

41. **[M]** Suppose an experiment leads to the following system of equations:

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.249 \\ 1.6x_1 + 1.1x_2 &= 6.843 \end{aligned} \tag{3}$$

a. Solve system (3), and then solve system (4), below, in which the data on the right have been rounded to two decimal places. In each case, find the *exact* solution.

$$\begin{aligned} 4.5x_1 + 3.1x_2 &= 19.25 \\ 1.6x_1 + 1.1x_2 &= 6.84 \end{aligned} \tag{4}$$

b. The entries in system (4) differ from those in system (3) by less than .05%. Find the percentage error when using the solution of (4) as an approximation for the solution of (3).

c. Use a matrix program to produce the condition number of the coefficient matrix in (3).

Exercises 42–44 show how to use the condition number of a matrix A to estimate the accuracy of a computed solution of $A\mathbf{x} = \mathbf{b}$. If the entries of A and $\mathbf{b}$ are accurate to about r significant digits and if the condition number of A is approximately 10^k (with k a positive integer), then the computed solution of $A\mathbf{x} = \mathbf{b}$ should usually be accurate to at least $r - k$ significant digits.

42. **[M]** Let A be the matrix in Exercise 9. Find the condition number of A. Construct a random vector $\mathbf{x}$ in $\mathbb{R}^4$ and compute $\mathbf{b} = A\mathbf{x}$. Then use a matrix program to compute the solution $\mathbf{x}_1$ of $A\mathbf{x} = \mathbf{b}$. To how many digits do $\mathbf{x}$ and $\mathbf{x}_1$ agree? Find out the number of digits the matrix program stores accurately, and report how many digits of accuracy are lost when $\mathbf{x}_1$ is used in place of the exact solution $\mathbf{x}$.

43. **[M]** Repeat Exercise 42 for the matrix in Exercise 10.

44. **[M]** Solve an equation $A\mathbf{x} = \mathbf{b}$ for a suitable $\mathbf{b}$ to find the last column of the inverse of the *fifth-order Hilbert matrix*

$$A = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1/2 & 1/3 & 1/4 & 1/5 & 1/6 \\ 1/3 & 1/4 & 1/5 & 1/6 & 1/7 \\ 1/4 & 1/5 & 1/6 & 1/7 & 1/8 \\ 1/5 & 1/6 & 1/7 & 1/8 & 1/9 \end{bmatrix}$$

How many digits in each entry of $\mathbf{x}$ do you expect to be correct? Explain. [*Note:* The exact solution is (630, −12600, 56700, −88200, 44100).]

45. **[M]** Some matrix programs, such as MATLAB, have a command to create Hilbert matrices of various sizes. If possible, use an inverse command to compute the inverse of a twelfth-order or larger Hilbert matrix, A. Compute AA^{-1}. Report what you find.

SG Mastering: Reviewing and Reflecting 2–13

SOLUTIONS TO PRACTICE PROBLEMS

1. The columns of A are obviously linearly dependent because columns 2 and 3 are multiples of column 1. Hence A cannot be invertible, by the Invertible Matrix Theorem.
2. If statement (g) is *not* true, then the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent for at least one $\mathbf{b}$ in $\mathbb{R}^n$.
3. Apply the Invertible Matrix Theorem to the matrix AB in place of A. Then statement (d) becomes: $AB\mathbf{x} = \mathbf{0}$ has only the trivial solution. This is not true. So AB is not invertible.

2.4 PARTITIONED MATRICES

A key feature of our work with matrices has been the ability to regard a matrix A as a list of column vectors rather than just a rectangular array of numbers. This point of view has been so useful that we wish to consider other **partitions** of A, indicated by horizontal and vertical dividing rules, as in Example 1 below. Partitioned matrices appear in most modern applications of linear algebra because the notation highlights essential structures in matrix analysis, as in the chapter introductory example on aircraft design. This section provides an opportunity to review matrix algebra and use the Invertible Matrix Theorem.

EXAMPLE 1 The matrix

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

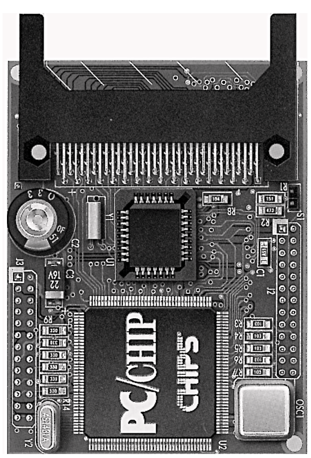
can also be written as the 2×3 **partitioned** (or **block**) **matrix**

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

whose entries are the *blocks* (or *submatrices*)

$$A_{11} = \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, \quad A_{13} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 7 \end{bmatrix}, \quad A_{23} = \begin{bmatrix} -4 \end{bmatrix}$$ ■



EXAMPLE 2 When a matrix A appears in a mathematical model of a physical system such as an electrical network, a transportation system, or a large corporation, it may be natural to regard A as a partitioned matrix. For instance, if a microcomputer circuit board consists mainly of three VLSI (very large-scale integrated) microchips, then the matrix for the circuit board might have the general form

$$A = \left[\begin{array}{c|c|c} A_{11} & A_{12} & A_{13} \\ \hline A_{21} & A_{22} & A_{23} \\ \hline A_{31} & A_{32} & A_{33} \end{array}\right]$$

The submatrices on the "diagonal" of A—namely, A_{11}, A_{22}, and A_{33}—concern the three VLSI chips, while the other submatrices depend on the interconnections among those microchips. ■

The exercises that follow give practice with matrix algebra and illustrate typical calculations found in applications.

PRACTICE PROBLEMS

1. Show that $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible and find its inverse.
2. Compute X^TX, where X is partitioned as $\begin{bmatrix} X_1 & X_2 \end{bmatrix}$.

2.4 EXERCISES

In Exercises 1–9, assume that the matrices are partitioned conformably for block multiplication. Compute the products shown in Exercises 1–4.

1. $\begin{bmatrix} I & 0 \\ E & I \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

2. $\begin{bmatrix} E & 0 \\ 0 & F \end{bmatrix}\begin{bmatrix} P & Q \\ R & S \end{bmatrix}$

3. $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

4. $\begin{bmatrix} I & 0 \\ -E & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$

In Exercises 5–8, find formulas for X, Y, and Z in terms of A, B, and C, and justify your calculations. In some cases, you may need to make assumptions about the size of a matrix in order to produce a formula. [*Hint:* Compute the product on the left, and set it equal to the right side.]

5. $\begin{bmatrix} A & B \\ C & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ X & Y \end{bmatrix} = \begin{bmatrix} 0 & I \\ Z & 0 \end{bmatrix}$

6. $\begin{bmatrix} X & 0 \\ Y & Z \end{bmatrix}\begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

7. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix}\begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

8. $\begin{bmatrix} A & B \\ 0 & I \end{bmatrix}\begin{bmatrix} X & Y & Z \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$

9. Suppose B_{11} is an invertible matrix. Find matrices A_{21} and A_{31} (in terms of the blocks of B) such that the product below has the form indicated. Also, compute C_{22} (in terms of the blocks of B). [*Hint:* Compute the product on the left, and set it equal to the right side.]

$$\begin{bmatrix} I & 0 & 0 \\ A_{21} & I & 0 \\ A_{31} & 0 & I \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ 0 & C_{22} \\ 0 & C_{32} \end{bmatrix}$$

10. The inverse of

$$\begin{bmatrix} I & 0 & 0 \\ A & I & 0 \\ B & D & I \end{bmatrix} \quad \text{is} \quad \begin{bmatrix} I & 0 & 0 \\ P & I & 0 \\ Q & R & I \end{bmatrix}.$$

Find P, Q, and R.

In Exercises 11 and 12, mark each statement True or False. Justify each answer.

11. a. If $A = [\,A_1 \quad A_2\,]$ and $B = [\,B_1 \quad B_2\,]$, with A_1 and A_2 the same sizes as B_1 and B_2, respectively, then $A + B = [\,A_1 + B_1 \quad A_2 + B_2\,]$.

b. If $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ and $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$, then the partitions of A and B are conformable for block multiplication.

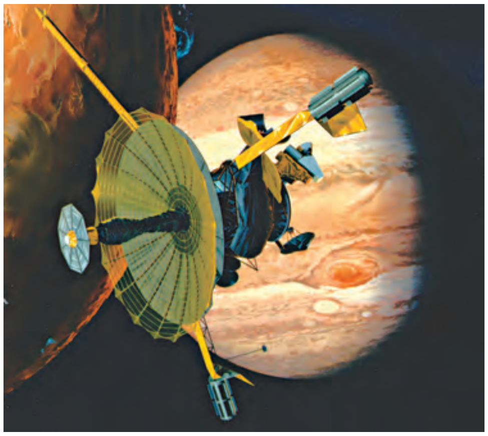
12. a. If A_1, A_2, B_1, and B_2 are $n \times n$ matrices, $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, and $B = [\,B_1 \quad B_2\,]$, then the product BA is defined, but AB is not.

b. If $A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$, then the transpose of A is $A^T = \begin{bmatrix} P^T & Q^T \\ R^T & S^T \end{bmatrix}$.

13. Let $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$, where B and C are square. Show that A is invertible if and only if both B and C are invertible.

14. Show that the block upper triangular matrix A in Example 5 is invertible if and only if both A_{11} and A_{22} are invertible. [*Hint:* If A_{11} and A_{22} are invertible, the formula for A^{-1} given in Example 5 actually works as the inverse of A.] This fact about A is an important part of several computer algorithms that estimate eigenvalues of matrices. Eigenvalues are discussed in Chapter 5.

15. When a deep space probe is launched, corrections may be necessary to place the probe on a precisely calculated trajectory. Radio telemetry provides a stream of vectors, $\mathbf{x}_1, \ldots, \mathbf{x}_k$, giving information at different times about how the probe's position compares with its planned trajectory. Let X_k be the matrix $[\mathbf{x}_1 \ \cdots \ \mathbf{x}_k]$. The matrix $G_k = X_k X_k^T$ is computed as the radar data are analyzed. When $\mathbf{x}_{k+1}$ arrives, a new G_{k+1} must be computed. Since the data vectors arrive at high speed, the computational burden could be severe. But partitioned matrix multiplication helps tremendously. Compute the column–row expansions of G_k and G_{k+1}, and describe what must be computed in order to *update* G_k to form G_{k+1}.



The probe Galileo was launched October 18, 1989, and arrived near Jupiter in early December 1995.

16. Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. If A_{11} is invertible, then the matrix $S = A_{22} - A_{21}A_{11}^{-1}A_{12}$ is called the **Schur complement** of A_{11}. Likewise, if A_{22} is invertible, the matrix $A_{11} - A_{12}A_{22}^{-1}A_{21}$ is called the Schur complement of A_{22}. Suppose A_{11} is invertible. Find X and Y such that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}\begin{bmatrix} A_{11} & 0 \\ 0 & S \end{bmatrix}\begin{bmatrix} I & Y \\ 0 & I \end{bmatrix} \tag{7}$$

17. Suppose the block matrix A on the left side of (7) is invertible and A_{11} is invertible. Show that the Schur complement S of A_{11} is invertible. [*Hint:* The outside factors on the right side of (7) are always invertible. Verify this.] When A and A_{11} are both invertible, (7) leads to a formula for A^{-1}, using S^{-1}, A_{11}^{-1}, and the other entries in A.

18. Let X be an $m \times n$ data matrix such that X^TX is invertible, and let $M = I_m - X(X^TX)^{-1}X^T$. Add a column $\mathbf{x}_0$ to the data and form

$$W = \begin{bmatrix} X & \mathbf{x}_0 \end{bmatrix}$$

Compute W^TW. The (1, 1)-entry is X^TX. Show that the Schur complement (Exercise 16) of X^TX can be written in the form $\mathbf{x}_0^TM\mathbf{x}_0$. It can be shown that the quantity $(\mathbf{x}_0^TM\mathbf{x}_0)^{-1}$ is the (2, 2)-entry in $(W^TW)^{-1}$. This entry has a useful statistical interpretation, under appropriate hypotheses.

In the study of engineering control of physical systems, a standard set of differential equations is transformed by Laplace transforms into the following system of linear equations:

$$\begin{bmatrix} A - sI_n & B \\ C & I_m \end{bmatrix}\begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y} \end{bmatrix} \tag{8}$$

where A is $n \times n$, B is $n \times m$, C is $m \times n$, and s is a variable. The vector $\mathbf{u}$ in $\mathbb{R}^m$ is the "input" to the system, $\mathbf{y}$ in $\mathbb{R}^m$ is the "output," and $\mathbf{x}$ in $\mathbb{R}^n$ is the "state" vector. (Actually, the vectors $\mathbf{x}$, $\mathbf{u}$, and $\mathbf{y}$ are functions of s, but this does not affect the algebraic calculations in Exercises 19 and 20.)

19. Assume $A - sI_n$ is invertible and view (8) as a system of two matrix equations. Solve the top equation for $\mathbf{x}$ and substitute into the bottom equation. The result is an equation of the form $W(s)\mathbf{u} = \mathbf{y}$, where $W(s)$ is a matrix that depends on s. $W(s)$ is called the *transfer function* of the system because it transforms the input $\mathbf{u}$ into the output $\mathbf{y}$. Find $W(s)$ and describe how it is related to the partitioned *system matrix* on the left side of (8). See Exercise 16.

20. Suppose the transfer function $W(s)$ in Exercise 19 is invertible for some s. It can be shown that the inverse transfer function $W(s)^{-1}$, which transforms outputs into inputs, is the Schur complement of $A - BC - sI_n$ for the matrix below. Find this Schur complement. See Exercise 16.

$$\begin{bmatrix} A - BC - sI_n & B \\ -C & I_m \end{bmatrix}$$

21. a. Verify that $A^2 = I$ when $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$.

b. Use partitioned matrices to show that $M^2 = I$ when

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

22. Generalize the idea of Exercise 21 by constructing a 6×6 matrix $M = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ C & 0 & D \end{bmatrix}$ such that $M^2 = I$. Make C a nonzero 2×2 matrix. Show that your construction works.

23. Use partitioned matrices to prove by induction that the product of two lower triangular matrices is also lower triangular. [*Hint*: A $(k+1) \times (k+1)$ matrix A_1 can be written in the form below, where a is a scalar, $\mathbf{v}$ is in $\mathbb{R}^k$, and A is a $k \times k$ lower triangular matrix. See the *Study Guide* for help with induction.]

$$A_1 = \begin{bmatrix} a & \mathbf{0}^T \\ \mathbf{v} & A \end{bmatrix}$$

SG The Principle of Induction 2–19

24. Use partitioned matrices to prove by induction that for $n = 2, 3, \ldots$, the $n \times n$ matrix A shown below is invertible and B is its inverse.

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

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

For the induction step, assume A and B are $(k+1) \times (k+1)$ matrices, and partition A and B in a form similar to that displayed in Exercise 23.

25. Without using row reduction, find the inverse of

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

26. [**M**] For block operations, it may be necessary to access or enter submatrices of a large matrix. Describe the functions or commands of a matrix program that accomplish the following tasks. Suppose A is a 20×30 matrix.

a. Display the submatrix of A from rows 5 to 10 and columns 15 to 20.

b. Insert a 5×10 matrix B into A, beginning at row 5 and column 10.

c. Create a 50×50 matrix of the form $C = \begin{bmatrix} A & 0 \\ 0 & A^T \end{bmatrix}$.

[*Note:* It may not be necessary to specify the zero blocks in C.]

27. [**M**] Suppose memory or size restrictions prevent a matrix program from working with matrices having more than 32 rows and 32 columns, and suppose some project involves 50×50 matrices A and B. Describe the commands or operations of the matrix program that accomplish the following tasks.

a. Compute $A + B$.

b. Compute AB.

c. Solve $A\mathbf{x} = \mathbf{b}$ for some vector $\mathbf{b}$ in $\mathbb{R}^{50}$, assuming that A can be partitioned into a 2×2 block matrix $[A_{ij}]$, with A_{11} an invertible 20×20 matrix, A_{22} an invertible 30×30 matrix, and A_{12} a zero matrix. [*Hint:* Describe appropriate smaller systems to solve, without using any matrix inverses.]

SOLUTIONS TO PRACTICE PROBLEMS

1. If $\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}$ is invertible, its inverse has the form $\begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$. Verify that

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix}\begin{bmatrix} W & X \\ Y & Z \end{bmatrix} = \begin{bmatrix} W & X \\ AW + Y & AX + Z \end{bmatrix}$$

So W, X, Y, Z must satisfy $W = I$, $X = 0$, $AW + Y = 0$, and $AX + Z = I$. It follows that $Y = -A$ and $Z = I$. Hence

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

The product in the reverse order is also the identity, so the block matrix is invertible, and its inverse is $\begin{bmatrix} I & 0 \\ -A & I \end{bmatrix}$. (You could also appeal to the Invertible Matrix Theorem.)

2. $X^TX = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix}\begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^TX_1 & X_1^TX_2 \\ X_2^TX_1 & X_2^TX_2 \end{bmatrix}$. The partitions of X^T and X are automatically conformable for block multiplication because the columns of X^T are the rows of X. This partition of X^TX is used in several computer algorithms for matrix computations.

2.5 MATRIX FACTORIZATIONS

A *factorization* of a matrix A is an equation that expresses A as a product of two or more matrices. Whereas matrix multiplication involves a *synthesis* of data (combining the effects of two or more linear transformations into a single matrix), matrix factorization is an *analysis* of data. In the language of computer science, the expression of A as a product amounts to a *preprocessing* of the data in A, organizing that data into two or more parts whose structures are more useful in some way, perhaps more accessible for computation.

b. To factor the matrix $\begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$ into the product of transfer matrices, as in equation (6), look for R_1 and R_2 in Fig. 4 to satisfy

$$\begin{bmatrix} 1 & -R_1 \\ -1/R_2 & 1 + R_1/R_2 \end{bmatrix} = \begin{bmatrix} 1 & -8 \\ -.5 & 5 \end{bmatrix}$$

From the $(1, 2)$-entries, $R_1 = 8$ ohms, and from the $(2, 1)$-entries, $1/R_2 = .5$ ohm and $R_2 = 1/.5 = 2$ ohms. With these values, the network in Fig. 4 has the desired transfer matrix. ■

A network transfer matrix summarizes the input–output behavior (the design specifications) of the network without reference to the interior circuits. To physically build a network with specified properties, an engineer first determines if such a network can be constructed (or *realized*). Then the engineer tries to factor the transfer matrix into matrices corresponding to smaller circuits that perhaps are already manufactured and ready for assembly. In the common case of alternating current, the entries in the transfer matrix are usually rational complex-valued functions. (See Exercises 19 and 20 in Section 2.4 and Example 2 in Section 3.3.) A standard problem is to find a *minimal realization* that uses the smallest number of electrical components.

PRACTICE PROBLEM

Find an LU factorization of $A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix}$. [*Note*: It will turn out that A has only three pivot columns, so the method of Example 2 will produce only the first three columns of L. The remaining two columns of L come from I_5.]

2.5 EXERCISES

In Exercises 1–6, solve the equation $A\mathbf{x} = \mathbf{b}$ by using the LU factorization given for A. In Exercises 1 and 2, also solve $A\mathbf{x} = \mathbf{b}$ by ordinary row reduction.

1. $A = \begin{bmatrix} 3 & -7 & -2 \\ -3 & 5 & 1 \\ 6 & -4 & 0 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -7 \\ 5 \\ 2 \end{bmatrix}$

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

2. $A = \begin{bmatrix} 2 & -6 & 4 \\ -4 & 8 & 0 \\ 0 & -4 & 6 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix}$

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

3. $A = \begin{bmatrix} 2 & -4 & 2 \\ -4 & 5 & 2 \\ 6 & -9 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}$

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

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

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

5. $A = \begin{bmatrix} 1 & -2 & -2 & -3 \\ 3 & -9 & 0 & -9 \\ -1 & 2 & 4 & 7 \\ -3 & -6 & 26 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 0 \\ 3 \end{bmatrix}$

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

6. $A = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -2 & -3 & -4 & 12 \\ 3 & 0 & 4 & -36 \\ -5 & -3 & -8 & 49 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$

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

Find an LU factorization of the matrices in Exercises 7–16 (with L unit lower triangular). Note that MATLAB will usually produce a permuted LU factorization because it uses partial pivoting for numerical accuracy.

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

8. $\begin{bmatrix} 6 & 4 \\ 12 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 3 & 1 & 2 \\ -9 & 0 & -4 \\ 9 & 9 & 14 \end{bmatrix}$

10. $\begin{bmatrix} -5 & 0 & 4 \\ 10 & 2 & -5 \\ 10 & 10 & 16 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 7 & 2 \\ 6 & 19 & 4 \\ -3 & -2 & 3 \end{bmatrix}$

12. $\begin{bmatrix} 2 & 3 & 2 \\ 4 & 13 & 9 \\ -6 & 5 & 4 \end{bmatrix}$

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

14. $\begin{bmatrix} 1 & 3 & 1 & 5 \\ 5 & 20 & 6 & 31 \\ -2 & -1 & -1 & -4 \\ -1 & 7 & 1 & 7 \end{bmatrix}$

15. $\begin{bmatrix} 2 & 0 & 5 & 2 \\ -6 & 3 & -13 & -3 \\ 4 & 9 & 16 & 17 \end{bmatrix}$

16. $\begin{bmatrix} 2 & -3 & 4 \\ -4 & 8 & -7 \\ 6 & -5 & 14 \\ -6 & 9 & -12 \\ 8 & -6 & 19 \end{bmatrix}$

17. When A is invertible, MATLAB finds A^{-1} by factoring $A = LU$ (where L may be permuted lower triangular), inverting L and U, and then computing $U^{-1}L^{-1}$. Use this method to compute the inverse of A in Exercise 2. (Apply the algorithm in Section 2.2 to L and to U.)

18. Find A^{-1} as in Exercise 17, using A from Exercise 3.

19. Let A be a lower triangular $n \times n$ matrix with nonzero entries on the diagonal. Show that A is invertible and A^{-1} is lower triangular. [*Hint:* Explain why A can be changed into I using only row replacements and scaling. (Where are the pivots?) Also, explain why the row operations that reduce A to I change I into a lower triangular matrix.]

20. Let $A = LU$ be an LU factorization. Explain why A can be row reduced to U using only replacement operations. (This fact is the converse of what was proved in the text.)

21. Suppose $A = BC$, where B is invertible. Show that any sequence of row operations that reduces B to I also reduces A to C. The converse is not true, since the zero matrix may be factored as $0 = B \cdot 0$.

Exercises 22–26 provide a glimpse of some widely used matrix factorizations, some of which are discussed later in the text.

22. (*Reduced LU Factorization*) With A as in the Practice Problem, find a 5×3 matrix B and a 3×4 matrix C such that $A = BC$. Generalize this idea to the case where A is $m \times n$, $A = LU$, and U has only three nonzero rows.

23. (*Rank Factorization*) Suppose an $m \times n$ matrix A admits a factorization $A = CD$ where C is $m \times 4$ and D is $4 \times n$.
    a. Show that A is the sum of four outer products. (See Section 2.4.)
    b. Let $m = 400$ and $n = 100$. Explain why a computer programmer might prefer to store the data from A in the form of two matrices C and D.

24. (*QR Factorization*) Suppose $A = QR$, where Q and R are $n \times n$, R is invertible and upper triangular, and Q has the property that $Q^TQ = I$. Show that for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. What computations with Q and R will produce the solution?

WEB

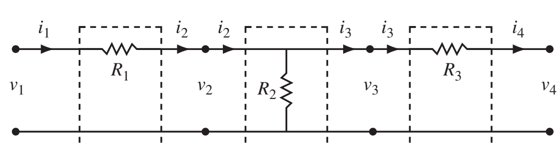
25. (*Singular Value Decomposition*) Suppose $A = UDV^T$, where U and V are $n \times n$ matrices with the property that $U^TU = I$ and $V^TV = I$, and where D is a diagonal matrix with positive numbers $\sigma_1, \dots, \sigma_n$ on the diagonal. Show that A is invertible, and find a formula for A^{-1}.

26. (*Spectral Factorization*) Suppose a 3×3 matrix A admits a factorization as $A = PDP^{-1}$, where P is some invertible 3×3 matrix and D is the diagonal matrix
$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
Show that this factorization is useful when computing high powers of A. Find fairly simple formulas for A^2, A^3, and A^k (k a positive integer), using P and the entries in D.

27. Design two different ladder networks that each output 9 volts and 4 amps when the input is 12 volts and 6 amps.

28. Show that if three shunt circuits (with resistances R_1, R_2, R_3) are connected in series, the resulting network has the same transfer matrix as a single shunt circuit. Find a formula for the resistance in that circuit.

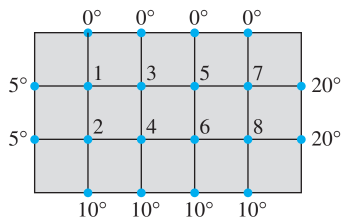
29. a. Compute the transfer matrix of the network in the figure below.



b. Let $A = \begin{bmatrix} 3 & -12 \\ -1/3 & 5/3 \end{bmatrix}$. Design a ladder network whose transfer matrix is A by finding a suitable matrix factorization of A.

30. Find a different factorization of the transfer matrix A in Exercise 29, and thereby design a different ladder network whose transfer matrix is A.

31. [**M**] Consider the heat plate in the following figure (refer to Exercise 33 in Section 1.1).



The solution to the steady-state heat flow problem for this plate is approximated by the solution to the equation $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b} = (5, 15, 0, 10, 0, 10, 20, 30)$ and

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

WEB

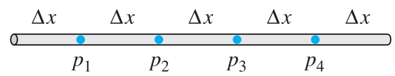
The missing entries in A are zeros. The nonzero entries of A lie within a band along the main diagonal. Such *band matrices* occur in a variety of applications and often are extremely large (with thousands of rows and columns but relatively narrow bands).

a. Use the method in Example 2 to construct an LU factorization of A, and note that both factors are band matrices (with two nonzero diagonals below or above the main diagonal). Compute $LU - A$ to check your work.

b. Use the LU factorization to solve $A\mathbf{x} = \mathbf{b}$.

c. Obtain A^{-1} and note that A^{-1} is a dense matrix with no band structure. When A is large, L and U can be stored in much less space than A^{-1}. This fact is another reason for preferring the LU factorization of A to A^{-1} itself.

32. [**M**] The band matrix A shown below can be used to estimate the unsteady conduction of heat in a rod when the temperatures at points $p_1, \dots, p_4$ on the rod change with time.[2]



The constant C in the matrix depends on the physical nature of the rod, the distance Δx between the points on the rod, and the length of time Δt between successive temperature measurements. Suppose that for $k = 0, 1, 2, \dots$, a vector $\mathbf{t}_k$ in $\mathbb{R}^4$ lists the temperatures at time $k\Delta t$. If the two ends of the rod are maintained at 0°, then the temperature vectors satisfy the equation $A\mathbf{t}_{k+1} = \mathbf{t}_k$ $(k = 0, 1, \dots)$, where

$$A = \begin{bmatrix} (1+2C) & -C & & \\ -C & (1+2C) & -C & \\ & -C & (1+2C) & -C \\ & & -C & (1+2C) \end{bmatrix}$$

a. Find the LU factorization of A when $C = 1$. A matrix such as A with three nonzero diagonals is called a *tridiagonal matrix*. The L and U factors are *bidiagonal matrices*.

b. Suppose $C = 1$ and $\mathbf{t}_0 = (10, 15, 15, 10)^T$. Use the LU factorization of A to find the temperature distributions $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$, and $\mathbf{t}_4$.

[2] See Biswa N. Datta, *Numerical Linear Algebra and Applications* (Pacific Grove, CA: Brooks/Cole, 1994), pp. 200–201.

SOLUTION TO PRACTICE PROBLEM

$$A = \begin{bmatrix} 2 & -4 & -2 & 3 \\ 6 & -9 & -5 & 8 \\ 2 & -7 & -3 & 9 \\ 4 & -2 & -2 & -1 \\ -6 & 3 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & -4 & -2 & 3 \\ 0 & 3 & 1 & -1 \\ 0 & -3 & -1 & 6 \\ 0 & 6 & 2 & -7 \\ 0 & -9 & -3 & 13 \end{bmatrix}$$

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

Divide the entries in each highlighted column by the pivot at the top. The resulting columns form the first three columns in the lower half of L. This suffices to make row reduction of L to I correspond to reduction of A to U. Use the last two columns of I_5

to make L unit lower triangular.

$$\begin{bmatrix} 2 \\ 6 \\ 2 \\ 4 \\ -6 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \\ 6 \\ -9 \end{bmatrix} \begin{bmatrix} 5 \\ -5 \\ 10 \end{bmatrix}$$
$$\div 2 \quad \div 3 \quad \div 5$$
$$\downarrow \quad \downarrow \quad \downarrow$$
$$\begin{bmatrix} 1 & & & \\ 3 & 1 & & \\ 1 & -1 & 1 & \cdots \\ 2 & 2 & -1 & \\ -3 & -3 & 2 & \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 \\ 2 & 2 & -1 & 1 & 0 \\ -3 & -3 & 2 & 0 & 1 \end{bmatrix}$$

2.6 THE LEONTIEF INPUT–OUTPUT MODEL

WEB

Linear algebra played an essential role in the Nobel prize–winning work of Wassily Leontief, as mentioned at the beginning of Chapter 1. The economic model described in this section is the basis for more elaborate models used in many parts of the world.

Suppose a nation's economy is divided into n sectors that produce goods or services, and let $\mathbf{x}$ be a **production vector** in $\mathbb{R}^n$ that lists the output of each sector for one year. Also, suppose another part of the economy (called the *open sector*) does not produce goods or services but only consumes them, and let $\mathbf{d}$ be a **final demand vector** (or **bill of final demands)** that lists the values of the goods and services demanded from the various sectors by the nonproductive part of the economy. The vector $\mathbf{d}$ can represent consumer demand, government consumption, surplus production, exports, or other external demands.

As the various sectors produce goods to meet consumer demand, the producers themselves create additional **intermediate demand** for goods they need as inputs for their own production. The interrelations between the sectors are very complex, and the connection between the final demand and the production is unclear. Leontief asked if there is a production level $\mathbf{x}$ such that the amounts produced (or "supplied") will exactly balance the total demand for that production, so that

$$\begin{Bmatrix} \text{amount} \\ \text{produced} \\ \mathbf{x} \end{Bmatrix} = \begin{Bmatrix} \text{intermediate} \\ \text{demand} \end{Bmatrix} + \begin{Bmatrix} \text{final} \\ \text{demand} \\ \mathbf{d} \end{Bmatrix} \tag{1}$$

The basic assumption of Leontief's input–output model is that for each sector, there is a **unit consumption vector** in $\mathbb{R}^n$ that lists the inputs needed *per unit of output* of the sector. All input and output units are measured in millions of dollars, rather than in quantities such as tons or bushels. (Prices of goods and services are held constant.)

As a simple example, suppose the economy consists of three sectors—manufacturing, agriculture, and services—with unit consumption vectors $\mathbf{c}_1$, $\mathbf{c}_2$, and $\mathbf{c}_3$, as shown in the table that follows.

NUMERICAL NOTE

Continuous movement of graphical 3D objects requires intensive computation with 4×4 matrices, particularly when the surfaces are *rendered* to appear realistic, with texture and appropriate lighting. *High-end computer graphics boards* have 4×4 matrix operations and graphics algorithms embedded in their microchips and circuitry. Such boards can perform the billions of matrix multiplications per second needed for realistic color animation in 3D gaming programs.[2]

Further Reading

James D. Foley, Andries van Dam, Steven K. Feiner, and John F. Hughes, *Computer Graphics: Principles and Practice*, 3rd ed. (Boston, MA: Addison-Wesley, 2002), Chapters 5 and 6.

PRACTICE PROBLEM

Rotation of a figure about a point $\mathbf{p}$ in $\mathbb{R}^2$ is accomplished by first translating the figure by $-\mathbf{p}$, rotating about the origin, and then translating back by $\mathbf{p}$. See Fig. 7. Construct the 3×3 matrix that rotates points $-30°$ about the point $(-2, 6)$, using homogeneous coordinates.

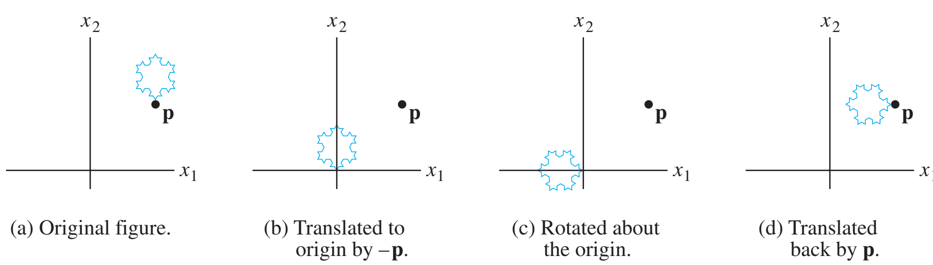


FIGURE 7 Rotation of figure about point $\mathbf{p}$.

2.7 EXERCISES

1. What 3×3 matrix will have the same effect on homogeneous coordinates for $\mathbb{R}^2$ that the shear matrix A has in Example 2?

2. Use matrix multiplication to find the image of the triangle with data matrix $D = \begin{bmatrix} 4 & 2 & 5 \\ 0 & 2 & 3 \end{bmatrix}$ under the transformation that reflects points through the y-axis. Sketch both the original triangle and its image.

In Exercises 3–8, find the 3×3 matrices that produce the described composite 2D transformations, using homogeneous coordinates.

3. Translate by $(2, 1)$, and then rotate $90°$ about the origin.

4. Translate by $(-1, 4)$, and then scale the x-coordinate by $1/2$ and the y-coordinate by $3/2$.

5. Reflect points through the x-axis, and then rotate $45°$ about the origin.

6. Rotate points $45°$ about the origin, then reflect through the x-axis.

7. Rotate points through $60°$ about the point $(6, 8)$.

8. Rotate points through $45°$ about the point $(3, 7)$.

9. A 2×100 data matrix D contains the coordinates of 100 points. Compute the number of multiplications required to transform these points using two arbitrary 2×2 matrices A and B. Consider the two possibilities $A(BD)$ and $(AB)D$. Discuss the implications of your results for computer graphics calculations.

[2] See Jan Ozer, "High-Performance Graphics Boards," *PC Magazine* **19**, 1 September 2000, pp. 187–200. Also, "The Ultimate Upgrade Guide: Moving On Up," *PC Magazine* **21**, 29 January 2002, pp. 82–91.

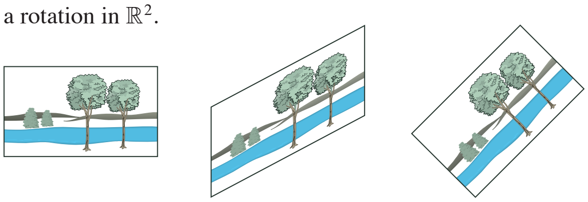
10. Consider the following geometric 2D transformations: D, a dilation (in which x-coordinates and y-coordinates are scaled by the same factor); R, a rotation; and T, a translation. Does D *commute* with R? That is, is $D(R(\mathbf{x})) = R(D(\mathbf{x}))$ for all $\mathbf{x}$ in $\mathbb{R}^2$? Does D commute with T? Does R commute with T?

11. A rotation on a computer screen is sometimes implemented as the product of two shear-and-scale transformations, which can speed up calculations that determine how a graphic image actually appears in terms of screen pixels. (The screen consists of rows and columns of small dots, called *pixels*.) The first transformation A_1 shears vertically and then compresses each column of pixels; the second transformation A_2 shears horizontally and then stretches each row of pixels. Let

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} \sec\varphi & -\tan\varphi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and show that the composition of the two transformations is a rotation in $\mathbb{R}^2$.



12. A rotation in $\mathbb{R}^2$ usually requires four multiplications. Compute the product below, and show that the matrix for a rotation can be factored into three shear transformations (each of which requires only one multiplication).

$$\begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \sin\varphi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\tan\varphi/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

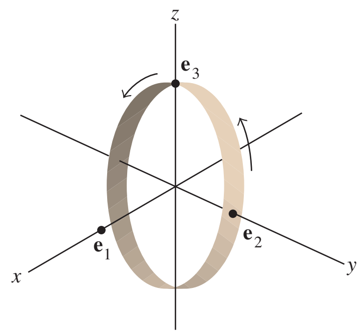
13. The usual transformations on homogeneous coordinates for 2D computer graphics involve 3×3 matrices of the form $\begin{bmatrix} A & \mathbf{p} \\ \mathbf{0}^T & 1 \end{bmatrix}$ where A is a 2×2 matrix and $\mathbf{p}$ is in $\mathbb{R}^2$. Show that such a transformation amounts to a linear transformation on $\mathbb{R}^2$ followed by a translation. [*Hint:* Find an appropriate matrix factorization involving partitioned matrices.]

14. Show that the transformation in Exercise 7 is equivalent to a rotation about the origin followed by a translation by $\mathbf{p}$. Find $\mathbf{p}$.

15. What vector in $\mathbb{R}^3$ has homogeneous coordinates $(\frac{1}{2}, -\frac{1}{4}, -\frac{1}{8}, \frac{1}{24})$?

16. Are $(1, -2, -3, 4)$ and $(10, -20, -30, 40)$ homogeneous coordinates for the same point in $\mathbb{R}^3$? Why or why not?

17. Give the 4×4 matrix that rotates points in $\mathbb{R}^3$ about the x-axis through an angle of $60°$. (See the figure.)



18. Give the 4×4 matrix that rotates points in $\mathbb{R}^3$ about the z-axis through an angle of $-30°$, and then translates by $\mathbf{p} = (5, -2, 1)$.

19. Let S be the triangle with vertices $(4.2, 1.2, 4)$, $(6, 4, 2)$, and $(2, 2, 6)$. Find the image of S under the perspective projection with center of projection at $(0, 0, 10)$.

20. Let S be the triangle with vertices $(7, 3, -5)$, $(12, 8, 2)$, and $(1, 2, 1)$. Find the image of S under the perspective projection with center of projection at $(0, 0, 10)$.

Exercises 21 and 22 concern the way in which color is specified for display in computer graphics. A color on a computer screen is encoded by three numbers (R, G, B) that list the amounts of energy an electron gun must transmit to red, green, and blue phosphor dots on the computer screen. (A fourth number specifies the luminance or intensity of the color.)

21. [**M**] The actual color a viewer sees on a screen is influenced by the specific type and amount of phosphors on the screen. So each computer screen manufacturer must convert between the (R, G, B) data and an international CIE standard for color, which uses three primary colors, called X, Y, and Z. A typical conversion for short-persistence phosphors is

$$\begin{bmatrix} .61 & .29 & .150 \\ .35 & .59 & .063 \\ .04 & .12 & .787 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

A computer program will send a stream of color information to the screen, using standard CIE data (X, Y, Z). Find the equation that converts these data to the (R, G, B) data needed for the screen's electron gun.

22. [**M**] The signal broadcast by commercial television describes each color by a vector (Y, I, Q). If the screen is black and white, only the Y-coordinate is used. (This gives a better monochrome picture than using CIE data for colors.) The correspondence between *YIQ* and a "standard" *RGB* color is given by

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} .299 & .587 & .114 \\ .596 & -.275 & -.321 \\ .212 & -.528 & .311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

(A screen manufacturer would change the matrix entries to work for its *RGB* screens.) Find the equation that converts the *YIQ* data transmitted by the television station to the *RGB* data needed for the television screen.

SOLUTION TO PRACTICE PROBLEM

Assemble the matrices right-to-left for the three operations. Using $\mathbf{p} = (-2, 6)$, $\cos(-30°) = \sqrt{3}/2$, and $\sin(-30°) = -.5$, we have

$$\underbrace{\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Translate back by } p} \underbrace{\begin{bmatrix} \sqrt{3}/2 & 1/2 & 0 \\ -1/2 & \sqrt{3}/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Rotate around the origin}} \underbrace{\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -6 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Translate by } -p}$$

$$= \begin{bmatrix} \sqrt{3}/2 & 1/2 & \sqrt{3}-5 \\ -1/2 & \sqrt{3}/2 & -3\sqrt{3}+5 \\ 0 & 0 & 1 \end{bmatrix}$$

2.8 SUBSPACES OF $\mathbb{R}^n$

This section focuses on important sets of vectors in $\mathbb{R}^n$ called *subspaces*. Often subspaces arise in connection with some matrix A, and they provide useful information about the equation $A\mathbf{x} = \mathbf{b}$. The concepts and terminology in this section will be used repeatedly throughout the rest of the book.[1]

DEFINITION

A **subspace** of $\mathbb{R}^n$ is any set H in $\mathbb{R}^n$ that has three properties:

a. The zero vector is in H.
b. For each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u} + \mathbf{v}$ is in H.
c. For each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

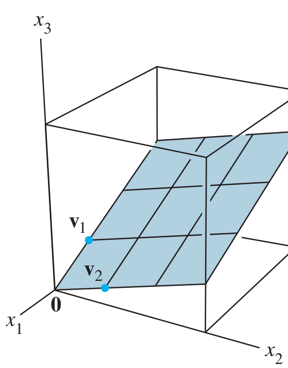


FIGURE 1
Span $\{\mathbf{v}_1, \mathbf{v}_2\}$ as a plane through the origin.

In words, a subspace is *closed* under addition and scalar multiplication. As you will see in the next few examples, most sets of vectors discussed in Chapter 1 are subspaces. For instance, a plane through the origin is the standard way to visualize the subspace in Example 1. See Fig. 1.

EXAMPLE 1 If $\mathbf{v}_1$ and $\mathbf{v}_2$ are in $\mathbb{R}^n$ and $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, then H is a subspace of $\mathbb{R}^n$. To verify this statement, note that the zero vector is in H (because $0\mathbf{v}_1 + 0\mathbf{v}_2$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$). Now take two arbitrary vectors in H, say,

$$\mathbf{u} = s_1\mathbf{v}_1 + s_2\mathbf{v}_2 \quad \text{and} \quad \mathbf{v} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

Then

$$\mathbf{u} + \mathbf{v} = (s_1 + t_1)\mathbf{v}_1 + (s_2 + t_2)\mathbf{v}_2$$

which shows that $\mathbf{u} + \mathbf{v}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ and hence is in H. Also, for any scalar c, the vector $c\mathbf{u}$ is in H, because $c\mathbf{u} = c(s_1\mathbf{v}_1 + s_2\mathbf{v}_2) = (cs_1)\mathbf{v}_1 + (cs_2)\mathbf{v}_2$. ■

[1] Sections 2.8 and 2.9 are included here to permit readers to postpone the study of most or all of the next two chapters and to skip directly to Chapter 5, if so desired. *Omit* these two sections if you plan to work through Chapter 4 before beginning Chapter 5.

The matrix B in Example 7 is in reduced echelon form. To handle a general matrix A, recall that linear dependence relations among the columns of A can be expressed in the form $A\mathbf{x} = \mathbf{0}$ for some $\mathbf{x}$. (If some columns are not involved in a particular dependence relation, then the corresponding entries in $\mathbf{x}$ are zero.) When A is row reduced to echelon form B, the columns are drastically changed, but the equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same set of solutions. That is, the columns of A have *exactly the same linear dependence relationships* as the columns of B.

EXAMPLE 8 It can be verified that the matrix

$$A = [\,\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_5\,] = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$$

is row equivalent to the matrix B in Example 7. Find a basis for Col A.

SOLUTION From Example 7, the pivot columns of A are columns 1, 2, and 5. Also, $\mathbf{b}_3 = -3\mathbf{b}_1 + 2\mathbf{b}_2$ and $\mathbf{b}_4 = 5\mathbf{b}_1 - \mathbf{b}_2$. Since row operations do not affect linear dependence relations among the columns of the matrix, we should have

$$\mathbf{a}_3 = -3\mathbf{a}_1 + 2\mathbf{a}_2 \quad \text{and} \quad \mathbf{a}_4 = 5\mathbf{a}_1 - \mathbf{a}_2$$

Check that this is true! By the argument in Example 7, $\mathbf{a}_3$ and $\mathbf{a}_4$ are not needed to generate the column space of A. Also, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ must be linearly independent, because any dependence relation among $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_5$ would imply the same dependence relation among $\mathbf{b}_1$, $\mathbf{b}_2$, and $\mathbf{b}_5$. Since $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5\}$ is linearly independent, $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_5\}$ is also linearly independent and hence is a basis for Col A. ■

The argument in Example 8 can be adapted to prove the following theorem.

THEOREM 13

The pivot columns of a matrix A form a basis for the column space of A.

Warning: Be careful to use *pivot columns of A itself* for the basis of Col A. The columns of an echelon form B are often not in the column space of A. (For instance, in Examples 7 and 8, the columns of B all have zeros in their last entries and cannot generate the columns of A.)

PRACTICE PROBLEMS

1. Let $A = \begin{bmatrix} 1 & -1 & 5 \\ 2 & 0 & 7 \\ -3 & -5 & -3 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} -7 \\ 3 \\ 2 \end{bmatrix}$. Is $\mathbf{u}$ in Nul A? Is $\mathbf{u}$ in Col A? Justify each answer.

2. Given $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, find a vector in Nul A and a vector in Col A.

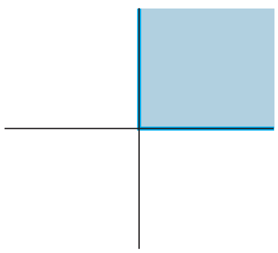
SG Mastering: Subspace, Col A, Nul A, Basis 2–37

3. Suppose an $n \times n$ matrix A is invertible. What can you say about Col A? About Nul A?

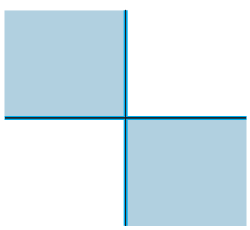
2.8 EXERCISES

Exercises 1–4 display sets in $\mathbb{R}^2$. Assume the sets include the bounding lines. In each case, give a specific reason why the set H is *not* a subspace of $\mathbb{R}^2$. (For instance, find two vectors in H whose sum is *not* in H, or find a vector in H with a scalar multiple that is not in H. Draw a picture.)

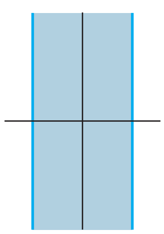
1.



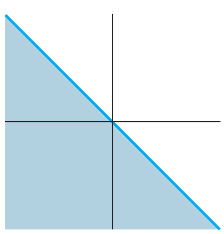
2.



3.



4.



5. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ -3 \\ 7 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} -3 \\ -3 \\ 10 \end{bmatrix}$. Determine if $\mathbf{w}$ is in the subspace of $\mathbb{R}^3$ generated by $\mathbf{v}_1$ and $\mathbf{v}_2$.

6. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ -4 \\ 5 \\ 7 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 5 \\ -3 \\ 6 \\ 5 \end{bmatrix}$, and $\mathbf{u} = \begin{bmatrix} -1 \\ -7 \\ -1 \\ 2 \end{bmatrix}$. Determine if $\mathbf{u}$ is in the subspace of $\mathbb{R}^4$ generated by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

7. Let

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 8 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -4 \\ 6 \\ -7 \end{bmatrix},$$

$$\mathbf{p} = \begin{bmatrix} 6 \\ -10 \\ 11 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}.$$

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

b. How many vectors are in Col A?

c. Is $\mathbf{p}$ in Col A? Why or why not?

8. Let

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 0 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -5 \\ 5 \end{bmatrix},$$

and $\mathbf{p} = \begin{bmatrix} -6 \\ 1 \\ 17 \end{bmatrix}$. Determine if $\mathbf{p}$ is in Col A, where $A = [\, \mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \mathbf{v}_3 \,]$.

9. With A and $\mathbf{p}$ as in Exercise 7, determine if $\mathbf{p}$ is in Nul A.

10. With $\mathbf{u} = \begin{bmatrix} -5 \\ 5 \\ 3 \end{bmatrix}$ and A as in Exercise 8, determine if $\mathbf{u}$ is in Nul A.

In Exercises 11 and 12, give integers p and q such that Nul A is a subspace of $\mathbb{R}^p$ and Col A is a subspace of $\mathbb{R}^q$.

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

12. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 7 \\ -5 & -1 & 0 \\ 2 & 7 & 11 \\ 3 & 3 & 4 \end{bmatrix}$

13. For A as in Exercise 11, find a nonzero vector in Nul A and a nonzero vector in Col A.

14. For A as in Exercise 12, find a nonzero vector in Nul A and a nonzero vector in Col A.

Determine which sets in Exercises 15–20 are bases for $\mathbb{R}^2$ or $\mathbb{R}^3$. Justify each answer.

15. $\begin{bmatrix} 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 16 \\ -3 \end{bmatrix}$

16. $\begin{bmatrix} -2 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ -10 \end{bmatrix}$

17. $\begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 3 \\ 2 \end{bmatrix}$

18. $\begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \end{bmatrix}$

19. $\begin{bmatrix} 3 \\ -8 \\ 1 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \\ -5 \end{bmatrix}$

20. $\begin{bmatrix} 1 \\ -6 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ -6 \\ 7 \end{bmatrix}, \begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 9 \end{bmatrix}$

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

21. a. A subspace of $\mathbb{R}^n$ is any set H such that (i) the zero vector is in H, (ii) $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u}+\mathbf{v}$ are in H, and (iii) c is a scalar and $c\mathbf{u}$ is in H.

b. If $\mathbf{v}_1, \dots, \mathbf{v}_p$ are in $\mathbb{R}^n$, then Span $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is the same as the column space of the matrix $\begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_p \end{bmatrix}$.

c. The set of all solutions of a system of m homogeneous equations in n unknowns is a subspace of $\mathbb{R}^m$.

d. The columns of an invertible $n \times n$ matrix form a basis for $\mathbb{R}^n$.

e. Row operations do not affect linear dependence relations among the columns of a matrix.

22. a. A subset H of $\mathbb{R}^n$ is a subspace if the zero vector is in H.

b. If B is an echelon form of a matrix A, then the pivot columns of B form a basis for Col A.

c. Given vectors $\mathbf{v}_1, \dots, \mathbf{v}_p$ in $\mathbb{R}^n$, the set of all linear combinations of these vectors is a subspace of $\mathbb{R}^n$.

d. Let H be a subspace of $\mathbb{R}^n$. If $\mathbf{x}$ is in H, and $\mathbf{y}$ is in $\mathbb{R}^n$, then $\mathbf{x}+\mathbf{y}$ is in H.

e. The column space of a matrix A is the set of solutions of $A\mathbf{x}=\mathbf{b}$.

Exercises 23–26 display a matrix A and an echelon form of A. Find a basis for Col A and a basis for Nul A.

23. $A = \begin{bmatrix} 4 & 5 & 9 & -2 \\ 6 & 5 & 1 & 12 \\ 3 & 4 & 8 & -3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 6 & -5 \\ 0 & 1 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

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

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

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

26. $A = \begin{bmatrix} 3 & -1 & -3 & -1 & 8 \\ 3 & 1 & 3 & 0 & 2 \\ 0 & 3 & 9 & -1 & -4 \\ 6 & 3 & 9 & -2 & 6 \end{bmatrix}$

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

27. Construct a 3×3 matrix A and a nonzero vector $\mathbf{b}$ such that $\mathbf{b}$ is in Col A, but $\mathbf{b}$ is not the same as any one of the columns of A.

28. Construct a 3×3 matrix A and a vector $\mathbf{b}$ such that $\mathbf{b}$ is *not* in Col A.

29. Construct a nonzero 3×3 matrix A and a nonzero vector $\mathbf{b}$ such that $\mathbf{b}$ is in Nul A.

30. Suppose the columns of a matrix $A = [\mathbf{a}_1 \cdots \mathbf{a}_p]$ are linearly independent. Explain why $\{\mathbf{a}_1, \dots, \mathbf{a}_p\}$ is a basis for Col A.

In Exercises 31–36, respond as comprehensively as possible, and justify your answer.

31. Suppose F is a 5×5 matrix whose column space is not equal to $\mathbb{R}^5$. What can be said about Nul F?

32. If B is a 7×7 matrix and Col $B = \mathbb{R}^7$, what can be said about solutions of equations of the form $B\mathbf{x} = \mathbf{b}$ for $\mathbf{b}$ in $\mathbb{R}^7$?

33. If C is a 6×6 matrix and Nul C is the zero subspace, what can be said about solutions of equations of the form $C\mathbf{x} = \mathbf{b}$ for $\mathbf{b}$ in $\mathbb{R}^6$?

34. What can be said about the shape of an $m \times n$ matrix A when the columns of A form a basis for $\mathbb{R}^m$?

35. If B is a 5×5 matrix and Nul B is *not* the zero subspace, what can be said about Col B?

36. What can be said about Nul C when C is a 6×4 matrix with linearly independent columns?

[M] In Exercises 37 and 38, construct bases for the column space and the null space of the given matrix A. Justify your work.

37. $A = \begin{bmatrix} 3 & -5 & 0 & -1 & 3 \\ -7 & 9 & -4 & 9 & -11 \\ -5 & 7 & -2 & 5 & -7 \\ 3 & -7 & -3 & 4 & 0 \end{bmatrix}$

38. $A = \begin{bmatrix} 5 & 3 & 2 & -6 & -8 \\ 4 & 1 & 3 & -8 & -7 \\ 5 & 1 & 4 & 5 & 19 \\ -7 & -5 & -2 & 8 & 5 \end{bmatrix}$

WEB Column Space and Null Space

WEB A Basis for Col A

SOLUTIONS TO PRACTICE PROBLEMS

1. To determine whether **u** is in Nul A, simply compute

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

The result shows that **u** is in Nul A. Deciding whether **u** is in Col A requires more work. Reduce the augmented matrix $[A \quad \mathbf{u}]$ to echelon form to determine whether the equation $A\mathbf{x} = \mathbf{u}$ is consistent:

$$\begin{bmatrix} 1 & -1 & 5 & -7 \\ 2 & 0 & 7 & 3 \\ -3 & -5 & -3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & -8 & 12 & -19 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 5 & -7 \\ 0 & 2 & -3 & 17 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

The equation $A\mathbf{x} = \mathbf{u}$ has no solution, so **u** is not in Col A.

2. In contrast to Practice Problem 1, finding a vector in Nul A requires more work than testing whether a specified vector is in Nul A. However, since A is already in reduced echelon form, the equation $A\mathbf{x} = \mathbf{0}$ shows that if $\mathbf{x} = (x_1, x_2, x_3)$, then $x_2 = 0$, $x_3 = 0$, and x_1 is a free variable. Thus, a basis for Nul A is $\mathbf{v} = (1, 0, 0)$. Finding just one vector in Col A is trivial, since each column of A is in Col A. In this particular case, the same vector **v** is in both Nul A *and* Col A. For most $n \times n$ matrices, the zero vector of $\mathbb{R}^n$ is the only vector in both Nul A and Col A.

3. If A is invertible, then the columns of A span $\mathbb{R}^n$, by the Invertible Matrix Theorem. By definition, the columns of any matrix always span the column space, so in this case Col A is all of $\mathbb{R}^n$. In symbols, Col $A = \mathbb{R}^n$. Also, since A is invertible, the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution. This means that Nul A is the zero subspace. In symbols, Nul $A = \{\mathbf{0}\}$.

2.9 DIMENSION AND RANK

This section continues the discussion of subspaces and bases for subspaces, beginning with the concept of a coordinate system. The definition and example below should make a useful new term, *dimension,* seem quite natural, at least for subspaces of $\mathbb{R}^3$.

Coordinate Systems

The main reason for selecting a basis for a subspace H, instead of merely a spanning set, is that each vector in H can be written in only one way as a linear combination of the basis vectors. To see why, suppose $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for H, and suppose a vector **x** in H can be generated in two ways, say,

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_p\mathbf{b}_p \quad \text{and} \quad \mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_p\mathbf{b}_p \tag{1}$$

Then, subtracting gives

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_p - d_p)\mathbf{b}_p \tag{2}$$

Since $\mathcal{B}$ is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \leq j \leq p$, which shows that the two representations in (1) are actually the same.

SG Expanded Table for the IMT 2-39

Also, statement (q) implies that the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution, which is statement (d). Since statements (d) and (g) are already known to be equivalent to the statement that A is invertible, the proof is complete. ■

NUMERICAL NOTES

Many algorithms discussed in this text are useful for understanding concepts and making simple computations by hand. However, the algorithms are often unsuitable for large-scale problems in real life.

Rank determination is a good example. It would seem easy to reduce a matrix to echelon form and count the pivots. But unless exact arithmetic is performed on a matrix whose entries are specified exactly, row operations can change the apparent rank of a matrix. For instance, if the value of x in the matrix $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ is not stored exactly as 7 in a computer, then the rank may be 1 or 2, depending on whether the computer treats $x - 7$ as zero.

WEB

In practical applications, the effective rank of a matrix A is often determined from the singular value decomposition of A, to be discussed in Section 7.4.

PRACTICE PROBLEMS

1. Determine the dimension of the subspace H of $\mathbb{R}^3$ spanned by the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. (First, find a basis for H.)

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -8 \\ 6 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 6 \\ -7 \end{bmatrix}$$

2. Consider the basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ .2 \end{bmatrix}, \begin{bmatrix} .2 \\ 1 \end{bmatrix} \right\}$$

for $\mathbb{R}^2$. If $[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, what is $\mathbf{x}$?

3. Could $\mathbb{R}^3$ possibly contain a four-dimensional subspace? Explain.

2.9 EXERCISES

In Exercises 1 and 2, find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ and the given basis $\mathcal{B}$. Illustrate your answer with a figure, as in the solution of Practice Problem 2.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

In Exercises 3–6, the vector $\mathbf{x}$ is in a subspace H with a basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$.

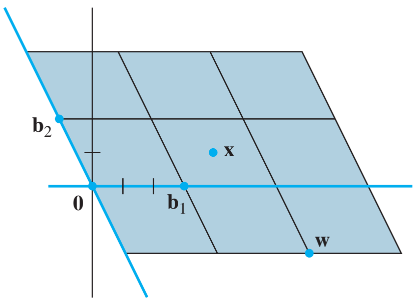
3. $\mathbf{b}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 0 \\ 7 \end{bmatrix}$

4. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1 \\ 9 \end{bmatrix}$

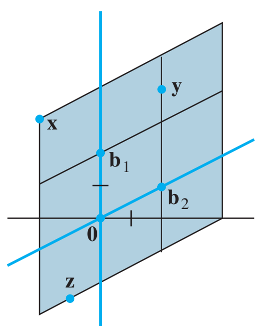
5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 4 \\ -3 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -2 \\ -7 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 2 \\ 9 \\ -7 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} -3 \\ 2 \\ -4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 7 \\ -3 \\ 5 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}$

7. Let $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use the figure to estimate $[\mathbf{w}]_\mathcal{B}$ and $[\mathbf{x}]_\mathcal{B}$. Confirm your estimate of $[\mathbf{x}]_\mathcal{B}$ by using it and $\{\mathbf{b}_1, \mathbf{b}_2\}$ to compute $\mathbf{x}$.



8. Let $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\mathbf{z} = \begin{bmatrix} -1 \\ -2.5 \end{bmatrix}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$. Use the figure to estimate $[\mathbf{x}]_\mathcal{B}$, $[\mathbf{y}]_\mathcal{B}$, and $[\mathbf{z}]_\mathcal{B}$. Confirm your estimates of $[\mathbf{y}]_\mathcal{B}$ and $[\mathbf{z}]_\mathcal{B}$ by using them and $\{\mathbf{b}_1, \mathbf{b}_2\}$ to compute $\mathbf{y}$ and $\mathbf{z}$.



Exercises 9–12 display a matrix A and an echelon form of A. Find bases for Col A and Nul A, and then state the dimensions of these subspaces.

9. $A = \begin{bmatrix} 1 & 3 & 2 & -6 \\ 3 & 9 & 1 & 5 \\ 2 & 6 & -1 & 9 \\ 5 & 15 & 0 & 14 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 5 & -7 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

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

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

11. $A = \begin{bmatrix} 2 & 4 & -5 & 2 & -3 \\ 3 & 6 & -8 & 3 & -5 \\ 0 & 0 & 9 & 0 & 9 \\ -3 & -6 & -7 & -3 & -10 \end{bmatrix}$

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

12. $A = \begin{bmatrix} 1 & 2 & -4 & 4 & 6 \\ 5 & 1 & -9 & 2 & 10 \\ 4 & 6 & -9 & 12 & 15 \\ 3 & 4 & -5 & 8 & 9 \end{bmatrix}$

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

In Exercises 13 and 14, find a basis for the subspace spanned by the given vectors. What is the dimension of the subspace?

13. $\begin{bmatrix} 1 \\ -3 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 9 \\ -6 \\ 12 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 5 \\ -3 \\ 7 \end{bmatrix}$

14. $\begin{bmatrix} 1 \\ -1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ -7 \\ 7 \end{bmatrix}, \begin{bmatrix} 3 \\ -7 \\ 6 \\ -9 \end{bmatrix}$

15. Suppose a 4×6 matrix A has four pivot columns. Is Col $A = \mathbb{R}^4$? Is Nul $A = \mathbb{R}^2$? Explain your answers.

16. Suppose a 4×7 matrix A has three pivot columns. Is Col $A = \mathbb{R}^3$? What is the dimension of Nul A? Explain your answers.

In Exercises 17 and 18, mark each statement True or False. Justify each answer. Here A is an $m \times n$ matrix.

17. a. If $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a basis for a subspace H and if $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, then $c_1, \dots, c_p$ are the coordinates of $\mathbf{x}$ relative to the basis $\mathcal{B}$.

b. Each line in $\mathbb{R}^n$ is a one-dimensional subspace of $\mathbb{R}^n$.

c. The dimension of Col A is the number of pivot columns in A.

d. The dimensions of Col A and Nul A add up to the number of columns in A.

e. If a set of p vectors spans a p-dimensional subspace H of $\mathbb{R}^n$, then these vectors form a basis for H.

18. a. If $\mathcal{B}$ is a basis for a subspace H, then each vector in H can be written in only one way as a linear combination of the vectors in $\mathcal{B}$.

b. The dimension of Nul A is the number of variables in the equation $A\mathbf{x} = \mathbf{0}$.

c. The dimension of the column space of A is rank A.

d. If $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a basis for a subspace H of $\mathbb{R}^n$, then the correspondence $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ makes H look and act the same as $\mathbb{R}^p$.

e. If H is a p-dimensional subspace of $\mathbb{R}^n$, then a linearly independent set of p vectors in H is a basis for H.

In Exercises 19–24, justify each answer or construction.

19. If the subspace of all solutions of $A\mathbf{x} = \mathbf{0}$ has a basis consisting of three vectors and if A is a 5×7 matrix, what is the rank of A?

20. What is the rank of a 6×8 matrix whose null space is three-dimensional?

21. If the rank of a 9×8 matrix A is 7, what is the dimension of the solution space of $A\mathbf{x} = \mathbf{0}$?

22. Show that a set $\{\mathbf{v}_1, \dots, \mathbf{v}_5\}$ in $\mathbb{R}^n$ is linearly dependent if $\dim \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_5\} = 4$.

23. If possible, construct a 3×5 matrix A such that $\dim \text{Nul}\, A = 3$ and $\dim \text{Col}\, A = 2$.

24. Construct a 3×4 matrix with rank 1.

25. Let A be an $n \times p$ matrix whose column space is p-dimensional. Explain why the columns of A must be linearly independent.

26. Suppose columns 1, 3, 4, 5, and 7 of a matrix A are linearly independent (but are not necessarily pivot columns) and the rank of A is 5. Explain why the five columns mentioned must be a basis for the column space of A.

27. Suppose vectors $\mathbf{b}_1, \dots, \mathbf{b}_p$ span a subspace W, and let $\{\mathbf{a}_1, \dots, \mathbf{a}_q\}$ be any set in W containing more than p vectors. Fill in the details of the following argument to show that $\{\mathbf{a}_1, \dots, \mathbf{a}_q\}$ must be linearly dependent. First, let $B = [\,\mathbf{b}_1 \;\cdots\; \mathbf{b}_p\,]$ and $A = [\,\mathbf{a}_1 \;\cdots\; \mathbf{a}_q\,]$.

a. Explain why for each vector $\mathbf{a}_j$, there exists a vector $\mathbf{c}_j$ in $\mathbb{R}^p$ such that $\mathbf{a}_j = B\mathbf{c}_j$.

b. Let $C = [\,\mathbf{c}_1 \;\cdots\; \mathbf{c}_q\,]$. Explain why there is a nonzero vector $\mathbf{u}$ such that $C\mathbf{u} = \mathbf{0}$.

c. Use B and C to show that $A\mathbf{u} = \mathbf{0}$. This shows that the columns of A are linearly dependent.

28. Use Exercise 27 to show that if $\mathcal{A}$ and $\mathcal{B}$ are bases for a subspace W of $\mathbb{R}^n$, then $\mathcal{A}$ cannot contain more vectors than $\mathcal{B}$, and, conversely, $\mathcal{B}$ cannot contain more vectors than $\mathcal{A}$.

29. **[M]** Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Show that $\mathbf{x}$ is in H, and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, when

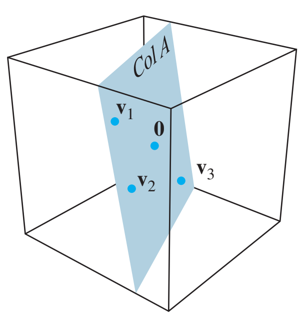
$$\mathbf{v}_1 = \begin{bmatrix} 15 \\ -5 \\ 12 \\ 7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 14 \\ -10 \\ 13 \\ 17 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 16 \\ 0 \\ 11 \\ -3 \end{bmatrix}$$

30. **[M]** Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Show that $\mathcal{B}$ is a basis for H and $\mathbf{x}$ is in H, and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, when

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 3 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ 0 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 4 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 11 \\ -2 \\ 17 \\ -8 \end{bmatrix}$$

SG Mastering: Dimension and Rank 2–41

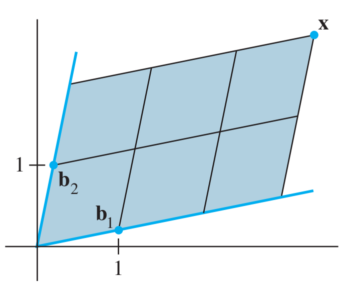
SOLUTIONS TO PRACTICE PROBLEMS



1. Construct $A = [\mathbf{v}_1 \;\; \mathbf{v}_2 \;\; \mathbf{v}_3]$ so that the subspace spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is the column space of A. A basis for this space is provided by the pivot columns of A.

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

The first two columns of A are pivot columns and form a basis for H. Thus $\dim H = 2$.



2. If $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, then $\mathbf{x}$ is formed from a linear combination of the basis vectors using weights 3 and 2:

$$\mathbf{x} = 3\mathbf{b}_1 + 2\mathbf{b}_2 = 3\begin{bmatrix} 1 \\ .2 \end{bmatrix} + 2\begin{bmatrix} .2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 2.6 \end{bmatrix}$$

The basis $\{\mathbf{b}_1, \mathbf{b}_2\}$ determines a *coordinate system* for $\mathbb{R}^2$, illustrated by the grid in the figure. Note how $\mathbf{x}$ is 3 units in the $\mathbf{b}_1$-direction and 2 units in the $\mathbf{b}_2$-direction.

3. A four-dimensional subspace would contain a basis of four linearly independent vectors. This is impossible inside $\mathbb{R}^3$. Since any linearly independent set in $\mathbb{R}^3$ has no more than three vectors, any subspace of $\mathbb{R}^3$ has dimension no more than 3. The space $\mathbb{R}^3$ itself is the only three-dimensional subspace of $\mathbb{R}^3$. Other subspaces of $\mathbb{R}^3$ have dimension 2, 1, or 0.

CHAPTER 2 SUPPLEMENTARY EXERCISES

1. Assume that the matrices mentioned in the statements below have appropriate sizes. Mark each statement True or False. Justify each answer.

a. If A and B are $m \times n$, then both AB^T and A^TB are defined.

b. If $AB = C$ and C has 2 columns, then A has 2 columns.

c. Left-multiplying a matrix B by a diagonal matrix A, with nonzero entries on the diagonal, scales the rows of B.

d. If $BC = BD$, then $C = D$.

e. If $AC = 0$, then either $A = 0$ or $C = 0$.

f. If A and B are $n \times n$, then $(A + B)(A - B) = A^2 - B^2$.

g. An elementary $n \times n$ matrix has either n or $n + 1$ nonzero entries.

h. The transpose of an elementary matrix is an elementary matrix.

i. An elementary matrix must be square.

j. Every square matrix is a product of elementary matrices.

k. If A is a 3×3 matrix with three pivot positions, there exist elementary matrices $E_1, \ldots, E_p$ such that $E_p \cdots E_1 A = I$.

l. If $AB = I$, then A is invertible.

m. If A and B are square and invertible, then AB is invertible, and $(AB)^{-1} = A^{-1}B^{-1}$.

n. If $AB = BA$ and if A is invertible, then $A^{-1}B = BA^{-1}$.

o. If A is invertible and if $r \neq 0$, then $(rA)^{-1} = rA^{-1}$.

p. If A is a 3×3 matrix and the equation $A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has a unique solution, then A is invertible.

2. Find the matrix C whose inverse is $C^{-1} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$.

3. Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Show that $A^3 = 0$. Use matrix algebra to compute the product $(I - A)(I + A + A^2)$.

4. Suppose $A^n = 0$ for some $n > 1$. Find an inverse for $I - A$.

5. Suppose an $n \times n$ matrix A satisfies the equation $A^2 - 2A + I = 0$. Show that $A^3 = 3A - 2I$ and $A^4 = 4A - 3I$.

6. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. These are *Pauli spin matrices* used in the study of electron spin in quantum mechanics. Show that $A^2 = I$, $B^2 = I$, and $AB = -BA$. Matrices such that $AB = -BA$ are said to *anticommute*.

7. Let $A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 4 & 11 \\ 1 & 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 5 \\ 1 & 5 \\ 3 & 4 \end{bmatrix}$. Compute $A^{-1}B$ without computing A^{-1}. [*Hint*: $A^{-1}B$ is the solution of the equation $AX = B$.]

8. Find a matrix A such that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ into $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectively. [*Hint*: Write a matrix equation involving A, and solve for A.]

9. Suppose $AB = \begin{bmatrix} 5 & 4 \\ -2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$. Find A.

10. Suppose A is invertible. Explain why A^TA is also invertible. Then show that $A^{-1} = (A^TA)^{-1}A^T$.

11. Let $x_1, \ldots, x_n$ be fixed numbers. The matrix below, called a *Vandermonde matrix,* occurs in applications such as signal processing, error-correcting codes, and polynomial interpolation.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Given $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, suppose $\mathbf{c} = (c_0, \ldots, c_{n-1})$ in $\mathbb{R}^n$ satisfies $V\mathbf{c} = \mathbf{y}$, and define the polynomial

$$p(t) = c_0 + c_1t + c_2t^2 + \cdots + c_{n-1}t^{n-1}$$

a. Show that $p(x_1) = y_1, \ldots, p(x_n) = y_n$. We call $p(t)$ an *interpolating polynomial for the points* $(x_1, y_1), \ldots, (x_n, y_n)$ because the graph of $p(t)$ passes through the points.

b. Suppose $x_1, \ldots, x_n$ are distinct numbers. Show that the columns of V are linearly independent. [*Hint*: How many zeros can a polynomial of degree $n - 1$ have?]

c. Prove: "If $x_1, \ldots, x_n$ are distinct numbers, and $y_1, \ldots, y_n$ are arbitrary numbers, then there is an interpolating polynomial of degree $\leq n - 1$ for $(x_1, y_1), \ldots, (x_n, y_n)$."

12. Let $A = LU$, where L is an invertible lower triangular matrix and U is upper triangular. Explain why the first column of A is a multiple of the first column of L. How is the second column of A related to the columns of L?

3. A four-dimensional subspace would contain a basis of four linearly independent vectors. This is impossible inside $\mathbb{R}^3$. Since any linearly independent set in $\mathbb{R}^3$ has no more than three vectors, any subspace of $\mathbb{R}^3$ has dimension no more than 3. The space $\mathbb{R}^3$ itself is the only three-dimensional subspace of $\mathbb{R}^3$. Other subspaces of $\mathbb{R}^3$ have dimension 2, 1, or 0.

CHAPTER 2 SUPPLEMENTARY EXERCISES

1. Assume that the matrices mentioned in the statements below have appropriate sizes. Mark each statement True or False. Justify each answer.
   a. If A and B are $m \times n$, then both AB^T and A^TB are defined.
   b. If $AB = C$ and C has 2 columns, then A has 2 columns.
   c. Left-multiplying a matrix B by a diagonal matrix A, with nonzero entries on the diagonal, scales the rows of B.
   d. If $BC = BD$, then $C = D$.
   e. If $AC = 0$, then either $A = 0$ or $C = 0$.
   f. If A and B are $n \times n$, then $(A + B)(A - B) = A^2 - B^2$.
   g. An elementary $n \times n$ matrix has either n or $n + 1$ nonzero entries.
   h. The transpose of an elementary matrix is an elementary matrix.
   i. An elementary matrix must be square.
   j. Every square matrix is a product of elementary matrices.
   k. If A is a 3×3 matrix with three pivot positions, there exist elementary matrices $E_1, \ldots, E_p$ such that $E_p \cdots E_1 A = I$.
   l. If $AB = I$, then A is invertible.
   m. If A and B are square and invertible, then AB is invertible, and $(AB)^{-1} = A^{-1}B^{-1}$.
   n. If $AB = BA$ and if A is invertible, then $A^{-1}B = BA^{-1}$.
   o. If A is invertible and if $r \neq 0$, then $(rA)^{-1} = rA^{-1}$.
   p. If A is a 3×3 matrix and the equation $A\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has a unique solution, then A is invertible.

2. Find the matrix C whose inverse is $C^{-1} = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix}$.

3. Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Show that $A^3 = 0$. Use matrix algebra to compute the product $(I - A)(I + A + A^2)$.

4. Suppose $A^n = 0$ for some $n > 1$. Find an inverse for $I - A$.

5. Suppose an $n \times n$ matrix A satisfies the equation $A^2 - 2A + I = 0$. Show that $A^3 = 3A - 2I$ and $A^4 = 4A - 3I$.

6. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. These are *Pauli spin matrices* used in the study of electron spin in quantum mechanics. Show that $A^2 = I$, $B^2 = I$, and $AB = -BA$. Matrices such that $AB = -BA$ are said to *anticommute*.

7. Let $A = \begin{bmatrix} 1 & 3 & 8 \\ 2 & 4 & 11 \\ 1 & 2 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} -3 & 5 \\ 1 & 5 \\ 3 & 4 \end{bmatrix}$. Compute $A^{-1}B$ without computing A^{-1}. [*Hint*: $A^{-1}B$ is the solution of the equation $AX = B$.]

8. Find a matrix A such that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$ into $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, respectively. [*Hint*: Write a matrix equation involving A, and solve for A.]

9. Suppose $AB = \begin{bmatrix} 5 & 4 \\ -2 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 7 & 3 \\ 2 & 1 \end{bmatrix}$. Find A.

10. Suppose A is invertible. Explain why A^TA is also invertible. Then show that $A^{-1} = (A^TA)^{-1}A^T$.

11. Let $x_1, \ldots, x_n$ be fixed numbers. The matrix below, called a *Vandermonde matrix,* occurs in applications such as signal processing, error-correcting codes, and polynomial interpolation.

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

Given $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$, suppose $\mathbf{c} = (c_0, \ldots, c_{n-1})$ in $\mathbb{R}^n$ satisfies $V\mathbf{c} = \mathbf{y}$, and define the polynomial

$$p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_{n-1}t^{n-1}$$

   a. Show that $p(x_1) = y_1, \ldots, p(x_n) = y_n$. We call $p(t)$ an *interpolating polynomial for the points* $(x_1, y_1), \ldots, (x_n, y_n)$ because the graph of $p(t)$ passes through the points.
   b. Suppose $x_1, \ldots, x_n$ are distinct numbers. Show that the columns of V are linearly independent. [*Hint*: How many zeros can a polynomial of degree $n - 1$ have?]
   c. Prove: "If $x_1, \ldots, x_n$ are distinct numbers, and $y_1, \ldots, y_n$ are arbitrary numbers, then there is an interpolating polynomial of degree $\leq n - 1$ for $(x_1, y_1), \ldots, (x_n, y_n)$."

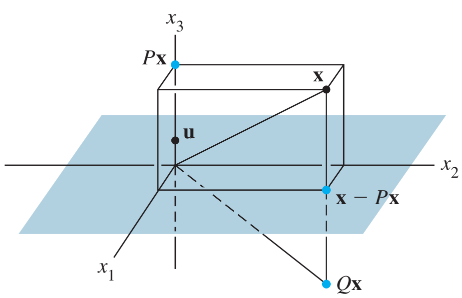
12. Let $A = LU$, where L is an invertible lower triangular matrix and U is upper triangular. Explain why the first column of A is a multiple of the first column of L. How is the second column of A related to the columns of L?

13. Given $\mathbf{u}$ in $\mathbb{R}^n$ with $\mathbf{u}^T\mathbf{u} = 1$, let $P = \mathbf{u}\mathbf{u}^T$ (an outer product) and $Q = I - 2P$. Justify statements (a), (b), and (c).

a. $P^2 = P$ b. $P^T = P$ c. $Q^2 = I$

The transformation $\mathbf{x} \mapsto P\mathbf{x}$ is called a *projection*, and $\mathbf{x} \mapsto Q\mathbf{x}$ is called a *Householder reflection*. Such reflections are used in computer programs to create multiple zeros in a vector (usually a column of a matrix).

14. Let $\mathbf{u} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$. Determine P and Q as in Exercise 13, and compute $P\mathbf{x}$ and $Q\mathbf{x}$. The figure shows that $Q\mathbf{x}$ is the reflection of $\mathbf{x}$ through the x_1x_2-plane.



A Householder reflection through the plane $x_3 = 0$.

15. Suppose $C = E_3E_2E_1B$, where E_1, E_2, and E_3 are elementary matrices. Explain why C is row equivalent to B.

16. Let A be an $n \times n$ singular matrix. Describe how to construct an $n \times n$ nonzero matrix B such that $AB = 0$.

17. Let A be a 6×4 matrix and B a 4×6 matrix. Show that the 6×6 matrix AB cannot be invertible.

18. Suppose A is a 5×3 matrix and there exists a 3×5 matrix C such that $CA = I_3$. Suppose further that for some given $\mathbf{b}$ in $\mathbb{R}^5$, the equation $A\mathbf{x} = \mathbf{b}$ has at least one solution. Show that this solution is unique.

19. [**M**] Certain dynamical systems can be studied by examining powers of a matrix, such as those below. Determine what happens to A^k and B^k as k increases (for example, try $k = 2, \dots, 16$). Try to identify what is special about A and B. Investigate large powers of other matrices of this type, and make a conjecture about such matrices.

$$A = \begin{bmatrix} .4 & .2 & .3 \\ .3 & .6 & .3 \\ .3 & .2 & .4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & .2 & .3 \\ .1 & .6 & .3 \\ .9 & .2 & .4 \end{bmatrix}$$

20. [**M**] Let A_n be the $n \times n$ matrix with 0's on the main diagonal and 1's elsewhere. Compute A_n^{-1} for $n = 4, 5,$ and 6, and make a conjecture about the general form of A_n^{-1} for larger values of n.

Henceforth we will omit the zero terms in the cofactor expansion. Next, expand this 4×4 determinant down the first column, in order to take advantage of the zeros there. We have

$$\det A = 3 \cdot 2 \cdot \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

This 3×3 determinant was computed in Example 1 and found to equal -2. Hence $\det A = 3 \cdot 2 \cdot (-2) = -12$. ■

The matrix in Example 3 was nearly triangular. The method in that example is easily adapted to prove the following theorem.

THEOREM 2

If A is a triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

The strategy in Example 3 of looking for zeros works extremely well when an entire row or column consists of zeros. In such a case, the cofactor expansion along such a row or column is a sum of zeros! So the determinant is zero. Unfortunately, most cofactor expansions are not so quickly evaluated.

NUMERICAL NOTE

By today's standards, a 25×25 matrix is small. Yet it would be impossible to calculate a 25×25 determinant by cofactor expansion. In general, a cofactor expansion requires over $n!$ multiplications, and $25!$ is approximately 1.5×10^{25}.

If a computer performs one trillion multiplications per second, it would have to run for over 500,000 years to compute a 25×25 determinant by this method. Fortunately, there are faster methods, as we'll soon discover.

Exercises 19–38 explore important properties of determinants, mostly for the 2×2 case. The results from Exercises 33–36 will be used in the next section to derive the analogous properties for $n \times n$ matrices.

PRACTICE PROBLEM

Compute $\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix}$.

3.1 EXERCISES

Compute the determinants in Exercises 1–8 using a cofactor expansion across the first row. In Exercises 1–4, also compute the determinant by a cofactor expansion down the second column.

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

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

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

4. $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

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

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

7. $\begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$ **8.** $\begin{vmatrix} 8 & 1 & 6 \\ 4 & 0 & 3 \\ 3 & -2 & 5 \end{vmatrix}$

Compute the determinants in Exercises 9–14 by cofactor expansions. At each step, choose a row or column that involves the least amount of computation.

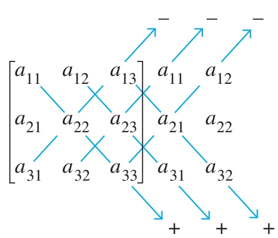
9. $\begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$ **10.** $\begin{vmatrix} 1 & -2 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 2 & -6 & -7 & 5 \\ 5 & 0 & 4 & 4 \end{vmatrix}$

11. $\begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$ **12.** $\begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$

13. $\begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$

14. $\begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$

The expansion of a 3×3 determinant can be remembered by the following device. Write a second copy of the first two columns to the right of the matrix, and compute the determinant by multiplying entries on six diagonals:



Add the downward diagonal products and subtract the upward products. Use this method to compute the determinants in Exercises 15–18. ***Warning:*** *This trick does not generalize in any reasonable way to* 4×4 *or larger matrices.*

15. $\begin{vmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{vmatrix}$ **16.** $\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix}$

17. $\begin{vmatrix} 2 & -4 & 3 \\ 3 & 1 & 2 \\ 1 & 4 & -1 \end{vmatrix}$ **18.** $\begin{vmatrix} 1 & 3 & 5 \\ 2 & 1 & 1 \\ 3 & 4 & 2 \end{vmatrix}$

In Exercises 19–24, explore the effect of an elementary row operation on the determinant of a matrix. In each case, state the row operation and describe how it affects the determinant.

19. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} c & d \\ a & b \end{bmatrix}$ **20.** $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a & b \\ kc & kd \end{bmatrix}$

21. $\begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 5+3k & 6+4k \end{bmatrix}$

22. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a+kc & b+kd \\ c & d \end{bmatrix}$

23. $\begin{bmatrix} 1 & 1 & 1 \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}, \begin{bmatrix} k & k & k \\ -3 & 8 & -4 \\ 2 & -3 & 2 \end{bmatrix}$

24. $\begin{bmatrix} a & b & c \\ 3 & 2 & 2 \\ 6 & 5 & 6 \end{bmatrix}, \begin{bmatrix} 3 & 2 & 2 \\ a & b & c \\ 6 & 5 & 6 \end{bmatrix}$

Compute the determinants of the elementary matrices given in Exercises 25–30. (See Section 2.2.)

25. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & k & 1 \end{bmatrix}$ **26.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{bmatrix}$

27. $\begin{bmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **28.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$

29. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **30.** $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

Use Exercises 25–28 to answer the questions in Exercises 31 and 32. Give reasons for your answers.

31. What is the determinant of an elementary row replacement matrix?

32. What is the determinant of an elementary scaling matrix with k on the diagonal?

In Exercises 33–36, verify that $\det EA = (\det E)(\det A)$, where E is the elementary matrix shown and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **34.** $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

35. $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ **36.** $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

37. Let $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$. Write $5A$. Is $\det 5A = 5 \det A$?

38. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and let k be a scalar. Find a formula that relates $\det kA$ to k and $\det A$.

In Exercises 39 and 40, A is an $n \times n$ matrix. Mark each statement True or False. Justify each answer.

39. a. An $n \times n$ determinant is defined by determinants of $(n-1) \times (n-1)$ submatrices.

b. The (i, j)-cofactor of a matrix A is the matrix A_{ij} obtained by deleting from A its ith row and jth column.

40. a. The cofactor expansion of $\det A$ down a column is the negative of the cofactor expansion along a row.

b. The determinant of a triangular matrix is the sum of the entries on the main diagonal.

41. Let $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute the area of the parallelogram determined by $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, and $\mathbf{0}$, and compute the determinant of $[\,\mathbf{u} \quad \mathbf{v}\,]$. How do they compare? Replace the first entry of $\mathbf{v}$ by an arbitrary number x, and repeat the problem. Draw a picture and explain what you find.

42. Let $u = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} c \\ 0 \end{bmatrix}$, where a, b, c are positive (for simplicity). Compute the area of the parallelogram determined by $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, and $\mathbf{0}$, and compute the determinants of the matrices $[\,\mathbf{u} \quad \mathbf{v}\,]$ and $[\,\mathbf{v} \quad \mathbf{u}\,]$. Draw a picture and explain what you find.

43. [**M**] Is it true that $\det(A + B) = \det A + \det B$? To find out, generate random 5×5 matrices A and B, and compute $\det(A + B) - \det A - \det B$. (Refer to Exercise 37 in Section 2.1.) Repeat the calculations for three other pairs of $n \times n$ matrices, for various values of n. Report your results.

44. [**M**] Is it true that $\det AB = (\det A)(\det B)$? Experiment with four pairs of random matrices as in Exercise 43, and make a conjecture.

45. [**M**] Construct a random 4×4 matrix A with integer entries between -9 and 9, and compare $\det A$ with $\det A^T$, $\det(-A)$, $\det(2A)$, and $\det(10A)$. Repeat with two other random 4×4 integer matrices, and make conjectures about how these determinants are related. (Refer to Exercise 36 in Section 2.1.) Then check your conjectures with several random 5×5 and 6×6 integer matrices. Modify your conjectures, if necessary, and report your results.

46. [**M**] How is $\det A^{-1}$ related to $\det A$? Experiment with random $n \times n$ integer matrices for $n = 4$, 5, and 6, and make a conjecture. *Note:* In the unlikely event that you encounter a matrix with a zero determinant, reduce it to echelon form and discuss what you find.

SOLUTION TO PRACTICE PROBLEM

Take advantage of the zeros. Begin with a cofactor expansion down the third column to obtain a 3×3 matrix, which may be evaluated by an expansion down its first column.

$$\begin{vmatrix} 5 & -7 & 2 & 2 \\ 0 & 3 & 0 & -4 \\ -5 & -8 & 0 & 3 \\ 0 & 5 & 0 & -6 \end{vmatrix} = (-1)^{1+3}2 \begin{vmatrix} 0 & 3 & -4 \\ -5 & -8 & 3 \\ 0 & 5 & -6 \end{vmatrix}$$

$$= 2 \cdot (-1)^{2+1}(-5) \begin{vmatrix} 3 & -4 \\ 5 & -6 \end{vmatrix} = 20$$

The $(-1)^{2+1}$ in the next-to-last calculation came from the $(2, 1)$-position of the -5 in the 3×3 determinant.

3.2 PROPERTIES OF DETERMINANTS

The secret of determinants lies in how they change when row operations are performed. The following theorem generalizes the results of Exercises 19–24 in Section 3.1. The proof is at the end of this section.

THEOREM 3

Row Operations

Let A be a square matrix.

a. If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.

b. If two rows of A are interchanged to produce B, then $\det B = -\det A$.

c. If one row of A is multiplied by k to produce B, then $\det B = k \cdot \det A$.

The following examples show how to use Theorem 3 to find determinants efficiently.

If A is an $n \times n$ matrix and E is an $n \times n$ elementary matrix, then

$$\det \mathrm{EA} = (\det E)(\det A)$$

where

$$\det E = \begin{cases} 1 & \text{if } E \text{ is a row replacement} \\ -1 & \text{if } E \text{ is an interchange} \\ r & \text{if } E \text{ is a scale by } r \end{cases}$$

PROOF OF THEOREM 3 The proof is by induction on the size of A. The case of a 2×2 matrix was verified in Exercises 33–36 of Section 3.1. Suppose the theorem has been verified for determinants of $k \times k$ matrices with $k \geq 2$, let $n = k + 1$, and let A be $n \times n$. The action of E on A involves either two rows or only one row. So we can expand $\det EA$ across a row that is unchanged by the action of E, say, row i. Let A_{ij} (respectively, B_{ij}) be the matrix obtained by deleting row i and column j from A (respectively, EA). Then the rows of B_{ij} are obtained from the rows of A_{ij} by the same type of elementary row operation that E performs on A. Since these submatrices are only $k \times k$, the induction assumption implies that

$$\det B_{ij} = \alpha \cdot \det A_{ij}$$

where $\alpha = 1, -1$, or r, depending on the nature of E. The cofactor expansion across row i is

$$\begin{aligned} \det EA &= a_{i1}(-1)^{i+1} \det B_{i1} + \cdots + a_{in}(-1)^{i+n} \det B_{in} \\ &= \alpha a_{i1}(-1)^{i+1} \det A_{i1} + \cdots + \alpha a_{in}(-1)^{i+n} \det A_{in} \\ &= \alpha \cdot \det A \end{aligned}$$

In particular, taking $A = I_n$, we see that $\det E = 1, -1$, or r, depending on the nature of E. Thus the theorem is true for $n = 2$, and the truth of the theorem for one value of n implies its truth for the next value of n. By the principle of induction, the theorem must be true for $n \geq 2$. The theorem is trivially true for $n = 1$. ■

PROOF OF THEOREM 6 If A is not invertible, then neither is AB, by Exercise 27 in Section 2.3. In this case, $\det AB = (\det A)(\det B)$, because both sides are zero, by Theorem 4. If A is invertible, then A and the identity matrix I_n are row equivalent by the Invertible Matrix Theorem. So there exist elementary matrices $E_1, \ldots, E_p$ such that

$$A = E_p E_{p-1} \cdots E_1 \cdot I_n = E_p E_{p-1} \cdots E_1$$

For brevity, write $|A|$ for $\det A$. Then repeated application of Theorem 3, as rephrased above, shows that

$$\begin{aligned} |AB| &= |E_p \cdots E_1 B| = |E_p||E_{p-1} \cdots E_1 B| = \cdots \\ &= |E_p| \cdots |E_1||B| = \cdots = |E_p \cdots E_1||B| \\ &= |A||B| \end{aligned}$$ ■

PRACTICE PROBLEMS

1. Compute $\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix}$ in as few steps as possible.

2. Use a determinant to decide if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent, when

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ -7 \\ 9 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -3 \\ 3 \\ -5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -7 \\ 5 \end{bmatrix}$$

3.2 EXERCISES

Each equation in Exercises 1–4 illustrates a property of determinants. State the property.

1. $\begin{vmatrix} 0 & 5 & -2 \\ 1 & -3 & 6 \\ 4 & -1 & 8 \end{vmatrix} = -\begin{vmatrix} 1 & -3 & 6 \\ 0 & 5 & -2 \\ 4 & -1 & 8 \end{vmatrix}$

2. $\begin{vmatrix} 2 & -6 & 4 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix} = 2\begin{vmatrix} 1 & -3 & 2 \\ 3 & 5 & -2 \\ 1 & 6 & 3 \end{vmatrix}$

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

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

Find the determinants in Exercises 5–10 by row reduction to echelon form.

5. $\begin{vmatrix} 1 & 5 & -6 \\ -1 & -4 & 4 \\ -2 & -7 & 9 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 5 & -3 \\ 3 & -3 & 3 \\ 2 & 13 & -7 \end{vmatrix}$

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

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

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

10. $\begin{vmatrix} 1 & 3 & -1 & 0 & -2 \\ 0 & 2 & -4 & -1 & -6 \\ -2 & -6 & 2 & 3 & 9 \\ 3 & 7 & -3 & 8 & -7 \\ 3 & 5 & 5 & 2 & 7 \end{vmatrix}$

Combine the methods of row reduction and cofactor expansion to compute the determinants in Exercises 11–14.

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

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

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

14. $\begin{vmatrix} -3 & -2 & 1 & -4 \\ 1 & 3 & 0 & -3 \\ -3 & 4 & -2 & 8 \\ 3 & -4 & 0 & 4 \end{vmatrix}$

Find the determinants in Exercises 15–20, where

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 7.$$

15. $\begin{vmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{vmatrix}$

16. $\begin{vmatrix} a & b & c \\ 3d & 3e & 3f \\ g & h & i \end{vmatrix}$

17. $\begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$

18. $\begin{vmatrix} g & h & i \\ a & b & c \\ d & e & f \end{vmatrix}$

19. $\begin{vmatrix} a & b & c \\ 2d+a & 2e+b & 2f+c \\ g & h & i \end{vmatrix}$

20. $\begin{vmatrix} a+d & b+e & c+f \\ d & e & f \\ g & h & i \end{vmatrix}$

In Exercises 21–23, use determinants to find out if the matrix is invertible.

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

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

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

In Exercises 24–26, use determinants to decide if the set of vectors is linearly independent.

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

25. $\begin{bmatrix} 7 \\ -4 \\ -6 \end{bmatrix}, \begin{bmatrix} -8 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \\ -5 \end{bmatrix}$

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

In Exercises 27 and 28, A and B are $n \times n$ matrices. Mark each statement True or False. Justify each answer.

27. a. A row replacement operation does not affect the determinant of a matrix.

b. The determinant of A is the product of the pivots in any echelon form U of A, multiplied by $(-1)^r$, where r is the number of row interchanges made during row reduction from A to U.

c. If the columns of A are linearly dependent, then $\det A = 0$.

d. $\det(A + B) = \det A + \det B$.

28. a. If two row interchanges are made in succession, then the new determinant equals the old determinant.

b. The determinant of A is the product of the diagonal entries in A.

c. If $\det A$ is zero, then two rows or two columns are the same, or a row or a column is zero.

d. $\det A^T = (-1)\det A$.

29. Compute $\det B^5$, where $B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$.

30. Use Theorem 3 (but not Theorem 4) to show that if two rows of a square matrix A are equal, then $\det A = 0$. The same is true for two columns. Why?

In Exercises 31–36, mention an appropriate theorem in your explanation.

31. Show that if A is invertible, then $\det A^{-1} = \dfrac{1}{\det A}$.

32. Find a formula for $\det(rA)$ when A is an $n \times n$ matrix.

33. Let A and B be square matrices. Show that even though AB and BA may not be equal, it is always true that $\det AB = \det BA$.

34. Let A and P be square matrices, with P invertible. Show that $\det(PAP^{-1}) = \det A$.

35. Let U be a square matrix such that $U^TU = I$. Show that $\det U = \pm 1$.

36. Suppose that A is a square matrix such that $\det A^4 = 0$. Explain why A cannot be invertible.

Verify that $\det AB = (\det A)(\det B)$ for the matrices in Exercises 37 and 38. (Do not use Theorem 6.)

37. $A = \begin{bmatrix} 3 & 0 \\ 6 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 5 & 4 \end{bmatrix}$

38. $A = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 2 \\ -1 & -1 \end{bmatrix}$

39. Let A and B be 3×3 matrices, with $\det A = 4$ and $\det B = -3$. Use properties of determinants (in the text and in the exercises above) to compute:

a. $\det AB$ b. $\det 5A$ c. $\det B^T$

d. $\det A^{-1}$ e. $\det A^3$

40. Let A and B be 4×4 matrices, with $\det A = -1$ and $\det B = 2$. Compute:

a. $\det AB$ b. $\det B^5$ c. $\det 2A$

d. $\det A^TA$ e. $\det B^{-1}AB$

41. Verify that $\det A = \det B + \det C$, where

$$A = \begin{bmatrix} a+e & b+f \\ c & d \end{bmatrix},\ B = \begin{bmatrix} a & b \\ c & d \end{bmatrix},\ C = \begin{bmatrix} e & f \\ c & d \end{bmatrix}$$

42. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Show that $\det(A + B) = \det A + \det B$ if and only if $a + d = 0$.

43. Verify that $\det A = \det B + \det C$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & u_1 + v_1 \\ a_{21} & a_{22} & u_2 + v_2 \\ a_{31} & a_{32} & u_3 + v_3 \end{bmatrix},$$

$$B = \begin{bmatrix} a_{11} & a_{12} & u_1 \\ a_{21} & a_{22} & u_2 \\ a_{31} & a_{32} & u_3 \end{bmatrix},\ C = \begin{bmatrix} a_{11} & a_{12} & v_1 \\ a_{21} & a_{22} & v_2 \\ a_{31} & a_{32} & v_3 \end{bmatrix}$$

Note, however, that A is *not* the same as $B + C$.

44. Right-multiplication by an elementary matrix E affects the *columns* of A in the same way that left-multiplication affects the *rows*. Use Theorems 5 and 3 and the obvious fact that E^T is another elementary matrix to show that

$\det AE = (\det E)(\det A)$

Do not use Theorem 6.

45. **[M]** Compute $\det A^TA$ and $\det AA^T$ for several random 4×5 matrices and several random 5×6 matrices. What can you say about A^TA and AA^T when A has more columns than rows?

46. **[M]** If $\det A$ is close to zero, is the matrix A nearly singular? Experiment with the nearly singular 4×4 matrix A in Exercise 9 of Section 2.3. Compute the determinants of A, $10A$, and $0.1A$. In contrast, compute the condition numbers of these matrices. Repeat these calculations when A is the 4×4 identity matrix. Discuss your results.

SOLUTIONS TO PRACTICE PROBLEMS

1. Perform row replacements to create zeros in the first column and then create a row of zeros.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$

2. $\det[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,] = \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix}$ Row 1 added to row 2

$= -(-3)\begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5)\begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix}$ Cofactors of column 2

$= 3 \cdot (35) + 5 \cdot (-21) = 0$

By Theorem 4, the matrix $[\,\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3\,]$ is not invertible. The columns are linearly dependent, by the Invertible Matrix Theorem.

3.3 CRAMER'S RULE, VOLUME, AND LINEAR TRANSFORMATIONS

This section applies the theory of the preceding sections to obtain important theoretical formulas and a geometric interpretation of the determinant.

Cramer's Rule

Cramer's rule is needed in a variety of theoretical calculations. For instance, it can be used to study how the solution of $A\mathbf{x} = \mathbf{b}$ is affected by changes in the entries of $\mathbf{b}$. However, the formula is inefficient for hand calculations, except for 2×2 or perhaps 3×3 matrices.

For any $n \times n$ matrix A and any $\mathbf{b}$ in $\mathbb{R}^n$, let $A_i(\mathbf{b})$ be the matrix obtained from A by replacing column i by the vector $\mathbf{b}$.

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \quad \cdots \quad \underset{\substack{\uparrow \\ \text{col } i}}{\mathbf{b}} \quad \cdots \quad \mathbf{a}_n]$$

THEOREM 7

Cramer's Rule

Let A be an invertible $n \times n$ matrix. For any $\mathbf{b}$ in $\mathbb{R}^n$, the unique solution $\mathbf{x}$ of $A\mathbf{x} = \mathbf{b}$ has entries given by

$$x_i = \frac{\det A_i(\mathbf{b})}{\det A}, \qquad i = 1, 2, \ldots, n \tag{1}$$

PROOF Denote the columns of A by $\mathbf{a}_1, \ldots, \mathbf{a}_n$ and the columns of the $n \times n$ identity matrix I by $\mathbf{e}_1, \ldots, \mathbf{e}_n$. If $A\mathbf{x} = \mathbf{b}$, the definition of matrix multiplication shows that

$$\begin{aligned} A \cdot I_i(\mathbf{x}) &= A\,[\,\mathbf{e}_1 \quad \cdots \quad \mathbf{x} \quad \cdots \quad \mathbf{e}_n\,] = [\,A\mathbf{e}_1 \quad \cdots \quad A\mathbf{x} \quad \cdots \quad A\mathbf{e}_n\,] \\ &= [\,\mathbf{a}_1 \quad \cdots \quad \mathbf{b} \quad \cdots \quad \mathbf{a}_n\,] = A_i(\mathbf{b}) \end{aligned}$$

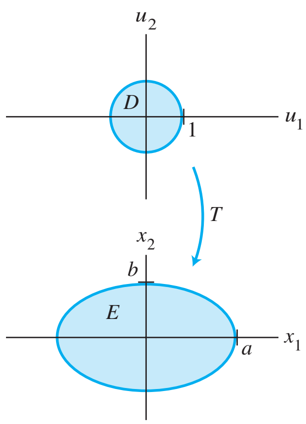
By the multiplicative property of determinants,

$$(\det A)(\det I_i(\mathbf{x})) = \det A_i(\mathbf{b})$$

The second determinant on the left is simply x_i. (Make a cofactor expansion along the ith row.) Hence $(\det A) \cdot x_i = \det A_i(\mathbf{b})$. This proves (1) because A is invertible and $\det A \neq 0$. ■

EXAMPLE 1 Use Cramer's rule to solve the system

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



SOLUTION We claim that E is the image of the unit disk D under the linear transformation T determined by the matrix $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, because if $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, and $\mathbf{x} = A\mathbf{u}$, then

$$u_1 = \frac{x_1}{a} \quad \text{and} \quad u_2 = \frac{x_2}{b}$$

It follows that $\mathbf{u}$ is in the unit disk, with $u_1^2 + u_2^2 \leq 1$, if and only if $\mathbf{x}$ is in E, with $(x_1/a)^2 + (x_2/b)^2 \leq 1$. By the generalization of Theorem 10,

$$\begin{aligned} \{\text{area of ellipse}\} &= \{\text{area of } T(D)\} \\ &= |\det A| \cdot \{\text{area of } D\} \\ &= ab \cdot \pi(1)^2 = \pi ab \end{aligned}$$

■

PRACTICE PROBLEM

Let S be the parallelogram determined by the vectors $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, and let $A = \begin{bmatrix} 1 & -.1 \\ 0 & 2 \end{bmatrix}$. Compute the area of the image of S under the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

3.3 EXERCISES

Use Cramer's rule to compute the solutions of the systems in Exercises 1–6.

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

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

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

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

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

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

In Exercises 7–10, determine the values of the parameter s for which the system has a unique solution, and describe the solution.

7. $\begin{aligned} 6sx_1 + 4x_2 &= 5 \\ 9x_1 + 2sx_2 &= -2 \end{aligned}$

8. $\begin{aligned} 3sx_1 - 5x_2 &= 3 \\ 9x_1 + 5sx_2 &= 2 \end{aligned}$

9. $\begin{aligned} sx_1 - 2sx_2 &= -1 \\ 3x_1 + 6sx_2 &= 4 \end{aligned}$

10. $\begin{aligned} 2sx_1 + x_2 &= 1 \\ 3sx_1 + 6sx_2 &= 2 \end{aligned}$

In Exercises 11–16, compute the adjugate of the given matrix, and then use Theorem 8 to give the inverse of the matrix.

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

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

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

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

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

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

17. Show that if A is 2×2, then Theorem 8 gives the same formula for A^{-1} as that given by Theorem 4 in Section 2.2.

18. Suppose that all the entries in A are integers and $\det A = 1$. Explain why all the entries in A^{-1} are integers.

In Exercises 19–22, find the area of the parallelogram whose vertices are listed.

19. $(0,0)$, $(5,2)$, $(6,4)$, $(11,6)$

20. $(0,0)$, $(-1,3)$, $(4,-5)$, $(3,-2)$

21. $(-1,0)$, $(0,5)$, $(1,-4)$, $(2,1)$

22. $(0,-2)$, $(6,-1)$, $(-3,1)$, $(3,2)$

23. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at $(1,0,-2)$, $(1,2,4)$, and $(7,1,0)$.

24. Find the volume of the parallelepiped with one vertex at the origin and adjacent vertices at $(1,4,0)$, $(-2,-5,2)$, and $(-1,2,-1)$.

25. Use the concept of volume to explain why the determinant of a 3×3 matrix A is zero if and only if A is not invertible. Do not appeal to Theorem 4 in Section 3.2. [*Hint:* Think about the columns of A.]

26. Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation, and let $\mathbf{p}$ be a vector and S a set in $\mathbb{R}^m$. Show that the image of $\mathbf{p} + S$ under T is the translated set $T(\mathbf{p}) + T(S)$ in $\mathbb{R}^n$.

27. Let S be the parallelogram determined by the vectors $\mathbf{b}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, and let $A = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Compute the area of the image of S under the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

28. Repeat Exercise 27 with $\mathbf{b}_1 = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $A = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$.

29. Find a formula for the area of the triangle whose vertices are $\mathbf{0}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ in $\mathbb{R}^2$.

30. Let R be the triangle with vertices at (x_1, y_1), (x_2, y_2), and (x_3, y_3). Show that

$$\{\text{area of triangle}\} = \frac{1}{2}\det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$$

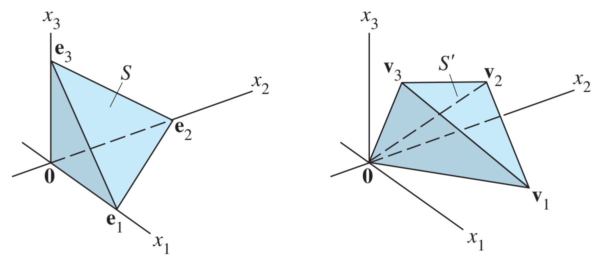
[*Hint:* Translate R to the origin by subtracting one of the vertices, and use Exercise 29.]

31. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation determined by the matrix $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, where a, b, and c are positive numbers. Let S be the unit ball, whose bounding surface has the equation $x_1^2 + x_2^2 + x_3^2 = 1$.

a. Show that $T(S)$ is bounded by the ellipsoid with the equation $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} + \dfrac{x_3^2}{c^2} = 1$.

b. Use the fact that the volume of the unit ball is $4\pi/3$ to determine the volume of the region bounded by the ellipsoid in part (a).

32. Let S be the tetrahedron in $\mathbb{R}^3$ with vertices at the vectors $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, and let S' be the tetrahedron with vertices at vectors $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. See the figure.



a. Describe a linear transformation that maps S onto S'.

b. Find a formula for the volume of the tetrahedron S' using the fact that

$$\{\text{volume of } S\} = (1/3)\{\text{area of base}\} \cdot \{\text{height}\}$$

33. **[M]** Test the inverse formula of Theorem 8 for a random 4×4 matrix A. Use your matrix program to compute the cofactors of the 3×3 submatrices, construct the adjugate, and set $B = (\text{adj } A)/(\det A)$. Then compute $B - \text{inv}(A)$, where $\text{inv}(A)$ is the inverse of A as computed by the matrix program. Use floating point arithmetic with the maximum possible number of decimal places. Report your results.

34. **[M]** Test Cramer's rule for a random 4×4 matrix A and a random 4×1 vector $\mathbf{b}$. Compute each entry in the solution of $A\mathbf{x} = \mathbf{b}$, and compare these entries with the entries in $A^{-1}\mathbf{b}$. Write the command (or keystrokes) for your matrix program that uses Cramer's rule to produce the second entry of $\mathbf{x}$.

35. **[M]** If your version of MATLAB has the `flops` command, use it to count the number of floating point operations to compute A^{-1} for a random 30×30 matrix. Compare this number with the number of flops needed to form $(\text{adj } A)/(\det A)$.

SOLUTION TO PRACTICE PROBLEM

The area of S is $\left|\det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right| = 14$, and $\det A = 2$. By Theorem 10, the area of the image of S under the mapping $\mathbf{x} \mapsto A\mathbf{x}$ is

$$|\det A| \cdot \{\text{area of } S\} = 2 \cdot 14 = 28$$

CHAPTER 3 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. Assume that all matrices here are square.

a. If A is a 2×2 matrix with a zero determinant, then one column of A is a multiple of the other.

b. If two rows of a 3×3 matrix A are the same, then $\det A = 0$.

c. If A is a 3×3 matrix, then $\det 5A = 5 \det A$.

d. If A and B are $n \times n$ matrices, with $\det A = 2$ and $\det B = 3$, then $\det(A + B) = 5$.

e. If A is $n \times n$ and $\det A = 2$, then $\det A^3 = 6$.

f. If B is produced by interchanging two rows of A, then $\det B = \det A$.

g. If B is produced by multiplying row 3 of A by 5, then $\det B = 5 \cdot \det A$.

27. Let S be the parallelogram determined by the vectors $\mathbf{b}_1 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$, and let $A = \begin{bmatrix} 6 & -2 \\ -3 & 2 \end{bmatrix}$. Compute the area of the image of S under the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

28. Repeat Exercise 27 with $\mathbf{b}_1 = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $A = \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}$.

29. Find a formula for the area of the triangle whose vertices are $\mathbf{0}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ in $\mathbb{R}^2$.

30. Let R be the triangle with vertices at (x_1, y_1), (x_2, y_2), and (x_3, y_3). Show that

$$\{\text{area of triangle}\} = \frac{1}{2} \det \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$$

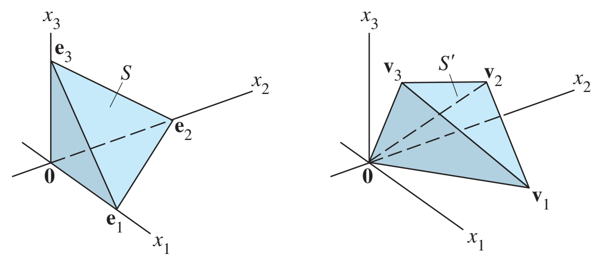
[*Hint:* Translate R to the origin by subtracting one of the vertices, and use Exercise 29.]

31. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation determined by the matrix $A = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$, where a, b, and c are positive numbers. Let S be the unit ball, whose bounding surface has the equation $x_1^2 + x_2^2 + x_3^2 = 1$.

a. Show that $T(S)$ is bounded by the ellipsoid with the equation $\dfrac{x_1^2}{a^2} + \dfrac{x_2^2}{b^2} + \dfrac{x_3^2}{c^2} = 1$.

b. Use the fact that the volume of the unit ball is $4\pi/3$ to determine the volume of the region bounded by the ellipsoid in part (a).

32. Let S be the tetrahedron in $\mathbb{R}^3$ with vertices at the vectors $\mathbf{0}$, $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$, and let S' be the tetrahedron with vertices at vectors $\mathbf{0}$, $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. See the figure.



a. Describe a linear transformation that maps S onto S'.

b. Find a formula for the volume of the tetrahedron S' using the fact that

$$\{\text{volume of } S\} = (1/3)\{\text{area of base}\} \cdot \{\text{height}\}$$

33. [**M**] Test the inverse formula of Theorem 8 for a random 4×4 matrix A. Use your matrix program to compute the cofactors of the 3×3 submatrices, construct the adjugate, and set $B = (\text{adj } A)/(\det A)$. Then compute $B - \text{inv}(A)$, where $\text{inv}(A)$ is the inverse of A as computed by the matrix program. Use floating point arithmetic with the maximum possible number of decimal places. Report your results.

34. [**M**] Test Cramer's rule for a random 4×4 matrix A and a random 4×1 vector $\mathbf{b}$. Compute each entry in the solution of $A\mathbf{x} = \mathbf{b}$, and compare these entries with the entries in $A^{-1}\mathbf{b}$. Write the command (or keystrokes) for your matrix program that uses Cramer's rule to produce the second entry of $\mathbf{x}$.

35. [**M**] If your version of MATLAB has the `flops` command, use it to count the number of floating point operations to compute A^{-1} for a random 30×30 matrix. Compare this number with the number of flops needed to form $(\text{adj } A)/(\det A)$.

SOLUTION TO PRACTICE PROBLEM

The area of S is $\left|\det \begin{bmatrix} 1 & 5 \\ 3 & 1 \end{bmatrix}\right| = 14$, and $\det A = 2$. By Theorem 10, the area of the image of S under the mapping $\mathbf{x} \mapsto A\mathbf{x}$ is

$$|\det A| \cdot \{\text{area of } S\} = 2 \cdot 14 = 28$$

CHAPTER 3 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. Assume that all matrices here are square.

a. If A is a 2×2 matrix with a zero determinant, then one column of A is a multiple of the other.

b. If two rows of a 3×3 matrix A are the same, then $\det A = 0$.

c. If A is a 3×3 matrix, then $\det 5A = 5 \det A$.

d. If A and B are $n \times n$ matrices, with $\det A = 2$ and $\det B = 3$, then $\det(A + B) = 5$.

e. If A is $n \times n$ and $\det A = 2$, then $\det A^3 = 6$.

f. If B is produced by interchanging two rows of A, then $\det B = \det A$.

g. If B is produced by multiplying row 3 of A by 5, then $\det B = 5 \cdot \det A$.

h. If B is formed by adding to one row of A a linear combination of the other rows, then $\det B = \det A$.

i. $\det A^T = -\det A$.

j. $\det(-A) = -\det A$.

k. $\det A^TA \geq 0$.

l. Any system of n linear equations in n variables can be solved by Cramer's rule.

m. If $\mathbf{u}$ and $\mathbf{v}$ are in $\mathbb{R}^2$ and $\det [\,\mathbf{u} \quad \mathbf{v}\,] = 10$, then the area of the triangle in the plane with vertices at $\mathbf{0}$, $\mathbf{u}$, and $\mathbf{v}$ is 10.

n. If $A^3 = 0$, then $\det A = 0$.

o. If A is invertible, then $\det A^{-1} = \det A$.

p. If A is invertible, then $(\det A)(\det A^{-1}) = 1$.

Use row operations to show that the determinants in Exercises 2–4 are all zero.

2. $\begin{vmatrix} 12 & 13 & 14 \\ 15 & 16 & 17 \\ 18 & 19 & 20 \end{vmatrix}$ **3.** $\begin{vmatrix} 1 & a & b+c \\ 1 & b & a+c \\ 1 & c & a+b \end{vmatrix}$

4. $\begin{vmatrix} a & b & c \\ a+x & b+x & c+x \\ a+y & b+y & c+y \end{vmatrix}$

Compute the determinants in Exercises 5 and 6.

5. $\begin{vmatrix} 9 & 1 & 9 & 9 & 9 \\ 9 & 0 & 9 & 9 & 2 \\ 4 & 0 & 0 & 5 & 0 \\ 9 & 0 & 3 & 9 & 0 \\ 6 & 0 & 0 & 7 & 0 \end{vmatrix}$

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

7. Show that the equation of the line in $\mathbb{R}^2$ through distinct points (x_1, y_1) and (x_2, y_2) can be written as

$$\det \begin{bmatrix} 1 & x & y \\ 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \end{bmatrix} = 0$$

8. Find a 3×3 determinant equation similar to that in Exercise 7 that describes the equation of the line through (x_1, y_1) with slope m.

Exercises 9 and 10 concern determinants of the following *Vandermonde matrices*.

$$T = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}, \quad V(t) = \begin{bmatrix} 1 & t & t^2 & t^3 \\ 1 & x_1 & x_1^2 & x_1^3 \\ 1 & x_2 & x_2^2 & x_2^3 \\ 1 & x_3 & x_3^2 & x_3^3 \end{bmatrix}$$

9. Use row operations to show that

$$\det T = (b-a)(c-a)(c-b)$$

10. Let $f(t) = \det V$, with x_1, x_2, x_3 all distinct. Explain why $f(t)$ is a cubic polynomial, show that the coefficient of t^3 is nonzero, and find three points on the graph of f.

11. Determine the area of the parallelogram determined by the points $(1, 4)$, $(-1, 5)$, $(3, 9)$, and $(5, 8)$. How can you tell that the quadrilateral determined by the points is actually a parallelogram?

12. Use the concept of area of a parallelogram to write a statement about a 2×2 matrix A that is true if and only if A is invertible.

13. Show that if A is invertible, then adj A is invertible, and

$$(\text{adj } A)^{-1} = \frac{1}{\det A} A$$

[*Hint:* Given matrices B and C, what calculation(s) would show that C is the inverse of B?]

14. Let A, B, C, D, and I be $n \times n$ matrices. Use the definition or properties of a determinant to justify the following formulas. Part (c) is useful in applications of eigenvalues (Chapter 5).

a. $\det \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} = \det A$ b. $\det \begin{bmatrix} I & 0 \\ C & D \end{bmatrix} = \det D$

c. $\det \begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = (\det A)(\det D) = \det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$

15. Let A, B, C, and D be $n \times n$ matrices with A invertible.

a. Find matrices X and Y to produce the block LU factorization

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & Y \end{bmatrix}$$

and then show that

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (\det A) \cdot \det(D - CA^{-1}B)$$

b. Show that if $AC = CA$, then

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(AD - CB)$$

16. Let J be the $n \times n$ matrix of all 1's, and consider $A = (a-b)I + bJ$; that is,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Confirm that $\det A = (a-b)^{n-1}[a + (n-1)b]$ as follows:

a. Subtract row 2 from row 1, row 3 from row 2, and so on, and explain why this does not change the determinant of the matrix.

b. With the resulting matrix from part (a), add column 1 to column 2, then add this new column 2 to column 3, and so on, and explain why this does not change the determinant.

c. Find the determinant of the resulting matrix from (b).

17. Let A be the original matrix given in Exercise 16, and let

$$B = \begin{bmatrix} a-b & b & b & \cdots & b \\ 0 & a & b & \cdots & b \\ 0 & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & b & b & \cdots & a \end{bmatrix},$$

$$C = \begin{bmatrix} b & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Notice that A, B, and C are nearly the same except that the first column of A equals the sum of the first columns of B and C. A *linearity property* of the determinant function, discussed in Section 3.2, says that $\det A = \det B + \det C$. Use this fact to prove the formula in Exercise 16 by induction on the size of matrix A.

18. **[M]** Apply the result of Exercise 16 to find the determinants of the following matrices, and confirm your answers using a matrix program.

$$\begin{bmatrix} 3 & 8 & 8 & 8 \\ 8 & 3 & 8 & 8 \\ 8 & 8 & 3 & 8 \\ 8 & 8 & 8 & 3 \end{bmatrix} \qquad \begin{bmatrix} 8 & 3 & 3 & 3 & 3 \\ 3 & 8 & 3 & 3 & 3 \\ 3 & 3 & 8 & 3 & 3 \\ 3 & 3 & 3 & 8 & 3 \\ 3 & 3 & 3 & 3 & 8 \end{bmatrix}$$

19. **[M]** Use a matrix program to compute the determinants of the following matrices.

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

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

Use the results to guess the determinant of the matrix below, and confirm your guess by using row operations to evaluate that determinant.

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{bmatrix}$$

20. **[M]** Use the method of Exercise 19 to guess the determinant of

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 3 & 3 & \cdots & 3 \\ 1 & 3 & 6 & \cdots & 6 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 6 & \cdots & 3(n-1) \end{bmatrix}$$

Justify your conjecture. [*Hint:* Use Exercise 14(c) and the result of Exercise 19.]

SOLUTION Write the vectors in H as column vectors. Then an arbitrary vector in H has the form

$$\begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\uparrow \mathbf{v}_1 \qquad \uparrow \mathbf{v}_2$$

This calculation shows that $H = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the vectors indicated above. Thus H is a subspace of $\mathbb{R}^4$ by Theorem 1. ■

Example 11 illustrates a useful technique of expressing a subspace H as the set of linear combinations of some small collection of vectors. If $H = \text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$, we can think of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ in the spanning set as "handles" that allow us to hold on to the subspace H. Calculations with the infinitely many vectors in H are often reduced to operations with the finite number of vectors in the spanning set.

EXAMPLE 12 For what value(s) of h will $\mathbf{y}$ be in the subspace of $\mathbb{R}^3$ spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, if

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

SOLUTION This question is Practice Problem 2 in Section 1.3, written here with the term *subspace* rather than $\text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. The solution there shows that $\mathbf{y}$ is in $\text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if $h = 5$. That solution is worth reviewing now, along with Exercises 11–16 and 19–21 in Section 1.3. ■

Although many vector spaces in this chapter will be subspaces of $\mathbb{R}^n$, it is important to keep in mind that the abstract theory applies to other vector spaces as well. Vector spaces of functions arise in many applications, and they will receive more attention later.

PRACTICE PROBLEMS

1. Show that the set H of all points in $\mathbb{R}^2$ of the form $(3s, 2+5s)$ is not a vector space, by showing that it is not closed under scalar multiplication. (Find a specific vector $\mathbf{u}$ in H and a scalar c such that $c\mathbf{u}$ is not in H.)
2. Let $W = \text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$, where $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are in a vector space V. Show that $\mathbf{v}_k$ is in W for $1 \le k \le p$. [*Hint:* First write an equation that shows that $\mathbf{v}_1$ is in W. Then adjust your notation for the general case.]

WEB

4.1 EXERCISES

1. Let V be the first quadrant in the xy-plane; that is, let
$$V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \ge 0, y \ge 0 \right\}$$
   a. If $\mathbf{u}$ and $\mathbf{v}$ are in V, is $\mathbf{u}+\mathbf{v}$ in V? Why?
   b. Find a specific vector $\mathbf{u}$ in V and a specific scalar c such that $c\mathbf{u}$ is *not* in V. (This is enough to show that V is *not* a vector space.)
2. Let W be the union of the first and third quadrants in the xy-plane. That is, let $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : xy \ge 0 \right\}$.
   a. If $\mathbf{u}$ is in W and c is any scalar, is $c\mathbf{u}$ in W? Why?

b. Find specific vectors **u** and **v** in W such that $\mathbf{u}+\mathbf{v}$ is not in W. This is enough to show that W is *not* a vector space.

3. Let H be the set of points inside and on the unit circle in the xy-plane. That is, let $H = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + y^2 \le 1 \right\}$. Find a specific example—two vectors or a vector and a scalar—to show that H is not a subspace of $\mathbb{R}^2$.

4. Construct a geometric figure that illustrates why a line in $\mathbb{R}^2$ *not* through the origin is not closed under vector addition.

In Exercises 5–8, determine if the given set is a subspace of $\mathbb{P}_n$ for an appropriate value of n. Justify your answers.

5. All polynomials of the form $\mathbf{p}(t) = at^2$, where a is in $\mathbb{R}$.

6. All polynomials of the form $\mathbf{p}(t) = a + t^2$, where a is in $\mathbb{R}$.

7. All polynomials of degree at most 3, with integers as coefficients.

8. All polynomials in $\mathbb{P}_n$ such that $\mathbf{p}(0) = 0$.

9. Let H be the set of all vectors of the form $\begin{bmatrix} -2t \\ 5t \\ 3t \end{bmatrix}$. Find a vector **v** in $\mathbb{R}^3$ such that $H = \text{Span}\,\{\mathbf{v}\}$. Why does this show that H is a subspace of $\mathbb{R}^3$?

10. Let H be the set of all vectors of the form $\begin{bmatrix} 3t \\ 0 \\ -7t \end{bmatrix}$, where t is any real number. Show that H is a subspace of $\mathbb{R}^3$. (Use the method of Exercise 9.)

11. Let W be the set of all vectors of the form $\begin{bmatrix} 2b+3c \\ -b \\ 2c \end{bmatrix}$, where b and c are arbitrary. Find vectors **u** and **v** such that $W = \text{Span}\,\{\mathbf{u}, \mathbf{v}\}$. Why does this show that W is a subspace of $\mathbb{R}^3$?

12. Let W be the set of all vectors of the form $\begin{bmatrix} 2s+4t \\ 2s \\ 2s-3t \\ 5t \end{bmatrix}$. Show that W is a subspace of $\mathbb{R}^4$. (Use the method of Exercise 11.)

13. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$.

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

b. How many vectors are in $\text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

c. Is **w** in the subspace spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? Why?

14. Let $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ be as in Exercise 13, and let $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 14 \end{bmatrix}$. Is **w** in the subspace spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$? Why?

In Exercises 15–18, let W be the set of all vectors of the form shown, where a, b, and c represent arbitrary real numbers. In each case, either find a set S of vectors that spans W or give an example to show that W is *not* a vector space.

15. $\begin{bmatrix} 2a+3b \\ -1 \\ 2a-5b \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 3a-5b \\ 3b+2a \end{bmatrix}$

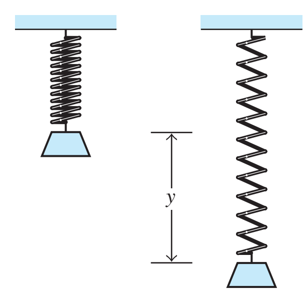
17. $\begin{bmatrix} 2a-b \\ 3b-c \\ 3c-a \\ 3b \end{bmatrix}$

18. $\begin{bmatrix} 4a+3b \\ 0 \\ a+3b+c \\ 3b-2c \end{bmatrix}$

19. If a mass m is placed at the end of a spring, and if the mass is pulled downward and released, the mass–spring system will begin to oscillate. The displacement y of the mass from its resting position is given by a function of the form

$$y(t) = c_1 \cos \omega t + c_2 \sin \omega t \qquad (5)$$

where ω is a constant that depends on the spring and the mass. (See the figure below.) Show that the set of all functions described in (5) (with ω fixed and c_1, c_2 arbitrary) is a vector space.



20. The set of all continuous real-valued functions defined on a closed interval $[a, b]$ in $\mathbb{R}$ is denoted by $C[a, b]$. This set is a subspace of the vector space of all real-valued functions defined on $[a, b]$.

a. What facts about continuous functions should be proved in order to demonstrate that $C[a, b]$ is indeed a subspace as claimed? (These facts are usually discussed in a calculus class.)

b. Show that $\{\mathbf{f} \text{ in } C[a, b] : \mathbf{f}(a) = \mathbf{f}(b)\}$ is a subspace of $C[a, b]$.

For fixed positive integers m and n, the set $M_{m \times n}$ of all $m \times n$ matrices is a vector space, under the usual operations of addition of matrices and multiplication by real scalars.

21. Determine if the set H of all matrices of the form $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$ is a subspace of $M_{2 \times 2}$.

22. Let F be a fixed 3×2 matrix, and let H be the set of all matrices A in $M_{2 \times 4}$ with the property that $FA = 0$ (the zero matrix in $M_{3 \times 4}$). Determine if H is a subspace of $M_{2 \times 4}$.

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

23. a. If $\mathbf{f}$ is a function in the vector space V of all real-valued functions on $\mathbb{R}$ and if $\mathbf{f}(t) = 0$ for some t, then $\mathbf{f}$ is the zero vector in V.

b. A vector is an arrow in three-dimensional space.

c. A subset H of a vector space V is a subspace of V if the zero vector is in H.

d. A subspace is also a vector space.

e. Analog signals are used in the major control systems for the space shuttle, mentioned in the introduction to the chapter.

24. a. A vector is any element of a vector space.

b. If $\mathbf{u}$ is a vector in a vector space V, then $(-1)\mathbf{u}$ is the same as the negative of $\mathbf{u}$.

c. A vector space is also a subspace.

d. $\mathbb{R}^2$ is a subspace of $\mathbb{R}^3$.

e. A subset H of a vector space V is a subspace of V if the following conditions are satisfied: (i) the zero vector of V is in H, (ii) $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u}+\mathbf{v}$ are in H, and (iii) c is a scalar and $c\mathbf{u}$ is in H.

Exercises 25–29 show how the axioms for a vector space V can be used to prove the elementary properties described after the definition of a vector space. Fill in the blanks with the appropriate axiom numbers. Because of Axiom 2, Axioms 4 and 5 imply, respectively, that $\mathbf{0}+\mathbf{u}=\mathbf{u}$ and $-\mathbf{u}+\mathbf{u}=\mathbf{0}$ for all $\mathbf{u}$.

25. Complete the following proof that the zero vector is unique. Suppose that $\mathbf{w}$ in V has the property that $\mathbf{u}+\mathbf{w}=\mathbf{w}+\mathbf{u}=\mathbf{u}$ for all $\mathbf{u}$ in V. In particular, $\mathbf{0}+\mathbf{w}=\mathbf{0}$. But $\mathbf{0}+\mathbf{w}=\mathbf{w}$, by Axiom ______. Hence $\mathbf{w}=\mathbf{0}+\mathbf{w}=\mathbf{0}$.

26. Complete the following proof that $-\mathbf{u}$ is the *unique vector* in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$. Suppose that $\mathbf{w}$ satisfies $\mathbf{u}+\mathbf{w}=\mathbf{0}$. Adding $-\mathbf{u}$ to both sides, we have

$(-\mathbf{u})+[\mathbf{u}+\mathbf{w}]=(-\mathbf{u})+\mathbf{0}$

$[(-\mathbf{u})+\mathbf{u}]+\mathbf{w}=(-\mathbf{u})+\mathbf{0}$ by Axiom ______ (a)

$\mathbf{0}+\mathbf{w}=(-\mathbf{u})+\mathbf{0}$ by Axiom ______ (b)

$\mathbf{w}=-\mathbf{u}$ by Axiom ______ (c)

27. Fill in the missing axiom numbers in the following proof that $0\mathbf{u}=\mathbf{0}$ for every $\mathbf{u}$ in V.

$0\mathbf{u}=(0+0)\mathbf{u}=0\mathbf{u}+0\mathbf{u}$ by Axiom ______ (a)

Add the negative of $0\mathbf{u}$ to both sides:

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

$0\mathbf{u}+(-0\mathbf{u})=0\mathbf{u}+[0\mathbf{u}+(-0\mathbf{u})]$ by Axiom ______ (b)

$\mathbf{0}=0\mathbf{u}+\mathbf{0}$ by Axiom ______ (c)

$\mathbf{0}=0\mathbf{u}$ by Axiom ______ (d)

28. Fill in the missing axiom numbers in the following proof that $c\mathbf{0}=\mathbf{0}$ for every scalar c.

$c\mathbf{0}=c(\mathbf{0}+\mathbf{0})$ by Axiom ______ (a)

$=c\mathbf{0}+c\mathbf{0}$ by Axiom ______ (b)

Add the negative of $c\mathbf{0}$ to both sides:

$c\mathbf{0}+(-c\mathbf{0})=[c\mathbf{0}+c\mathbf{0}]+(-c\mathbf{0})$

$c\mathbf{0}+(-c\mathbf{0})=c\mathbf{0}+[c\mathbf{0}+(-c\mathbf{0})]$ by Axiom ______ (c)

$\mathbf{0}=c\mathbf{0}+\mathbf{0}$ by Axiom ______ (d)

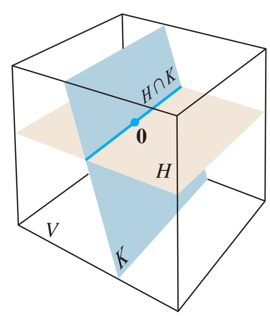
$\mathbf{0}=c\mathbf{0}$ by Axiom ______ (e)

29. Prove that $(-1)\mathbf{u}=-\mathbf{u}$. [*Hint:* Show that $\mathbf{u}+(-1)\mathbf{u}=\mathbf{0}$. Use some axioms and the results of Exercises 27 and 26.]

30. Suppose $c\mathbf{u}=\mathbf{0}$ for some nonzero scalar c. Show that $\mathbf{u}=\mathbf{0}$. Mention the axioms or properties you use.

31. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in a vector space V, and let H be any subspace of V that contains both $\mathbf{u}$ and $\mathbf{v}$. Explain why H also contains Span $\{\mathbf{u},\mathbf{v}\}$. This shows that Span $\{\mathbf{u},\mathbf{v}\}$ is the smallest subspace of V that contains both $\mathbf{u}$ and $\mathbf{v}$.

32. Let H and K be subspaces of a vector space V. The **intersection** of H and K, written as $H\cap K$, is the set of $\mathbf{v}$ in V that belong to both H and K. Show that $H\cap K$ is a subspace of V. (See the figure.) Give an example in $\mathbb{R}^2$ to show that the union of two subspaces is not, in general, a subspace.



33. Given subspaces H and K of a vector space V, the **sum** of H and K, written as $H+K$, is the set of all vectors in V that can be written as the sum of two vectors, one in H and the other in K; that is,

$$H+K=\{\mathbf{w}:\mathbf{w}=\mathbf{u}+\mathbf{v}\text{ for some }\mathbf{u}\text{ in }H\text{ and some }\mathbf{v}\text{ in }K\}$$

a. Show that $H+K$ is a subspace of V.

b. Show that H is a subspace of $H+K$ and K is a subspace of $H+K$.

34. Suppose $\mathbf{u}_1,\ldots,\mathbf{u}_p$ and $\mathbf{v}_1,\ldots,\mathbf{v}_q$ are vectors in a vector space V, and let

$$H=\text{Span}\,\{\mathbf{u}_1,\ldots,\mathbf{u}_p\}\text{ and }K=\text{Span}\,\{\mathbf{v}_1,\ldots,\mathbf{v}_q\}$$

Show that $H+K=\text{Span}\,\{\mathbf{u}_1,\ldots,\mathbf{u}_p,\mathbf{v}_1,\ldots,\mathbf{v}_q\}$.

35. [**M**] Show that $\mathbf{w}$ is in the subspace of $\mathbb{R}^4$ spanned by $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, where

$$\mathbf{w} = \begin{bmatrix} 9 \\ -4 \\ -4 \\ 7 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 8 \\ -4 \\ -3 \\ 9 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \\ -2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 6 \\ -5 \\ -18 \end{bmatrix}$$

36. [**M**] Determine if $\mathbf{y}$ is in the subspace of $\mathbb{R}^4$ spanned by the columns of A, where

$$\mathbf{y} = \begin{bmatrix} -4 \\ -8 \\ 6 \\ -5 \end{bmatrix}, \quad A = \begin{bmatrix} 3 & -5 & -9 \\ 8 & 7 & -6 \\ -5 & -8 & 3 \\ 2 & -2 & -9 \end{bmatrix}$$

37. [**M**] The vector space $H = \text{Span}\,\{1, \cos^2 t, \cos^4 t, \cos^6 t\}$ contains at least two interesting functions that will be used in a later exercise:

$$\mathbf{f}(t) = 1 - 8\cos^2 t + 8\cos^4 t$$
$$\mathbf{g}(t) = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$$

Study the graph of $\mathbf{f}$ for $0 \le t \le 2\pi$, and guess a simple formula for $\mathbf{f}(t)$. Verify your conjecture by graphing the difference between $1 + \mathbf{f}(t)$ and your formula for $\mathbf{f}(t)$. (Hopefully, you will see the constant function 1.) Repeat for $\mathbf{g}$.

38. [**M**] Repeat Exercise 37 for the functions

$$\mathbf{f}(t) = 3\sin t - 4\sin^3 t$$
$$\mathbf{g}(t) = 1 - 8\sin^2 t + 8\sin^4 t$$
$$\mathbf{h}(t) = 5\sin t - 20\sin^3 t + 16\sin^5 t$$

in the vector space $\text{Span}\,\{1, \sin t, \sin^2 t, \dots, \sin^5 t\}$.

SOLUTIONS TO PRACTICE PROBLEMS

1. Take any $\mathbf{u}$ in H—say, $\mathbf{u} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$—and take any $c \neq 1$—say, $c = 2$. Then $c\mathbf{u} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$. If this is in H, then there is some s such that

$$\begin{bmatrix} 3s \\ 2 + 5s \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

That is, $s = 2$ and $s = 12/5$, which is impossible. So $2\mathbf{u}$ is not in H and H is not a vector space.

2. $\mathbf{v}_1 = 1\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_p$. This expresses $\mathbf{v}_1$ as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_p$, so $\mathbf{v}_1$ is in W. In general, $\mathbf{v}_k$ is in W because

$$\mathbf{v}_k = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{k-1} + 1\mathbf{v}_k + 0\mathbf{v}_{k+1} + \cdots + 0\mathbf{v}_p$$

4.2 NULL SPACES, COLUMN SPACES, AND LINEAR TRANSFORMATIONS

In applications of linear algebra, subspaces of $\mathbb{R}^n$ usually arise in one of two ways: (1) as the set of all solutions to a system of homogeneous linear equations or (2) as the set of all linear combinations of certain specified vectors. In this section, we compare and contrast these two descriptions of subspaces, allowing us to practice using the concept of a subspace. Actually, as you will soon discover, we have been working with subspaces ever since Section 1.3. The main new feature here is the terminology. The section concludes with a discussion of the kernel and range of a linear transformation.

The Null Space of a Matrix

Consider the following system of homogeneous equations:

$$\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -5x_1 + 9x_2 + x_3 &= 0 \end{aligned} \tag{1}$$

Typically, such a linear transformation is described in terms of one or more derivatives of a function. To explain this in any detail would take us too far afield at this point. So we consider only two examples. The first explains why the operation of differentiation is a linear transformation.

EXAMPLE 8 (Calculus required) Let V be the vector space of all real-valued functions f defined on an interval $[a, b]$ with the property that they are differentiable and their derivatives are continuous functions on $[a, b]$. Let W be the vector space $C[a, b]$ of all continuous functions on $[a, b]$, and let $D : V \to W$ be the transformation that changes f in V into its derivative f'. In calculus, two simple differentiation rules are

$$D(f + g) = D(f) + D(g) \quad \text{and} \quad D(cf) = cD(f)$$

That is, D is a linear transformation. It can be shown that the kernel of D is the set of constant functions on $[a, b]$ and the range of D is the set W of all continuous functions on $[a, b]$. ■

EXAMPLE 9 (Calculus required) The differential equation

$$y'' + \omega^2 y = 0 \tag{4}$$

where ω is a constant, is used to describe a variety of physical systems, such as the vibration of a weighted spring, the movement of a pendulum, and the voltage in an inductance-capacitance electrical circuit. The set of solutions of (4) is precisely the kernel of the linear transformation that maps a function $y = f(t)$ into the function $f''(t) + \omega^2 f(t)$. Finding an explicit description of this vector space is a problem in differential equations. The solution set turns out to be the space described in Exercise 19 in Section 4.1. ■

PRACTICE PROBLEMS

1. Let $W = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a - 3b - c = 0 \right\}$. Show in two different ways that W is a subspace of $\mathbb{R}^3$. (Use two theorems.)

2. Let $A = \begin{bmatrix} 7 & -3 & 5 \\ -4 & 1 & -5 \\ -5 & 2 & -4 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 7 \\ 6 \\ -3 \end{bmatrix}$. Suppose you know that the equations $A\mathbf{x} = \mathbf{v}$ and $A\mathbf{x} = \mathbf{w}$ are both consistent. What can you say about the equation $A\mathbf{x} = \mathbf{v} + \mathbf{w}$?

4.2 EXERCISES

1. Determine if $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$ is in Nul A, where

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

2. Determine if $\mathbf{w} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ is in Nul A, where

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

In Exercises 3–6, find an explicit description of Nul A, by listing vectors that span the null space.

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

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

5. $A = \begin{bmatrix} 1 & -4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}$

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

In Exercises 7–14, either use an appropriate theorem to show that the given set, W, is a vector space, or find a specific example to the contrary.

7. $\left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} : a + b + c = 2 \right\}$ **8.** $\left\{ \begin{bmatrix} r \\ s \\ t \end{bmatrix} : 3r - 2 = 3s + t \right\}$

9. $\left\{ \begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} : \begin{matrix} p - 3q = 4s \\ 2p = s + 5r \end{matrix} \right\}$ **10.** $\left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} : \begin{matrix} 3a + b = c \\ a + b + 2c = 2d \end{matrix} \right\}$

11. $\left\{ \begin{bmatrix} s - 2t \\ 3 + 3s \\ 3s + t \\ 2s \end{bmatrix} : s, t \text{ real} \right\}$ **12.** $\left\{ \begin{bmatrix} 3p - 5q \\ 4q \\ p \\ q + 1 \end{bmatrix} : p, q \text{ real} \right\}$

13. $\left\{ \begin{bmatrix} c - 6d \\ d \\ c \end{bmatrix} : c, d \text{ real} \right\}$ **14.** $\left\{ \begin{bmatrix} -s + 3t \\ s - 2t \\ 5s - t \end{bmatrix} : s, t \text{ real} \right\}$

In Exercises 15 and 16, find A such that the given set is Col A.

15. $\left\{ \begin{bmatrix} 2s + t \\ r - s + 2t \\ 3r + s \\ 2r - s - t \end{bmatrix} : r, s, t \text{ real} \right\}$

16. $\left\{ \begin{bmatrix} b - c \\ 2b + 3d \\ b + 3c - 3d \\ c + d \end{bmatrix} : b, c, d \text{ real} \right\}$

For the matrices in Exercises 17–20, (a) find k such that Nul A is a subspace of $\mathbb{R}^k$, and (b) find k such that Col A is a subspace of $\mathbb{R}^k$.

17. $A = \begin{bmatrix} 6 & -4 \\ -3 & 2 \\ -9 & 6 \\ 9 & -6 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 5 & -2 & 3 \\ -1 & 0 & -1 \\ 0 & -2 & -2 \\ -5 & 7 & 2 \end{bmatrix}$

19. $A = \begin{bmatrix} 4 & 5 & -2 & 6 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}$

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

21. With A as in Exercise 17, find a nonzero vector in Nul A and a nonzero vector in Col A.

22. With A as in Exercise 18, find a nonzero vector in Nul A and a nonzero vector in Col A.

23. Let $A = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Determine if $\mathbf{w}$ is in Col A. Is $\mathbf{w}$ in Nul A?

24. Let $A = \begin{bmatrix} 10 & -8 & -2 & -2 \\ 0 & 2 & 2 & -2 \\ 1 & -1 & 6 & 0 \\ 1 & 1 & 0 & -2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 2 \\ 2 \\ 0 \\ 2 \end{bmatrix}$. Determine if $\mathbf{w}$ is in Col A. Is $\mathbf{w}$ in Nul A?

In Exercises 25 and 26, A denotes an $m \times n$ matrix. Mark each statement True or False. Justify each answer.

25. a. The null space of A is the solution set of the equation $A\mathbf{x} = \mathbf{0}$.

b. The null space of an $m \times n$ matrix is in $\mathbb{R}^m$.

c. The column space of A is the range of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

d. If the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then Col A is $\mathbb{R}^m$.

e. The kernel of a linear transformation is a vector space.

f. Col A is the set of all vectors that can be written as $A\mathbf{x}$ for some $\mathbf{x}$.

26. a. A null space is a vector space.

b. The column space of an $m \times n$ matrix is in $\mathbb{R}^m$.

c. Col A is the set of all solutions of $A\mathbf{x} = \mathbf{b}$.

d. Nul A is the kernel of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

e. The range of a linear transformation is a vector space.

f. The set of all solutions of a homogeneous linear differential equation is the kernel of a linear transformation.

27. It can be shown that a solution of the system below is $x_1 = 3$, $x_2 = 2$, and $x_3 = -1$. Use this fact and the theory from this section to explain why another solution is $x_1 = 30$, $x_2 = 20$, and $x_3 = -10$. (Observe how the solutions are related, but make no other calculations.)

$$\begin{aligned} x_1 - 3x_2 - 3x_3 &= 0 \\ -2x_1 + 4x_2 + 2x_3 &= 0 \\ -x_1 + 5x_2 + 7x_3 &= 0 \end{aligned}$$

28. Consider the following two systems of equations:

$$\begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 1 \\ 4x_1 + x_2 - 6x_3 &= 9 \end{aligned} \qquad \begin{aligned} 5x_1 + x_2 - 3x_3 &= 0 \\ -9x_1 + 2x_2 + 5x_3 &= 5 \\ 4x_1 + x_2 - 6x_3 &= 45 \end{aligned}$$

It can be shown that the first system has a solution. Use this fact and the theory from this section to explain why the second system must also have a solution. (Make no row operations.)

29. Prove Theorem 3 as follows: Given an $m \times n$ matrix A, an element in Col A has the form $A\mathbf{x}$ for some $\mathbf{x}$ in $\mathbb{R}^n$. Let $A\mathbf{x}$ and $A\mathbf{w}$ represent any two vectors in Col A.

a. Explain why the zero vector is in Col A.

b. Show that the vector $A\mathbf{x} + A\mathbf{w}$ is in Col A.

c. Given a scalar c, show that $c(A\mathbf{x})$ is in Col A.

30. Let $T : V \to W$ be a linear transformation from a vector space V into a vector space W. Prove that the range of T is a subspace of W. [*Hint:* Typical elements of the range have the form $T(\mathbf{x})$ and $T(\mathbf{w})$ for some $\mathbf{x}$, $\mathbf{w}$ in V.]

31. Define $T : \mathbb{P}_2 \to \mathbb{R}^2$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$. For instance, if $\mathbf{p}(t) = 3 + 5t + 7t^2$, then $T(\mathbf{p}) = \begin{bmatrix} 3 \\ 15 \end{bmatrix}$.

a. Show that T is a linear transformation. [*Hint:* For arbitrary polynomials $\mathbf{p}$, $\mathbf{q}$ in $\mathbb{P}_2$, compute $T(\mathbf{p} + \mathbf{q})$ and $T(c\mathbf{p})$.]

b. Find a polynomial $\mathbf{p}$ in $\mathbb{P}_2$ that spans the kernel of T, and describe the range of T.

32. Define a linear transformation $T : \mathbb{P}_2 \to \mathbb{R}^2$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(0) \\ \mathbf{p}(0) \end{bmatrix}$. Find polynomials $\mathbf{p}_1$ and $\mathbf{p}_2$ in $\mathbb{P}_2$ that span the kernel of T, and describe the range of T.

33. Let $M_{2\times 2}$ be the vector space of all 2×2 matrices, and define $T : M_{2\times 2} \to M_{2\times 2}$ by $T(A) = A + A^T$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

a. Show that T is a linear transformation.

b. Let B be any element of $M_{2\times 2}$ such that $B^T = B$. Find an A in $M_{2\times 2}$ such that $T(A) = B$.

c. Show that the range of T is the set of B in $M_{2\times 2}$ with the property that $B^T = B$.

d. Describe the kernel of T.

34. (*Calculus required*) Define $T : C[0, 1] \to C[0, 1]$ as follows: For $\mathbf{f}$ in $C[0, 1]$, let $T(\mathbf{f})$ be the antiderivative $\mathbf{F}$ of $\mathbf{f}$ such that $\mathbf{F}(0) = 0$. Show that T is a linear transformation, and describe the kernel of T. (See the notation in Exercise 20 of Section 4.1.)

35. Let V and W be vector spaces, and let $T : V \to W$ be a linear transformation. Given a subspace U of V, let $T(U)$ denote the set of all images of the form $T(\mathbf{x})$, where $\mathbf{x}$ is in U. Show that $T(U)$ is a subspace of W.

36. Given $T : V \to W$ as in Exercise 35, and given a subspace Z of W, let U be the set of all $\mathbf{x}$ in V such that $T(\mathbf{x})$ is in Z. Show that U is a subspace of V.

37. [**M**] Determine whether $\mathbf{w}$ is in the column space of A, the null space of A, or both, where

$$\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -3 \end{bmatrix}, \quad A = \begin{bmatrix} 7 & 6 & -4 & 1 \\ -5 & -1 & 0 & -2 \\ 9 & -11 & 7 & -3 \\ 19 & -9 & 7 & 1 \end{bmatrix}$$

38. [**M**] Determine whether $\mathbf{w}$ is in the column space of A, the null space of A, or both, where

$$\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$$

39. [**M**] Let $\mathbf{a}_1, \ldots, \mathbf{a}_5$ denote the columns of the matrix A, where

$$A = \begin{bmatrix} 5 & 1 & 2 & 2 & 0 \\ 3 & 3 & 2 & -1 & -12 \\ 8 & 4 & 4 & -5 & 12 \\ 2 & 1 & 1 & 0 & -2 \end{bmatrix}, \quad B = [\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_4 \,]$$

a. Explain why $\mathbf{a}_3$ and $\mathbf{a}_5$ are in the column space of B.

b. Find a set of vectors that spans Nul A.

c. Let $T : \mathbb{R}^5 \to \mathbb{R}^4$ be defined by $T(\mathbf{x}) = A\mathbf{x}$. Explain why T is neither one-to-one nor onto.

40. [**M**] Let $H = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ and $K = \text{Span}\,\{\mathbf{v}_3, \mathbf{v}_4\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 3 \\ 8 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 0 \\ -12 \\ -28 \end{bmatrix}.$$

Then H and K are subspaces of $\mathbb{R}^3$. In fact, H and K are planes in $\mathbb{R}^3$ through the origin, and they intersect in a line through $\mathbf{0}$. Find a nonzero vector $\mathbf{w}$ that generates that line. [*Hint:* $\mathbf{w}$ can be written as $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ and also as $c_3\mathbf{v}_3 + c_4\mathbf{v}_4$. To build $\mathbf{w}$, solve the equation $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = c_3\mathbf{v}_3 + c_4\mathbf{v}_4$ for the unknown c_j's.]

SG Mastering: Vector Space, Subspace, Col A, and Nul A 4–6

SOLUTIONS TO PRACTICE PROBLEMS

1. *First method:* W is a subspace of $\mathbb{R}^3$ by Theorem 2 because W is the set of all solutions to a system of homogeneous linear equations (where the system has only one equation). Equivalently, W is the null space of the 1×3 matrix $A = [\, 1 \quad -3 \quad -1 \,]$.

Second method: Solve the equation $a - 3b - c = 0$ for the leading variable a in terms of the free variables b and c. Any solution has the form $\begin{bmatrix} 3b+c \\ b \\ c \end{bmatrix}$, where b and c are arbitrary, and

$$\begin{bmatrix} 3b+c \\ b \\ c \end{bmatrix} = b\underset{\substack{\uparrow \\ \mathbf{v}_1}}{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}} + c\underset{\substack{\uparrow \\ \mathbf{v}_2}}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}$$

This calculation shows that $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Thus W is a subspace of $\mathbb{R}^3$ by Theorem 1. We could also solve the equation $a - 3b - c = 0$ for b or c and get alternative descriptions of W as a set of linear combinations of two vectors.

2. Both $\mathbf{v}$ and $\mathbf{w}$ are in Col A. Since Col A is a vector space, $\mathbf{v} + \mathbf{w}$ must be in Col A. That is, the equation $A\mathbf{x} = \mathbf{v} + \mathbf{w}$ is consistent.

4.3 LINEARLY INDEPENDENT SETS; BASES

In this section we identify and study the subsets that span a vector space V or a subspace H as "efficiently" as possible. The key idea is that of linear independence, defined as in $\mathbb{R}^n$.

An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in V is said to be **linearly independent** if the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0} \tag{1}$$

has *only* the trivial solution, $c_1 = 0, \dots, c_p = 0$.[1]

The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be **linearly dependent** if (1) has a nontrivial solution, that is, if there are some weights, $c_1, \dots, c_p$, *not all zero*, such that (1) holds. In such a case, (1) is called a **linear dependence relation** among $\mathbf{v}_1, \dots, \mathbf{v}_p$.

Just as in $\mathbb{R}^n$, a set containing a single vector $\mathbf{v}$ is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$. Also, a set of two vectors is linearly dependent if and only if one of the vectors is a multiple of the other. And any set containing the zero vector is linearly dependent. The following theorem has the same proof as Theorem 7 in Section 1.7.

THEOREM 4

An indexed set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ of two or more vectors, with $\mathbf{v}_1 \neq \mathbf{0}$, is linearly dependent if and only if some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors, $\mathbf{v}_1, \dots, \mathbf{v}_{j-1}$.

The main difference between linear dependence in $\mathbb{R}^n$ and in a general vector space is that when the vectors are not n-tuples, the homogeneous equation (1) usually cannot be written as a system of n linear equations. That is, the vectors cannot be made into the columns of a matrix A in order to study the equation $A\mathbf{x} = \mathbf{0}$. We must rely instead on the definition of linear dependence and on Theorem 4.

EXAMPLE 1 Let $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$, and $\mathbf{p}_3(t) = 4 - t$. Then $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is linearly dependent in $\mathbb{P}$ because $\mathbf{p}_3 = 4\mathbf{p}_1 - \mathbf{p}_2$. ■

[1] It is convenient to use $c_1, \dots, c_p$ in (1) for the scalars instead of $x_1, \dots, x_p$, as we did in Chapter 1.

the spanning property.

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix}\right\} \qquad \left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\3\\0\end{bmatrix},\begin{bmatrix}4\\5\\6\end{bmatrix},\begin{bmatrix}7\\8\\9\end{bmatrix}\right\}$$

Linearly independent but does not span $\mathbb{R}^3$ | A basis for $\mathbb{R}^3$ | Spans $\mathbb{R}^3$ but is linearly dependent ■

PRACTICE PROBLEMS

1. Let $\mathbf{v}_1 = \begin{bmatrix}1\\-2\\3\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}-2\\7\\-9\end{bmatrix}$. Determine if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^3$. Is $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for $\mathbb{R}^2$?

2. Let $\mathbf{v}_1 = \begin{bmatrix}1\\-3\\4\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}6\\2\\-1\end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix}2\\-2\\3\end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix}-4\\-8\\9\end{bmatrix}$. Find a basis for the subspace W spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

3. Let $\mathbf{v}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, and $H = \left\{\begin{bmatrix}s\\s\\0\end{bmatrix} : s \text{ in } \mathbb{R}\right\}$. Then every vector in H is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ because

$$\begin{bmatrix}s\\s\\0\end{bmatrix} = s\begin{bmatrix}1\\0\\0\end{bmatrix} + s\begin{bmatrix}0\\1\\0\end{bmatrix}$$

SG Mastering: Basis 4–9

Is $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for H?

4.3 EXERCISES

Determine whether the sets in Exercises 1–8 are bases for $\mathbb{R}^3$. Of the sets that are *not* bases, determine which ones are linearly independent and which ones span $\mathbb{R}^3$. Justify your answers.

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

2. $\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}$

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

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

5. $\begin{bmatrix}3\\-3\\0\end{bmatrix},\begin{bmatrix}-3\\7\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}0\\-3\\5\end{bmatrix}$

6. $\begin{bmatrix}1\\2\\-4\end{bmatrix},\begin{bmatrix}-4\\3\\6\end{bmatrix}$

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

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

Find bases for the null spaces of the matrices given in Exercises 9 and 10. Refer to the remarks that follow Example 3 in Section 4.2.

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

10. $\begin{bmatrix}1 & 1 & -2 & 1 & 5\\0 & 1 & 0 & -1 & -2\\0 & 0 & -8 & 0 & 16\end{bmatrix}$

11. Find a basis for the set of vectors in $\mathbb{R}^3$ in the plane $x - 3y + 2z = 0$. [*Hint:* Think of the equation as a "system" of homogeneous equations.]

12. Find a basis for the set of vectors in $\mathbb{R}^2$ on the line $y = -3x$.

In Exercises 13 and 14, assume that A is row equivalent to B. Find bases for Nul A and Col A.

13. $A = \begin{bmatrix}-2 & 4 & -2 & -4\\2 & -6 & -3 & 1\\-3 & 8 & 2 & -3\end{bmatrix}$, $B = \begin{bmatrix}1 & 0 & 6 & 5\\0 & 2 & 5 & 3\\0 & 0 & 0 & 0\end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 2 & 3 & -4 & 8 \\ 1 & 2 & 0 & 2 & 8 \\ 2 & 4 & -3 & 10 & 9 \\ 3 & 6 & 0 & 6 & 9 \end{bmatrix}$,

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

In Exercises 15–18, find a basis for the space spanned by the given vectors, $\mathbf{v}_1, \dots, \mathbf{v}_5$.

15. $\begin{bmatrix} 1 \\ 0 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -8 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 10 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -6 \\ 9 \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$

17. [**M**] $\begin{bmatrix} 2 \\ 0 \\ -4 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 2 \\ -4 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ -4 \\ 0 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} 8 \\ 4 \\ 8 \\ -3 \\ 15 \end{bmatrix}, \begin{bmatrix} -8 \\ 4 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

18. [**M**] $\begin{bmatrix} -3 \\ 2 \\ 6 \\ 0 \\ -7 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -9 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -4 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 6 \\ -2 \\ -14 \\ 0 \\ 13 \end{bmatrix}, \begin{bmatrix} -6 \\ 3 \\ 0 \\ -1 \\ 0 \end{bmatrix}$

19. Let $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -3 \\ 7 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 9 \\ -2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 7 \\ 11 \\ 6 \end{bmatrix}$, and also let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. It can be verified that $4\mathbf{v}_1 + 5\mathbf{v}_2 - 3\mathbf{v}_3 = \mathbf{0}$. Use this information to find a basis for H. There is more than one answer.

20. Let $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \\ -2 \\ -5 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 4 \\ 3 \\ 2 \\ 4 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 2 \\ 5 \\ -6 \\ -14 \end{bmatrix}$. It can be verified that $2\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$. Use this information to find a basis for $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

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

21. a. A single vector by itself is linearly dependent.

b. If $H = \text{Span}\{\mathbf{b}_1, \dots, \mathbf{b}_p\}$, then $\{\mathbf{b}_1, \dots, \mathbf{b}_p\}$ is a basis for H.

c. The columns of an invertible $n \times n$ matrix form a basis for $\mathbb{R}^n$.

d. A basis is a spanning set that is as large as possible.

e. In some cases, the linear dependence relations among the columns of a matrix can be affected by certain elementary row operations on the matrix.

22. a. A linearly independent set in a subspace H is a basis for H.

b. If a finite set S of nonzero vectors spans a vector space V, then some subset of S is a basis for V.

c. A basis is a linearly independent set that is as large as possible.

d. The standard method for producing a spanning set for Nul A, described in Section 4.2, sometimes fails to produce a basis for Nul A.

e. If B is an echelon form of a matrix A, then the pivot columns of B form a basis for Col A.

23. Suppose $\mathbb{R}^4 = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$. Explain why $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is a basis for $\mathbb{R}^4$.

24. Let $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a linearly independent set in $\mathbb{R}^n$. Explain why $\mathcal{B}$ must be a basis for $\mathbb{R}^n$.

25. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and let H be the set of vectors in $\mathbb{R}^3$ whose second and third entries are equal. Then every vector in H has a unique expansion as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, because

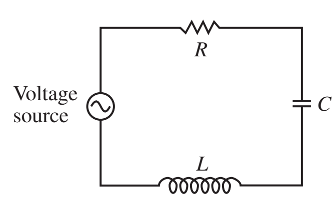
$$\begin{bmatrix} s \\ t \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (t - s)\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

for any s and t. Is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a basis for H? Why or why not?

26. In the vector space of all real-valued functions, find a basis for the subspace spanned by $\{\sin t, \sin 2t, \sin t \cos t\}$.

27. Let V be the vector space of functions that describe the vibration of a mass–spring system. (Refer to Exercise 19 in Section 4.1.) Find a basis for V.

28. (*RLC circuit*) The circuit in the figure consists of a resistor (R ohms), an inductor (L henrys), a capacitor (C farads), and an initial voltage source. Let $b = R/(2L)$, and suppose R, L, and C have been selected so that b also equals $1/\sqrt{LC}$. (This is done, for instance, when the circuit is used in a voltmeter.) Let $v(t)$ be the voltage (in volts) at time t, measured across the capacitor. It can be shown that v is in the null space H of the linear transformation that maps $v(t)$ into $Lv''(t) + Rv'(t) + (1/C)v(t)$, and H consists of all functions of the form $v(t) = e^{-bt}(c_1 + c_2 t)$. Find a basis for H.



Exercises 29 and 30 show that every basis for $\mathbb{R}^n$ must contain exactly n vectors.

29. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of k vectors in $\mathbb{R}^n$, with $k < n$. Use a theorem from Section 1.4 to explain why S cannot be a basis for $\mathbb{R}^n$.

30. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of k vectors in $\mathbb{R}^n$, with $k > n$. Use a theorem from Chapter 1 to explain why S cannot be a basis for $\mathbb{R}^n$.

Exercises 31 and 32 reveal an important connection between linear independence and linear transformations and provide practice using the definition of linear dependence. Let V and W be vector spaces, let $T : V \to W$ be a linear transformation, and let $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a subset of V.

31. Show that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly dependent in V, then the set of images, $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$, is linearly dependent in W. This fact shows that if a linear transformation maps a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ onto a linearly *independent* set $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$, then the original set is linearly independent, too (because it cannot be linearly dependent).

32. Suppose that T is a one-to-one transformation, so that an equation $T(\mathbf{u}) = T(\mathbf{v})$ always implies $\mathbf{u} = \mathbf{v}$. Show that if the set of images $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_p)\}$ is linearly dependent, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly dependent. This fact shows that *a one-to-one linear transformation maps a linearly independent set onto a linearly independent set* (because in this case the set of images cannot be linearly dependent).

33. Consider the polynomials $\mathbf{p}_1(t) = 1 + t^2$ and $\mathbf{p}_2(t) = 1 - t^2$. Is $\{\mathbf{p}_1, \mathbf{p}_2\}$ a linearly independent set in $\mathbb{P}_3$? Why or why not?

34. Consider the polynomials $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, and $\mathbf{p}_3(t) = 2$ (for all t). By inspection, write a linear dependence relation among $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$. Then find a basis for Span $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$.

35. Let V be a vector space that contains a linearly independent set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$. Describe how to construct a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in V such that $\{\mathbf{v}_1, \mathbf{v}_3\}$ is a basis for Span $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

36. [**M**] Let $H = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $K = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 3 \\ 4 \\ 1 \\ -4 \end{bmatrix},$$

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ -2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 2 \\ -6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 6 \\ -2 \end{bmatrix}$$

Find bases for H, K, and $H + K$. (See Exercises 33 and 34 in Section 4.1.)

37. [**M**] Show that $\{t, \sin t, \cos 2t, \sin t \cos t\}$ is a linearly independent set of functions defined on $\mathbb{R}$. Start by assuming that

$$c_1 \cdot t + c_2 \cdot \sin t + c_3 \cdot \cos 2t + c_4 \cdot \sin t \cos t = 0 \qquad (5)$$

Equation (5) must hold for all real t, so choose several specific values of t (say, $t = 0, .1, .2$) until you get a system of enough equations to determine that all the c_j must be zero.

38. [**M**] Show that $\{1, \cos t, \cos^2 t, \ldots, \cos^6 t\}$ is a linearly independent set of functions defined on $\mathbb{R}$. Use the method of Exercise 37. (This result will be needed in Exercise 34 in Section 4.5.)

WEB

SOLUTIONS TO PRACTICE PROBLEMS

1. Let $A = [\,\mathbf{v}_1 \;\; \mathbf{v}_2\,]$. Row operations show that

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

Not every row of A contains a pivot position. So the columns of A do not span $\mathbb{R}^3$, by Theorem 4 in Section 1.4. Hence $\{\mathbf{v}_1, \mathbf{v}_2\}$ is not a basis for $\mathbb{R}^3$. Since $\mathbf{v}_1$ and $\mathbf{v}_2$ are not in $\mathbb{R}^2$, they cannot possibly be a basis for $\mathbb{R}^2$. However, since $\mathbf{v}_1$ and $\mathbf{v}_2$ are obviously linearly independent, they are a basis for a subspace of $\mathbb{R}^3$, namely, Span $\{\mathbf{v}_1, \mathbf{v}_2\}$.

2. Set up a matrix A whose column space is the space spanned by $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, and then row reduce A to find its pivot columns.

$$A = \begin{bmatrix} 1 & 6 & 2 & -4 \\ -3 & 2 & -2 & -8 \\ 4 & -1 & 3 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 20 & 4 & -20 \\ 0 & -25 & -5 & 25 \end{bmatrix} \sim \begin{bmatrix} 1 & 6 & 2 & -4 \\ 0 & 5 & 1 & -5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first two columns of A are the pivot columns and hence form a basis of Col $A = W$. Hence $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for W. Note that the reduced echelon form of A is not needed in order to locate the pivot columns.

3. Neither $\mathbf{v}_1$ nor $\mathbf{v}_2$ is in H, so $\{\mathbf{v}_1, \mathbf{v}_2\}$ cannot be a basis for H. In fact, $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for the *plane* of all vectors of the form $(c_1, c_2, 0)$, but H is only a *line*.

4.4 COORDINATE SYSTEMS

An important reason for specifying a basis $\mathcal{B}$ for a vector space V is to impose a "coordinate system" on V. This section will show that if $\mathcal{B}$ contains n vectors, then the coordinate system will make V act like $\mathbb{R}^n$. If V is already $\mathbb{R}^n$ itself, then $\mathcal{B}$ will determine a coordinate system that gives a new "view" of V.

The existence of coordinate systems rests on the following fundamental result.

THEOREM 7

The Unique Representation Theorem

Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis for a vector space V. Then for each $\mathbf{x}$ in V, there exists a unique set of scalars $c_1, \dots, c_n$ such that

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n \tag{1}$$

PROOF Since $\mathcal{B}$ spans V, there exist scalars such that (1) holds. Suppose $\mathbf{x}$ also has the representation

$$\mathbf{x} = d_1\mathbf{b}_1 + \cdots + d_n\mathbf{b}_n$$

for scalars $d_1, \dots, d_n$. Then, subtracting, we have

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = (c_1 - d_1)\mathbf{b}_1 + \cdots + (c_n - d_n)\mathbf{b}_n \tag{2}$$

Since $\mathcal{B}$ is linearly independent, the weights in (2) must all be zero. That is, $c_j = d_j$ for $1 \le j \le n$. ■

DEFINITION

Suppose $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for V and $\mathbf{x}$ is in V. The **coordinates of x relative to the basis $\mathcal{B}$** (or the **$\mathcal{B}$-coordinates of x**) are the weights $c_1, \dots, c_n$ such that $\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n$.

If $c_1, \dots, c_n$ are the $\mathcal{B}$-coordinates of $\mathbf{x}$, then the vector in $\mathbb{R}^n$

$$[\,\mathbf{x}\,]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

is the **coordinate vector of x (relative to $\mathcal{B}$)**, or the **$\mathcal{B}$-coordinate vector of x**. The mapping $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$ is the **coordinate mapping (determined by $\mathcal{B}$)**.[1]

[1] The concept of a coordinate mapping assumes that the basis $\mathcal{B}$ is an indexed set whose vectors are listed in some fixed preassigned order. This property makes the definition of $[\,\mathbf{x}\,]_{\mathcal{B}}$ unambiguous.

Thus $c_1 = 2$, $c_2 = 3$, and $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. The coordinate system on H determined by $\mathcal{B}$ is shown in Fig. 7. ■

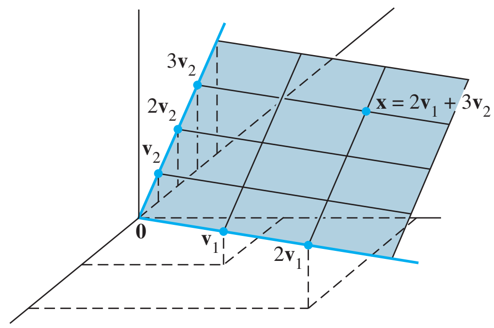


FIGURE 7 A coordinate system on a plane H in $\mathbb{R}^3$.

If a different basis for H were chosen, would the associated coordinate system also make H isomorphic to $\mathbb{R}^2$? Surely, this must be true. We shall prove it in the next section.

PRACTICE PROBLEMS

1. Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -6 \\ 3 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -8 \\ 2 \\ 3 \end{bmatrix}$.

a. Show that the set $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is a basis of $\mathbb{R}^3$.

b. Find the change-of-coordinates matrix from $\mathcal{B}$ to the standard basis.

c. Write the equation that relates $\mathbf{x}$ in $\mathbb{R}^3$ to $[\mathbf{x}]_{\mathcal{B}}$.

d. Find $[\mathbf{x}]_{\mathcal{B}}$, for the $\mathbf{x}$ given above.

2. The set $\mathcal{B} = \{1 + t, 1 + t^2, t + t^2\}$ is a basis for $\mathbb{P}_2$. Find the coordinate vector of $\mathbf{p}(t) = 6 + 3t - t^2$ relative to $\mathcal{B}$.

4.4 EXERCISES

In Exercises 1–4, find the vector $\mathbf{x}$ determined by the given coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ and the given basis $\mathcal{B}$.

1. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ -5 \end{bmatrix}, \begin{bmatrix} -4 \\ 6 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

2. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$

3. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -3 \\ 0 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$

4. $\mathcal{B} = \left\{ \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} \right\}$, $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$

In Exercises 5–8, find the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ of $\mathbf{x}$ relative to the given basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$.

5. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

6. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} -1 \\ -6 \end{bmatrix}$

7. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -3 \\ 4 \\ 9 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 8 \\ -9 \\ 6 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 8 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ -2 \end{bmatrix}$

In Exercises 9 and 10, find the change-of-coordinates matrix from $\mathcal{B}$ to the standard basis in $\mathbb{R}^n$.

9. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \end{bmatrix} \right\}$

10. $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} \right\}$

In Exercises 11 and 12, use an inverse matrix to find $[\,\mathbf{x}\,]_{\mathcal{B}}$ for the given $\mathbf{x}$ and $\mathcal{B}$.

11. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -3 \\ 5 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}$

12. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}, \mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

13. The set $\mathcal{B} = \{1 + t^2, t + t^2, 1 + 2t + t^2\}$ is a basis for $\mathbb{P}_2$. Find the coordinate vector of $\mathbf{p}(t) = 1 + 4t + 7t^2$ relative to $\mathcal{B}$.

14. The set $\mathcal{B} = \{1 - t^2, t - t^2, 2 - t + t^2\}$ is a basis for $\mathbb{P}_2$. Find the coordinate vector of $\mathbf{p}(t) = 1 + 3t - 6t^2$ relative to $\mathcal{B}$.

In Exercises 15 and 16, mark each statement True or False. Justify each answer. Unless stated otherwise, $\mathcal{B}$ is a basis for a vector space V.

15. a. If $\mathbf{x}$ is in V and if $\mathcal{B}$ contains n vectors, then the $\mathcal{B}$-coordinate vector of $\mathbf{x}$ is in $\mathbb{R}^n$.

b. If $P_{\mathcal{B}}$ is the change-of-coordinates matrix, then $[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}\mathbf{x}$, for $\mathbf{x}$ in V.

c. The vector spaces $\mathbb{P}_3$ and $\mathbb{R}^3$ are isomorphic.

16. a. If $\mathcal{B}$ is the standard basis for $\mathbb{R}^n$, then the $\mathcal{B}$-coordinate vector of an $\mathbf{x}$ in $\mathbb{R}^n$ is $\mathbf{x}$ itself.

b. The correspondence $[\,\mathbf{x}\,]_{\mathcal{B}} \mapsto \mathbf{x}$ is called the coordinate mapping.

c. In some cases, a plane in $\mathbb{R}^3$ can be isomorphic to $\mathbb{R}^2$.

17. The vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -8 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -3 \\ 7 \end{bmatrix}$ span $\mathbb{R}^2$ but do not form a basis. Find two different ways to express $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

18. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis for a vector space V. Explain why the $\mathcal{B}$-coordinate vectors of $\mathbf{b}_1, \dots, \mathbf{b}_n$ are the columns $\mathbf{e}_1, \dots, \mathbf{e}_n$ of the $n \times n$ identity matrix.

19. Let S be a finite set in a vector space V with the property that every $\mathbf{x}$ in V has a unique representation as a linear combination of elements of S. Show that S is a basis of V.

20. Suppose $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is a linearly dependent spanning set for a vector space V. Show that each $\mathbf{w}$ in V can be expressed in more than one way as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_4$. [*Hint:* Let $\mathbf{w} = k_1\mathbf{v}_1 + \cdots + k_4\mathbf{v}_4$ be an arbitrary vector in V. Use the linear dependence of $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ to produce another representation of $\mathbf{w}$ as a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_4$.]

21. Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \begin{bmatrix} -2 \\ 9 \end{bmatrix} \right\}$. Since the coordinate mapping determined by $\mathcal{B}$ is a linear transformation from $\mathbb{R}^2$ into $\mathbb{R}^2$, this mapping must be implemented by some 2×2 matrix A. Find it. [*Hint:* Multiplication by A should transform a vector $\mathbf{x}$ into its coordinate vector $[\,\mathbf{x}\,]_{\mathcal{B}}$.]

22. Let $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ be a basis for $\mathbb{R}^n$. Produce a description of an $n \times n$ matrix A that implements the coordinate mapping $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$. (See Exercise 21.)

Exercises 23–26 concern a vector space V, a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, and the coordinate mapping $\mathbf{x} \mapsto [\,\mathbf{x}\,]_{\mathcal{B}}$.

23. Show that the coordinate mapping is one-to-one. (*Hint:* Suppose $[\,\mathbf{u}\,]_{\mathcal{B}} = [\,\mathbf{w}\,]_{\mathcal{B}}$ for some $\mathbf{u}$ and $\mathbf{w}$ in V, and show that $\mathbf{u} = \mathbf{w}$.)

24. Show that the coordinate mapping is *onto* $\mathbb{R}^n$. That is, given any $\mathbf{y}$ in $\mathbb{R}^n$, with entries $y_1, \dots, y_n$, produce $\mathbf{u}$ in V such that $[\,\mathbf{u}\,]_{\mathcal{B}} = \mathbf{y}$.

25. Show that a subset $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in V is linearly independent if and only if the set of coordinate vectors $\{[\,\mathbf{u}_1\,]_{\mathcal{B}}, \dots, [\,\mathbf{u}_p\,]_{\mathcal{B}}\}$ is linearly independent in $\mathbb{R}^n$. *Hint:* Since the coordinate mapping is one-to-one, the following equations have the same solutions, $c_1, \dots, c_p$.

$$c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = \mathbf{0} \qquad \text{The zero vector in } V$$
$$[\,c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p\,]_{\mathcal{B}} = [\,\mathbf{0}\,]_{\mathcal{B}} \qquad \text{The zero vector in } \mathbb{R}^n$$

26. Given vectors $\mathbf{u}_1, \dots, \mathbf{u}_p$, and $\mathbf{w}$ in V, show that $\mathbf{w}$ is a linear combination of $\mathbf{u}_1, \dots, \mathbf{u}_p$ if and only if $[\,\mathbf{w}\,]_{\mathcal{B}}$ is a linear combination of the coordinate vectors $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \dots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$.

In Exercises 27–30, use coordinate vectors to test the linear independence of the sets of polynomials. Explain your work.

27. $1 + 2t^3,\ 2 + t - 3t^2,\ -t + 2t^2 - t^3$

28. $1 - 2t^2 - t^3,\ t + 2t^3,\ 1 + t - 2t^2$

29. $(1 - t)^2,\ t - 2t^2 + t^3,\ (1 - t)^3$

30. $(2 - t)^3,\ (3 - t)^2,\ 1 + 6t - 5t^2 + t^3$

31. Use coordinate vectors to test whether the following sets of polynomials span $\mathbb{P}_2$. Justify your conclusions.

a. $1 - 3t + 5t^2, -3 + 5t - 7t^2, -4 + 5t - 6t^2, 1 - t^2$

b. $5t + t^2, 1 - 8t - 2t^2, -3 + 4t + 2t^2, 2 - 3t$

32. Let $\mathbf{p}_1(t) = 1 + t^2, \mathbf{p}_2(t) = t - 3t^2, \mathbf{p}_3(t) = 1 + t - 3t^2$.

a. Use coordinate vectors to show that these polynomials form a basis for $\mathbb{P}_2$.

b. Consider the basis $\mathcal{B} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ for $\mathbb{P}_2$. Find $\mathbf{q}$ in $\mathbb{P}_2$, given that $[\mathbf{q}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$.

In Exercises 33 and 34, determine whether the sets of polynomials form a basis for $\mathbb{P}_3$. Justify your conclusions.

33. [**M**] $3+7t, 5+t-2t^3, t-2t^2, 1+16t-6t^2+2t^3$

34. [**M**] $5-3t+4t^2+2t^3, 9+t+8t^2-6t^3, 6-2t+5t^2, t^3$

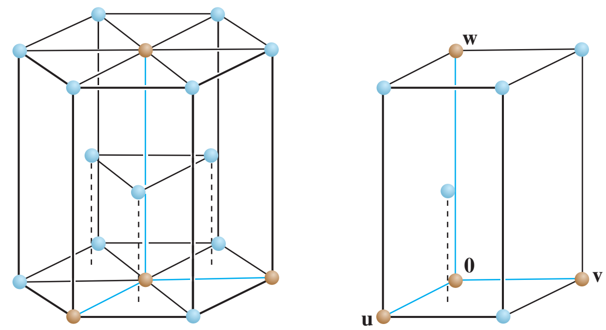
35. [**M**] Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. Show that $\mathbf{x}$ is in H and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, for

$$\mathbf{v}_1 = \begin{bmatrix} 11 \\ -5 \\ 10 \\ 7 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 14 \\ -8 \\ 13 \\ 10 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 19 \\ -13 \\ 18 \\ 15 \end{bmatrix}$$

36. [**M**] Let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. Show that $\mathcal{B}$ is a basis for H and $\mathbf{x}$ is in H, and find the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, for

$$\mathbf{v}_1 = \begin{bmatrix} -6 \\ 4 \\ -9 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 8 \\ -3 \\ 7 \\ -3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -9 \\ 5 \\ -8 \\ 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 4 \\ 7 \\ -8 \\ 3 \end{bmatrix}$$

[**M**] Exercises 37 and 38 concern the crystal lattice for titanium, which has the hexagonal structure shown on the left in the accompanying figure. The vectors $\begin{bmatrix} 2.6 \\ -1.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4.8 \end{bmatrix}$ in $\mathbb{R}^3$ form a basis for the unit cell shown on the right. The numbers here are Ångstrom units (1 Å $= 10^{-8}$ cm). In alloys of titanium, some additional atoms may be in the unit cell at the *octahedral* and *tetrahedral* sites (so named because of the geometric objects formed by atoms at these locations).



The hexagonal close-packed lattice and its unit cell.

37. One of the octahedral sites is $\begin{bmatrix} 1/2 \\ 1/4 \\ 1/6 \end{bmatrix}$, relative to the lattice basis. Determine the coordinates of this site relative to the standard basis of $\mathbb{R}^3$.

38. One of the tetrahedral sites is $\begin{bmatrix} 1/2 \\ 1/2 \\ 1/3 \end{bmatrix}$. Determine the coordinates of this site relative to the standard basis of $\mathbb{R}^3$.

SOLUTIONS TO PRACTICE PROBLEMS

1. a. It is evident that the matrix $P_{\mathcal{B}} = [\,\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3\,]$ is row-equivalent to the identity matrix. By the Invertible Matrix Theorem, $P_{\mathcal{B}}$ is invertible and its columns form a basis for $\mathbb{R}^3$.

b. From part (a), the change-of-coordinates matrix is $P_{\mathcal{B}} = \begin{bmatrix} 1 & -3 & 3 \\ 0 & 4 & -6 \\ 0 & 0 & 3 \end{bmatrix}$.

c. $\mathbf{x} = P_{\mathcal{B}}[\,\mathbf{x}\,]_{\mathcal{B}}$

d. To solve the equation in (c), it is probably easier to row reduce an augmented matrix than to compute $P_{\mathcal{B}}^{-1}$:

$$\underset{P_{\mathcal{B}} \qquad\qquad \mathbf{x}}{\begin{bmatrix} 1 & -3 & 3 & -8 \\ 0 & 4 & -6 & 2 \\ 0 & 0 & 3 & 3 \end{bmatrix}} \sim \underset{I \qquad\qquad [\,\mathbf{x}\,]_{\mathcal{B}}}{\begin{bmatrix} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}}$$

Hence

$$[\,\mathbf{x}_{\mathcal{B}}\,] = \begin{bmatrix} -5 \\ 2 \\ 1 \end{bmatrix}$$

2. The coordinates of $\mathbf{p}(t) = 6 + 3t - t^2$ with respect to $\mathcal{B}$ satisfy

$$c_1(1+t) + c_2(1+t^2) + c_3(t+t^2) = 6 + 3t - t^2$$

Equating coefficients of like powers of t, we have

$$\begin{aligned} c_1 + c_2 \qquad &= \ \ 6 \\ c_1 \qquad + c_3 &= \ \ 3 \\ c_2 + c_3 &= -1 \end{aligned}$$

Solving, we find that $c_1 = 5$, $c_2 = 1$, $c_3 = -2$, and $[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$.

4.5 THE DIMENSION OF A VECTOR SPACE

Theorem 8 in Section 4.4 implies that a vector space V with a basis $\mathcal{B}$ containing n vectors is isomorphic to $\mathbb{R}^n$. This section shows that this number n is an intrinsic property (called the dimension) of the space V that does not depend on the particular choice of basis. The discussion of dimension will give additional insight into properties of bases.

The first theorem generalizes a well-known result about the vector space $\mathbb{R}^n$.

THEOREM 9

If a vector space V has a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, then any set in V containing more than n vectors must be linearly dependent.

PROOF Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be a set in V with more than n vectors. The coordinate vectors $[\,\mathbf{u}_1\,]_{\mathcal{B}}, \dots, [\,\mathbf{u}_p\,]_{\mathcal{B}}$ form a linearly dependent set in $\mathbb{R}^n$, because there are more vectors (p) than entries (n) in each vector. So there exist scalars $c_1, \dots, c_p$, not all zero, such that

$$c_1[\,\mathbf{u}_1\,]_{\mathcal{B}} + \cdots + c_p[\,\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{The zero vector in } \mathbb{R}^n$$

Since the coordinate mapping is a linear transformation,

$$[\,c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p\,]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

The zero vector on the right displays the n weights needed to build the vector $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ from the basis vectors in $\mathcal{B}$. That is, $c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p = 0 \cdot \mathbf{b}_1 + \cdots + 0 \cdot \mathbf{b}_n = \mathbf{0}$. Since the c_i are not all zero, $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is linearly dependent.[1] ■

Theorem 9 implies that if a vector space V has a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, then each linearly independent set in V has no more than n vectors.

[1] Theorem 9 also applies to infinite sets in V. An infinite set is said to be linearly dependent if some finite subset is linearly dependent; otherwise, the set is linearly independent. If S is an infinite set in V, take any subset $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ of S, with $p > n$. The proof above shows that this subset is linearly dependent, and hence so is S.

PROOF By Theorem 11, a linearly independent set S of p elements can be extended to a basis for V. But that basis must contain exactly p elements, since $\dim V = p$. So S must already be a basis for V. Now suppose that S has p elements and spans V. Since V is nonzero, the Spanning Set Theorem implies that a subset S' of S is a basis of V. Since $\dim V = p$, S' must contain p vectors. Hence $S = S'$. ■

The Dimensions of Nul A and Col A

Since the pivot columns of a matrix A form a basis for Col A, we know the dimension of Col A as soon as we know the pivot columns. The dimension of Nul A might seem to require more work, since finding a basis for Nul A usually takes more time than a basis for Col A. But there is a shortcut!

Let A be an $m \times n$ matrix, and suppose the equation $A\mathbf{x} = \mathbf{0}$ has k free variables. From Section 4.2, we know that the standard method of finding a spanning set for Nul A will produce exactly k linearly independent vectors—say, $\mathbf{u}_1, \dots, \mathbf{u}_k$—one for each free variable. So $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ is a basis for Nul A, and the number of free variables determines the size of the basis. Let us summarize these facts for future reference.

The dimension of Nul A is the number of free variables in the equation $A\mathbf{x} = \mathbf{0}$, and the dimension of Col A is the number of pivot columns in A.

EXAMPLE 5 Find the dimensions of the null space and the column space of

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

SOLUTION Row reduce the augmented matrix $[\, A \quad \mathbf{0} \,]$ to echelon form:

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

There are three free variables—x_2, x_4, and x_5. Hence the dimension of Nul A is 3. Also, $\dim \text{Col}\, A = 2$ because A has two pivot columns. ■

PRACTICE PROBLEMS

Decide whether each statement is True or False, and give a reason for each answer. Here V is a nonzero finite-dimensional vector space.

1. If $\dim V = p$ and if S is a linearly dependent subset of V, then S contains more than p vectors.
2. If S spans V and if T is a subset of V that contains more vectors than S, then T is linearly dependent.

4.5 EXERCISES

For each subspace in Exercises 1–8, (a) find a basis for the subspace, and (b) state the dimension.

1. $\left\{ \begin{bmatrix} s-2t \\ s+t \\ 3t \end{bmatrix} : s, t \text{ in } \mathbb{R} \right\}$

2. $\left\{ \begin{bmatrix} 2a \\ -4b \\ -2a \end{bmatrix} : a, b \text{ in } \mathbb{R} \right\}$

3. $\left\{ \begin{bmatrix} 2c \\ a-b \\ b-3c \\ a+2b \end{bmatrix} : a, b, c \text{ in } \mathbb{R} \right\}$

4. $\left\{ \begin{bmatrix} p+2q \\ -p \\ 3p-q \\ p+q \end{bmatrix} : p, q \text{ in } \mathbb{R} \right\}$

5. $\left\{ \begin{bmatrix} p-2q \\ 2p+5r \\ -2q+2r \\ -3p+6r \end{bmatrix} : p, q, r \text{ in } \mathbb{R} \right\}$

6. $\left\{ \begin{bmatrix} 3a-c \\ -b-3c \\ -7a+6b+5c \\ -3a+c \end{bmatrix} : a, b, c \text{ in } \mathbb{R} \right\}$

7. $\{(a,b,c) : a-3b+c=0, b-2c=0, 2b-c=0\}$

8. $\{(a,b,c,d) : a-3b+c=0\}$

9. Find the dimension of the subspace of all vectors in $\mathbb{R}^3$ whose first and third entries are equal.

10. Find the dimension of the subspace H of $\mathbb{R}^2$ spanned by $\begin{bmatrix} 1 \\ -5 \end{bmatrix}, \begin{bmatrix} -2 \\ 10 \end{bmatrix}, \begin{bmatrix} -3 \\ 15 \end{bmatrix}$.

In Exercises 11 and 12, find the dimension of the subspace spanned by the given vectors.

11. $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$

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

Determine the dimensions of Nul A and Col A for the matrices shown in Exercises 13–18.

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

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

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

16. $A = \begin{bmatrix} 3 & 2 \\ -6 & 5 \end{bmatrix}$

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

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

In Exercises 19 and 20, V is a vector space. Mark each statement True or False. Justify each answer.

19. a. The number of pivot columns of a matrix equals the dimension of its column space.
    b. A plane in $\mathbb{R}^3$ is a two-dimensional subspace of $\mathbb{R}^3$.
    c. The dimension of the vector space $\mathbb{P}_4$ is 4.
    d. If $\dim V = n$ and S is a linearly independent set in V, then S is a basis for V.
    e. If a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ spans a finite-dimensional vector space V and if T is a set of more than p vectors in V, then T is linearly dependent.

20. a. $\mathbb{R}^2$ is a two-dimensional subspace of $\mathbb{R}^3$.
    b. The number of variables in the equation $A\mathbf{x} = \mathbf{0}$ equals the dimension of Nul A.
    c. A vector space is infinite-dimensional if it is spanned by an infinite set.
    d. If $\dim V = n$ and if S spans V, then S is a basis of V.
    e. The only three-dimensional subspace of $\mathbb{R}^3$ is $\mathbb{R}^3$ itself.

21. The first four Hermite polynomials are $1, 2t, -2+4t^2$, and $-12t+8t^3$. These polynomials arise naturally in the study of certain important differential equations in mathematical physics.[2] Show that the first four Hermite polynomials form a basis of $\mathbb{P}_3$.

22. The first four Laguerre polynomials are $1, 1-t, 2-4t+t^2$, and $6-18t+9t^2-t^3$. Show that these polynomials form a basis of $\mathbb{P}_3$.

23. Let $\mathcal{B}$ be the basis of $\mathbb{P}_3$ consisting of the Hermite polynomials in Exercise 21, and let $\mathbf{p}(t) = -1+8t^2+8t^3$. Find the coordinate vector of $\mathbf{p}$ relative to $\mathcal{B}$.

24. Let $\mathcal{B}$ be the basis of $\mathbb{P}_2$ consisting of the first three Laguerre polynomials listed in Exercise 22, and let $\mathbf{p}(t) = 5+5t-2t^2$. Find the coordinate vector of $\mathbf{p}$ relative to $\mathcal{B}$.

25. Let S be a subset of an n-dimensional vector space V, and suppose S contains fewer than n vectors. Explain why S cannot span V.

26. Let H be an n-dimensional subspace of an n-dimensional vector space V. Show that $H = V$.

27. Explain why the space $\mathbb{P}$ of all polynomials is an infinite-dimensional space.

[2] See *Introduction to Functional Analysis*, 2d ed., by A. E. Taylor and David C. Lay (New York: John Wiley & Sons, 1980), pp. 92–93. Other sets of polynomials are discussed there, too.

28. Show that the space $C(\mathbb{R})$ of all continuous functions defined on the real line is an infinite-dimensional space.

In Exercises 29 and 30, V is a nonzero finite-dimensional vector space, and the vectors listed belong to V. Mark each statement True or False. Justify each answer. (These questions are more difficult than those in Exercises 19 and 20.)

29. a. If there exists a set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ that spans V, then $\dim V \le p$.

b. If there exists a linearly independent set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in V, then $\dim V \ge p$.

c. If $\dim V = p$, then there exists a spanning set of $p + 1$ vectors in V.

30. a. If there exists a linearly dependent set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in V, then $\dim V \le p$.

b. If every set of p elements in V fails to span V, then $\dim V > p$.

c. If $p \ge 2$ and $\dim V = p$, then every set of $p - 1$ nonzero vectors is linearly independent.

Exercises 31 and 32 concern finite-dimensional vector spaces V and W and a linear transformation $T : V \to W$.

31. Let H be a nonzero subspace of V, and let $T(H)$ be the set of images of vectors in H. Then $T(H)$ is a subspace of W, by Exercise 35 in Section 4.2. Prove that $\dim T(H) \le \dim H$.

32. Let H be a nonzero subspace of V, and suppose T is a one-to-one (linear) mapping of V into W. Prove that $\dim T(H) = \dim H$. If T happens to be a one-to-one mapping of V *onto* W, then $\dim V = \dim W$. Isomorphic finite-dimensional vector spaces have the same dimension.

33. **[M]** According to Theorem 11, a linearly independent set $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ in $\mathbb{R}^n$ can be expanded to a basis for $\mathbb{R}^n$. One way to do this is to create $A = [\,\mathbf{v}_1 \ \cdots \ \mathbf{v}_k \ \mathbf{e}_1 \ \cdots \ \mathbf{e}_n\,]$, with $\mathbf{e}_1, \dots, \mathbf{e}_n$ the columns of the identity matrix; the pivot columns of A form a basis for $\mathbb{R}^n$.

a. Use the method described to extend the following vectors to a basis for $\mathbb{R}^5$:

$$\mathbf{v}_1 = \begin{bmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{bmatrix}$$

b. Explain why the method works in general: Why are the original vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ included in the basis found for Col A? Why is Col $A = \mathbb{R}^n$?

34. **[M]** Let $\mathcal{B} = \{1, \cos t, \cos^2 t, \dots, \cos^6 t\}$ and $\mathcal{C} = \{1, \cos t, \cos 2t, \dots, \cos 6t\}$. Assume the following trigonometric identities (see Exercise 37 in Section 4.1).

$$\begin{aligned} \cos 2t &= -1 + 2\cos^2 t \\ \cos 3t &= -3\cos t + 4\cos^3 t \\ \cos 4t &= 1 - 8\cos^2 t + 8\cos^4 t \\ \cos 5t &= 5\cos t - 20\cos^3 t + 16\cos^5 t \\ \cos 6t &= -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t \end{aligned}$$

Let H be the subspace of functions spanned by the functions in $\mathcal{B}$. Then $\mathcal{B}$ is a basis for H, by Exercise 38 in Section 4.3.

a. Write the $\mathcal{B}$-coordinate vectors of the vectors in $\mathcal{C}$, and use them to show that $\mathcal{C}$ is a linearly independent set in H.

b. Explain why $\mathcal{C}$ is a basis for H.

SOLUTIONS TO PRACTICE PROBLEMS

1. False. Consider the set $\{\mathbf{0}\}$.

2. True. By the Spanning Set Theorem, S contains a basis for V; call that basis S'. Then T will contain more vectors than S'. By Theorem 9, T is linearly dependent.

4.6 RANK

With the aid of vector space concepts, this section takes a look *inside* a matrix and reveals several interesting and useful relationships hidden in its rows and columns.

For instance, imagine placing 2000 random numbers into a 40×50 matrix A and then determining both the maximum number of linearly independent columns in A and the maximum number of linearly independent columns in A^T (rows in A). Remarkably, the two numbers are the same. As we'll soon see, their common value is the *rank* of the matrix. To explain why, we need to examine the subspace spanned by the rows of A.

NUMERICAL NOTE

Many algorithms discussed in this text are useful for understanding concepts and making simple computations by hand. However, the algorithms are often unsuitable for large-scale problems in real life.

Rank determination is a good example. It would seem easy to reduce a matrix to echelon form and count the pivots. But unless exact arithmetic is performed on a matrix whose entries are specified exactly, row operations can change the apparent rank of a matrix. For instance, if the value of x in the matrix $\begin{bmatrix} 5 & 7 \\ 5 & x \end{bmatrix}$ is not stored exactly as 7 in a computer, then the rank may be 1 or 2, depending on whether the computer treats $x - 7$ as zero.

In practical applications, the effective rank of a matrix A is often determined from the singular value decomposition of A, to be discussed in Section 7.4. This decomposition is also a reliable source of bases for Col A, Row A, Nul A, and Nul A^T.

WEB

PRACTICE PROBLEMS

The matrices below are row equivalent.

$$A = \begin{bmatrix} 2 & -1 & 1 & -6 & 8 \\ 1 & -2 & -4 & 3 & -2 \\ -7 & 8 & 10 & 3 & -10 \\ 4 & -5 & -7 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -2 & -4 & 3 & -2 \\ 0 & 3 & 9 & -12 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

1. Find rank A and dim Nul A.
2. Find bases for Col A and Row A.
3. What is the next step to perform to find a basis for Nul A?
4. How many pivot columns are in a row echelon form of A^T?

4.6 EXERCISES

In Exercises 1–4, assume that the matrix A is row equivalent to B. Without calculations, list rank A and dim Nul A. Then find bases for Col A, Row A, and Nul A.

1. $A = \begin{bmatrix} 1 & -4 & 9 & -7 \\ -1 & 2 & -4 & 1 \\ 5 & -6 & 10 & 7 \end{bmatrix}$,

$B = \begin{bmatrix} 1 & 0 & -1 & 5 \\ 0 & -2 & 5 & -6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

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

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

3. $A = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ -2 & -3 & 6 & -3 & 0 & -6 \\ 4 & 9 & -12 & 9 & 3 & 12 \\ -2 & 3 & 6 & 3 & 3 & -6 \end{bmatrix}$,

$B = \begin{bmatrix} 2 & 6 & -6 & 6 & 3 & 6 \\ 0 & 3 & 0 & 3 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

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

$B = \begin{bmatrix} 1 & 1 & -2 & 0 & 1 & -2 \\ 0 & 1 & -1 & 0 & -3 & -1 \\ 0 & 0 & 1 & 1 & -13 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

5. If a 4×7 matrix A has rank 3, find $\dim \text{Nul } A$, $\dim \text{Row } A$, and rank A^T.

6. If a 7×5 matrix A has rank 2, find $\dim \text{Nul } A$, $\dim \text{Row } A$, and rank A^T.

7. Suppose a 4×7 matrix A has four pivot columns. Is $\text{Col } A = \mathbb{R}^4$? Is $\text{Nul } A = \mathbb{R}^3$? Explain your answers.

8. Suppose a 6×8 matrix A has four pivot columns. What is $\dim \text{Nul } A$? Is $\text{Col } A = \mathbb{R}^4$? Why or why not?

9. If the null space of a 4×6 matrix A is 3-dimensional, what is the dimension of the column space of A? Is $\text{Col } A = \mathbb{R}^3$? Why or why not?

10. If the null space of an 8×7 matrix A is 5-dimensional, what is the dimension of the column space of A?

11. If the null space of an 8×5 matrix A is 3-dimensional, what is the dimension of the row space of A?

12. If the null space of a 5×4 matrix A is 2-dimensional, what is the dimension of the row space of A?

13. If A is a 7×5 matrix, what is the largest possible rank of A? If A is a 5×7 matrix, what is the largest possible rank of A? Explain your answers.

14. If A is a 5×4 matrix, what is the largest possible dimension of the row space of A? If A is a 4×5 matrix, what is the largest possible dimension of the row space of A? Explain.

15. If A is a 3×7 matrix, what is the smallest possible dimension of $\text{Nul } A$?

16. If A is a 7×5 matrix, what is the smallest possible dimension of $\text{Nul } A$?

In Exercises 17 and 18, A is an $m \times n$ matrix. Mark each statement True or False. Justify each answer.

17. a. The row space of A is the same as the column space of A^T.

b. If B is any echelon form of A, and if B has three nonzero rows, then the first three rows of A form a basis for Row A.

c. The dimensions of the row space and the column space of A are the same, even if A is not square.

d. The sum of the dimensions of the row space and the null space of A equals the number of rows in A.

e. On a computer, row operations can change the apparent rank of a matrix.

18. a. If B is any echelon form of A, then the pivot columns of B form a basis for the column space of A.

b. Row operations preserve the linear dependence relations among the rows of A.

c. The dimension of the null space of A is the number of columns of A that are *not* pivot columns.

d. The row space of A^T is the same as the column space of A.

e. If A and B are row equivalent, then their row spaces are the same.

19. Suppose the solutions of a homogeneous system of five linear equations in six unknowns are all multiples of one nonzero solution. Will the system necessarily have a solution for every possible choice of constants on the right sides of the equations? Explain.

20. Suppose a nonhomogeneous system of six linear equations in eight unknowns has a solution, with two free variables. Is it possible to change some constants on the equations' right sides to make the new system inconsistent? Explain.

21. Suppose a nonhomogeneous system of nine linear equations in ten unknowns has a solution for all possible constants on the right sides of the equations. Is it possible to find two nonzero solutions of the associated homogeneous system that are *not* multiples of each other? Discuss.

22. Is is possible that all solutions of a homogeneous system of ten linear equations in twelve variables are multiples of one fixed nonzero solution? Discuss.

23. A homogeneous system of twelve linear equations in eight unknowns has two fixed solutions that are not multiples of each other, and all other solutions are linear combinations of these two solutions. Can the set of all solutions be described with fewer than twelve homogeneous linear equations? If so, how many? Discuss.

24. Is it possible for a nonhomogeneous system of seven equations in six unknowns to have a unique solution for some right-hand side of constants? Is it possible for such a system to have a unique solution for every right-hand side? Explain.

25. A scientist solves a nonhomogeneous system of ten linear equations in twelve unknowns and finds that three of the unknowns are free variables. Can the scientist be certain that, if the right sides of the equations are changed, the new nonhomogeneous system will have a solution? Discuss.

26. In statistical theory, a common requirement is that a matrix be of *full rank*. That is, the rank should be as large as possible. Explain why an $m \times n$ matrix with more rows than columns has full rank if and only if its columns are linearly independent.

Exercises 27–29 concern an $m \times n$ matrix A and what are often called the *fundamental subspaces* determined by A.

27. Which of the subspaces Row A, Col A, Nul A, Row A^T, Col A^T, and Nul A^T are in $\mathbb{R}^m$ and which are in $\mathbb{R}^n$? How many distinct subspaces are in this list?

28. Justify the following equalities:

a. $\dim \text{Row } A + \dim \text{Nul } A = n$ Number of columns of A

b. $\dim \text{Col } A + \dim \text{Nul } A^T = m$ Number of rows of A

29. Use Exercise 28 to explain why the equation $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b}$ in $\mathbb{R}^m$ if and only if the equation $A^T\mathbf{x} = \mathbf{0}$ has only the trivial solution.

30. Suppose A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$. What has to be true about the two numbers rank $[\,A \quad \mathbf{b}\,]$ and rank A in order for the equation $A\mathbf{x} = \mathbf{b}$ to be consistent?

Rank 1 matrices are important in some computer algorithms and several theoretical contexts, including the singular value decomposition in Chapter 7. It can be shown that an $m \times n$ matrix A has rank 1 if and only if it is an outer product; that is, $A = \mathbf{u}\mathbf{v}^T$ for some $\mathbf{u}$ in $\mathbb{R}^m$ and $\mathbf{v}$ in $\mathbb{R}^n$. Exercises 31–33 suggest why this property is true.

31. Verify that rank $\mathbf{u}\mathbf{v}^T \le 1$ if $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

32. Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Find $\mathbf{v}$ in $\mathbb{R}^3$ such that $\begin{bmatrix} 1 & -3 & 4 \\ 2 & -6 & 8 \end{bmatrix} = \mathbf{u}\mathbf{v}^T$.

33. Let A be any 2×3 matrix such that rank $A = 1$, let $\mathbf{u}$ be the first column of A, and suppose $\mathbf{u} \neq \mathbf{0}$. Explain why there is a vector $\mathbf{v}$ in $\mathbb{R}^3$ such that $A = \mathbf{u}\mathbf{v}^T$. How could this construction be modified if the first column of A were zero?

34. Let A be an $m \times n$ matrix of rank $r > 0$ and let U be an echelon form of A. Explain why there exists an invertible matrix E such that $A = EU$, and use this factorization to write A as the sum of r rank 1 matrices. [*Hint:* See Theorem 10 in Section 2.4.]

35. **[M]** Let $A = \begin{bmatrix} 7 & -9 & -4 & 5 & 3 & -3 & -7 \\ -4 & 6 & 7 & -2 & -6 & -5 & 5 \\ 5 & -7 & -6 & 5 & -6 & 2 & 8 \\ -3 & 5 & 8 & -1 & -7 & -4 & 8 \\ 6 & -8 & -5 & 4 & 4 & 9 & 3 \end{bmatrix}$.

a. Construct matrices C and N whose columns are bases for Col A and Nul A, respectively, and construct a matrix R whose *rows* form a basis for Row A.

b. Construct a matrix M whose columns form a basis for Nul A^T, form the matrices $S = [\,R^T \quad N\,]$ and $T = [\,C \quad M\,]$, and explain why S and T should be square. Verify that both S and T are invertible.

36. **[M]** Repeat Exercise 35 for a random integer-valued 6×7 matrix A whose rank is at most 4. One way to make A is to create a random integer-valued 6×4 matrix J and a random integer-valued 4×7 matrix K, and set $A = JK$. (See Supplementary Exercise 12 at the end of the chapter; and see the *Study Guide* for matrix-generating programs.)

37. **[M]** Let A be the matrix in Exercise 35. Construct a matrix C whose columns are the pivot columns of A, and construct a matrix R whose rows are the nonzero rows of the reduced echelon form of A. Compute CR, and discuss what you see.

38. **[M]** Repeat Exercise 37 for three random integer-valued 5×7 matrices A whose ranks are 5, 4, and 3. Make a conjecture about how CR is related to A for any matrix A. Prove your conjecture.

SOLUTIONS TO PRACTICE PROBLEMS

1. A has two pivot columns, so rank $A = 2$. Since A has 5 columns altogether, dim Nul $A = 5 - 2 = 3$.

2. The pivot columns of A are the first two columns. So a basis for Col A is

$$\{\mathbf{a}_1, \mathbf{a}_2\} = \left\{ \begin{bmatrix} 2 \\ 1 \\ -7 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 8 \\ -5 \end{bmatrix} \right\}$$

The nonzero rows of B form a basis for Row A, namely, $\{(1, -2, -4, 3, -2), (0, 3, 9, -12, 12)\}$. In this particular example, it happens that any two rows of A form a basis for the row space, because the row space is two-dimensional and none of the rows of A is a multiple of another row. In general, the nonzero rows of an echelon form of A should be used as a basis for Row A, not the rows of A itself.

3. For Nul A, the next step is to perform row operations on B to obtain the reduced echelon form of A.

4. Rank A^T = rank A, by the Rank Theorem, because Col A^T = Row A. So A^T has two pivot positions.

SG Major Review of Key Concepts 4–22

EXAMPLE 3 Let $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} -7 \\ 9 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} -5 \\ 7 \end{bmatrix}$, and consider the bases for $\mathbb{R}^2$ given by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$.

a. Find the change-of-coordinates matrix from $\mathcal{C}$ to $\mathcal{B}$.

b. Find the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.

SOLUTION

a. Notice that $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ is needed rather than $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, and compute

$$\left[\begin{array}{cc:cc} \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{c}_1 & \mathbf{c}_2 \end{array}\right] = \left[\begin{array}{cc:cc} 1 & -2 & -7 & -5 \\ -3 & 4 & 9 & 7 \end{array}\right] \sim \left[\begin{array}{cc:cc} 1 & 0 & 5 & 3 \\ 0 & 1 & 6 & 4 \end{array}\right]$$

So

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 5 & 3 \\ 6 & 4 \end{bmatrix}$$

b. By part (a) and property (6) above (with $\mathcal{B}$ and $\mathcal{C}$ interchanged),

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = (\underset{\mathcal{B}\leftarrow\mathcal{C}}{P})^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -3 \\ -6 & 5 \end{bmatrix} = \begin{bmatrix} 2 & -3/2 \\ -3 & 5/2 \end{bmatrix} \quad \blacksquare$$

Another description of the change-of-coordinates matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ uses the change-of-coordinate matrices $P_{\mathcal{B}}$ and $P_{\mathcal{C}}$ that convert $\mathcal{B}$-coordinates and $\mathcal{C}$-coordinates, respectively, into standard coordinates. Recall that for each $\mathbf{x}$ in $\mathbb{R}^n$,

$$P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}, \quad P_{\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} = \mathbf{x}, \quad \text{and} \quad [\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x}$$

Thus

$$[\mathbf{x}]_{\mathcal{C}} = P_{\mathcal{C}}^{-1}\mathbf{x} = P_{\mathcal{C}}^{-1}P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$$

In $\mathbb{R}^n$, the change-of-coordinates matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ may be computed as $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. Actually, for matrices larger than 2×2, an algorithm analogous to the one in Example 3 is faster than computing $P_{\mathcal{C}}^{-1}$ and then $P_{\mathcal{C}}^{-1}P_{\mathcal{B}}$. See Exercise 12 in Section 2.2.

PRACTICE PROBLEMS

1. Let $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2\}$ and $\mathcal{G} = \{\mathbf{g}_1, \mathbf{g}_2\}$ be bases for a vector space V, and let P be a matrix whose columns are $[\,\mathbf{f}_1\,]_{\mathcal{G}}$ and $[\,\mathbf{f}_2\,]_{\mathcal{G}}$. Which of the following equations is satisfied by P for all $\mathbf{v}$ in V?

   (i) $[\,\mathbf{v}\,]_{\mathcal{F}} = P[\,\mathbf{v}\,]_{\mathcal{G}}$ (ii) $[\,\mathbf{v}\,]_{\mathcal{G}} = P[\,\mathbf{v}\,]_{\mathcal{F}}$

2. Let $\mathcal{B}$ and $\mathcal{C}$ be as in Example 1. Use the results of that example to find the change-of-coordinates matrix from $\mathcal{C}$ to $\mathcal{B}$.

4.7 EXERCISES

1. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for a vector space V, and suppose $\mathbf{b}_1 = 6\mathbf{c}_1 - 2\mathbf{c}_2$ and $\mathbf{b}_2 = 9\mathbf{c}_1 - 4\mathbf{c}_2$.

   a. Find the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.

   b. Find $[\,\mathbf{x}\,]_{\mathcal{C}}$ for $\mathbf{x} = -3\mathbf{b}_1 + 2\mathbf{b}_2$. Use part (a).

2. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for a vector space V, and suppose $\mathbf{b}_1 = -2\mathbf{c}_1 + 4\mathbf{c}_2$ and $\mathbf{b}_2 = 3\mathbf{c}_1 - 6\mathbf{c}_2$.

   a. Find the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$.

   b. Find $[\,\mathbf{x}\,]_{\mathcal{C}}$ for $\mathbf{x} = 2\mathbf{b}_1 + 3\mathbf{b}_2$.

3. Let $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2\}$ be bases for V, and let P be a matrix whose columns are $[\mathbf{u}_1]_{\mathcal{W}}$ and $[\mathbf{u}_2]_{\mathcal{W}}$. Which of the following equations is satisfied by P for all $\mathbf{x}$ in V?

(i) $[\mathbf{x}]_{\mathcal{U}} = P[\mathbf{x}]_{\mathcal{W}}$ (ii) $[\mathbf{x}]_{\mathcal{W}} = P[\mathbf{x}]_{\mathcal{U}}$

4. Let $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ be bases for V, and let $P = [\,[\mathbf{d}_1]_{\mathcal{A}} \;\; [\mathbf{d}_2]_{\mathcal{A}} \;\; [\mathbf{d}_3]_{\mathcal{A}}\,]$. Which of the following equations is satisfied by P for all $\mathbf{x}$ in V?

(i) $[\mathbf{x}]_{\mathcal{A}} = P[\mathbf{x}]_{\mathcal{D}}$ (ii) $[\mathbf{x}]_{\mathcal{D}} = P[\mathbf{x}]_{\mathcal{A}}$

5. Let $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be bases for a vector space V, and suppose $\mathbf{a}_1 = 4\mathbf{b}_1 - \mathbf{b}_2$, $\mathbf{a}_2 = -\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3$, and $\mathbf{a}_3 = \mathbf{b}_2 - 2\mathbf{b}_3$.

a. Find the change-of-coordinates matrix from $\mathcal{A}$ to $\mathcal{B}$.

b. Find $[\mathbf{x}]_{\mathcal{B}}$ for $\mathbf{x} = 3\mathbf{a}_1 + 4\mathbf{a}_2 + \mathbf{a}_3$.

6. Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ and $\mathcal{F} = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ be bases for a vector space V, and suppose $\mathbf{f}_1 = 2\mathbf{d}_1 - \mathbf{d}_2 + \mathbf{d}_3$, $\mathbf{f}_2 = 3\mathbf{d}_2 + \mathbf{d}_3$, and $\mathbf{f}_3 = -3\mathbf{d}_1 + 2\mathbf{d}_3$.

a. Find the change-of-coordinates matrix from $\mathcal{F}$ to $\mathcal{D}$.

b. Find $[\mathbf{x}]_{\mathcal{D}}$ for $\mathbf{x} = \mathbf{f}_1 - 2\mathbf{f}_2 + 2\mathbf{f}_3$.

In Exercises 7–10, let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for $\mathbb{R}^2$. In each exercise, find the change-of-coordinates matrix from $\mathcal{B}$ to $\mathcal{C}$ and the change-of-coordinates matrix from $\mathcal{C}$ to $\mathcal{B}$.

7. $\mathbf{b}_1 = \begin{bmatrix} 7 \\ 5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ -5 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

8. $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 8 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -7 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

9. $\mathbf{b}_1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$

10. $\mathbf{b}_1 = \begin{bmatrix} 6 \\ -12 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_1 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \mathbf{c}_2 = \begin{bmatrix} 3 \\ 9 \end{bmatrix}$

In Exercises 11 and 12, $\mathcal{B}$ and $\mathcal{C}$ are bases for a vector space V. Mark each statement True or False. Justify each answer.

11. a. The columns of the change-of-coordinates matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ are $\mathcal{B}$-coordinate vectors of the vectors in $\mathcal{C}$.

b. If $V = \mathbb{R}^n$ and $\mathcal{C}$ is the *standard* basis for V, then $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ is the same as the change-of-coordinates matrix $P_{\mathcal{B}}$ introduced in Section 4.4.

12. a. The columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ are linearly independent.

b. If $V = \mathbb{R}^2$, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, then row reduction of $[\,\mathbf{c}_1 \;\; \mathbf{c}_2 \;\; \mathbf{b}_1 \;\; \mathbf{b}_2\,]$ to $[\,I \;\; P\,]$ produces a matrix P that satisfies $[\mathbf{x}]_{\mathcal{B}} = P[\mathbf{x}]_{\mathcal{C}}$ for all $\mathbf{x}$ in V.

13. In $\mathbb{P}_2$, find the change-of-coordinates matrix from the basis $\mathcal{B} = \{1 - 2t + t^2, 3 - 5t + 4t^2, 2t + 3t^2\}$ to the standard basis $\mathcal{C} = \{1, t, t^2\}$. Then find the $\mathcal{B}$-coordinate vector for $-1 + 2t$.

14. In $\mathbb{P}_2$, find the change-of-coordinates matrix from the basis $\mathcal{B} = \{1 - 3t^2, 2 + t - 5t^2, 1 + 2t\}$ to the standard basis. Then write t^2 as a linear combination of the polynomials in $\mathcal{B}$.

Exercises 15 and 16 provide a proof of Theorem 15. Fill in a justification for each step.

15. Given $\mathbf{v}$ in V, there exist scalars $x_1, \ldots, x_n$, such that

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

because (a) _____. Apply the coordinate mapping determined by the basis $\mathcal{C}$, and obtain

$$[\mathbf{v}]_{\mathcal{C}} = x_1[\mathbf{b}_1]_{\mathcal{C}} + x_2[\mathbf{b}_2]_{\mathcal{C}} + \cdots + x_n[\mathbf{b}_n]_{\mathcal{C}}$$

because (b) _____. This equation may be written in the form

$$[\mathbf{v}]_{\mathcal{C}} = \big[\,[\mathbf{b}_1]_{\mathcal{C}} \;\; [\mathbf{b}_2]_{\mathcal{C}} \;\; \cdots \;\; [\mathbf{b}_n]_{\mathcal{C}}\,\big] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{8}$$

by the definition of (c) _____. This shows that the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ shown in (5) satisfies $[\mathbf{v}]_{\mathcal{C}} = \underset{\mathcal{C}\leftarrow\mathcal{B}}{P}[\mathbf{v}]_{\mathcal{B}}$ for each $\mathbf{v}$ in V, because the vector on the right side of (8) is (d) _____.

16. Suppose Q is any matrix such that

$$[\mathbf{v}]_{\mathcal{C}} = Q[\mathbf{v}]_{\mathcal{B}} \quad \text{for each } \mathbf{v} \text{ in } V \tag{9}$$

Set $\mathbf{v} = \mathbf{b}_1$ in (9). Then (9) shows that $[\mathbf{b}_1]_{\mathcal{C}}$ is the first column of Q because (a) _____. Similarly, for $k = 2, \ldots, n$, the kth column of Q is (b) _____ because (c) _____. This shows that the matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ defined by (5) in Theorem 15 is the only matrix that satisfies condition (4).

17. **[M]** Let $\mathcal{B} = \{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$ and $C = \{\mathbf{y}_0, \ldots, \mathbf{y}_6\}$, where $\mathbf{x}_k$ is the function $\cos^k t$ and $\mathbf{y}_k$ is the function $\cos kt$. Exercise 34 in Section 4.5 showed that both $\mathcal{B}$ and $\mathcal{C}$ are bases for the vector space $H = \text{Span}\{\mathbf{x}_0, \ldots, \mathbf{x}_6\}$.

a. Set $P = \big[\,[\mathbf{y}_0]_{\mathcal{B}} \;\; \cdots \;\; [\mathbf{y}_6]_{\mathcal{B}}\,\big]$, and calculate P^{-1}.

b. Explain why the columns of P^{-1} are the $\mathcal{C}$-coordinate vectors of $\mathbf{x}_0, \ldots, \mathbf{x}_6$. Then use these coordinate vectors to write trigonometric identities that express powers of $\cos t$ in terms of the functions in $\mathcal{C}$.

See the *Study Guide*.

18. **[M]** (*Calculus required*)[3] Recall from calculus that integrals such as

$$\int (5\cos^3 t - 6\cos^4 t + 5\cos^5 t - 12\cos^6 t)\, dt \tag{10}$$

are tedious to compute. (The usual method is to apply integration by parts repeatedly and use the half-angle formula.) Use the matrix P or P^{-1} from Exercise 17 to transform (10); then compute the integral.

[3] The idea for Exercises 17 and 18 and five related exercises in earlier sections came from a paper by Jack W. Rogers, Jr., of Auburn University, presented at a meeting of the International Linear Algebra Society, August 1995. See "Applications of Linear Algebra in Calculus," *American Mathematical Monthly* **104** (1), 1997.

19. [**M**] Let

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

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -8 \\ 5 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -7 \\ 2 \\ 6 \end{bmatrix}$$

a. Find a basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ for $\mathbb{R}^3$ such that P is the change-of-coordinates matrix from $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. [*Hint:* What do the columns of $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ represent?]

b. Find a basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ for $\mathbb{R}^3$ such that P is the change-of-coordinates matrix from $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ to $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

20. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$, and $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ be bases for a two-dimensional vector space.

a. Write an equation that relates the matrices $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$, $\underset{\mathcal{D}\leftarrow\mathcal{C}}{P}$, and $\underset{\mathcal{D}\leftarrow\mathcal{B}}{P}$. Justify your result.

b. [**M**] Use a matrix program either to help you find the equation or to check the equation you write. Work with three bases for $\mathbb{R}^2$. (See Exercises 7–10.)

SOLUTIONS TO PRACTICE PROBLEMS

1. Since the columns of P are $\mathcal{G}$-coordinate vectors, a vector of the form $P\mathbf{x}$ must be a $\mathcal{G}$-coordinate vector. Thus P satisfies equation (ii).

2. The coordinate vectors found in Example 1 show that

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{P} = \begin{bmatrix} [\mathbf{b}_1]_\mathcal{C} & [\mathbf{b}_2]_\mathcal{C} \end{bmatrix} = \begin{bmatrix} 4 & -6 \\ 1 & 1 \end{bmatrix}$$

Hence

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = (\underset{\mathcal{C}\leftarrow\mathcal{B}}{P})^{-1} = \frac{1}{10}\begin{bmatrix} 1 & 6 \\ -1 & 4 \end{bmatrix} = \begin{bmatrix} .1 & .6 \\ -.1 & .4 \end{bmatrix}$$

4.8 APPLICATIONS TO DIFFERENCE EQUATIONS

Now that powerful computers are widely available, more and more scientific and engineering problems are being treated in a way that uses discrete, or digital, data rather than continuous data. Difference equations are often the appropriate tool to analyze such data. Even when a differential equation is used to model a continuous process, a numerical solution is often produced from a related difference equation.

This section highlights some fundamental properties of linear difference equations that are best explained using linear algebra.

Discrete-Time Signals

The vector space $\mathbb{S}$ of discrete-time signals was introduced in Section 4.1. A **signal** in $\mathbb{S}$ is a function defined only on the integers and is visualized as a sequence of numbers, say, $\{y_k\}$. Figure 1 shows three typical signals whose general terms are $(.7)^k$, 1^k, and $(-1)^k$, respectively.

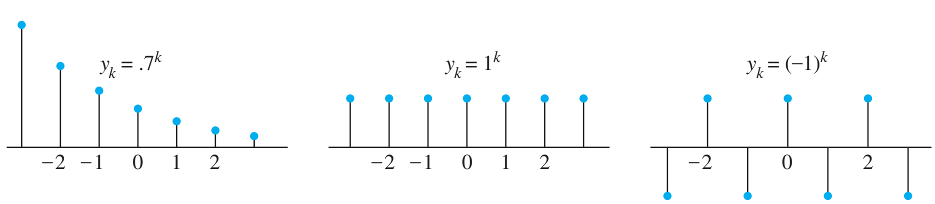


FIGURE 1 Three signals in $\mathbb{S}$.

Further Reading

Hamming, R. W., *Digital Filters*, 3rd ed. (Englewood Cliffs, NJ: Prentice-Hall, 1989), pp. 1–37.

Kelly, W. G., and A. C. Peterson, *Difference Equations*, 2nd ed. (San Diego: Harcourt-Academic Press, 2001).

Mickens, R. E., *Difference Equations*, 2nd ed. (New York: Van Nostrand Reinhold, 1990), pp. 88–141.

Oppenheim, A. V., and A. S. Willsky, *Signals and Systems*, 2nd ed. (Upper Saddle River, NJ: Prentice-Hall, 1997), pp. 1–14, 21–30, 38–43.

PRACTICE PROBLEM

It can be shown that the signals 2^k, $3^k \sin \frac{k\pi}{2}$, and $3^k \cos \frac{k\pi}{2}$ are solutions of

$$y_{k+3} - 2y_{k+2} + 9y_{k+1} - 18y_k = 0$$

Show that these signals form a basis for the set of all solutions of the difference equation.

4.8 EXERCISES

Verify that the signals in Exercises 1 and 2 are solutions of the accompanying difference equation.

1. $2^k, (-4)^k;\ y_{k+2} + 2y_{k+1} - 8y_k = 0$

2. $5^k, (-5)^k;\ y_{k+2} - 25y_k = 0$

Show that the signals in Exercises 3–6 form a basis for the solution set of the accompanying difference equation.

3. The signals and equation in Exercise 1

4. The signals and equation in Exercise 2

5. $(-2)^k, k(-2)^k;\ y_{k+2} + 4y_{k+1} + 4y_k = 0$

6. $4^k \cos\left(\frac{k\pi}{2}\right), 4^k \sin\left(\frac{k\pi}{2}\right);\ y_{k+2} + 16y_k = 0$

In Exercises 7–12, assume the signals listed are solutions of the given difference equation. Do the signals form a basis for the solution space of the equation? Justify your answers using appropriate theorems.

7. $1^k, 2^k, (-2)^k;\ y_{k+3} - y_{k+2} - 4y_{k+1} + 4y_k = 0$

8. $(-1)^k, 2^k, 3^k;\ y_{k+3} - 4y_{k+2} + 1y_{k+1} + 6y_k = 0$

9. $2^k, 5^k \cos\left(\frac{k\pi}{2}\right), 5^k \sin\left(\frac{k\pi}{2}\right);$
$y_{k+3} - 2y_{k+2} + 25y_{k+1} - 50y_k = 0$

10. $(-2)^k, k(-2)^k, 3^k;\ y_{k+3} + y_{k+2} - 8y_{k+1} - 12y_k = 0$

11. $(-1)^k, 2^k;\ y_{k+3} - 3y_{k+2} + 4y_k = 0$

12. $3^k, (-2)^k;\ y_{k+4} - 13y_{k+2} + 36y_k = 0$

In Exercises 13–16, find a basis for the solution space of the difference equation. Prove that the solutions you find span the solution set.

13. $y_{k+2} - y_{k+1} + \frac{2}{9}y_k = 0$

14. $y_{k+2} - 5y_{k+1} + 6y_k = 0$

15. $6y_{k+2} + y_{k+1} - 2y_k = 0$

16. $y_{k+2} - 25y_k = 0$

Exercises 17 and 18 concern a simple model of the national economy described by the difference equation

$$Y_{k+2} - a(1+b)Y_{k+1} + abY_k = 1 \tag{14}$$

Here Y_k is the total national income during year k, a is a constant less than 1, called the *marginal propensity to consume*, and b is a positive *constant of adjustment* that describes how changes in consumer spending affect the annual rate of private investment.[1]

17. Find the general solution of equation (14) when $a = .9$ and $b = \frac{4}{9}$. What happens to Y_k as k increases? [*Hint:* First find a particular solution of the form $Y_k = T$, where T is a constant, called the equilibrium level of national income.]

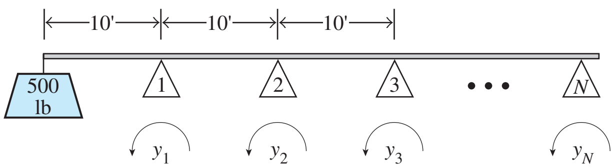
18. Find the general solution of equation (14) when $a = .9$ and $b = .5$.

---

[1] For example, see *Discrete Dynamical Systems*, by James T. Sandefur (Oxford: Clarendon Press, 1990), pp. 267–276. The original *accelerator-multiplier model* is attributed to the economist P. A. Samuelson.

A lightweight cantilevered beam is supported at N points spaced 10 ft apart, and a weight of 500 lb is placed at the end of the beam, 10 ft from the first support, as in the figure. Let y_k be the bending moment at the kth support. Then $y_1 = 5000$ ft-lb. Suppose the beam is rigidly attached at the Nth support and the bending moment there is zero. In between, the moments satisfy the *three-moment equation*

$$y_{k+2} + 4y_{k+1} + y_k = 0 \quad \text{for } k = 1, 2, \ldots, N-2 \tag{15}$$



Bending moments on a cantilevered beam.

19. Find the general solution of difference equation (15). Justify your answer.

20. Find the particular solution of (15) that satisfies the *boundary conditions* $y_1 = 5000$ and $y_N = 0$. (The answer involves N.)

21. When a signal is produced from a sequence of measurements made on a process (a chemical reaction, a flow of heat through a tube, a moving robot arm, etc.), the signal usually contains random *noise* produced by measurement errors. A standard method of preprocessing the data to reduce the noise is to smooth or filter the data. One simple filter is a *moving average* that replaces each y_k by its average with the two adjacent values:

$$\tfrac{1}{3}y_{k+1} + \tfrac{1}{3}y_k + \tfrac{1}{3}y_{k-1} = z_k \quad \text{for } k = 1, 2, \ldots$$

Suppose a signal y_k, for $k = 0, \ldots, 14$, is

9, 5, 7, 3, 2, 4, 6, 5, 7, 6, 8, 10, 9, 5, 7

Use the filter to compute $z_1, \ldots, z_{13}$. Make a broken-line graph that superimposes the original signal and the smoothed signal.

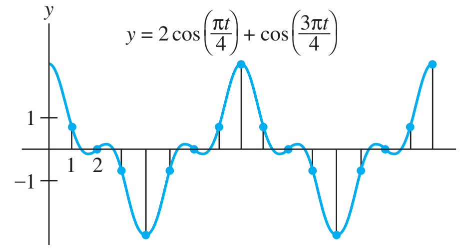
22. Let $\{y_k\}$ be the sequence produced by sampling the continuous signal $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$ at $t = 0, 1, 2, \ldots$, as shown in the figure. The values of y_k, beginning with $k = 0$, are

3, .7, 0, −.7, −3, −.7, 0, .7, 3, .7, 0, ...

where .7 is an abbreviation for $\sqrt{2}/2$.

a. Compute the output signal $\{z_k\}$ when $\{y_k\}$ is fed into the filter in Example 3.

b. Explain how and why the output in part (a) is related to the calculations in Example 3.



Sampled data from $2\cos\frac{\pi t}{4} + \cos\frac{3\pi t}{4}$.

Exercises 23 and 24 refer to a difference equation of the form $y_{k+1} - ay_k = b$, for suitable constants a and b.

23. A loan of $10,000 has an interest rate of 1% per month and a monthly payment of $450. The loan is made at month $k = 0$, and the first payment is made one month later, at $k = 1$. For $k = 0, 1, 2, \ldots$, let y_k be the unpaid balance of the loan just after the kth monthly payment. Thus

$$y_1 = 10{,}000 + (.01)10{,}000 - 450$$

y_1	= 10,000	+ (.01)10,000	− 450
New balance	Balance due	Interest added	Payment

a. Write a difference equation satisfied by $\{y_k\}$.

b. **[M]** Create a table showing k and the balance y_k at month k. List the program or the keystrokes you used to create the table.

c. **[M]** What will k be when the last payment is made? How much will the last payment be? How much money did the borrower pay in total?

24. At time $k = 0$, an initial investment of $1000 is made into a savings account that pays 6% interest per year compounded monthly. (The interest rate per month is .005.) Each month after the initial investment, an additional $200 is added to the account. For $k = 0, 1, 2, \ldots$, let y_k be the amount in the account at time k, just after a deposit has been made.

a. Write a difference equation satisfied by $\{y_k\}$.

b. **[M]** Create a table showing k and the total amount in the savings account at month k, for $k = 0$ through 60. List your program or the keystrokes you used to create the table.

c. **[M]** How much will be in the account after two years (that is, 24 months), four years, and five years? How much of the five-year total is interest?

In Exercises 25–28, show that the given signal is a solution of the difference equation. Then find the general solution of that difference equation.

25. $y_k = k^2$; $y_{k+2} + 3y_{k+1} - 4y_k = 7 + 10k$

26. $y_k = 1 + k$; $y_{k+2} - 6y_{k+1} + 5y_k = -4$

27. $y_k = k - 2$; $y_{k+2} - 4y_k = 8 - 3k$

28. $y_k = 1 + 2k$; $y_{k+2} - 25y_k = -48k - 20$

Write the difference equations in Exercises 29 and 30 as first-order systems, $\mathbf{x}_{k+1} = A\mathbf{x}_k$, for all k.

29. $y_{k+4} + 3y_{k+3} - 8y_{k+2} + 6y_{k+1} - 2y_k = 0$

30. $y_{k+3} - 5y_{k+2} + 8y_k = 0$

31. Is the following difference equation of order 3? Explain.

$$y_{k+3} + 5y_{k+2} + 6y_{k+1} = 0$$

32. What is the order of the following difference equation? Explain your answer.

$$y_{k+3} + a_1 y_{k+2} + a_2 y_{k+1} + a_3 y_k = 0$$

33. Let $y_k = k^2$ and $z_k = 2k|k|$. Are the signals $\{y_k\}$ and $\{z_k\}$ linearly independent? Evaluate the associated Casorati matrix $C(k)$ for $k = 0$, $k = -1$, and $k = -2$, and discuss your results.

34. Let f, g, and h be linearly independent functions defined for all real numbers, and construct three signals by sampling the values of the functions at the integers:

$$u_k = f(k), \qquad v_k = g(k), \qquad w_k = h(k)$$

Must the signals be linearly independent in $\mathbb{S}$? Discuss.

35. Let a and b be nonzero numbers. Show that the mapping T defined by $T\{y_k\} = \{w_k\}$, where

$$w_k = y_{k+2} + ay_{k+1} + by_k$$

is a linear transformation from $\mathbb{S}$ into $\mathbb{S}$.

36. Let V be a vector space, and let $T : V \to V$ be a linear transformation. Given $\mathbf{z}$ in V, suppose $\mathbf{x}_p$ in V satisfies $T(\mathbf{x}_p) = \mathbf{z}$, and let $\mathbf{u}$ be any vector in the kernel of T. Show that $\mathbf{u} + \mathbf{x}_p$ satisfies the nonhomogeneous equation $T(\mathbf{x}) = \mathbf{z}$.

37. Let $\mathbb{S}_0$ be the vector space of all sequences of the form $(y_0, y_1, y_2, \ldots)$, and define linear transformations T and D from $\mathbb{S}_0$ into $\mathbb{S}_0$ by

$$T(y_0, y_1, y_2, \ldots) = (y_1, y_2, y_3, \ldots)$$
$$D(y_0, y_1, y_2, \ldots) = (0, y_0, y_1, y_2, \ldots)$$

Show that $TD = I$ (the identity transformation on $\mathbb{S}_0$) and yet $DT \neq I$.

SOLUTION TO PRACTICE PROBLEM

Examine the Casorati matrix:

$$C(k) = \begin{bmatrix} 2^k & 3^k \sin\frac{k\pi}{2} & 3^k \cos\frac{k\pi}{2} \\ 2^{k+1} & 3^{k+1} \sin\frac{(k+1)\pi}{2} & 3^{k+1} \cos\frac{(k+1)\pi}{2} \\ 2^{k+2} & 3^{k+2} \sin\frac{(k+2)\pi}{2} & 3^{k+2} \cos\frac{(k+2)\pi}{2} \end{bmatrix}$$

Set $k = 0$ and row reduce the matrix to verify that it has three pivot positions and hence is invertible:

$$C(0) = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 3 & 0 \\ 4 & 0 & -9 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 3 & -2 \\ 0 & 0 & -13 \end{bmatrix}$$

The Casorati matrix is invertible at $k = 0$, so the signals are linearly independent. Since there are three signals, and the solution space H of the difference equation has dimension 3 (Theorem 17), the signals form a basis for H, by the Basis Theorem.

4.9 APPLICATIONS TO MARKOV CHAINS

The Markov chains described in this section are used as mathematical models of a wide variety of situations in biology, business, chemistry, engineering, physics, and elsewhere. In each case, the model is used to describe an experiment or measurement that is performed many times in the same way, where the outcome of each trial of the experiment will be one of several specified possible outcomes, and where the outcome of one trial depends only on the immediately preceding trial.

For example, if the population of a city and its suburbs were measured each year, then a vector such as

$$\mathbf{x}_0 = \begin{bmatrix} .60 \\ .40 \end{bmatrix} \tag{1}$$

21. [**M**] Examine powers of a regular stochastic matrix.

a. Compute P^k for $k = 2, 3, 4, 5$, when

$$P = \begin{bmatrix} .3355 & .3682 & .3067 & .0389 \\ .2663 & .2723 & .3277 & .5451 \\ .1935 & .1502 & .1589 & .2395 \\ .2047 & .2093 & .2067 & .1765 \end{bmatrix}$$

Display calculations to four decimal places. What happens to the columns of P^k as k increases? Compute the steady-state vector for P.

b. Compute Q^k for $k = 10, 20, \ldots, 80$, when

$$Q = \begin{bmatrix} .97 & .05 & .10 \\ 0 & .90 & .05 \\ .03 & .05 & .85 \end{bmatrix}$$

(Stability for Q^k to four decimal places may require $k = 116$ or more.) Compute the steady-state vector for Q. Conjecture what might be true for any regular stochastic matrix.

c. Use Theorem 18 to explain what you found in parts (a) and (b).

22. [**M**] Compare two methods for finding the steady-state vector $\mathbf{q}$ of a regular stochastic matrix P: (1) computing $\mathbf{q}$ as in Example 5, or (2) computing P^k for some large value of k and using one of the columns of P^k as an approximation for $\mathbf{q}$. [The *Study Guide* describes a program *nulbasis* that almost automates method (1).]

Experiment with the largest random stochastic matrices your matrix program will allow, and use $k = 100$ or some other large value. For each method, describe the time *you* need to enter the keystrokes and run your program. (Some versions of MATLAB have commands `flops` and `tic ... toc` that record the number of floating point operations and the total elapsed time MATLAB uses.) Contrast the advantages of each method, and state which you prefer.

SOLUTIONS TO PRACTICE PROBLEMS

1. a. Since 5% of the city residents will move to the suburbs within one year, there is a 5% chance of choosing such a person. Without further knowledge about the person, we say that there is a 5% chance the person will move to the suburbs. This fact is contained in the second entry of the state vector $\mathbf{x}_1$, where

$$\mathbf{x}_1 = M\mathbf{x}_0 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .95 \\ .05 \end{bmatrix}$$

b. The likelihood that the person will be living in the suburbs after two years is 9.6%, because

$$\mathbf{x}_2 = M\mathbf{x}_1 = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}\begin{bmatrix} .95 \\ .05 \end{bmatrix} = \begin{bmatrix} .904 \\ .096 \end{bmatrix}$$

2. The steady-state vector satisfies $P\mathbf{x} = \mathbf{x}$. Since

$$P\mathbf{q} = \begin{bmatrix} .6 & .2 \\ .4 & .8 \end{bmatrix}\begin{bmatrix} .3 \\ .7 \end{bmatrix} = \begin{bmatrix} .32 \\ .68 \end{bmatrix} \neq \mathbf{q}$$

we conclude that $\mathbf{q}$ is *not* the steady-state vector for P.

3. M in Example 1 is a regular stochastic matrix because its entries are all strictly positive. So we may use Theorem 18. We already know the steady-state vector from Example 4. Thus the population distribution vectors $\mathbf{x}_k$ converge to

$$\mathbf{q} = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$$

WEB

Eventually 62.5% of the population will live in the suburbs.

CHAPTER 4 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. (If true, cite appropriate facts or theorems. If false, explain why or give a counterexample that shows why the statement is not true in every case.) In parts (a)–(f), $\mathbf{v}_1, \ldots, \mathbf{v}_p$ are vectors in a nonzero finite-dimensional vector space V, and $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$.

a. The set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_p$ is a vector space.

b. If $\{\mathbf{v}_1, \dots, \mathbf{v}_{p-1}\}$ spans V, then S spans V.

c. If $\{\mathbf{v}_1, \dots, \mathbf{v}_{p-1}\}$ is linearly independent, then so is S.

d. If S is linearly independent, then S is a basis for V.

e. If Span $S = V$, then some subset of S is a basis for V.

f. If dim $V = p$ and Span $S = V$, then S cannot be linearly dependent.

g. A plane in $\mathbb{R}^3$ is a two-dimensional subspace.

h. The nonpivot columns of a matrix are always linearly dependent.

i. Row operations on a matrix A can change the linear dependence relations among the rows of A.

j. Row operations on a matrix can change the null space.

k. The rank of a matrix equals the number of nonzero rows.

l. If an $m \times n$ matrix A is row equivalent to an echelon matrix U and if U has k nonzero rows, then the dimension of the solution space of $A\mathbf{x} = \mathbf{0}$ is $m - k$.

m. If B is obtained from a matrix A by several elementary row operations, then rank B = rank A.

n. The nonzero rows of a matrix A form a basis for Row A.

o. If matrices A and B have the same reduced echelon form, then Row A = Row B.

p. If H is a subspace of $\mathbb{R}^3$, then there is a 3×3 matrix A such that H = Col A.

q. If A is $m \times n$ and rank $A = m$, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

r. If A is $m \times n$ and the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto, then rank $A = m$.

s. A change-of-coordinates matrix is always invertible.

t. If $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ are bases for a vector space V, then the jth column of the change-of-coordinates matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{P}$ is the coordinate vector $[\mathbf{c}_j]_{\mathcal{B}}$.

2. Find a basis for the set of all vectors of the form
$$\begin{bmatrix} a - 2b + 5c \\ 2a + 5b - 8c \\ -a - 4b + 7c \\ 3a + b + c \end{bmatrix}. \quad \text{(Be careful.)}$$

3. Let $\mathbf{u}_1 = \begin{bmatrix} -2 \\ 4 \\ -6 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, and $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Find an *implicit* description of W; that is, find a set of one or more homogeneous equations that characterize the points of W. [*Hint:* When is $\mathbf{b}$ in W?]

4. Explain what is wrong with the following discussion: Let $\mathbf{f}(t) = 3 + t$ and $\mathbf{g}(t) = 3t + t^2$, and note that $\mathbf{g}(t) = t\mathbf{f}(t)$. Then $\{\mathbf{f}, \mathbf{g}\}$ is linearly dependent because $\mathbf{g}$ is a multiple of $\mathbf{f}$.

5. Consider the polynomials $\mathbf{p}_1(t) = 1 + t$, $\mathbf{p}_2(t) = 1 - t$, $\mathbf{p}_3(t) = 4$, $\mathbf{p}_4(t) = t + t^2$, and $\mathbf{p}_5(t) = 1 + 2t + t^2$, and let H be the subspace of $\mathbb{P}_5$ spanned by the set $S = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4, \mathbf{p}_5\}$. Use the method described in the proof of the Spanning Set Theorem (Section 4.3) to produce a basis for H. (Explain how to select appropriate members of S.)

6. Suppose $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, $\mathbf{p}_4$ are specific polynomials that span a two-dimensional subspace H of $\mathbb{P}_5$. Describe how one can find a basis for H by examining the four polynomials and making almost no computations.

7. What would you have to know about the solution set of a homogeneous system of 18 linear equations in 20 variables in order to know that every associated nonhomogeneous equation has a solution? Discuss.

8. Let H be an n-dimensional subspace of an n-dimensional vector space V. Explain why $H = V$.

9. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

a. What is the dimension of the range of T if T is a one-to-one mapping? Explain.

b. What is the dimension of the kernel of T (see Section 4.2) if T maps $\mathbb{R}^n$ onto $\mathbb{R}^m$? Explain.

10. Let S be a maximal linearly independent subset of a vector space V. That is, S has the property that if a vector not in S is adjoined to S, then the new set will no longer be linearly independent. Prove that S must be a basis for V. [*Hint:* What if S were linearly independent but not a basis of V?]

11. Let S be a finite minimal spanning set of a vector space V. That is, S has the property that if a vector is removed from S, then the new set will no longer span V. Prove that S must be a basis for V.

Exercises 12–17 develop properties of rank that are sometimes needed in applications. Assume the matrix A is $m \times n$.

12. Show from parts (a) and (b) that rank AB cannot exceed the rank of A or the rank of B. (In general, the rank of a product of matrices cannot exceed the rank of any factor in the product.)

a. Show that if B is $n \times p$, then rank $AB \le$ rank A. [*Hint:* Explain why every vector in the column space of AB is in the column space of A.]

b. Show that if B is $n \times p$, then rank $AB \le$ rank B. [*Hint:* Use part (a) to study rank$(AB)^T$.]

13. Show that if P is an invertible $m \times m$ matrix, then rank PA = rank A. [*Hint:* Apply Exercise 12 to PA and $P^{-1}(PA)$.]

14. Show that if Q is invertible, then rank AQ = rank A. [*Hint:* Use Exercise 13 to study rank$(AQ)^T$.]

15. Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix such that $AB = 0$. Show that rank A + rank $B \le n$. [*Hint:* One of the four subspaces Nul A, Col A, Nul B, and Col B is contained in one of the other three subspaces.]

16. If A is an $m \times n$ matrix of rank r, then a *rank factorization* of A is an equation of the form $A = CR$, where C is an $m \times r$ matrix of rank r and R is an $r \times n$ matrix of rank r.

Such a factorization always exists (Exercise 38 in Section 4.6). Given any two $m \times n$ matrices A and B, use rank factorizations of A and B to prove that

$$\text{rank}(A + B) \leq \text{rank}\, A + \text{rank}\, B$$

[*Hint:* Write $A + B$ as the product of two partitioned matrices.]

17. A **submatrix** of a matrix A is any matrix that results from deleting some (or no) rows and/or columns of A. It can be shown that A has rank r if and only if A contains an invertible $r \times r$ submatrix and no larger square submatrix is invertible. Demonstrate part of this statement by explaining (a) why an $m \times n$ matrix A of rank r has an $m \times r$ submatrix A_1 of rank r, and (b) why A_1 has an invertible $r \times r$ submatrix A_2.

The concept of rank plays an important role in the design of engineering control systems, such as the space shuttle system mentioned in this chapter's introductory example. A *state-space model* of a control system includes a difference equation of the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \quad \text{for } k = 0, 1, \ldots \tag{1}$$

where A is $n \times n$, B is $n \times m$, $\{\mathbf{x}_k\}$ is a sequence of "state vectors" in $\mathbb{R}^n$ that describe the state of the system at discrete times, and $\{\mathbf{u}_k\}$ is a *control*, or *input*, sequence. The pair (A, B) is said to be **controllable** if

$$\text{rank}\,[\,B \quad AB \quad A^2B \quad \cdots \quad A^{n-1}B\,] = n \tag{2}$$

The matrix that appears in (2) is called the **controllability matrix** for the system. If (A, B) is controllable, then the system can be controlled, or driven from the state $\mathbf{0}$ to any specified state $\mathbf{v}$ (in $\mathbb{R}^n$) in at most n steps, simply by choosing an appropriate control sequence in $\mathbb{R}^m$. This fact is illustrated in Exercise 18 for $n = 4$ and $m = 2$. For a further discussion of controllability, see this text's web site (Case Study for Chapter 4).

WEB

18. Suppose A is a 4×4 matrix and B is a 4×2 matrix, and let $\mathbf{u}_0, \ldots, \mathbf{u}_3$ represent a sequence of input vectors in $\mathbb{R}^2$.

a. Set $\mathbf{x}_0 = \mathbf{0}$, compute $\mathbf{x}_1, \ldots, \mathbf{x}_4$ from equation (1), and write a formula for $\mathbf{x}_4$ involving the controllability matrix M appearing in equation (2). (*Note:* The matrix M is constructed as a partitioned matrix. Its overall size here is 4×8.)

b. Suppose (A, B) is controllable and $\mathbf{v}$ is any vector in $\mathbb{R}^4$. Explain why there exists a control sequence $\mathbf{u}_0, \ldots, \mathbf{u}_3$ in $\mathbb{R}^2$ such that $\mathbf{x}_4 = \mathbf{v}$.

Determine if the matrix pairs in Exercises 19–22 are controllable.

19. $A = \begin{bmatrix} .9 & 1 & 0 \\ 0 & -.9 & 0 \\ 0 & 0 & .5 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

20. $A = \begin{bmatrix} .8 & -.3 & 0 \\ .2 & .5 & 1 \\ 0 & 0 & -.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

21. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & -4.2 & -4.8 & -3.6 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

22. [**M**] $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & -13 & -12.2 & -1.5 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

Eigenvectors and Difference Equations

This section concludes by showing how to construct solutions of the first-order difference equation discussed in the chapter introductory example:

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \quad (k = 0, 1, 2, \ldots) \tag{8}$$

If A is an $n \times n$ matrix, then (8) is a *recursive* description of a sequence $\{\mathbf{x}_k\}$ in $\mathbb{R}^n$. A **solution** of (8) is an explicit description of $\{\mathbf{x}_k\}$ whose formula for each $\mathbf{x}_k$ does not depend directly on A or on the preceding terms in the sequence other than the initial term $\mathbf{x}_0$.

The simplest way to build a solution of (8) is to take an eigenvector $\mathbf{x}_0$ and its corresponding eigenvalue λ and let

$$\mathbf{x}_k = \lambda^k \mathbf{x}_0 \quad (k = 1, 2, \ldots) \tag{9}$$

This sequence is a solution because

$$A\mathbf{x}_k = A(\lambda^k \mathbf{x}_0) = \lambda^k (A\mathbf{x}_0) = \lambda^k (\lambda \mathbf{x}_0) = \lambda^{k+1}\mathbf{x}_0 = \mathbf{x}_{k+1}$$

Linear combinations of solutions in the form of equation (9) are solutions, too! See Exercise 33.

PRACTICE PROBLEMS

1. Is 5 an eigenvalue of $A = \begin{bmatrix} 6 & -3 & 1 \\ 3 & 0 & 5 \\ 2 & 2 & 6 \end{bmatrix}$?
2. If $\mathbf{x}$ is an eigenvector of A corresponding to λ, what is $A^3\mathbf{x}$?
3. Suppose that $\mathbf{b}_1$ and $\mathbf{b}_2$ are eigenvectors corresponding to distinct eigenvalues λ_1 and λ_2, respectively, and suppose that $\mathbf{b}_3$ and $\mathbf{b}_4$ are linearly independent eigenvectors corresponding to a third distinct eigenvalue λ_3. Does it necessarily follow that $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$ is a linearly independent set? [*Hint*: Consider the equation $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$.]

5.1 EXERCISES

1. Is $\lambda = 2$ an eigenvalue of $\begin{bmatrix} 3 & 2 \\ 3 & 8 \end{bmatrix}$? Why or why not?

2. Is $\lambda = -3$ an eigenvalue of $\begin{bmatrix} -1 & 4 \\ 6 & 9 \end{bmatrix}$? Why or why not?

3. Is $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 1 & -1 \\ 6 & -4 \end{bmatrix}$? If so, find the eigenvalue.

4. Is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 5 & 2 \\ 3 & 6 \end{bmatrix}$? If so, find the eigenvalue.

5. Is $\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} -4 & 3 & 3 \\ 2 & -3 & -2 \\ -1 & 0 & -2 \end{bmatrix}$? If so, find the eigenvalue.

6. Is $\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$ an eigenvector of $\begin{bmatrix} 3 & 6 & 7 \\ 3 & 2 & 7 \\ 5 & 6 & 4 \end{bmatrix}$? If so, find the eigenvalue.

7. Is $\lambda = 4$ an eigenvalue of $\begin{bmatrix} 3 & 0 & -1 \\ 2 & 3 & 1 \\ -3 & 4 & 5 \end{bmatrix}$? If so, find one corresponding eigenvector.

8. Is $\lambda = 1$ an eigenvalue of $\begin{bmatrix} 4 & -2 & 3 \\ 0 & -1 & 3 \\ -1 & 2 & -2 \end{bmatrix}$? If so, find one corresponding eigenvector.

In Exercises 9–16, find a basis for the eigenspace corresponding to each listed eigenvalue.

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

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

11. $A = \begin{bmatrix} 1 & -3 \\ -4 & 5 \end{bmatrix}, \lambda = -1, 7$

12. $A = \begin{bmatrix} 4 & 1 \\ 3 & 6 \end{bmatrix}, \lambda = 3, 7$

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

14. $A = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 0 & 3 \\ 2 & -2 & 5 \end{bmatrix}, \lambda = 3$

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

16. $A = \begin{bmatrix} 5 & 0 & -1 & 0 \\ 1 & 3 & 0 & 0 \\ 2 & -1 & 3 & 0 \\ 4 & -2 & -2 & 4 \end{bmatrix}, \lambda = 4$

Find the eigenvalues of the matrices in Exercises 17 and 18.

17. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & -2 \end{bmatrix}$ **18.** $\begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 3 \end{bmatrix}$

19. For $A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}$, find one eigenvalue, with no calculation. Justify your answer.

20. Without calculation, find one eigenvalue and two linearly independent eigenvectors of $A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$. Justify your answer.

In Exercises 21 and 22, A is an $n \times n$ matrix. Mark each statement True or False. Justify each answer

.

21. a. If $A\mathbf{x} = \lambda\mathbf{x}$ for some vector $\mathbf{x}$, then λ is an eigenvalue of A.

b. A matrix A is not invertible if and only if 0 is an eigenvalue of A.

c. A number c is an eigenvalue of A if and only if the equation $(A - cI)\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

d. Finding an eigenvector of A may be difficult, but checking whether a given vector is in fact an eigenvector is easy.

e. To find the eigenvalues of A, reduce A to echelon form.

22. a. If $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ, then $\mathbf{x}$ is an eigenvector of A.

b. If $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent eigenvectors, then they correspond to distinct eigenvalues.

c. A steady-state vector for a stochastic matrix is actually an eigenvector.

d. The eigenvalues of a matrix are on its main diagonal.

e. An eigenspace of A is a null space of a certain matrix.

23. Explain why a 2×2 matrix can have at most two distinct eigenvalues. Explain why an $n \times n$ matrix can have at most n distinct eigenvalues.

24. Construct an example of a 2×2 matrix with only one distinct eigenvalue.

25. Let λ be an eigenvalue of an invertible matrix A. Show that λ^{-1} is an eigenvalue of A^{-1}. [*Hint:* Suppose a nonzero $\mathbf{x}$ satisfies $A\mathbf{x} = \lambda\mathbf{x}$.]

26. Show that if A^2 is the zero matrix, then the only eigenvalue of A is 0.

27. Show that λ is an eigenvalue of A if and only if λ is an eigenvalue of A^T. [*Hint:* Find out how $A - \lambda I$ and $A^T - \lambda I$ are related.]

28. Use Exercise 27 to complete the proof of Theorem 1 for the case in which A is lower triangular.

29. Consider an $n \times n$ matrix A with the property that the row sums all equal the same number s. Show that s is an eigenvalue of A. [*Hint:* Find an eigenvector.]

30. Consider an $n \times n$ matrix A with the property that the column sums all equal the same number s. Show that s is an eigenvalue of A. [*Hint:* Use Exercises 27 and 29.]

In Exercises 31 and 32, let A be the matrix of the linear transformation T. Without writing A, find an eigenvalue of A and describe the eigenspace.

31. T is the transformation on $\mathbb{R}^2$ that reflects points across some line through the origin.

32. T is the transformation on $\mathbb{R}^3$ that rotates points about some line through the origin.

33. Let $\mathbf{u}$ and $\mathbf{v}$ be eigenvectors of a matrix A, with corresponding eigenvalues λ and μ, and let c_1 and c_2 be scalars. Define

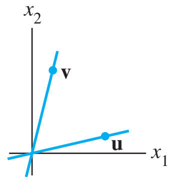
$$\mathbf{x}_k = c_1\lambda^k\mathbf{u} + c_2\mu^k\mathbf{v} \quad (k = 0, 1, 2, \ldots)$$

a. What is $\mathbf{x}_{k+1}$, by definition?

b. Compute $A\mathbf{x}_k$ from the formula for $\mathbf{x}_k$, and show that $A\mathbf{x}_k = \mathbf{x}_{k+1}$. This calculation will prove that the sequence $\{\mathbf{x}_k\}$ defined above satisfies the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$.

34. Describe how you might try to build a solution of a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$ if you were given the initial $\mathbf{x}_0$ and this vector did not happen to be an eigenvector of A. [*Hint:* How might you relate $\mathbf{x}_0$ to eigenvectors of A?]

35. Let $\mathbf{u}$ and $\mathbf{v}$ be the vectors shown in the figure, and suppose $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors of a 2×2 matrix A that correspond to eigenvalues 2 and 3, respectively. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by $T(\mathbf{x}) = A\mathbf{x}$ for each $\mathbf{x}$ in $\mathbb{R}^2$, and let $\mathbf{w} = \mathbf{u} + \mathbf{v}$. Make a copy of the figure, and on the same coordinate system, carefully plot the vectors $T(\mathbf{u})$, $T(\mathbf{v})$, and $T(\mathbf{w})$.



36. Repeat Exercise 35, assuming $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors of A that correspond to eigenvalues -1 and 3, respectively.

[**M**] In Exercises 37–40, use a matrix program to find the eigenvalues of the matrix. Then use the method of Example 4 with a row reduction routine to produce a basis for each eigenspace.

37. $\begin{bmatrix} 12 & 1 & 4 \\ 2 & 11 & 4 \\ 1 & 3 & 7 \end{bmatrix}$ **38.** $\begin{bmatrix} 5 & -2 & 2 & -4 \\ 7 & -4 & 2 & -4 \\ 4 & -4 & 2 & 0 \\ 3 & -1 & 1 & -3 \end{bmatrix}$

39. $\begin{bmatrix} 12 & -90 & 30 & 30 & 30 \\ 8 & -49 & 15 & 15 & 15 \\ 16 & -52 & 12 & 0 & 20 \\ 0 & -30 & 10 & 22 & 10 \\ 8 & -41 & 15 & 15 & 7 \end{bmatrix}$

40. $\begin{bmatrix} -23 & 57 & -9 & -15 & -59 \\ -10 & 12 & -10 & 2 & -22 \\ 11 & 5 & -3 & -19 & -15 \\ -27 & 31 & -27 & 25 & -37 \\ -5 & -15 & -5 & 1 & 31 \end{bmatrix}$

SOLUTIONS TO PRACTICE PROBLEMS

1. The number 5 is an eigenvalue of A if and only if the equation $(A - 5I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution. Form

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

and row reduce the augmented matrix:

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

At this point, it is clear that the homogeneous system has no free variables. Thus $A - 5I$ is an invertible matrix, which means that 5 is *not* an eigenvalue of A.

2. If $\mathbf{x}$ is an eigenvector of A corresponding to λ, then $A\mathbf{x} = \lambda\mathbf{x}$ and so

$$A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$$

Again, $A^3\mathbf{x} = A(A^2\mathbf{x}) = A(\lambda^2\mathbf{x}) = \lambda^2 A\mathbf{x} = \lambda^3\mathbf{x}$. The general pattern, $A^k\mathbf{x} = \lambda^k\mathbf{x}$, is proved by induction.

3. Yes. Suppose $c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + c_3\mathbf{b}_3 + c_4\mathbf{b}_4 = \mathbf{0}$. Since any linear combination of eigenvectors from the same eigenvalue is again an eigenvector for that eigenvalue, $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ is an eigenvector for λ_3. By Theorem 2, the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, and $c_3\mathbf{b}_3 + c_4\mathbf{b}_4$ are linearly independent, so

$$c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + (c_3\mathbf{b}_3 + c_4\mathbf{b}_4) = \mathbf{0}$$

implies $c_1 = c_2 = 0$. But then, c_3 and c_4 must also be zero since $\mathbf{b}_3$ and $\mathbf{b}_4$ are linearly independent. Hence all the coefficients in the original equation must be zero, and the vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$, and $\mathbf{b}_4$ are linearly independent.

5.2 THE CHARACTERISTIC EQUATION

Useful information about the eigenvalues of a square matrix A is encoded in a special scalar equation called the characteristic equation of A. A simple example will lead to the general case.

This explicit formula for $\mathbf{x}_k$ gives the solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. As $k \to \infty$, $(.92)^k$ tends to zero and $\mathbf{x}_k$ tends to $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$. ■

The calculations in Example 5 have an interesting application to a Markov chain discussed in Section 4.9. Those who read that section may recognize that matrix A in Example 5 above is the same as the migration matrix M in Section 4.9, $\mathbf{x}_0$ is the initial population distribution between city and suburbs, and $\mathbf{x}_k$ represents the population distribution after k years.

Theorem 18 in Section 4.9 stated that for a matrix such as A, the sequence $\mathbf{x}_k$ tends to a steady-state vector. Now we know *why* the $\mathbf{x}_k$ behave this way, at least for the migration matrix. The steady-state vector is $.125\mathbf{v}_1$, a multiple of the eigenvector $\mathbf{v}_1$, and formula (5) for $\mathbf{x}_k$ shows precisely why $\mathbf{x}_k \to .125\mathbf{v}_1$.

NUMERICAL NOTES

1. Computer software such as Mathematica and Maple can use symbolic calculations to find the characteristic polynomial of a moderate-sized matrix. But there is no formula or finite algorithm to solve the characteristic equation of a general $n \times n$ matrix for $n \geq 5$.
2. The best numerical methods for finding eigenvalues avoid the characteristic polynomial entirely. In fact, MATLAB finds the characteristic polynomial of a matrix A by first computing the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A and then expanding the product $(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$.
3. Several common algorithms for estimating the eigenvalues of a matrix A are based on Theorem 4. The powerful *QR algorithm* is discussed in the exercises. Another technique, called *Jacobi's method*, works when $A = A^T$ and computes a sequence of matrices of the form
$$A_1 = A \quad \text{and} \quad A_{k+1} = P_k^{-1} A_k P_k \quad (k = 1, 2, \ldots)$$
Each matrix in the sequence is similar to A and so has the same eigenvalues as A. The nondiagonal entries of A_{k+1} tend to zero as k increases, and the diagonal entries tend to approach the eigenvalues of A.
4. Other methods of estimating eigenvalues are discussed in Section 5.8.

PRACTICE PROBLEM

Find the characteristic equation and eigenvalues of $A = \begin{bmatrix} 1 & -4 \\ 4 & 2 \end{bmatrix}$.

5.2 EXERCISES

Find the characteristic polynomial and the real eigenvalues of the matrices in Exercises 1–8.

1. $\begin{bmatrix} 2 & 7 \\ 7 & 2 \end{bmatrix}$

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

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

4. $\begin{bmatrix} 8 & 2 \\ 3 & 3 \end{bmatrix}$

5. $\begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}$

6. $\begin{bmatrix} 9 & -2 \\ 2 & 5 \end{bmatrix}$

7. $\begin{bmatrix} 5 & 3 \\ -4 & 4 \end{bmatrix}$

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

Exercises 9–14 require techniques from Section 3.1. Find the characteristic polynomial of each matrix, using either a cofactor expansion or the special formula for 3×3 determinants described prior to Exercises 15–18 in Section 3.1. [*Note*: Finding the characteristic polynomial of a 3×3 matrix is not easy to do with just row operations, because the variable λ is involved.]

9. $\begin{bmatrix} 4 & 0 & -1 \\ 0 & 4 & -1 \\ 1 & 0 & 2 \end{bmatrix}$ **10.** $\begin{bmatrix} 3 & 1 & 1 \\ 0 & 5 & 0 \\ -2 & 0 & 7 \end{bmatrix}$

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

13. $\begin{bmatrix} 6 & -2 & 0 \\ -2 & 9 & 0 \\ 5 & 8 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 4 & 0 & -1 \\ -1 & 0 & 4 \\ 0 & 2 & 3 \end{bmatrix}$

For the matrices in Exercises 15–17, list the real eigenvalues, repeated according to their multiplicities.

15. $\begin{bmatrix} 5 & 5 & 0 & 2 \\ 0 & 2 & -3 & 6 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 5 \end{bmatrix}$ **16.** $\begin{bmatrix} 3 & 0 & 0 & 0 \\ 6 & 2 & 0 & 0 \\ 0 & 3 & 6 & 0 \\ 2 & 3 & 3 & -5 \end{bmatrix}$

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

18. It can be shown that the algebraic multiplicity of an eigenvalue λ is always greater than or equal to the dimension of the eigenspace corresponding to λ. Find h in the matrix A below such that the eigenspace for $\lambda = 4$ is two-dimensional:

$$A = \begin{bmatrix} 4 & 2 & 3 & 3 \\ 0 & 2 & h & 3 \\ 0 & 0 & 4 & 14 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

19. Let A be an $n \times n$ matrix, and suppose A has n real eigenvalues, $\lambda_1, \dots, \lambda_n$, repeated according to multiplicities, so that

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$$

Explain why $\det A$ is the product of the n eigenvalues of A. (This result is true for any square matrix when complex eigenvalues are considered.)

20. Use a property of determinants to show that A and A^T have the same characteristic polynomial.

In Exercises 21 and 22, A and B are $n \times n$ matrices. Mark each statement True or False. Justify each answer.

21. a. The determinant of A is the product of the diagonal entries in A.

b. An elementary row operation on A does not change the determinant.

c. $(\det A)(\det B) = \det AB$

d. If $\lambda + 5$ is a factor of the characteristic polynomial of A, then 5 is an eigenvalue of A.

22. a. If A is 3×3, with columns $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, then $\det A$ equals the volume of the parallelepiped determined by $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$.

b. $\det A^T = (-1)\det A$.

c. The multiplicity of a root r of the characteristic equation of A is called the algebraic multiplicity of r as an eigenvalue of A.

d. A row replacement operation on A does not change the eigenvalues.

A widely used method for estimating eigenvalues of a general matrix A is the *QR algorithm*. Under suitable conditions, this algorithm produces a sequence of matrices, all similar to A, that become almost upper triangular, with diagonal entries that approach the eigenvalues of A. The main idea is to factor A (or another matrix similar to A) in the form $A = Q_1 R_1$, where $Q_1^T = Q_1^{-1}$ and R_1 is upper triangular. The factors are interchanged to form $A_1 = R_1 Q_1$, which is again factored as $A_1 = Q_2 R_2$; then to form $A_2 = R_2 Q_2$, and so on. The similarity of $A, A_1, \dots$ follows from the more general result in Exercise 23.

23. Show that if $A = QR$ with Q invertible, then A is similar to $A_1 = RQ$.

24. Show that if A and B are similar, then $\det A = \det B$.

25. Let $A = \begin{bmatrix} .6 & .3 \\ .4 & .7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3/7 \\ 4/7 \end{bmatrix}$, and $\mathbf{x}_0 = \begin{bmatrix} .5 \\ .5 \end{bmatrix}$. [*Note:* A is the stochastic matrix studied in Example 5 in Section 4.9.]

a. Find a basis for $\mathbb{R}^2$ consisting of $\mathbf{v}_1$ and another eigenvector $\mathbf{v}_2$ of A.

b. Verify that $\mathbf{x}_0$ may be written in the form $\mathbf{x}_0 = \mathbf{v}_1 + c\mathbf{v}_2$.

c. For $k = 1, 2, \dots$, define $\mathbf{x}_k = A^k \mathbf{x}_0$. Compute $\mathbf{x}_1$ and $\mathbf{x}_2$, and write a formula for $\mathbf{x}_k$. Then show that $\mathbf{x}_k \to \mathbf{v}_1$ as k increases.

26. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Use formula (1) for a determinant (given before Example 2) to show that $\det A = ad - bc$. Consider two cases: $a \neq 0$ and $a = 0$.

27. Let $A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

a. Show that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are eigenvectors of A. [*Note*: A is the stochastic matrix studied in Example 3 of Section 4.9.]

b. Let $\mathbf{x}_0$ be any vector in $\mathbb{R}^3$ with nonnegative entries whose sum is 1. (In Section 4.9, $\mathbf{x}_0$ was called a probability vector.) Explain why there are constants c_1, c_2, c_3 such that $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$. Compute $\mathbf{w}^T\mathbf{x}_0$, and deduce that $c_1 = 1$.

c. For $k = 1, 2, \dots$, define $\mathbf{x}_k = A^k\mathbf{x}_0$, with $\mathbf{x}_0$ as in part (b). Show that $\mathbf{x}_k \to \mathbf{v}_1$ as k increases.

28. [**M**] Construct a random integer-valued 4×4 matrix A, and verify that A and A^T have the same characteristic polynomial (the same eigenvalues with the same multiplicities). Do A and A^T have the same eigenvectors? Make the same analysis of a 5×5 matrix. Report the matrices and your conclusions.

29. [**M**] Construct a random integer-valued 4×4 matrix A.

a. Reduce A to echelon form U with no row scaling, and use U in formula (1) (before Example 2) to compute det A. (If A happens to be singular, start over with a new random matrix.)

b. Compute the eigenvalues of A and the product of these eigenvalues (as accurately as possible).

c. List the matrix A, and, to four decimal places, list the pivots in U and the eigenvalues of A. Compute det A with your matrix program, and compare it with the products you found in (a) and (b).

30. [**M**] Let $A = \begin{bmatrix} -6 & 28 & 21 \\ 4 & -15 & -12 \\ -8 & a & 25 \end{bmatrix}$. For each value of a in the set $\{32, 31.9, 31.8, 32.1, 32.2\}$, compute the characteristic polynomial of A and the eigenvalues. In each case, create a graph of the characteristic polynomial $p(t) = \det(A - tI)$ for $0 \le t \le 3$. If possible, construct all graphs on one coordinate system. Describe how the graphs reveal the changes in the eigenvalues as a changes.

SOLUTION TO PRACTICE PROBLEM

The characteristic equation is

$$0 = \det(A - \lambda I) = \det \begin{bmatrix} 1-\lambda & -4 \\ 4 & 2-\lambda \end{bmatrix}$$

$$= (1-\lambda)(2-\lambda) - (-4)(4) = \lambda^2 - 3\lambda + 18$$

From the quadratic formula,

$$\lambda = \frac{3 \pm \sqrt{(-3)^2 - 4(18)}}{2} = \frac{3 \pm \sqrt{-63}}{2}$$

It is clear that the characteristic equation has no real solutions, so A has no real eigenvalues. The matrix A is acting on the real vector space $\mathbb{R}^2$, and there is no nonzero vector $\mathbf{v}$ in $\mathbb{R}^2$ such that $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar λ.

5.3 DIAGONALIZATION

In many cases, the eigenvalue–eigenvector information contained within a matrix A can be displayed in a useful factorization of the form $A = PDP^{-1}$ where D is a diagonal matrix. In this section, the factorization enables us to compute A^k quickly for large values of k, a fundamental idea in several applications of linear algebra. Later, in Sections 5.6 and 5.7, the factorization will be used to analyze (and *decouple*) dynamical systems.

The following example illustrates that powers of a diagonal matrix are easy to compute.

EXAMPLE 1 If $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, then $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$

and

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

In general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{for } k \ge 1$$

■

If $A = PDP^{-1}$ for some invertible P and diagonal D, then A^k is also easy to compute, as the next example shows.

SOLUTION Since A is a triangular matrix, the eigenvalues are 5 and -3, each with multiplicity 2. Using the method in Section 5.1, we find a basis for each eigenspace.

$$\text{Basis for } \lambda = 5: \quad \mathbf{v}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -3: \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

The set $\{\mathbf{v}_1, \dots, \mathbf{v}_4\}$ is linearly independent, by Theorem 7. So the matrix $P = [\,\mathbf{v}_1 \ \cdots \ \mathbf{v}_4\,]$ is invertible, and $A = PDP^{-1}$, where

$$P = \begin{bmatrix} -8 & -16 & 0 & 0 \\ 4 & 4 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$ ■

PRACTICE PROBLEMS

1. Compute A^8, where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.
2. Let $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of A. Use this information to diagonalize A.
3. Let A be a 4×4 matrix with eigenvalues 5, 3, and -2, and suppose you know that the eigenspace for $\lambda = 3$ is two-dimensional. Do you have enough information to determine if A is diagonalizable?

WEB

5.3 EXERCISES

In Exercises 1 and 2, let $A = PDP^{-1}$ and compute A^4.

1. $P = \begin{bmatrix} 5 & 7 \\ 2 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

2. $P = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

In Exercises 3 and 4, use the factorization $A = PDP^{-1}$ to compute A^k, where k represents an arbitrary positive integer.

3. $\begin{bmatrix} a & 0 \\ 2(a-b) & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}$

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

In Exercises 5 and 6, the matrix A is factored in the form PDP^{-1}. Use the Diagonalization Theorem to find the eigenvalues of A and a basis for each eigenspace.

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

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

Diagonalize the matrices in Exercises 7–20, if possible. The real eigenvalues for Exercises 11–16 and 18 are included below the matrix.

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

8. $\begin{bmatrix} 3 & 2 \\ 0 & 3 \end{bmatrix}$

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

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

11. $\begin{bmatrix} 0 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 3 & 2 \end{bmatrix}$
$\lambda = -1, 5$

12. $\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$
$\lambda = 2, 5$

13. $\begin{bmatrix} 2 & 2 & -1 \\ 1 & 3 & -1 \\ -1 & -2 & 2 \end{bmatrix}$
$\lambda = 1, 5$

14. $\begin{bmatrix} 2 & 0 & -2 \\ 1 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$
$\lambda = 2, 3$

15. $\begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$
$\lambda = 0, 1$

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

17. $\begin{bmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 2 & 2 & 2 \end{bmatrix}$

18. $\begin{bmatrix} 2 & -2 & -2 \\ 3 & -3 & -2 \\ 2 & -2 & -2 \end{bmatrix}$
$\lambda = -2, -1, 0$

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

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

In Exercises 21 and 22, A, B, P, and D are $n \times n$ matrices. Mark each statement True or False. Justify each answer. (Study Theorems 5 and 6 and the examples in this section carefully before you try these exercises.)

21. a. A is diagonalizable if $A = PDP^{-1}$ for some matrix D and some invertible matrix P.

b. If $\mathbb{R}^n$ has a basis of eigenvectors of A, then A is diagonalizable.

c. A is diagonalizable if and only if A has n eigenvalues, counting multiplicities.

d. If A is diagonalizable, then A is invertible.

22. a. A is diagonalizable if A has n eigenvectors.

b. If A is diagonalizable, then A has n distinct eigenvalues.

c. If $AP = PD$, with D diagonal, then the nonzero columns of P must be eigenvectors of A.

d. If A is invertible, then A is diagonalizable.

23. A is a 5×5 matrix with two eigenvalues. One eigenspace is three-dimensional, and the other eigenspace is two-dimensional. Is A diagonalizable? Why?

24. A is a 3×3 matrix with two eigenvalues. Each eigenspace is one-dimensional. Is A diagonalizable? Why?

25. A is a 4×4 matrix with three eigenvalues. One eigenspace is one-dimensional, and one of the other eigenspaces is two-dimensional. Is it possible that A is *not* diagonalizable? Justify your answer.

26. A is a 7×7 matrix with three eigenvalues. One eigenspace is two-dimensional, and one of the other eigenspaces is three-dimensional. Is it possible that A is *not* diagonalizable? Justify your answer.

27. Show that if A is both diagonalizable and invertible, then so is A^{-1}.

28. Show that if A has n linearly independent eigenvectors, then so does A^T. [*Hint:* Use the Diagonalization Theorem.]

29. A factorization $A = PDP^{-1}$ is not unique. Demonstrate this for the matrix A in Example 2. With $D_1 = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}$, use the information in Example 2 to find a matrix P_1 such that $A = P_1 D_1 P_1^{-1}$.

30. With A and D as in Example 2, find an invertible P_2 unequal to the P in Example 2, such that $A = P_2 D P_2^{-1}$.

31. Construct a nonzero 2×2 matrix that is invertible but not diagonalizable.

32. Construct a nondiagonal 2×2 matrix that is diagonalizable but not invertible.

[**M**] Diagonalize the matrices in Exercises 33–36. Use your matrix program's eigenvalue command to find the eigenvalues, and then compute bases for the eigenspaces as in Section 5.1.

33. $\begin{bmatrix} 9 & -4 & -2 & -4 \\ -56 & 32 & -28 & 44 \\ -14 & -14 & 6 & -14 \\ 42 & -33 & 21 & -45 \end{bmatrix}$

34. $\begin{bmatrix} 4 & -9 & -7 & 8 & 2 \\ -7 & -9 & 0 & 7 & 14 \\ 5 & 10 & 5 & -5 & -10 \\ -2 & 3 & 7 & 0 & 4 \\ -3 & -13 & -7 & 10 & 11 \end{bmatrix}$

35. $\begin{bmatrix} 13 & -12 & 9 & -15 & 9 \\ 6 & -5 & 9 & -15 & 9 \\ 6 & -12 & -5 & 6 & 9 \\ 6 & -12 & 9 & -8 & 9 \\ -6 & 12 & 12 & -6 & -2 \end{bmatrix}$

36. $\begin{bmatrix} 24 & -6 & 2 & 6 & 2 \\ 72 & 51 & 9 & -99 & 9 \\ 0 & -63 & 15 & 63 & 63 \\ 72 & 15 & 9 & -63 & 9 \\ 0 & 63 & 21 & -63 & -27 \end{bmatrix}$

SOLUTIONS TO PRACTICE PROBLEMS

1. $\det(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. The eigenvalues are 2 and 1, and the corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Next, form

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Since $A = PDP^{-1}$,

$$\begin{aligned} A^8 = PD^8P^{-1} &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

2. Compute $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1$, and

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

So, $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors for the eigenvalues 1 and 3, respectively. Thus

$$A = PDP^{-1}, \quad \text{where} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

3. Yes, A is diagonalizable. There is a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for the eigenspace corresponding to $\lambda = 3$. In addition, there will be at least one eigenvector for $\lambda = 5$ and one for $\lambda = -2$. Call them $\mathbf{v}_3$ and $\mathbf{v}_4$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly independent by Theorem 2 and Practice Problem 3 in Section 5.1. There can be no additional eigenvectors that are linearly independent from $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4$, because the vectors are all in $\mathbb{R}^4$. Hence the eigenspaces for $\lambda = 5$ and $\lambda = -2$ are both one-dimensional. It follows that A is diagonalizable by Theorem 7(b).

SG Mastering: Eigenvalue and Eigenspace 5–14

5.4 EIGENVECTORS AND LINEAR TRANSFORMATIONS

The goal of this section is to understand the matrix factorization $A = PDP^{-1}$ as a statement about linear transformations. We shall see that the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is essentially the same as the very simple mapping $\mathbf{u} \mapsto D\mathbf{u}$, when viewed from the proper perspective. A similar interpretation will apply to A and D even when D is not a diagonal matrix.

Recall from Section 1.9 that any linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ can be implemented via left-multiplication by a matrix A, called the *standard matrix* of T. Now we need the same sort of representation for any linear transformation between two finite-dimensional vector spaces.

NUMERICAL NOTE

An efficient way to compute a $\mathcal{B}$-matrix $P^{-1}AP$ is to compute AP and then to row reduce the augmented matrix $[\,P \quad AP\,]$ to $[\,I \quad P^{-1}AP\,]$. A separate computation of P^{-1} is unnecessary. See Exercise 15 in Section 2.2.

PRACTICE PROBLEMS

1. Find $T(a_0 + a_1 t + a_2 t^2)$, if T is the linear transformation from $\mathbb{P}_2$ to $\mathbb{P}_2$ whose matrix relative to $\mathcal{B} = \{1, t, t^2\}$ is

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}$$

2. Let A, B, and C be $n \times n$ matrices. The text has shown that if A is similar to B, then B is similar to A. This property, together with the statements below, shows that "similar to" is an *equivalence relation*. (Row equivalence is another example of an equivalence relation.) Verify parts (a) and (b).
   a. A is similar to A.
   b. If A is similar to B and B is similar to C, then A is similar to C.

5.4 EXERCISES

1. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ be bases for vector spaces V and W, respectively. Let $T : V \to W$ be a linear transformation with the property that

$$T(\mathbf{b}_1) = 3\mathbf{d}_1 - 5\mathbf{d}_2, \quad T(\mathbf{b}_2) = -\mathbf{d}_1 + 6\mathbf{d}_2, \quad T(\mathbf{b}_3) = 4\mathbf{d}_2$$

Find the matrix for T relative to $\mathcal{B}$ and $\mathcal{D}$.

2. Let $\mathcal{D} = \{\mathbf{d}_1, \mathbf{d}_2\}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ be bases for vector spaces V and W, respectively. Let $T : V \to W$ be a linear transformation with the property that

$$T(\mathbf{d}_1) = 3\mathbf{b}_1 - 3\mathbf{b}_2, \quad T(\mathbf{d}_2) = -2\mathbf{b}_1 + 5\mathbf{b}_2$$

Find the matrix for T relative to $\mathcal{D}$ and $\mathcal{B}$.

3. Let $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the standard basis for $\mathbb{R}^3$, let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V, and let $T : \mathbb{R}^3 \to V$ be a linear transformation with the property that

$$T(x_1, x_2, x_3) = (2x_3 - x_2)\mathbf{b}_1 - (2x_2)\mathbf{b}_2 + (x_1 + 3x_3)\mathbf{b}_3$$

   a. Compute $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$.
   b. Compute $[T(\mathbf{e}_1)]_{\mathcal{B}}$, $[T(\mathbf{e}_2)]_{\mathcal{B}}$, and $[T(\mathbf{e}_3)]_{\mathcal{B}}$.
   c. Find the matrix for T relative to $\mathcal{E}$ and $\mathcal{B}$.

4. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V and let $T : V \to \mathbb{R}^2$ be a linear transformation with the property that

$$T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3) = \begin{bmatrix} 2x_1 - 3x_2 + x_3 \\ -2x_1 + 5x_3 \end{bmatrix}$$

Find the matrix for T relative to $\mathcal{B}$ and the standard basis for $\mathbb{R}^2$.

5. Let $T : \mathbb{P}_2 \to \mathbb{P}_3$ be the transformation that maps a polynomial $\mathbf{p}(t)$ into the polynomial $(t + 3)\mathbf{p}(t)$.
   a. Find the image of $\mathbf{p}(t) = 3 - 2t + t^2$.
   b. Show that T is a linear transformation.
   c. Find the matrix for T relative to the bases $\{1, t, t^2\}$ and $\{1, t, t^2, t^3\}$.

6. Let $T : \mathbb{P}_2 \to \mathbb{P}_4$ be the transformation that maps a polynomial $\mathbf{p}(t)$ into the polynomial $\mathbf{p}(t) + 2t^2\mathbf{p}(t)$.
   a. Find the image of $\mathbf{p}(t) = 3 - 2t + t^2$.
   b. Show that T is a linear transformation.
   c. Find the matrix for T relative to the bases $\{1, t, t^2\}$ and $\{1, t, t^2, t^3, t^4\}$.

7. Assume the mapping $T : \mathbb{P}_2 \to \mathbb{P}_2$ defined by

$$T(a_0 + a_1 t + a_2 t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$$

is linear. Find the matrix representation of T relative to the basis $\mathcal{B} = \{1, t, t^2\}$.

8. Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ be a basis for a vector space V. Find $T(4\mathbf{b}_1 - 3\mathbf{b}_2)$ when T is a linear transformation from V to V whose matrix relative to $\mathcal{B}$ is

$$[\,T\,]_{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 1 \\ 2 & 1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$$

9. Define $T : \mathbb{P}_2 \to \mathbb{R}^3$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-1) \\ \mathbf{p}(0) \\ \mathbf{p}(1) \end{bmatrix}$.

a. Find the image under T of $\mathbf{p}(t) = 5 + 3t$.

b. Show that T is a linear transformation.

c. Find the matrix for T relative to the basis $\{1, t, t^2\}$ for $\mathbb{P}_2$ and the standard basis for $\mathbb{R}^3$.

10. Define $T : \mathbb{P}_3 \to \mathbb{R}^4$ by $T(\mathbf{p}) = \begin{bmatrix} \mathbf{p}(-2) \\ \mathbf{p}(3) \\ \mathbf{p}(1) \\ \mathbf{p}(0) \end{bmatrix}$.

a. Show that T is a linear transformation.

b. Find the matrix for T relative to the basis $\{1, t, t^2, t^3\}$ for $\mathbb{P}_3$ and the standard basis for $\mathbb{R}^4$.

In Exercises 11 and 12, find the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$, where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$.

11. $A = \begin{bmatrix} -4 & -1 \\ 6 & 1 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

12. $A = \begin{bmatrix} -6 & -2 \\ 4 & 0 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

In Exercises 13–16, define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$. Find a basis $\mathcal{B}$ for $\mathbb{R}^2$ with the property that $[T]_{\mathcal{B}}$ is diagonal.

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

14. $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$

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

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

17. Let $A = \begin{bmatrix} 4 & 1 \\ -1 & 2 \end{bmatrix}$ and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$, for $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$. Define $T : \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$.

a. Verify that $\mathbf{b}_1$ is an eigenvector of A but that A is not diagonalizable.

b. Find the $\mathcal{B}$-matrix for T.

18. Define $T : \mathbb{R}^3 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$, where A is a 3×3 matrix with eigenvalues 5, 5, and -2. Does there exist a basis $\mathcal{B}$ for $\mathbb{R}^3$ such that the $\mathcal{B}$-matrix for T is a diagonal matrix? Discuss.

Verify the statements in Exercises 19–24. The matrices are square.

19. If A is invertible and similar to B, then B is invertible and A^{-1} is similar to B^{-1}. [*Hint:* $P^{-1}AP = B$ for some invertible P. Explain why B is invertible. Then find an invertible Q such that $Q^{-1}A^{-1}Q = B^{-1}$.]

20. If A is similar to B, then A^2 is similar to B^2.

21. If B is similar to A and C is similar to A, then B is similar to C.

22. If A is diagonalizable and B is similar to A, then B is also diagonalizable.

23. If $B = P^{-1}AP$ and $\mathbf{x}$ is an eigenvector of A corresponding to an eigenvalue λ, then $P^{-1}\mathbf{x}$ is an eigenvector of B corresponding also to λ.

24. If A and B are similar, then they have the same rank. [*Hint:* Refer to Supplementary Exercises 13 and 14 in Chapter 4.]

25. The *trace* of a square matrix A is the sum of the diagonal entries in A and is denoted by tr A. It can be verified that $\text{tr}(FG) = \text{tr}(GF)$ for any two $n \times n$ matrices F and G. Show that if A and B are similar, then tr A = tr B.

26. It can be shown that the trace of a matrix A equals the sum of the eigenvalues of A. Verify this statement for the case when A is diagonalizable.

27. Let V be $\mathbb{R}^n$ with a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$; let W be $\mathbb{R}^n$ with the standard basis, denoted here by $\mathcal{E}$; and consider the identity transformation $I : \mathbb{R}^n \to \mathbb{R}^n$, where $I(\mathbf{x}) = \mathbf{x}$. Find the matrix for I relative to $\mathcal{B}$ and $\mathcal{E}$. What was this matrix called in Section 4.4?

28. Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, let W be the same space V with a basis $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$, and let I be the identity transformation $I : V \to W$. Find the matrix for I relative to $\mathcal{B}$ and $\mathcal{C}$. What was this matrix called in Section 4.7?

29. Let V be a vector space with a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. Find the $\mathcal{B}$-matrix for the identity transformation $I : V \to V$.

[M] In Exercises 30 and 31, find the $\mathcal{B}$-matrix for the transformation $\mathbf{x} \mapsto A\mathbf{x}$ where $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

30. $A = \begin{bmatrix} 6 & -2 & -2 \\ 3 & 1 & -2 \\ 2 & -2 & 2 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$

31. $A = \begin{bmatrix} -7 & -48 & -16 \\ 1 & 14 & 6 \\ -3 & -45 & -19 \end{bmatrix}$,

$\mathbf{b}_1 = \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{b}_3 = \begin{bmatrix} 3 \\ -1 \\ 0 \end{bmatrix}$

32. **[M]** Let T be the transformation whose standard matrix is given below. Find a basis for $\mathbb{R}^4$ with the property that $[\,T\,]_{\mathcal{B}}$ is diagonal.

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

SOLUTIONS TO PRACTICE PROBLEMS

1. Let $\mathbf{p}(t) = a_0 + a_1 t + a_2 t^2$ and compute

$$[\,T(\mathbf{p})\,]_{\mathcal{B}} = [\,T\,]_{\mathcal{B}}[\,\mathbf{p}\,]_{\mathcal{B}} = \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & -1 \\ 1 & -2 & 7 \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 3a_0 + 4a_1 \\ 5a_1 - a_2 \\ a_0 - 2a_1 + 7a_2 \end{bmatrix}$$

So $T(\mathbf{p}) = (3a_0 + 4a_1) + (5a_1 - a_2)t + (a_0 - 2a_1 + 7a_2)t^2$.

2. a. $A = (I)^{-1}AI$, so A is similar to A.

b. By hypothesis, there exist invertible matrices P and Q with the property that $B = P^{-1}AP$ and $C = Q^{-1}BQ$. Substitute the formula for B into the formula for C, and use a fact about the inverse of a product:

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$$

This equation has the proper form to show that A is similar to C.

5.5 COMPLEX EIGENVALUES

Since the characteristic equation of an $n \times n$ matrix involves a polynomial of degree n, the equation always has exactly n roots, counting multiplicities, *provided that possibly complex roots are included.* This section shows that if the characteristic equation of a real matrix A has some complex roots, then these roots provide critical information about A. The key is to let A act on the space $\mathbb{C}^n$ of n-tuples of complex numbers.[1]

Our interest in $\mathbb{C}^n$ does not arise from a desire to "generalize" the results of the earlier chapters, although that would in fact open up significant new applications of linear algebra.[2] Rather, this study of complex eigenvalues is essential in order to uncover "hidden" information about certain matrices with real entries that arise in a variety of real-life problems. Such problems include many real dynamical systems that involve periodic motion, vibration, or some type of rotation in space.

The matrix eigenvalue–eigenvector theory already developed for $\mathbb{R}^n$ applies equally well to $\mathbb{C}^n$. So a complex scalar λ satisfies $\det(A - \lambda I) = 0$ if and only if there is a nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$. We call λ a (**complex**) **eigenvalue** and $\mathbf{x}$ a (**complex**) **eigenvector** corresponding to λ.

EXAMPLE 1 If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ on $\mathbb{R}^2$ rotates the plane counterclockwise through a quarter-turn. The action of A is periodic, since after four quarter-turns, a vector is back where it started. Obviously, no nonzero vector is mapped into a multiple of itself, so A has no eigenvectors in $\mathbb{R}^2$ and hence no real eigenvalues. In fact, the characteristic equation of A is

$$\lambda^2 + 1 = 0$$

[1] Refer to Appendix B for a brief discussion of complex numbers. Matrix algebra and concepts about real vector spaces carry over to the case with complex entries and scalars. In particular, $A(c\mathbf{x} + d\mathbf{y}) = cA\mathbf{x} + dA\mathbf{y}$, for A an $m \times n$ matrix with complex entries, $\mathbf{x}, \mathbf{y}$ in $\mathbb{C}^n$, and c, d in $\mathbb{C}$.

[2] A second course in linear algebra often discusses such topics. They are of particular importance in electrical engineering.

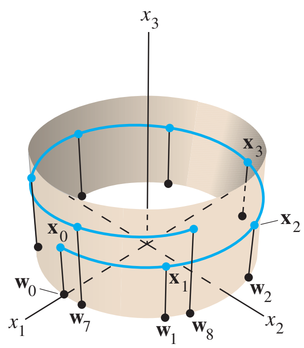


FIGURE 5
Iterates of two points under the action of a 3×3 matrix with a complex eigenvalue.

The phenomenon displayed in Example 7 persists in higher dimensions. For instance, if A is a 3×3 matrix with a complex eigenvalue, then there is a plane in $\mathbb{R}^3$ on which A acts as a rotation (possibly combined with scaling). Every vector in that plane is rotated into another point on the same plane. We say that the plane is **invariant** under A.

EXAMPLE 8 The matrix $A = \begin{bmatrix} .8 & -.6 & 0 \\ .6 & .8 & 0 \\ 0 & 0 & 1.07 \end{bmatrix}$ has eigenvalues $.8 \pm .6i$ and 1.07. Any vector $\mathbf{w}_0$ in the x_1x_2-plane (with third coordinate 0) is rotated by A into another point in the plane. Any vector $\mathbf{x}_0$ not in the plane has its x_3-coordinate multiplied by 1.07. The iterates of the points $\mathbf{w}_0 = (2, 0, 0)$ and $\mathbf{x}_0 = (2, 0, 1)$ under multiplication by A are shown in Fig. 5. ∎

PRACTICE PROBLEM

Show that if a and b are real, then the eigenvalues of $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ are $a \pm bi$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

5.5 EXERCISES

Let each matrix in Exercises 1–6 act on $\mathbb{C}^2$. Find the eigenvalues and a basis for each eigenspace in $\mathbb{C}^2$.

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

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

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

In Exercises 7–12, use Example 6 to list the eigenvalues of A. In each case, the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is the composition of a rotation and a scaling. Give the angle φ of the rotation, where $-\pi < \varphi \le \pi$, and give the scale factor r.

7. $\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$ **8.** $\begin{bmatrix} 3 & 3\sqrt{3} \\ -3\sqrt{3} & 3 \end{bmatrix}$

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

11. $\begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$ **12.** $\begin{bmatrix} 3 & -\sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix}$

In Exercises 13–20, find an invertible matrix P and a matrix C of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ such that the given matrix has the form $A = PCP^{-1}$.

13. $\begin{bmatrix} 1 & -2 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 3 & -3 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 5 \\ -2 & 2 \end{bmatrix}$ **16.** $\begin{bmatrix} 4 & -2 \\ 1 & 6 \end{bmatrix}$

17. $\begin{bmatrix} -11 & -4 \\ 20 & 5 \end{bmatrix}$ **18.** $\begin{bmatrix} 3 & -5 \\ 2 & 5 \end{bmatrix}$

19. $\begin{bmatrix} 1.52 & -.7 \\ .56 & .4 \end{bmatrix}$ **20.** $\begin{bmatrix} -3 & -8 \\ 4 & 5 \end{bmatrix}$

21. In Example 2, solve the first equation in (2) for x_2 in terms of x_1, and from that produce the eigenvector $\mathbf{y} = \begin{bmatrix} 2 \\ -1 + 2i \end{bmatrix}$ for the matrix A. Show that this $\mathbf{y}$ is a (complex) multiple of the vector $\mathbf{v}_1$ used in Example 2.

22. Let A be a complex (or real) $n \times n$ matrix, and let $\mathbf{x}$ in $\mathbb{C}^n$ be an eigenvector corresponding to an eigenvalue λ in $\mathbb{C}$. Show that for each nonzero complex scalar μ, the vector $\mu\mathbf{x}$ is an eigenvector of A.

Chapter 7 will focus on matrices A with the property that $A^T = A$. Exercises 23 and 24 show that every eigenvalue of such a matrix is necessarily real.

23. Let A be an $n \times n$ real matrix with the property that $A^T = A$, let $\mathbf{x}$ be any vector in $\mathbb{C}^n$, and let $q = \bar{\mathbf{x}}^T A\mathbf{x}$. The equalities below show that q is a real number by verifying that $\bar{q} = q$. Give a reason for each step.

$$\bar{q} = \overline{\bar{\mathbf{x}}^T A\mathbf{x}} \underset{(a)}{=} \mathbf{x}^T \overline{A\mathbf{x}} \underset{(b)}{=} \mathbf{x}^T A\bar{\mathbf{x}} \underset{(c)}{=} (\mathbf{x}^T A\bar{\mathbf{x}})^T \underset{(d)}{=} \bar{\mathbf{x}}^T A^T \mathbf{x} \underset{(e)}{=} q$$

24. Let A be an $n \times n$ real matrix with the property that $A^T = A$. Show that if $A\mathbf{x} = \lambda\mathbf{x}$ for some nonzero vector $\mathbf{x}$ in $\mathbb{C}^n$, then, in fact, λ is real and the real part of $\mathbf{x}$ is an eigenvector of A. [*Hint:* Compute $\overline{\mathbf{x}}^T A\mathbf{x}$, and use Exercise 23. Also, examine the real and imaginary parts of $A\mathbf{x}$.]

25. Let A be a real $n \times n$ matrix, and let $\mathbf{x}$ be a vector in $\mathbb{C}^n$. Show that $\text{Re}(A\mathbf{x}) = A(\text{Re}\,\mathbf{x})$ and $\text{Im}(A\mathbf{x}) = A(\text{Im}\,\mathbf{x})$.

26. Let A be a real 2×2 matrix with a complex eigenvalue $\lambda = a - bi$ ($b \neq 0$) and an associated eigenvector $\mathbf{v}$ in $\mathbb{C}^2$.

a. Show that $A(\text{Re}\,\mathbf{v}) = a\,\text{Re}\,\mathbf{v} + b\,\text{Im}\,\mathbf{v}$ and $A(\text{Im}\,\mathbf{v}) = -b\,\text{Re}\,\mathbf{v} + a\,\text{Im}\,\mathbf{v}$. [*Hint:* Write $\mathbf{v} = \text{Re}\,\mathbf{v} + i\,\text{Im}\,\mathbf{v}$, and compute $A\mathbf{v}$.]

b. Verify that if P and C are given as in Theorem 9, then $AP = PC$.

[**M**] In Exercises 27 and 28, find a factorization of the given matrix A in the form $A = PCP^{-1}$, where C is a block-diagonal matrix with 2×2 blocks of the form shown in Example 6. (For each conjugate pair of eigenvalues, use the real and imaginary parts of one eigenvector in $\mathbb{C}^4$ to create two columns of P.)

27. $A = \begin{bmatrix} 26 & 33 & 23 & 20 \\ -6 & -8 & -1 & -13 \\ -14 & -19 & -16 & 3 \\ -20 & -20 & -20 & -14 \end{bmatrix}$

28. $A = \begin{bmatrix} 7 & 11 & 20 & 17 \\ -20 & -40 & -86 & -74 \\ 0 & -5 & -10 & -10 \\ 10 & 28 & 60 & 53 \end{bmatrix}$

SOLUTION TO PRACTICE PROBLEM

Remember that it is easy to test whether a vector is an eigenvector. There is no need to examine the characteristic equation. Compute

$$A\mathbf{x} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} a + bi \\ b - ai \end{bmatrix} = (a + bi)\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

Thus $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to $\lambda = a + bi$. From the discussion in this section, $\begin{bmatrix} 1 \\ i \end{bmatrix}$ must be an eigenvector corresponding to $\overline{\lambda} = a - bi$.

5.6 DISCRETE DYNAMICAL SYSTEMS

Eigenvalues and eigenvectors provide the key to understanding the long-term behavior, or *evolution*, of a dynamical system described by a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Such an equation was used to model population movement in Section 1.10, various Markov chains in Section 4.9, and the spotted owl population in the introductory example for this chapter. The vectors $\mathbf{x}_k$ give information about the system as time (denoted by k) passes. In the spotted owl example, for instance, $\mathbf{x}_k$ listed the numbers of owls in three age classes at time k.

The applications in this section focus on ecological problems because they are easier to state and explain than, say, problems in physics or engineering. However, dynamical systems arise in many scientific fields. For instance, standard undergraduate courses in control systems discuss several aspects of dynamical systems. The modern *state-space* design method in such courses relies heavily on matrix algebra.[1] The *steady-state response* of a control system is the engineering equivalent of what we call here the "long-term behavior" of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

[1] See G. F. Franklin, J. D. Powell, and A. Emami-Naeimi, *Feedback Control of Dynamic Systems*, 5th ed. (Upper Saddle River, NJ: Prentice-Hall, 2006). This undergraduate text has a nice introduction to dynamic models (Chapter 2). State-space design is covered in Chapters 7 and 8.

(11) becomes

$$\mathbf{x}_k = c_1(1.01)^k\mathbf{v}_1 + c_2(-.03 + .26i)^k\mathbf{v}_2 + c_3(-.03 - .26i)^k\mathbf{v}_3$$

As $k \to \infty$, the second two vectors tend to zero. So $\mathbf{x}_k$ becomes more and more like the (real) vector $c_1(1.01)^k\mathbf{v}_1$. The approximations in equations (6) and (7), following Example 1, apply here. Also, it can be shown that the constant c_1 in the initial decomposition of $\mathbf{x}_0$ is positive when the entries in $\mathbf{x}_0$ are nonnegative. Thus the owl population will grow slowly, with a long-term growth rate of 1.01. The eigenvector $\mathbf{v}_1$ describes the eventual distribution of the owls by life stages: for every 31 adults, there will be about 10 juveniles and 3 subadults. ■

Further Reading

Franklin, G. F., J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*, 3rd ed. Reading, MA: Addison-Wesley, 1998.

Sandefur, James T. *Discrete Dynamical Systems—Theory and Applications*. Oxford: Oxford University Press, 1990.

Tuchinsky, Philip. *Management of a Buffalo Herd*, UMAP Module 207. Lexington, MA: COMAP, 1980.

PRACTICE PROBLEMS

1. The matrix A below has eigenvalues 1, $\frac{2}{3}$, and $\frac{1}{3}$, with corresponding eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$:

$$A = \frac{1}{9}\begin{bmatrix} 7 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 5 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

Find the general solution of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ if $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 11 \\ -2 \end{bmatrix}$.

2. What happens to the sequence $\{\mathbf{x}_k\}$ in Practice Problem 1 as $k \to \infty$?

5.6 EXERCISES

1. Let A be a 2×2 matrix with eigenvalues 3 and $1/3$ and corresponding eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Let $\{\mathbf{x}_k\}$ be a solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$, $\mathbf{x}_0 = \begin{bmatrix} 9 \\ 1 \end{bmatrix}$.

a. Compute $\mathbf{x}_1 = A\mathbf{x}_0$. [*Hint:* You do not need to know A itself.]

b. Find a formula for $\mathbf{x}_k$ involving k and the eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

2. Suppose the eigenvalues of a 3×3 matrix A are 3, $4/5$, and $3/5$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 1 \\ -5 \end{bmatrix}$, and $\begin{bmatrix} -3 \\ -3 \\ 7 \end{bmatrix}$. Let $\mathbf{x}_0 = \begin{bmatrix} -2 \\ -5 \\ 3 \end{bmatrix}$. Find the solution of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for the specified $\mathbf{x}_0$, and describe what happens as $k \to \infty$.

In Exercises 3–6, assume that any initial vector $\mathbf{x}_0$ has an eigenvector decomposition such that the coefficient c_1 in equation (1) of this section is positive.[3]

3. Determine the evolution of the dynamical system in Example 1 when the predation parameter p is .2 in equation (3). (Give a formula for $\mathbf{x}_k$.) Does the owl population grow or decline? What about the wood rat population?

4. Determine the evolution of the dynamical system in Example 1 when the predation parameter p is .125. (Give a formula for $\mathbf{x}_k$.) As time passes, what happens to the sizes of the owl and wood rat populations? The system tends toward what is sometimes called an unstable equilibrium. What do you think might happen to the system if some aspect of the model (such as birth rates or the predation rate) were to change slightly?

5. In old-growth forests of Douglas fir, the spotted owl dines mainly on flying squirrels. Suppose the predator–prey matrix for these two populations is $A = \begin{bmatrix} .4 & .3 \\ -p & 1.2 \end{bmatrix}$. Show that if the predation parameter p is .325, both populations grow. Estimate the long-term growth rate and the eventual ratio of owls to flying squirrels.

6. Show that if the predation parameter p in Exercise 5 is .5, both the owls and the squirrels will eventually perish. Find a value of p for which populations of both owls and squirrels tend toward constant levels. What are the relative population sizes in this case?

7. Let A have the properties described in Exercise 1.

a. Is the origin an attractor, a repeller, or a saddle point of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$?

b. Find the directions of greatest attraction and/or repulsion for this dynamical system.

c. Make a graphical description of the system, showing the directions of greatest attraction or repulsion. Include a rough sketch of several typical trajectories (without computing specific points).

8. Determine the nature of the origin (attractor, repeller, or saddle point) for the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ if A has the properties described in Exercise 2. Find the directions of greatest attraction or repulsion.

In Exercises 9–14, classify the origin as an attractor, repeller, or saddle point of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$. Find the directions of greatest attraction and/or repulsion.

9. $A = \begin{bmatrix} 1.7 & -.3 \\ -1.2 & .8 \end{bmatrix}$ **10.** $A = \begin{bmatrix} .3 & .4 \\ -.3 & 1.1 \end{bmatrix}$

11. $A = \begin{bmatrix} .4 & .5 \\ -.4 & 1.3 \end{bmatrix}$ **12.** $A = \begin{bmatrix} .5 & .6 \\ -.3 & 1.4 \end{bmatrix}$

13. $A = \begin{bmatrix} .8 & .3 \\ -.4 & 1.5 \end{bmatrix}$ **14.** $A = \begin{bmatrix} 1.7 & .6 \\ -.4 & .7 \end{bmatrix}$

15. Let $A = \begin{bmatrix} .4 & 0 & .2 \\ .3 & .8 & .3 \\ .3 & .2 & .5 \end{bmatrix}$. The vector $\mathbf{v}_1 = \begin{bmatrix} .1 \\ .6 \\ .3 \end{bmatrix}$ is an eigenvector for A, and two eigenvalues are .5 and .2. Construct the solution of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ that satisfies $\mathbf{x}_0 = (0, .3, .7)$. What happens to $\mathbf{x}_k$ as $k \to \infty$?

16. **[M]** Produce the general solution of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$ when A is the stochastic matrix for the Hertz Rent A Car model in Exercise 16 of Section 4.9.

17. Construct a stage-matrix model for an animal species that has two life stages: juvenile (up to 1 year old) and adult. Suppose the female adults give birth each year to an average of 1.6 female juveniles. Each year, 30% of the juveniles survive to become adults and 80% of the adults survive. For $k \geq 0$, let $\mathbf{x}_k = (j_k, a_k)$, where the entries in $\mathbf{x}_k$ are the numbers of female juveniles and female adults in year k.

a. Construct the stage-matrix A such that $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for $k \geq 0$.

b. Show that the population is growing, compute the eventual growth rate of the population, and give the eventual ratio of juveniles to adults.

c. **[M]** Suppose that initially there are 15 juveniles and 10 adults in the population. Produce four graphs that show how the population changes over eight years: (a) the number of juveniles, (b) the number of adults, (c) the total population, and (d) the ratio of juveniles to adults (each year). When does the ratio in (d) seem to stabilize? Include a listing of the program or keystrokes used to produce the graphs for (c) and (d).

18. A herd of American buffalo (bison) can be modeled by a stage matrix similar to that for the spotted owls. The females can be divided into calves (up to 1 year old), yearlings (1 to 2 years), and adults. Suppose an average of 42 female calves are born each year per 100 adult females. (Only adults produce offspring.) Each year, about 60% of the calves survive, 75% of the yearlings survive, and 95% of the adults survive. For $k \geq 0$, let $\mathbf{x}_k = (c_k, y_k, a_k)$, where the entries in $\mathbf{x}_k$ are the numbers of females in each life stage at year k.

a. Construct the stage-matrix A for the buffalo herd, such that $\mathbf{x}_{k+1} = A\mathbf{x}_k$ for $k \geq 0$.

b. **[M]** Show that the buffalo herd is growing, determine the expected growth rate after many years, and give the expected numbers of calves and yearlings present per 100 adults.

---

[3] One of the limitations of the model in Example 1 is that there always exist initial population vectors $\mathbf{x}_0$ with positive entries such that the coefficient c_1 is negative. The approximation (7) is still valid, but the entries in $\mathbf{x}_k$ eventually become negative.

SOLUTIONS TO PRACTICE PROBLEMS

1. The first step is to write $\mathbf{x}_0$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Row reduction of $[\,\mathbf{v}_1\quad \mathbf{v}_2\quad \mathbf{v}_3\quad \mathbf{x}_0\,]$ produces the weights $c_1 = 2$, $c_2 = 1$, and $c_3 = 3$, so that

$$\mathbf{x}_0 = 2\mathbf{v}_1 + 1\mathbf{v}_2 + 3\mathbf{v}_3$$

Since the eigenvalues are 1, $\frac{2}{3}$, and $\frac{1}{3}$, the general solution is

$$\begin{aligned}\mathbf{x}_k &= 2\cdot 1^k\mathbf{v}_1 + 1\cdot\left(\frac{2}{3}\right)^k\mathbf{v}_2 + 3\cdot\left(\frac{1}{3}\right)^k\mathbf{v}_3\\ &= 2\begin{bmatrix}-2\\2\\1\end{bmatrix} + \left(\frac{2}{3}\right)^k\begin{bmatrix}2\\1\\2\end{bmatrix} + 3\cdot\left(\frac{1}{3}\right)^k\begin{bmatrix}1\\2\\-2\end{bmatrix}\end{aligned} \tag{12}$$

2. As $k \to \infty$, the second and third terms in (12) tend to the zero vector, and

$$\mathbf{x}_k = 2\mathbf{v}_1 + \left(\frac{2}{3}\right)^k\mathbf{v}_2 + 3\left(\frac{1}{3}\right)^k\mathbf{v}_3 \to 2\mathbf{v}_1 = \begin{bmatrix}-4\\4\\2\end{bmatrix}$$

5.7 APPLICATIONS TO DIFFERENTIAL EQUATIONS

This section describes continuous analogues of the difference equations studied in Section 5.6. In many applied problems, several quantities are varying continuously in time, and they are related by a system of differential equations:

$$\begin{aligned}x_1' &= a_{11}x_1 + \cdots + a_{1n}x_n\\ x_2' &= a_{21}x_1 + \cdots + a_{2n}x_n\\ &\vdots\\ x_n' &= a_{n1}x_1 + \cdots + a_{nn}x_n\end{aligned}$$

Here $x_1, \ldots, x_n$ are differentiable functions of t, with derivatives $x_1', \ldots, x_n'$, and the a_{ij} are constants. The crucial feature of this system is that it is *linear*. To see this, write the system as a matrix differential equation

$$\mathbf{x}'(t) = A\mathbf{x}(t) \tag{1}$$

where

$$\mathbf{x}(t) = \begin{bmatrix}x_1(t)\\ \vdots\\ x_n(t)\end{bmatrix},\quad \mathbf{x}'(t) = \begin{bmatrix}x_1'(t)\\ \vdots\\ x_n'(t)\end{bmatrix},\quad \text{and}\quad A = \begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \cdots & a_{nn}\end{bmatrix}$$

A **solution** of equation (1) is a vector-valued function that satisfies (1) for all t in some interval of real numbers, such as $t \geq 0$.

Equation (1) is *linear* because both differentiation of functions and multiplication of vectors by a matrix are linear transformations. Thus, if $\mathbf{u}$ and $\mathbf{v}$ are solutions of $\mathbf{x}' = A\mathbf{x}$, then $c\mathbf{u} + d\mathbf{v}$ is also a solution, because

$$\begin{aligned}(c\mathbf{u} + d\mathbf{v})' &= c\mathbf{u}' + d\mathbf{v}'\\ &= cA\mathbf{u} + dA\mathbf{v} = A(c\mathbf{u} + d\mathbf{v})\end{aligned}$$

CHAPTER 5 SUPPLEMENTARY EXERCISES

Throughout these supplementary exercises, A and B represent square matrices of appropriate sizes.

1. Mark each statement as True or False. Justify each answer.
   a. If A is invertible and 1 is an eigenvalue for A, then 1 is also an eigenvalue of A^{-1}.
   b. If A is row equivalent to the identity matrix I, then A is diagonalizable.
   c. If A contains a row or column of zeros, then 0 is an eigenvalue of A.
   d. Each eigenvalue of A is also an eigenvalue of A^2.
   e. Each eigenvector of A is also an eigenvector of A^2.
   f. Each eigenvector of an invertible matrix A is also an eigenvector of A^{-1}.
   g. Eigenvalues must be nonzero scalars.
   h. Eigenvectors must be nonzero vectors.
   i. Two eigenvectors corresponding to the same eigenvalue are always linearly dependent.
   j. Similar matrices always have exactly the same eigenvalues.
   k. Similar matrices always have exactly the same eigenvectors.
   l. The sum of two eigenvectors of a matrix A is also an eigenvector of A.
   m. The eigenvalues of an upper triangular matrix A are exactly the nonzero entries on the diagonal of A.
   n. The matrices A and A^T have the same eigenvalues, counting multiplicities.
   o. If a 5×5 matrix A has fewer than 5 distinct eigenvalues, then A is not diagonalizable.
   p. There exists a 2×2 matrix that has no eigenvectors in $\mathbb{R}^2$.
   q. If A is diagonalizable, then the columns of A are linearly independent.
   r. A nonzero vector cannot correspond to two different eigenvalues of A.
   s. A (square) matrix A is invertible if and only if there is a coordinate system in which the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is represented by a diagonal matrix.
   t. If each vector $\mathbf{e}_j$ in the standard basis for $\mathbb{R}^n$ is an eigenvector of A, then A is a diagonal matrix.
   u. If A is similar to a diagonalizable matrix B, then A is also diagonalizable.
   v. If A and B are invertible $n \times n$ matrices, then AB is similar to BA.
   w. An $n \times n$ matrix with n linearly independent eigenvectors is invertible.
   x. If A is an $n \times n$ diagonalizable matrix, then each vector in $\mathbb{R}^n$ can be written as a linear combination of eigenvectors of A.

2. Show that if $\mathbf{x}$ is an eigenvector of the matrix product AB and $B\mathbf{x} \neq \mathbf{0}$, then $B\mathbf{x}$ is an eigenvector of BA.

3. Suppose $\mathbf{x}$ is an eigenvector of A corresponding to an eigenvalue λ.
   a. Show that $\mathbf{x}$ is an eigenvector of $5I - A$. What is the corresponding eigenvalue?
   b. Show that $\mathbf{x}$ is an eigenvector of $5I - 3A + A^2$. What is the corresponding eigenvalue?

4. Use mathematical induction to show that if λ is an eigenvalue of an $n \times n$ matrix A, with $\mathbf{x}$ a corresponding eigenvector, then, for each positive integer m, λ^m is an eigenvalue of A^m, with $\mathbf{x}$ a corresponding eigenvector.

5. If $p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$, define $p(A)$ to be the matrix formed by replacing each power of t in $p(t)$ by the corresponding power of A (with $A^0 = I$). That is,

$$p(A) = c_0 I + c_1 A + c_2 A^2 + \cdots + c_n A^n$$

   Show that if λ is an eigenvalue of A, then one eigenvalue of $p(A)$ is $p(\lambda)$.

6. Suppose $A = PDP^{-1}$, where P is 2×2 and $D = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.
   a. Let $B = 5I - 3A + A^2$. Show that B is diagonalizable by finding a suitable factorization of B.
   b. Given $p(t)$ and $p(A)$ as in Exercise 5, show that $p(A)$ is diagonalizable.

7. Suppose A is diagonalizable and $p(t)$ is the characteristic polynomial of A. Define $p(A)$ as in Exercise 5, and show that $p(A)$ is the zero matrix. This fact, which is also true for *any* square matrix, is called the *Cayley–Hamilton theorem*.

8. a. Let A be a diagonalizable $n \times n$ matrix. Show that if the multiplicity of an eigenvalue λ is n, then $A = \lambda I$.
   b. Use part (a) to show that the matrix $A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$ is not diagonalizable.

9. Show that $I - A$ is invertible when all the eigenvalues of A are less than 1 in magnitude. [*Hint:* What would be true if $I - A$ were not invertible?]

10. Show that if A is diagonalizable, with all eigenvalues less than 1 in magnitude, then A^k tends to the zero matrix as $k \to \infty$. [*Hint:* Consider $A^k\mathbf{x}$ where $\mathbf{x}$ represents any one of the columns of I.]

11. Let $\mathbf{u}$ be an eigenvector of A corresponding to an eigenvalue λ, and let H be the line in $\mathbb{R}^n$ through $\mathbf{u}$ and the origin.
    a. Explain why H is invariant under A in the sense that $A\mathbf{x}$ is in H whenever $\mathbf{x}$ is in H.

b. Let K be a one-dimensional subspace of $\mathbb{R}^n$ that is invariant under A. Explain why K contains an eigenvector of A.

12. Let $G = \begin{bmatrix} A & X \\ 0 & B \end{bmatrix}$. Use formula (1) for the determinant in Section 5.2 to explain why $\det G = (\det A)(\det B)$. From this, deduce that the characteristic polynomial of G is the product of the characteristic polynomials of A and B.

Use Exercise 12 to find the eigenvalues of the matrices in Exercises 13 and 14.

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

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

15. Let J be the $n \times n$ matrix of all 1's, and consider $A = (a-b)I + bJ$; that is,

$$A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & b & \cdots & a \end{bmatrix}$$

Use the results of Exercise 16 in the Supplementary Exercises for Chapter 3 to show that the eigenvalues of A are $a-b$ and $a+(n-1)b$. What are the multiplicities of these eigenvalues?

16. Apply the result of Exercise 15 to find the eigenvalues of the matrices $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ and $\begin{bmatrix} 7 & 3 & 3 & 3 & 3 \\ 3 & 7 & 3 & 3 & 3 \\ 3 & 3 & 7 & 3 & 3 \\ 3 & 3 & 3 & 7 & 3 \\ 3 & 3 & 3 & 3 & 7 \end{bmatrix}$.

17. Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Recall from Exercise 25 in Section 5.4 that $\operatorname{tr} A$ (the trace of A) is the sum of the diagonal entries in A. Show that the characteristic polynomial of A is

$$\lambda^2 - (\operatorname{tr} A)\lambda + \det A$$

Then show that the eigenvalues of a 2×2 matrix A are both real if and only if $\det A \le \left(\frac{\operatorname{tr} A}{2}\right)^2$.

18. Let $A = \begin{bmatrix} .4 & -.3 \\ .4 & 1.2 \end{bmatrix}$. Explain why A^k approaches $\begin{bmatrix} -.5 & -.75 \\ 1.0 & 1.50 \end{bmatrix}$ as $k \to \infty$.

Exercises 19–23 concern the polynomial

$$p(t) = a_0 + a_1 t + \cdots + a_{n-1}t^{n-1} + t^n$$

and an $n \times n$ matrix C_p called the **companion matrix** of p:

$$C_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}$$

19. Write the companion matrix C_p for $p(t) = 6 - 5t + t^2$, and then find the characteristic polynomial of C_p.

20. Let $p(t) = (t-2)(t-3)(t-4) = -24 + 26t - 9t^2 + t^3$. Write the companion matrix for $p(t)$, and use techniques from Chapter 3 to find its characteristic polynomial.

21. Use mathematical induction to prove that for $n \ge 2$,

$$\begin{aligned}\det(C_p - \lambda I) &= (-1)^n(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) \\ &= (-1)^n p(\lambda)\end{aligned}$$

[*Hint:* Expanding by cofactors down the first column, show that $\det(C_p - \lambda I)$ has the form $(-\lambda)B + (-1)^n a_0$, where B is a certain polynomial (by the induction assumption).]

22. Let $p(t) = a_0 + a_1 t + a_2 t^2 + t^3$, and let λ be a zero of p.

a. Write the companion matrix for p.

b. Explain why $\lambda^3 = -a_0 - a_1\lambda - a_2\lambda^2$, and show that $(1, \lambda, \lambda^2)$ is an eigenvector of the companion matrix for p.

23. Let p be the polynomial in Exercise 22, and suppose the equation $p(t) = 0$ has distinct roots λ_1, λ_2, λ_3. Let V be the Vandermonde matrix

$$V = \begin{bmatrix} 1 & 1 & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 \end{bmatrix}$$

(The transpose of V was considered in Supplementary Exercise 11 in Chapter 2.) Use Exercise 22 and a theorem from this chapter to deduce that V is invertible (but do not compute V^{-1}). Then explain why $V^{-1}C_pV$ is a diagonal matrix.

24. **[M]** The MATLAB command `roots(p)` computes the roots of the polynomial equation $p(t) = 0$. Read a MATLAB manual, and then describe the basic idea behind the algorithm for the `roots` command.

25. **[M]** Use a matrix program to diagonalize

$$A = \begin{bmatrix} -3 & -2 & 0 \\ 14 & 7 & -1 \\ -6 & -3 & 1 \end{bmatrix}$$

if possible. Use the eigenvalue command to create the diagonal matrix D. If the program has a command that produces eigenvectors, use it to create an invertible matrix P. Then compute $AP - PD$ and PDP^{-1}. Discuss your results.

26. **[M]** Repeat Exercise 25 for $A = \begin{bmatrix} -8 & 5 & -2 & 0 \\ -5 & 2 & 1 & -2 \\ 10 & -8 & 6 & -3 \\ 3 & -2 & 1 & 0 \end{bmatrix}$.

which can be rearranged to produce

$$\begin{aligned}\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\vartheta &= \frac{1}{2}\left[\|\mathbf{u}\|^2+\|\mathbf{v}\|^2-\|\mathbf{u}-\mathbf{v}\|^2\right]\\ &= \frac{1}{2}\left[u_1^2+u_2^2+v_1^2+v_2^2-(u_1-v_1)^2-(u_2-v_2)^2\right]\\ &= u_1v_1+u_2v_2\\ &= \mathbf{u}\cdot\mathbf{v}\end{aligned}$$

The verification for $\mathbb{R}^3$ is similar. When $n > 3$, formula (2) may be used to *define* the angle between two vectors in $\mathbb{R}^n$. In statistics, for instance, the value of $\cos\vartheta$ defined by (2) for suitable vectors **u** and **v** is what statisticians call a *correlation coefficient*.

PRACTICE PROBLEMS

Let $\mathbf{a} = \begin{bmatrix}-2\\1\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}-3\\1\end{bmatrix}$, $\mathbf{c} = \begin{bmatrix}4/3\\-1\\2/3\end{bmatrix}$, and $\mathbf{d} = \begin{bmatrix}5\\6\\-1\end{bmatrix}$.

1. Compute $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}$ and $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a}$.
2. Find a unit vector **u** in the direction of **c**.
3. Show that **d** is orthogonal to **c**.
4. Use the results of Practice Problems 2 and 3 to explain why **d** must be orthogonal to the unit vector **u**.

6.1 EXERCISES

Compute the quantities in Exercises 1–8 using the vectors

$$\mathbf{u} = \begin{bmatrix}-1\\2\end{bmatrix},\quad \mathbf{v} = \begin{bmatrix}4\\6\end{bmatrix},\quad \mathbf{w} = \begin{bmatrix}3\\-1\\-5\end{bmatrix},\quad \mathbf{x} = \begin{bmatrix}6\\-2\\3\end{bmatrix}$$

1. $\mathbf{u}\cdot\mathbf{u}$, $\mathbf{v}\cdot\mathbf{u}$, and $\dfrac{\mathbf{v}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$

2. $\mathbf{w}\cdot\mathbf{w}$, $\mathbf{x}\cdot\mathbf{w}$, and $\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}$

3. $\dfrac{1}{\mathbf{w}\cdot\mathbf{w}}\mathbf{w}$

4. $\dfrac{1}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$

5. $\left(\dfrac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$

6. $\left(\dfrac{\mathbf{x}\cdot\mathbf{w}}{\mathbf{x}\cdot\mathbf{x}}\right)\mathbf{x}$

7. $\|\mathbf{w}\|$

8. $\|\mathbf{x}\|$

In Exercises 9–12, find a unit vector in the direction of the given vector.

9. $\begin{bmatrix}-30\\40\end{bmatrix}$

10. $\begin{bmatrix}-6\\4\\-3\end{bmatrix}$

11. $\begin{bmatrix}7/4\\1/2\\1\end{bmatrix}$

12. $\begin{bmatrix}8/3\\2\end{bmatrix}$

13. Find the distance between $\mathbf{x} = \begin{bmatrix}10\\-3\end{bmatrix}$ and $\mathbf{y} = \begin{bmatrix}-1\\-5\end{bmatrix}$.

14. Find the distance between $\mathbf{u} = \begin{bmatrix}0\\-5\\2\end{bmatrix}$ and $\mathbf{z} = \begin{bmatrix}-4\\-1\\8\end{bmatrix}$.

Determine which pairs of vectors in Exercises 15–18 are orthogonal.

15. $\mathbf{a} = \begin{bmatrix}8\\-5\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}-2\\-3\end{bmatrix}$

16. $\mathbf{u} = \begin{bmatrix}12\\3\\-5\end{bmatrix}$, $\mathbf{v} = \begin{bmatrix}2\\-3\\3\end{bmatrix}$

17. $\mathbf{u} = \begin{bmatrix}3\\2\\-5\\0\end{bmatrix}$, $\mathbf{v} = \begin{bmatrix}-4\\1\\-2\\6\end{bmatrix}$

18. $\mathbf{y} = \begin{bmatrix}-3\\7\\4\\0\end{bmatrix}$, $\mathbf{z} = \begin{bmatrix}1\\-8\\15\\-7\end{bmatrix}$

In Exercises 19 and 20, all vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

19. a. $\mathbf{v}\cdot\mathbf{v} = \|\mathbf{v}\|^2$.
    b. For any scalar c, $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$.
    c. If the distance from **u** to **v** equals the distance from **u** to $-\mathbf{v}$, then **u** and **v** are orthogonal.
    d. For a square matrix A, vectors in Col A are orthogonal to vectors in Nul A.
    e. If vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$ span a subspace W and if **x** is orthogonal to each $\mathbf{v}_j$ for $j = 1, \ldots, p$, then **x** is in $W^\perp$.

20. a. $\mathbf{u}\cdot\mathbf{v} - \mathbf{v}\cdot\mathbf{u} = 0$.

b. For any scalar c, $\|c\mathbf{v}\| = c\|\mathbf{v}\|$.

c. If $\mathbf{x}$ is orthogonal to every vector in a subspace W, then $\mathbf{x}$ is in $W^\perp$.

d. If $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

e. For an $m \times n$ matrix A, vectors in the null space of A are orthogonal to vectors in the row space of A.

21. Use the transpose definition of the inner product to verify parts (b) and (c) of Theorem 1. Mention the appropriate facts from Chapter 2.

22. Let $\mathbf{u} = (u_1, u_2, u_3)$. Explain why $\mathbf{u}\cdot\mathbf{u} \geq 0$. When is $\mathbf{u}\cdot\mathbf{u} = 0$?

23. Let $\mathbf{u} = \begin{bmatrix} 2 \\ -5 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -7 \\ -4 \\ 6 \end{bmatrix}$. Compute and compare $\mathbf{u}\cdot\mathbf{v}$, $\|\mathbf{u}\|^2$, $\|\mathbf{v}\|^2$, and $\|\mathbf{u}+\mathbf{v}\|^2$. Do not use the Pythagorean Theorem.

24. Verify the *parallelogram law* for vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$:

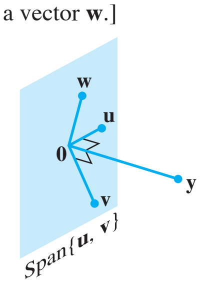
$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

25. Let $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$. Describe the set H of vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ that are orthogonal to $\mathbf{v}$. [*Hint:* Consider $\mathbf{v} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$.]

26. Let $\mathbf{u} = \begin{bmatrix} 5 \\ -6 \\ 7 \end{bmatrix}$, and let W be the set of all $\mathbf{x}$ in $\mathbb{R}^3$ such that $\mathbf{u}\cdot\mathbf{x} = 0$. What theorem in Chapter 4 can be used to show that W is a subspace of $\mathbb{R}^3$? Describe W in geometric language.

27. Suppose a vector $\mathbf{y}$ is orthogonal to vectors $\mathbf{u}$ and $\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to the vector $\mathbf{u}+\mathbf{v}$.

28. Suppose $\mathbf{y}$ is orthogonal to $\mathbf{u}$ and $\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to every $\mathbf{w}$ in Span $\{\mathbf{u}, \mathbf{v}\}$. [*Hint:* An arbitrary $\mathbf{w}$ in Span $\{\mathbf{u}, \mathbf{v}\}$ has the form $\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$. Show that $\mathbf{y}$ is orthogonal to such a vector $\mathbf{w}$.]



29. Let $W = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$. Show that if $\mathbf{x}$ is orthogonal to each $\mathbf{v}_j$, for $1 \leq j \leq p$, then $\mathbf{x}$ is orthogonal to every vector in W.

30. Let W be a subspace of $\mathbb{R}^n$, and let $W^\perp$ be the set of all vectors orthogonal to W. Show that $W^\perp$ is a subspace of $\mathbb{R}^n$ using the following steps.

a. Take $\mathbf{z}$ in $W^\perp$, and let $\mathbf{u}$ represent any element of W. Then $\mathbf{z}\cdot\mathbf{u} = 0$. Take any scalar c and show that $c\mathbf{z}$ is orthogonal to $\mathbf{u}$. (Since $\mathbf{u}$ was an arbitrary element of W, this will show that $c\mathbf{z}$ is in $W^\perp$.)

b. Take $\mathbf{z}_1$ and $\mathbf{z}_2$ in $W^\perp$, and let $\mathbf{u}$ be any element of W. Show that $\mathbf{z}_1 + \mathbf{z}_2$ is orthogonal to $\mathbf{u}$. What can you conclude about $\mathbf{z}_1 + \mathbf{z}_2$? Why?

c. Finish the proof that $W^\perp$ is a subspace of $\mathbb{R}^n$.

31. Show that if $\mathbf{x}$ is in both W and $W^\perp$, then $\mathbf{x} = \mathbf{0}$.

32. [**M**] Construct a pair $\mathbf{u}$, $\mathbf{v}$ of random vectors in $\mathbb{R}^4$, and let

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

a. Denote the columns of A by $\mathbf{a}_1, \ldots, \mathbf{a}_4$. Compute the length of each column, and compute $\mathbf{a}_1\cdot\mathbf{a}_2$, $\mathbf{a}_1\cdot\mathbf{a}_3$, $\mathbf{a}_1\cdot\mathbf{a}_4$, $\mathbf{a}_2\cdot\mathbf{a}_3$, $\mathbf{a}_2\cdot\mathbf{a}_4$, and $\mathbf{a}_3\cdot\mathbf{a}_4$.

b. Compute and compare the lengths of $\mathbf{u}$, $A\mathbf{u}$, $\mathbf{v}$, and $A\mathbf{v}$.

c. Use equation (2) in this section to compute the cosine of the angle between $\mathbf{u}$ and $\mathbf{v}$. Compare this with the cosine of the angle between $A\mathbf{u}$ and $A\mathbf{v}$.

d. Repeat parts (b) and (c) for two other pairs of random vectors. What do you conjecture about the effect of A on vectors?

33. [**M**] Generate random vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{v}$ in $\mathbb{R}^4$ with integer entries (and $\mathbf{v} \neq \mathbf{0}$), and compute the quantities

$$\left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \left(\frac{\mathbf{y}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v},\ \frac{(\mathbf{x}+\mathbf{y})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v},\ \frac{(10\mathbf{x})\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}$$

Repeat the computations with new random vectors $\mathbf{x}$ and $\mathbf{y}$. What do you conjecture about the mapping $\mathbf{x} \mapsto T(\mathbf{x}) = \left(\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\right)\mathbf{v}$ (for $\mathbf{v} \neq \mathbf{0}$)? Verify your conjecture algebraically.

34. [**M**] Let $A = \begin{bmatrix} -6 & 3 & -27 & -33 & -13 \\ 6 & -5 & 25 & 28 & 14 \\ 8 & -6 & 34 & 38 & 18 \\ 12 & -10 & 50 & 41 & 23 \\ 14 & -21 & 49 & 29 & 33 \end{bmatrix}$. Construct a matrix N whose columns form a basis for Nul A, and construct a matrix R whose *rows* form a basis for Row A (see Section 4.6 for details). Perform a matrix computation with N and R that illustrates a fact from Theorem 3.

SOLUTIONS TO PRACTICE PROBLEMS

1. $\mathbf{a}\cdot\mathbf{b} = 7$, $\mathbf{a}\cdot\mathbf{a} = 5$. Hence $\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}} = \dfrac{7}{5}$, and $\left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{\mathbf{a}\cdot\mathbf{a}}\right)\mathbf{a} = \dfrac{7}{5}\mathbf{a} = \begin{bmatrix} -14/5 \\ 7/5 \end{bmatrix}$.

2. Scale $\mathbf{c}$, multiplying by 3 to get $\mathbf{y} = \begin{bmatrix} 4 \\ -3 \\ 2 \end{bmatrix}$. Compute $\|\mathbf{y}\|^2 = 29$ and $\|\mathbf{y}\| = \sqrt{29}$.

The unit vector in the direction of both $\mathbf{c}$ and $\mathbf{y}$ is $\mathbf{u} = \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} = \begin{bmatrix} 4/\sqrt{29} \\ -3/\sqrt{29} \\ 2/\sqrt{29} \end{bmatrix}$.

3. $\mathbf{d}$ is orthogonal to $\mathbf{c}$, because

$$\mathbf{d}\cdot\mathbf{c} = \begin{bmatrix} 5 \\ 6 \\ -1 \end{bmatrix}\cdot\begin{bmatrix} 4/3 \\ -1 \\ 2/3 \end{bmatrix} = \frac{20}{3} - 6 - \frac{2}{3} = 0$$

4. $\mathbf{d}$ is orthogonal to $\mathbf{u}$ because $\mathbf{u}$ has the form $k\mathbf{c}$ for some k, and

$$\mathbf{d}\cdot\mathbf{u} = \mathbf{d}\cdot(k\mathbf{c}) = k(\mathbf{d}\cdot\mathbf{c}) = k(0) = 0$$

6.2 ORTHOGONAL SETS

A set of vectors $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i\cdot\mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{u}_2 &= 3(-1) + 1(2) + 1(1) = 0 \\ \mathbf{u}_1\cdot\mathbf{u}_3 &= 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0 \\ \mathbf{u}_2\cdot\mathbf{u}_3 &= -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0 \end{aligned}$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Fig. 1; the three line segments there are mutually perpendicular. ■

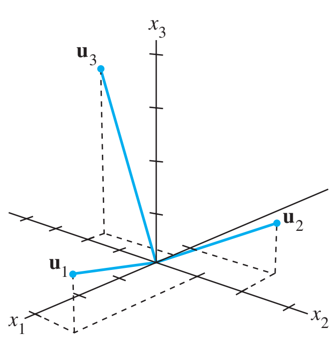


FIGURE 1

THEOREM 4

> If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \dots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \dots, c_p$ must be zero. Thus S is linearly independent. ■

SOLUTION

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$

Theorems 6 and 7 are particularly useful when applied to *square* matrices. An **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$. By Theorem 6, such a matrix has orthonormal columns.[1] It is easy to see that any *square* matrix with orthonormal columns is an orthogonal matrix. Surprisingly, such a matrix must have orthonormal *rows*, too. See Exercises 27 and 28. Orthogonal matrices will appear frequently in Chapter 7.

EXAMPLE 7 The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix because it is square and because its columns are orthonormal, by Example 5. Verify that the rows are orthonormal, too!

PRACTICE PROBLEMS

1. Let $\mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$. Show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for $\mathbb{R}^2$.
2. Let $\mathbf{y}$ and L be as in Example 3 and Fig. 3. Compute the orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto L using $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ instead of the $\mathbf{u}$ in Example 3.
3. Let U and $\mathbf{x}$ be as in Example 6, and let $\mathbf{y} = \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix}$. Verify that $U\mathbf{x}\cdot U\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$.

6.2 EXERCISES

In Exercises 1–6, determine which sets of vectors are orthogonal.

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

2. $\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -5 \\ -2 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 2 \\ -7 \\ -1 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ 9 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}$

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

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

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

In Exercises 7–10, show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ or $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbb{R}^2$ or $\mathbb{R}^3$, respectively. Then express $\mathbf{x}$ as a linear combination of the $\mathbf{u}$'s.

7. $\mathbf{u}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 9 \\ -7 \end{bmatrix}$

[1] A better name might be *orthonormal matrix*, and this term is found in some statistics texts. However, *orthogonal matrix* is the standard term in linear algebra.

8. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} -6 \\ 3 \end{bmatrix}$

9. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 8 \\ -4 \\ -3 \end{bmatrix}$

10. $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

11. Compute the orthogonal projection of $\begin{bmatrix} 1 \\ 7 \end{bmatrix}$ onto the line through $\begin{bmatrix} -4 \\ 2 \end{bmatrix}$ and the origin.

12. Compute the orthogonal projection of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ onto the line through $\begin{bmatrix} -1 \\ 3 \end{bmatrix}$ and the origin.

13. Let $\mathbf{y} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 4 \\ -7 \end{bmatrix}$. Write $\mathbf{y}$ as the sum of two orthogonal vectors, one in Span $\{\mathbf{u}\}$ and one orthogonal to $\mathbf{u}$.

14. Let $\mathbf{y} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$. Write $\mathbf{y}$ as the sum of a vector in Span $\{\mathbf{u}\}$ and a vector orthogonal to $\mathbf{u}$.

15. Let $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$. Compute the distance from $\mathbf{y}$ to the line through $\mathbf{u}$ and the origin.

16. Let $\mathbf{y} = \begin{bmatrix} -3 \\ 9 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Compute the distance from $\mathbf{y}$ to the line through $\mathbf{u}$ and the origin.

In Exercises 17–22, determine which sets of vectors are orthonormal. If a set is only orthogonal, normalize the vectors to produce an orthonormal set.

17. $\begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix}$, $\begin{bmatrix} -1/2 \\ 0 \\ 1/2 \end{bmatrix}$

18. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$

19. $\begin{bmatrix} -.6 \\ .8 \end{bmatrix}$, $\begin{bmatrix} .8 \\ .6 \end{bmatrix}$

20. $\begin{bmatrix} -2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\begin{bmatrix} 1/3 \\ 2/3 \\ 0 \end{bmatrix}$

21. $\begin{bmatrix} 1/\sqrt{10} \\ 3/\sqrt{20} \\ 3/\sqrt{20} \end{bmatrix}$, $\begin{bmatrix} 3/\sqrt{10} \\ -1/\sqrt{20} \\ -1/\sqrt{20} \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

22. $\begin{bmatrix} 1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$, $\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$, $\begin{bmatrix} -2/3 \\ 1/3 \\ -2/3 \end{bmatrix}$

In Exercises 23 and 24, all vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

23. a. Not every linearly independent set in $\mathbb{R}^n$ is an orthogonal set.

b. If $\mathbf{y}$ is a linear combination of nonzero vectors from an orthogonal set, then the weights in the linear combination can be computed without row operations on a matrix.

c. If the vectors in an orthogonal set of nonzero vectors are normalized, then some of the new vectors may not be orthogonal.

d. A matrix with orthonormal columns is an orthogonal matrix.

e. If L is a line through $\mathbf{0}$ and if $\hat{\mathbf{y}}$ is the orthogonal projection of $\mathbf{y}$ onto L, then $\|\hat{\mathbf{y}}\|$ gives the distance from $\mathbf{y}$ to L.

24. a. Not every orthogonal set in $\mathbb{R}^n$ is linearly independent.

b. If a set $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ has the property that $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$, then S is an orthonormal set.

c. If the columns of an $m \times n$ matrix A are orthonormal, then the linear mapping $\mathbf{x} \mapsto A\mathbf{x}$ preserves lengths.

d. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{v}$ is the same as the orthogonal projection of $\mathbf{y}$ onto $c\mathbf{v}$ whenever $c \neq 0$.

e. An orthogonal matrix is invertible.

25. Prove Theorem 7. [*Hint:* For (a), compute $\|U\mathbf{x}\|^2$, or prove (b) first.]

26. Suppose W is a subspace of $\mathbb{R}^n$ spanned by n nonzero orthogonal vectors. Explain why $W = \mathbb{R}^n$.

27. Let U be a square matrix with orthonormal columns. Explain why U is invertible. (Mention the theorems you use.)

28. Let U be an $n \times n$ orthogonal matrix. Show that the rows of U form an orthonormal basis of $\mathbb{R}^n$.

29. Let U and V be $n \times n$ orthogonal matrices. Explain why UV is an orthogonal matrix. [That is, explain why UV is invertible and its inverse is $(UV)^T$.]

30. Let U be an orthogonal matrix, and construct V by interchanging some of the columns of U. Explain why V is an orthogonal matrix.

31. Show that the orthogonal projection of a vector $\mathbf{y}$ onto a line L through the origin in $\mathbb{R}^2$ does not depend on the choice of the nonzero $\mathbf{u}$ in L used in the formula for $\hat{\mathbf{y}}$. To do this, suppose $\mathbf{y}$ and $\mathbf{u}$ are given and $\hat{\mathbf{y}}$ has been computed by formula (2) in this section. Replace $\mathbf{u}$ in that formula by $c\mathbf{u}$, where c is an unspecified nonzero scalar. Show that the new formula gives the same $\hat{\mathbf{y}}$.

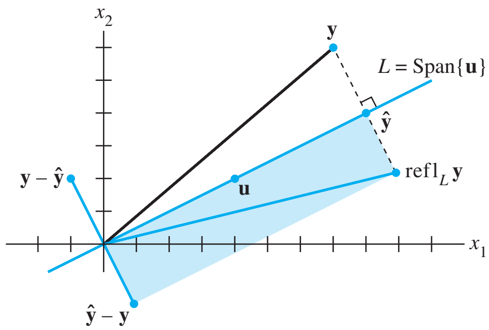
32. Let $\{\mathbf{v}_1, \mathbf{v}_2\}$ be an orthogonal set of nonzero vectors, and let c_1, c_2 be any nonzero scalars. Show that $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2\}$ is also an orthogonal set. Since orthogonality of a set is defined in terms of pairs of vectors, this shows that if the vectors in an orthogonal set are normalized, the new set will still be orthogonal.

33. Given $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^n$, let $L = \text{Span}\,\{\mathbf{u}\}$. Show that the mapping $\mathbf{x} \mapsto \text{proj}_L\, \mathbf{x}$ is a linear transformation.

34. Given $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^n$, let $L = \text{Span}\,\{\mathbf{u}\}$. For $\mathbf{y}$ in $\mathbb{R}^n$, the **reflection of y in** L is the point $\text{refl}_L\, \mathbf{y}$ defined by

$\text{refl}_L\,\mathbf{y} = 2\cdot \text{proj}_L\,\mathbf{y} - \mathbf{y}$

See the figure, which shows that $\text{refl}_L\,\mathbf{y}$ is the sum of $\hat{\mathbf{y}} = \text{proj}_L\,\mathbf{y}$ and $\hat{\mathbf{y}} - \mathbf{y}$. Show that the mapping $\mathbf{y} \mapsto \text{refl}_L\,\mathbf{y}$ is a linear transformation.



The reflection of $\mathbf{y}$ in a line through the origin.

35. [**M**] Show that the columns of the matrix A are orthogonal by making an appropriate matrix calculation. State the calculation you use.

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

36. [**M**] In parts (a)–(d), let U be the matrix formed by normalizing each column of the matrix A in Exercise 35.

a. Compute U^TU and UU^T. How do they differ?

b. Generate a random vector $\mathbf{y}$ in $\mathbb{R}^8$, and compute $\mathbf{p} = UU^T\mathbf{y}$ and $\mathbf{z} = \mathbf{y} - \mathbf{p}$. Explain why $\mathbf{p}$ is in Col A. Verify that $\mathbf{z}$ is orthogonal to $\mathbf{p}$.

c. Verify that $\mathbf{z}$ is orthogonal to each column of U.

d. Notice that $\mathbf{y} = \mathbf{p} + \mathbf{z}$, with $\mathbf{p}$ in Col A. Explain why $\mathbf{z}$ is in $(\text{Col } A)^{\perp}$. (The significance of this decomposition of $\mathbf{y}$ will be explained in the next section.)

SOLUTIONS TO PRACTICE PROBLEMS

1. The vectors are orthogonal because

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = -2/5 + 2/5 = 0$$

They are unit vectors because

$$\|\mathbf{u}_1\|^2 = (-1/\sqrt{5})^2 + (2/\sqrt{5})^2 = 1/5 + 4/5 = 1$$
$$\|\mathbf{u}_2\|^2 = (2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 4/5 + 1/5 = 1$$

In particular, the set $\{\mathbf{u}_1, \mathbf{u}_2\}$ is linearly independent, and hence is a basis for $\mathbb{R}^2$ since there are two vectors in the set.

2. When $\mathbf{y} = \begin{bmatrix} 7 \\ 6 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u} = \frac{20}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

This is the same $\hat{\mathbf{y}}$ found in Example 3. The orthogonal projection does not seem to depend on the $\mathbf{u}$ chosen on the line. See Exercise 31.

3. $U\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} -3\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ -7 \\ 2 \end{bmatrix}$

Also, from Example 6, $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$ and $U\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$. Hence

SG Mastering: Orthogonal Basis 6-4

$$U\mathbf{x}\cdot U\mathbf{y} = 3 + 7 + 2 = 12, \quad \text{and} \quad \mathbf{x}\cdot\mathbf{y} = -6 + 18 = 12$$

PRACTICE PROBLEM

Let $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Use the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $\text{proj}_W \mathbf{y}$.

6.3 EXERCISES

In Exercises 1 and 2, you may assume that $\{\mathbf{u}_1, \ldots, \mathbf{u}_4\}$ is an orthogonal basis for $\mathbb{R}^4$.

1. $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ -4 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -4 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} 5 \\ -3 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} 10 \\ -8 \\ 2 \\ 0 \end{bmatrix}$. Write $\mathbf{x}$ as the sum of two vectors, one in $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and the other in $\text{Span}\{\mathbf{u}_4\}$.

2. $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ -1 \end{bmatrix}$, $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 \\ 5 \\ -3 \\ 3 \end{bmatrix}$. Write $\mathbf{v}$ as the sum of two vectors, one in $\text{Span}\{\mathbf{u}_1\}$ and the other in $\text{Span}\{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$.

In Exercises 3–6, verify that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal set, and then find the orthogonal projection of $\mathbf{y}$ onto $\text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

3. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

4. $\mathbf{y} = \begin{bmatrix} 6 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -4 \\ 3 \\ 0 \end{bmatrix}$

5. $\mathbf{y} = \begin{bmatrix} -1 \\ 2 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$

6. $\mathbf{y} = \begin{bmatrix} 6 \\ 4 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -4 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

In Exercises 7–10, let W be the subspace spanned by the $\mathbf{u}$'s, and write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

7. $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$

8. $\mathbf{y} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ -2 \end{bmatrix}$

9. $\mathbf{y} = \begin{bmatrix} 4 \\ 3 \\ 3 \\ -1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{y} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

In Exercises 11 and 12, find the closest point to $\mathbf{y}$ in the subspace W spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$.

11. $\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$

12. $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$

In Exercises 13 and 14, find the best approximation to $\mathbf{z}$ by vectors of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

13. $\mathbf{z} = \begin{bmatrix} 3 \\ -7 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}$

14. $\mathbf{z} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 5 \\ -2 \\ 4 \\ 2 \end{bmatrix}$

15. Let $\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Find the distance from $\mathbf{y}$ to the plane in $\mathbb{R}^3$ spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$.

16. Let $\mathbf{y}$, $\mathbf{v}_1$, and $\mathbf{v}_2$ be as in Exercise 12. Find the distance from $\mathbf{y}$ to the subspace of $\mathbb{R}^4$ spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$.

17. Let $\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

a. Let $U = [\,\mathbf{u}_1 \quad \mathbf{u}_2\,]$. Compute U^TU and UU^T.

b. Compute $\text{proj}_W\,\mathbf{y}$ and $(UU^T)\mathbf{y}$.

18. Let $\mathbf{y} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$, and $W = \text{Span}\,\{\mathbf{u}_1\}$.

a. Let U be the 2×1 matrix whose only column is $\mathbf{u}_1$. Compute U^TU and UU^T.

b. Compute $\text{proj}_W\,\mathbf{y}$ and $(UU^T)\mathbf{y}$.

19. Let $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Note that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal but that $\mathbf{u}_3$ is not orthogonal to $\mathbf{u}_1$ or $\mathbf{u}_2$. It can be shown that $\mathbf{u}_3$ is not in the subspace W spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Use this fact to construct a nonzero vector $\mathbf{v}$ in $\mathbb{R}^3$ that is orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

20. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be as in Exercise 19, and let $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. It can be shown that $\mathbf{u}_4$ is not in the subspace W spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Use this fact to construct a nonzero vector $\mathbf{v}$ in $\mathbb{R}^3$ that is orthogonal to $\mathbf{u}_1$ and $\mathbf{u}_2$.

In Exercises 21 and 22, all vectors and subspaces are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

21. a. If $\mathbf{z}$ is orthogonal to $\mathbf{u}_1$ and to $\mathbf{u}_2$ and if $W = \text{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$, then $\mathbf{z}$ must be in $W^\perp$.

b. For each $\mathbf{y}$ and each subspace W, the vector $\mathbf{y} - \text{proj}_W\,\mathbf{y}$ is orthogonal to W.

c. The orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto a subspace W can sometimes depend on the orthogonal basis for W used to compute $\hat{\mathbf{y}}$.

d. If $\mathbf{y}$ is in a subspace W, then the orthogonal projection of $\mathbf{y}$ onto W is $\mathbf{y}$ itself.

e. If the columns of an $n \times p$ matrix U are orthonormal, then $UU^T\mathbf{y}$ is the orthogonal projection of $\mathbf{y}$ onto the column space of U.

22. a. If W is a subspace of $\mathbb{R}^n$ and if $\mathbf{v}$ is in both W and $W^\perp$, then $\mathbf{v}$ must be the zero vector.

b. In the Orthogonal Decomposition Theorem, each term in formula (2) for $\hat{\mathbf{y}}$ is itself an orthogonal projection of $\mathbf{y}$ onto a subspace of W.

c. If $\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$, where $\mathbf{z}_1$ is in a subspace W and $\mathbf{z}_2$ is in $W^\perp$, then $\mathbf{z}_1$ must be the orthogonal projection of $\mathbf{y}$ onto W.

d. The best approximation to $\mathbf{y}$ by elements of a subspace W is given by the vector $\mathbf{y} - \text{proj}_W\,\mathbf{y}$.

e. If an $n \times p$ matrix U has orthonormal columns, then $UU^T\mathbf{x} = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

23. Let A be an $m \times n$ matrix. Prove that every vector $\mathbf{x}$ in $\mathbb{R}^n$ can be written in the form $\mathbf{x} = \mathbf{p} + \mathbf{u}$, where $\mathbf{p}$ is in Row A and $\mathbf{u}$ is in Nul A. Also, show that if the equation $A\mathbf{x} = \mathbf{b}$ is consistent, then there is a unique $\mathbf{p}$ in Row A such that $A\mathbf{p} = \mathbf{b}$.

24. Let W be a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{\mathbf{w}_1, \dots, \mathbf{w}_p\}$, and let $\{\mathbf{v}_1, \dots, \mathbf{v}_q\}$ be an orthogonal basis for $W^\perp$.

a. Explain why $\{\mathbf{w}_1, \dots, \mathbf{w}_p, \mathbf{v}_1, \dots, \mathbf{v}_q\}$ is an orthogonal set.

b. Explain why the set in part (a) spans $\mathbb{R}^n$.

c. Show that $\dim W + \dim W^\perp = n$.

25. **[M]** Let U be the 8×4 matrix in Exercise 36 in Section 6.2. Find the closest point to $\mathbf{y} = (1, 1, 1, 1, 1, 1, 1, 1)$ in Col U. Write the keystrokes or commands you use to solve this problem.

26. **[M]** Let U be the matrix in Exercise 25. Find the distance from $\mathbf{b} = (1, 1, 1, 1, -1, -1, -1, -1)$ to Col U.

SOLUTION TO PRACTICE PROBLEM

Compute

$$\text{proj}_W\,\mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 = \frac{88}{66}\mathbf{u}_1 + \frac{-2}{6}\mathbf{u}_2$$

$$= \frac{4}{3}\begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix} - \frac{1}{3}\begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix} = \mathbf{y}$$

In this case, $\mathbf{y}$ happens to be a linear combination of $\mathbf{u}_1$ and $\mathbf{u}_2$, so $\mathbf{y}$ is in W. The closest point in W to $\mathbf{y}$ is $\mathbf{y}$ itself.

By construction, the first k columns of Q are an orthonormal basis of Span $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. From the proof of Theorem 12, $A = QR$ for some R. To find R, observe that $Q^TQ = I$, because the columns of Q are orthonormal. Hence

$$Q^TA = Q^T(QR) = IR = R$$

and

$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$$

■

NUMERICAL NOTES

1. When the Gram–Schmidt process is run on a computer, roundoff error can build up as the vectors $\mathbf{u}_k$ are calculated, one by one. For j and k large but unequal, the inner products $\mathbf{u}_j^T\mathbf{u}_k$ may not be sufficiently close to zero. This loss of orthogonality can be reduced substantially by rearranging the order of the calculations.[1] However, a different computer-based QR factorization is usually preferred to this modified Gram–Schmidt method because it yields a more accurate orthonormal basis, even though the factorization requires about twice as much arithmetic.
2. To produce a QR factorization of a matrix A, a computer program usually left-multiplies A by a sequence of orthogonal matrices until A is transformed into an upper triangular matrix. This construction is analogous to the left-multiplication by elementary matrices that produces an LU factorization of A.

PRACTICE PROBLEM

Let $W = \text{Span}\,\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Construct an orthonormal basis for W.

6.4 EXERCISES

In Exercises 1–6, the given set is a basis for a subspace W. Use the Gram–Schmidt process to produce an orthogonal basis for W.

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

2. $\begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \\ -7 \end{bmatrix}$

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

4. $\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$

5. $\begin{bmatrix} 1 \\ -4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -7 \\ -4 \\ 1 \end{bmatrix}$

6. $\begin{bmatrix} 3 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -5 \\ 9 \\ -9 \\ 3 \end{bmatrix}$

[1] See *Fundamentals of Matrix Computations*, by David S. Watkins (New York: John Wiley & Sons, 1991), pp. 167–180.

7. Find an orthonormal basis of the subspace spanned by the vectors in Exercise 3.

8. Find an orthonormal basis of the subspace spanned by the vectors in Exercise 4.

Find an orthogonal basis for the column space of each matrix in Exercises 9–12.

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

10. $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

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

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

In Exercises 13 and 14, the columns of Q were obtained by applying the Gram–Schmidt process to the columns of A. Find an upper triangular matrix R such that $A = QR$. Check your work.

13. $A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$

14. $A = \begin{bmatrix} -2 & 3 \\ 5 & 7 \\ 2 & -2 \\ 4 & 6 \end{bmatrix}, Q = \begin{bmatrix} -2/7 & 5/7 \\ 5/7 & 2/7 \\ 2/7 & -4/7 \\ 4/7 & 2/7 \end{bmatrix}$

15. Find a QR factorization of the matrix in Exercise 11.

16. Find a QR factorization of the matrix in Exercise 12.

In Exercises 17 and 18, all vectors and subspaces are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

17. a. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W, then multiplying $\mathbf{v}_3$ by a scalar c gives a new orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, c\mathbf{v}_3\}$.

b. The Gram–Schmidt process produces from a linearly independent set $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ an orthogonal set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ with the property that for each k, the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ span the same subspace as that spanned by $\mathbf{x}_1, \ldots, \mathbf{x}_k$.

c. If $A = QR$, where Q has orthonormal columns, then $R = Q^T A$.

18. a. If $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ with $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ linearly independent, and if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set in W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for W.

b. If $\mathbf{x}$ is not in a subspace W, then $\mathbf{x} - \text{proj}_W\, \mathbf{x}$ is not zero.

c. In a QR factorization, say $A = QR$ (when A has linearly independent columns), the columns of Q form an orthonormal basis for the column space of A.

19. Suppose $A = QR$, where Q is $m \times n$ and R is $n \times n$. Show that if the columns of A are linearly independent, then R must be invertible. [*Hint:* Study the equation $R\mathbf{x} = \mathbf{0}$ and use the fact that $A = QR$.]

20. Suppose $A = QR$, where R is an invertible matrix. Show that A and Q have the same column space. [*Hint:* Given $\mathbf{y}$ in Col A, show that $\mathbf{y} = Q\mathbf{x}$ for some $\mathbf{x}$. Also, given $\mathbf{y}$ in Col Q, show that $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x}$.]

21. Given $A = QR$ as in Theorem 12, describe how to find an orthogonal $m \times m$ (square) matrix Q_1 and an invertible $n \times n$ upper triangular matrix R such that

$$A = Q_1 \begin{bmatrix} R \\ 0 \end{bmatrix}$$

The MATLAB `qr` command supplies this "full" QR factorization when rank $A = n$.

22. Let $\mathbf{u}_1, \ldots, \mathbf{u}_p$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\mathbf{x}) = \text{proj}_W\, \mathbf{x}$. Show that T is a linear transformation.

23. Suppose $A = QR$ is a QR factorization of an $m \times n$ matrix A (with linearly independent columns). Partition A as $[A_1 \quad A_2]$, where A_1 has p columns. Show how to obtain a QR factorization of A_1, and explain why your factorization has the appropriate properties.

24. [**M**] Use the Gram–Schmidt process as in Example 2 to produce an orthogonal basis for the column space of

$$A = \begin{bmatrix} -10 & 13 & 7 & -11 \\ 2 & 1 & -5 & 3 \\ -6 & 3 & 13 & -3 \\ 16 & -16 & -2 & 5 \\ 2 & 1 & -5 & -7 \end{bmatrix}$$

25. [**M**] Use the method in this section to produce a QR factorization of the matrix in Exercise 24.

26. [**M**] For a matrix program, the Gram–Schmidt process works better with orthonormal vectors. Starting with $\mathbf{x}_1, \ldots, \mathbf{x}_p$ as in Theorem 11, let $A = [\,\mathbf{x}_1 \quad \cdots \quad x_p\,]$. Suppose Q is an $n \times k$ matrix whose columns form an orthonormal basis for the subspace W_k spanned by the first k columns of A. Then for $\mathbf{x}$ in $\mathbb{R}^n$, $QQ^T\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto W_k (Theorem 10 in Section 6.3). If $\mathbf{x}_{k+1}$ is the next column of A, then equation (2) in the proof of Theorem 11 becomes

$$\mathbf{v}_{k+1} = \mathbf{x}_{k+1} - Q(Q^T\mathbf{x}_{k+1})$$

(The parentheses above reduce the number of arithmetic operations.) Let $\mathbf{u}_{k+1} = \mathbf{v}_{k+1}/\|\mathbf{v}_{k+1}\|$. The new Q for the next step is $[\,Q \quad \mathbf{u}_{k+1}\,]$. Use this procedure to compute the QR factorization of the matrix in Exercise 24. Write the keystrokes or commands you use.

WEB

SOLUTION TO PRACTICE PROBLEM

Let $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 0\mathbf{v}_1 = \mathbf{x}_2$. So $\{\mathbf{x}_1, \mathbf{x}_2\}$ is already orthogonal. All that is needed is to normalize the vectors. Let

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Instead of normalizing $\mathbf{v}_2$ directly, normalize $\mathbf{v}_2' = 3\mathbf{v}_2$ instead:

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2'\|}\mathbf{v}_2' = \frac{1}{\sqrt{1^2 + 1^2 + (-2)^2}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for W.

6.5 LEAST-SQUARES PROBLEMS

The chapter's introductory example described a massive problem $A\mathbf{x} = \mathbf{b}$ that had no solution. Inconsistent systems arise often in applications, though usually not with such an enormous coefficient matrix. When a solution is demanded and none exists, the best one can do is to find an $\mathbf{x}$ that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$.

Think of $A\mathbf{x}$ as an *approximation* to $\mathbf{b}$. The smaller the distance between $\mathbf{b}$ and $A\mathbf{x}$, given by $\|\mathbf{b} - A\mathbf{x}\|$, the better the approximation. The **general least-squares problem** is to find an $\mathbf{x}$ that makes $\|\mathbf{b} - A\mathbf{x}\|$ as small as possible. The adjective "least-squares" arises from the fact that $\|\mathbf{b} - A\mathbf{x}\|$ is the square root of a sum of squares.

DEFINITION

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

The most important aspect of the least-squares problem is that no matter what $\mathbf{x}$ we select, the vector $A\mathbf{x}$ will necessarily be in the column space, Col A. So we seek an $\mathbf{x}$ that makes $A\mathbf{x}$ the closest point in Col A to $\mathbf{b}$. See Fig. 1. (Of course, if $\mathbf{b}$ happens to be in Col A, then $\mathbf{b}$ *is* $A\mathbf{x}$ for some $\mathbf{x}$, and such an $\mathbf{x}$ is a "least-squares solution.")

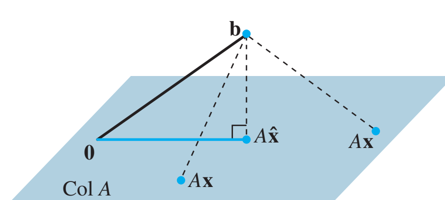


FIGURE 1 The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

PRACTICE PROBLEMS

1. Let $A = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$. Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$, and compute the associated least-squares error.
2. What can you say about the least-squares solution of $A\mathbf{x} = \mathbf{b}$ when $\mathbf{b}$ is orthogonal to the columns of A?

6.5 EXERCISES

In Exercises 1–4, find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ by (a) constructing the normal equations for $\hat{\mathbf{x}}$ and (b) solving for $\hat{\mathbf{x}}$.

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

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

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

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

In Exercises 5 and 6, describe all least-squares solutions of the equation $A\mathbf{x} = \mathbf{b}$.

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

6. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ 2 \\ 3 \\ 6 \\ 5 \\ 4 \end{bmatrix}$

7. Compute the least-squares error associated with the least-squares solution found in Exercise 3.

8. Compute the least-squares error associated with the least-squares solution found in Exercise 4.

In Exercises 9–12, find (a) the orthogonal projection of $\mathbf{b}$ onto Col A and (b) a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

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

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

11. $A = \begin{bmatrix} 4 & 0 & 1 \\ 1 & -5 & 1 \\ 6 & 1 & 0 \\ 1 & -1 & -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

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

13. Let $A = \begin{bmatrix} 3 & 4 \\ -2 & 1 \\ 3 & 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 11 \\ -9 \\ 5 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$. Compute $A\mathbf{u}$ and $A\mathbf{v}$, and compare them with $\mathbf{b}$. Could $\mathbf{u}$ possibly be a least-squares solution of $A\mathbf{x} = \mathbf{b}$? (Answer this without computing a least-squares solution.)

14. Let $A = \begin{bmatrix} 2 & 1 \\ -3 & -4 \\ 3 & 2 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$. Compute $A\mathbf{u}$ and $A\mathbf{v}$, and compare them with $\mathbf{b}$. Is it possible that at least one of $\mathbf{u}$ or $\mathbf{v}$ could be a least-squares solution of $A\mathbf{x} = \mathbf{b}$? (Answer this without computing a least-squares solution.)

In Exercises 15 and 16, use the factorization $A = QR$ to find the least-squares solution of $A\mathbf{x} = \mathbf{b}$.

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

16. $A = \begin{bmatrix} 1 & -1 \\ 1 & 4 \\ 1 & -1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \\ 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} -1 \\ 6 \\ 5 \\ 7 \end{bmatrix}$

In Exercises 17 and 18, A is an $m \times n$ matrix and $\mathbf{b}$ is in $\mathbb{R}^m$. Mark each statement True or False. Justify each answer.

17. a. The general least-squares problem is to find an $\mathbf{x}$ that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$.

b. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}}$ that satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is the orthogonal projection of $\mathbf{b}$ onto Col A.

c. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a vector $\hat{\mathbf{x}}$ such that $\|\mathbf{b} - A\mathbf{x}\| \leq \|\mathbf{b} - A\hat{\mathbf{x}}\|$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

d. Any solution of $A^TA\mathbf{x} = A^T\mathbf{b}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$.

e. If the columns of A are linearly independent, then the equation $A\mathbf{x} = \mathbf{b}$ has exactly one least-squares solution.

18. a. If $\mathbf{b}$ is in the column space of A, then every solution of $A\mathbf{x} = \mathbf{b}$ is a least-squares solution.

b. The least-squares solution of $A\mathbf{x} = \mathbf{b}$ is the point in the column space of A closest to $\mathbf{b}$.

c. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is a list of weights that, when applied to the columns of A, produces the orthogonal projection of $\mathbf{b}$ onto Col A.

d. If $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$, then $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

e. The normal equations always provide a reliable method for computing least-squares solutions.

f. If A has a QR factorization, say $A = QR$, then the best way to find the least-squares solution of $A\mathbf{x} = \mathbf{b}$ is to compute $\hat{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$.

19. Let A be an $m \times n$ matrix. Use the steps below to show that a vector $\mathbf{x}$ in $\mathbb{R}^n$ satisfies $A\mathbf{x} = \mathbf{0}$ if and only if $A^TA\mathbf{x} = \mathbf{0}$. This will show that Nul A = Nul A^TA.

a. Show that if $A\mathbf{x} = \mathbf{0}$, then $A^TA\mathbf{x} = \mathbf{0}$.

b. Suppose $A^TA\mathbf{x} = \mathbf{0}$. Explain why $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{0}$, and use this to show that $A\mathbf{x} = \mathbf{0}$.

20. Let A be an $m \times n$ matrix such that A^TA is invertible. Show that the columns of A are linearly independent. [*Careful:* You may not assume that A is invertible; it may not even be square.]

21. Let A be an $m \times n$ matrix whose columns are linearly independent. [*Careful:* A need not be square.]

a. Use Exercise 19 to show that A^TA is an invertible matrix.

b. Explain why A must have at least as many rows as columns.

c. Determine the rank of A.

22. Use Exercise 19 to show that rank A^TA = rank A. [*Hint:* How many columns does A^TA have? How is this connected with the rank of A^TA?]

23. Suppose A is $m \times n$ with linearly independent columns and $\mathbf{b}$ is in $\mathbb{R}^m$. Use the normal equations to produce a formula for $\hat{\mathbf{b}}$, the projection of $\mathbf{b}$ onto Col A. [*Hint:* Find $\hat{\mathbf{x}}$ first. The formula does not require an orthogonal basis for Col A.]

24. Find a formula for the least-squares solution of $A\mathbf{x} = \mathbf{b}$ when the columns of A are orthonormal.

25. Describe all least-squares solutions of the system

$$x + y = 2$$
$$x + y = 4$$

26. [**M**] Example 3 in Section 4.8 displayed a low-pass linear filter that changed a signal $\{y_k\}$ into $\{y_{k+1}\}$ and changed a higher-frequency signal $\{w_k\}$ into the zero signal, where $y_k = \cos(\pi k/4)$ and $w_k = \cos(3\pi k/4)$. The following calculations will design a filter with approximately those properties. The filter equation is

$$a_0 y_{k+2} + a_1 y_{k+1} + a_2 y_k = z_k \qquad \text{for all } k \tag{8}$$

Because the signals are periodic, with period 8, it suffices to study equation (8) for $k = 0, \dots, 7$. The action on the two signals described above translates into two sets of eight equations, shown below:

$$\begin{array}{l} k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\overset{\begin{array}{ccc} y_{k+2} & y_{k+1} & y_k \end{array}}{\begin{bmatrix} 0 & .7 & 1 \\ -.7 & 0 & .7 \\ -1 & -.7 & 0 \\ -.7 & -1 & -.7 \\ 0 & -.7 & -1 \\ .7 & 0 & -.7 \\ 1 & .7 & 0 \\ .7 & 1 & .7 \end{bmatrix}}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\overset{y_{k+1}}{\begin{bmatrix} .7 \\ 0 \\ -.7 \\ -1 \\ -.7 \\ 0 \\ .7 \\ 1 \end{bmatrix}}$$

$$\begin{array}{l} k=0 \\ k=1 \\ \vdots \\ \\ \\ \\ \\ k=7 \end{array}
\overset{\begin{array}{ccc} w_{k+2} & w_{k+1} & w_k \end{array}}{\begin{bmatrix} 0 & -.7 & 1 \\ .7 & 0 & -.7 \\ -1 & .7 & 0 \\ .7 & -1 & .7 \\ 0 & .7 & -1 \\ -.7 & 0 & .7 \\ 1 & -.7 & 0 \\ -.7 & 1 & -.7 \end{bmatrix}}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =
\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Write an equation $A\mathbf{x} = \mathbf{b}$, where A is a 16×3 matrix formed from the two coefficient matrices above and where $\mathbf{b}$ in $\mathbb{R}^{16}$ is formed from the two right sides of the equations. Find a_0, a_1, and a_2 given by the least-squares solution of $A\mathbf{x} = \mathbf{b}$. (The .7 in the data above was used as an approximation for $\sqrt{2}/2$, to illustrate how a typical computation in an applied problem might proceed. If .707 were used instead, the resulting filter coefficients would agree to at least seven decimal places with $\sqrt{2}/4$, $1/2$, and $\sqrt{2}/4$, the values produced by exact arithmetic calculations.)

WEB

SOLUTIONS TO PRACTICE PROBLEMS

1. First, compute

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 0 \\ 9 & 83 & 28 \\ 0 & 28 & 14 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & 5 & 7 \\ -3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix} = \begin{bmatrix} -3 \\ -65 \\ -28 \end{bmatrix}$$

Next, row reduce the augmented matrix for the normal equations, $A^TA\mathbf{x} = A^T\mathbf{b}$:

$$\begin{bmatrix} 3 & 9 & 0 & -3 \\ 9 & 83 & 28 & -65 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 56 & 28 & -56 \\ 0 & 28 & 14 & -28 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 0 & -3/2 & 2 \\ 0 & 1 & 1/2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The general least-squares solution is $x_1 = 2 + \frac{3}{2}x_3$, $x_2 = -1 - \frac{1}{2}x_3$, with x_3 free. For one specific solution, take $x_3 = 0$ (for example), and get

$$\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}$$

To find the least-squares error, compute

$$\hat{\mathbf{b}} = A\hat{\mathbf{x}} = \begin{bmatrix} 1 & -3 & -3 \\ 1 & 5 & 1 \\ 1 & 7 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \\ -5 \end{bmatrix}$$

It turns out that $\hat{\mathbf{b}} = \mathbf{b}$, so $\|\mathbf{b} - \hat{\mathbf{b}}\| = 0$. The least-squares error is zero because $\mathbf{b}$ happens to be in Col A.

2. If $\mathbf{b}$ is orthogonal to the columns of A, then the projection of $\mathbf{b}$ onto the column space of A is $\mathbf{0}$. In this case, a least-squares solution $\hat{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$ satisfies $A\hat{\mathbf{x}} = \mathbf{0}$.

6.6 APPLICATIONS TO LINEAR MODELS

A common task in science and engineering is to analyze and understand relationships among several quantities that vary. This section describes a variety of situations in which data are used to build or verify a formula that predicts the value of one variable as a function of other variables. In each case, the problem will amount to solving a least-squares problem.

For easy application of the discussion to real problems that you may encounter later in your career, we choose notation that is commonly used in the statistical analysis of scientific and engineering data. Instead of $A\mathbf{x} = \mathbf{b}$, we write $X\boldsymbol{\beta} = \mathbf{y}$ and refer to X as the **design matrix**, $\boldsymbol{\beta}$ as the **parameter vector**, and $\mathbf{y}$ as the **observation vector**.

Least-Squares Lines

The simplest relation between two variables x and y is the linear equation $y = \beta_0 + \beta_1 x$.[1] Experimental data often produce points $(x_1, y_1), \dots, (x_n, y_n)$ that,

[1] This notation is commonly used for least-squares lines instead of $y = mx + b$.

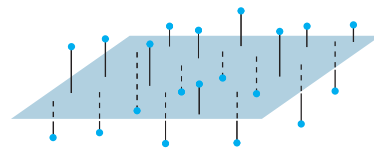


FIGURE 6 A least-squares plane.

SOLUTION We expect the data to satisfy the following equations:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 u_1 + \beta_2 v_1 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 u_2 + \beta_2 v_2 + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 u_n + \beta_2 v_n + \epsilon_n \end{aligned}$$

This system has the matrix form $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where

Observation vector, Design matrix, Parameter vector, Residual vector

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & u_1 & v_1 \\ 1 & u_2 & v_2 \\ \vdots & \vdots & \vdots \\ 1 & u_n & v_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

■

Example 4 shows that the linear model for multiple regression has the same abstract form as the model for the simple regression in the earlier examples. Linear algebra gives us the power to understand the general principle behind all the linear models. Once X is defined properly, the normal equations for $\boldsymbol{\beta}$ have the same matrix form, no matter how many variables are involved. Thus, for any linear model where X^TX is invertible, the least-squares $\hat{\boldsymbol{\beta}}$ is given by $(X^TX)^{-1}X^T\mathbf{y}$.

SG The Geometry of a Linear Model 6-19

Further Reading

Ferguson, J., *Introduction to Linear Algebra in Geology* (New York: Chapman & Hall, 1994).

Krumbein, W. C., and F. A. Graybill, *An Introduction to Statistical Models in Geology* (New York: McGraw-Hill, 1965).

Legendre, P., and L. Legendre, *Numerical Ecology* (Amsterdam: Elsevier, 1998).

Unwin, David J., *An Introduction to Trend Surface Analysis*, Concepts and Techniques in Modern Geography, No. 5 (Norwich, England: Geo Books, 1975).

PRACTICE PROBLEM

When the monthly sales of a product are subject to seasonal fluctuations, a curve that approximates the sales data might have the form

$$y = \beta_0 + \beta_1 x + \beta_2 \sin(2\pi x/12)$$

where x is the time in months. The term $\beta_0 + \beta_1 x$ gives the basic sales trend, and the sine term reflects the seasonal changes in sales. Give the design matrix and the parameter vector for the linear model that leads to a least-squares fit of the equation above. Assume the data are $(x_1, y_1), \dots, (x_n, y_n)$.

6.6 EXERCISES

In Exercises 1–4, find the equation $y = \beta_0 + \beta_1 x$ of the least-squares line that best fits the given data points.

1. $(0, 1), (1, 1), (2, 2), (3, 2)$

2. $(1, 0), (2, 1), (4, 2), (5, 3)$

3. $(-1, 0), (0, 1), (1, 2), (2, 4)$

4. $(2, 3), (3, 2), (5, 1), (6, 0)$

5. Let X be the design matrix used to find the least-squares line to fit data $(x_1, y_1), \ldots, (x_n, y_n)$. Use a theorem in Section 6.5 to show that the normal equations have a unique solution if and only if the data include at least two data points with different x-coordinates.

6. Let X be the design matrix in Example 2 corresponding to a least-squares fit of a parabola to data $(x_1, y_1), \ldots, (x_n, y_n)$. Suppose x_1, x_2, and x_3 are distinct. Explain why there is only one parabola that fits the data best, in a least-squares sense. (See Exercise 5.)

7. A certain experiment produces the data $(1, 1.8)$, $(2, 2.7)$, $(3, 3.4)$, $(4, 3.8)$, $(5, 3.9)$. Describe the model that produces a least-squares fit of these points by a function of the form

$$y = \beta_1 x + \beta_2 x^2$$

Such a function might arise, for example, as the revenue from the sale of x units of a product, when the amount offered for sale affects the price to be set for the product.

a. Give the design matrix, the observation vector, and the unknown parameter vector.

b. **[M]** Find the associated least-squares curve for the data.

8. A simple curve that often makes a good model for the variable costs of a company, as a function of the sales level x, has the form $y = \beta_1 x + \beta_2 x^2 + \beta_3 x^3$. There is no constant term because fixed costs are not included.

a. Give the design matrix and the parameter vector for the linear model that leads to a least-squares fit of the equation above, with data $(x_1, y_1), \ldots, (x_n, y_n)$.

b. **[M]** Find the least-squares curve of the form above to fit the data $(4, 1.58)$, $(6, 2.08)$, $(8, 2.5)$, $(10, 2.8)$, $(12, 3.1)$, $(14, 3.4)$, $(16, 3.8)$, and $(18, 4.32)$, with values in thousands. If possible, produce a graph that shows the data points and the graph of the cubic approximation.

9. A certain experiment produces the data $(1, 7.9)$, $(2, 5.4)$, and $(3, -.9)$. Describe the model that produces a least-squares fit of these points by a function of the form

$$y = A \cos x + B \sin x$$

10. Suppose radioactive substances A and B have decay constants of .02 and .07, respectively. If a mixture of these two substances at time $t = 0$ contains $M_{\rm A}$ grams of A and $M_{\rm B}$ grams of B, then a model for the total amount y of the mixture present at time t is

$$y = M_{\rm A} e^{-.02t} + M_{\rm B} e^{-.07t} \qquad (6)$$

Suppose the initial amounts $M_{\rm A}$ and $M_{\rm B}$ are unknown, but a scientist is able to measure the total amounts present at several times and records the following points (t_i, y_i): $(10, 21.34)$, $(11, 20.68)$, $(12, 20.05)$, $(14, 18.87)$, and $(15, 18.30)$.

a. Describe a linear model that can be used to estimate $M_{\rm A}$ and $M_{\rm B}$.

b. **[M]** Find the least-squares curve based on (6).

Halley's Comet last appeared in 1986 and will reappear in 2061.

11. **[M]** According to Kepler's first law, a comet should have an elliptic, parabolic, or hyperbolic orbit (with gravitational attractions from the planets ignored). In suitable polar coordinates, the position (r, ϑ) of a comet satisfies an equation of the form

$$r = \beta + e(r \cdot \cos \vartheta)$$

where β is a constant and e is the *eccentricity* of the orbit, with $0 \le e < 1$ for an ellipse, $e = 1$ for a parabola, and $e > 1$ for a hyperbola. Suppose observations of a newly discovered comet provide the data below. Determine the type of orbit, and predict where the comet will be when $\vartheta = 4.6$ (radians).[3]

ϑ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

12. **[M]** A healthy child's systolic blood pressure p (in millimeters of mercury) and weight w (in pounds) are approximately related by the equation

$$\beta_0 + \beta_1 \ln w = p$$

Use the following experimental data to estimate the systolic blood pressure of a healthy child weighing 100 pounds.

[3] The basic idea of least-squares fitting of data is due to K. F. Gauss (and, independently, to A. Legendre), whose initial rise to fame occurred in 1801 when he used the method to determine the path of the asteroid *Ceres*. Forty days after the asteroid was discovered, it disappeared behind the sun. Gauss predicted it would appear ten months later and gave its location. The accuracy of the prediction astonished the European scientific community.

w	44	61	81	113	131
$\ln w$	3.78	4.11	4.39	4.73	4.88
p	91	98	103	110	112

13. [**M**] To measure the takeoff performance of an airplane, the horizontal position of the plane was measured every second, from $t = 0$ to $t = 12$. The positions (in feet) were: 0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, and 809.2.

a. Find the least-squares cubic curve $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ for these data.

b. Use the result of part (a) to estimate the velocity of the plane when $t = 4.5$ seconds.

14. Let $\overline{x} = \frac{1}{n}(x_1 + \cdots + x_n)$ and $\overline{y} = \frac{1}{n}(y_1 + \cdots + y_n)$. Show that the least-squares line for the data $(x_1, y_1), \ldots, (x_n, y_n)$ must pass through $(\overline{x}, \overline{y})$. That is, show that $\overline{x}$ and $\overline{y}$ satisfy the linear equation $\overline{y} = \hat{\beta}_0 + \hat{\beta}_1\overline{x}$. [*Hint:* Derive this equation from the vector equation $\mathbf{y} = X\hat{\boldsymbol{\beta}} + \boldsymbol{\epsilon}$. Denote the first column of X by $\mathbf{1}$. Use the fact that the residual vector $\boldsymbol{\epsilon}$ is orthogonal to the column space of X and hence is orthogonal to $\mathbf{1}$.]

Given data for a least-squares problem, $(x_1, y_1), \ldots, (x_n, y_n)$, the following abbreviations are helpful:

$$\textstyle\sum x = \sum_{i=1}^n x_i, \quad \sum x^2 = \sum_{i=1}^n x_i^2,$$
$$\textstyle\sum y = \sum_{i=1}^n y_i, \quad \sum xy = \sum_{i=1}^n x_i y_i$$

The normal equations for a least-squares line $y = \hat{\beta}_0 + \hat{\beta}_1 x$ may be written in the form

$$\begin{aligned} n\hat{\beta}_0 + \hat{\beta}_1 \textstyle\sum x &= \textstyle\sum y \\ \hat{\beta}_0 \textstyle\sum x + \hat{\beta}_1 \textstyle\sum x^2 &= \textstyle\sum xy \end{aligned} \tag{7}$$

15. Derive the normal equations (7) from the matrix form given in this section.

16. Use a matrix inverse to solve the system of equations in (7) and thereby obtain formulas for $\hat{\beta}_0$ and $\hat{\beta}_1$ that appear in many statistics texts.

17. a. Rewrite the data in Example 1 with new x-coordinates in mean deviation form. Let X be the associated design matrix. Why are the columns of X orthogonal?

b. Write the normal equations for the data in part (a), and solve them to find the least-squares line, $y = \beta_0 + \beta_1 x^*$, where $x^* = x - 5.5$.

18. Suppose the x-coordinates of the data $(x_1, y_1), \ldots, (x_n, y_n)$ are in mean deviation form, so that $\sum x_i = 0$. Show that if X is the design matrix for the least-squares line in this case, then X^TX is a diagonal matrix.

Exercises 19 and 20 involve a design matrix X with two or more columns and a least-squares solution $\hat{\boldsymbol{\beta}}$ of $\mathbf{y} = X\boldsymbol{\beta}$. Consider the following numbers.

(i) $\|X\hat{\boldsymbol{\beta}}\|^2$—the sum of the squares of the "regression term." Denote this number by SS(R).

(ii) $\|\mathbf{y} - X\hat{\boldsymbol{\beta}}\|^2$—the sum of the squares for error term. Denote this number by SS(E).

(iii) $\|\mathbf{y}\|^2$—the "total" sum of the squares of the y-values. Denote this number by SS(T).

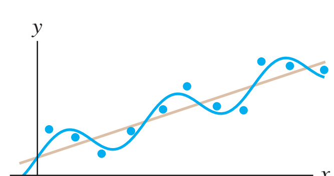
Every statistics text that discusses regression and the linear model $\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ introduces these numbers, though terminology and notation vary somewhat. To simplify matters, assume that the mean of the y-values is zero. In this case, SS(T) is proportional to what is called the *variance* of the set of y-values.

19. Justify the equation SS(T) = SS(R) + SS(E). [*Hint:* Use a theorem, and explain why the hypotheses of the theorem are satisfied.] This equation is extremely important in statistics, both in regression theory and in the analysis of variance.

20. Show that $\|X\hat{\boldsymbol{\beta}}\|^2 = \hat{\boldsymbol{\beta}}^T X^T\mathbf{y}$. [*Hint:* Rewrite the left side and use the fact that $\hat{\boldsymbol{\beta}}$ satisfies the normal equations.] This formula for SS(R) is used in statistics. From this and from Exercise 19, obtain the standard formula for SS(E):

$$\text{SS(E)} = \mathbf{y}^T\mathbf{y} - \hat{\boldsymbol{\beta}}^T X^T\mathbf{y}$$

SOLUTION TO PRACTICE PROBLEM



Sales trend with seasonal fluctuations.

Construct X and $\boldsymbol{\beta}$ so that the kth row of $X\boldsymbol{\beta}$ is the predicted y-value that corresponds to the data point (x_k, y_k), namely,

$$\beta_0 + \beta_1 x_k + \beta_2 \sin(2\pi x_k/12)$$

It should be clear that

$$X = \begin{bmatrix} 1 & x_1 & \sin(2\pi x_1/12) \\ \vdots & \vdots & \vdots \\ 1 & x_n & \sin(2\pi x_n/12) \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

So p_2 is already orthogonal to q_1, and we can take $q_2 = p_2$. For the projection of p_3 onto $W_2 = \text{Span}\{q_1, q_2\}$, compute

$$\langle p_3, q_1\rangle = \int_0^1 12t^2 \cdot 1\, dt = 4t^3\Big|_0^1 = 4$$

$$\langle q_1, q_1\rangle = \int_0^1 1 \cdot 1\, dt = t\Big|_0^1 = 1$$

$$\langle p_3, q_2\rangle = \int_0^1 12t^2(2t-1)\, dt = \int_0^1 (24t^3 - 12t^2)\, dt = 2$$

$$\langle q_2, q_2\rangle = \int_0^1 (2t-1)^2\, dt = \frac{1}{6}(2t-1)^3\Big|_0^1 = \frac{1}{3}$$

Then

$$\text{proj}_{W_2}\, p_3 = \frac{\langle p_3, q_1\rangle}{\langle q_1, q_1\rangle} q_1 + \frac{\langle p_3, q_2\rangle}{\langle q_2, q_2\rangle} q_2 = \frac{4}{1} q_1 + \frac{2}{1/3} q_2 = 4q_1 + 6q_2$$

and

$$q_3 = p_3 - \text{proj}_{W_2}\, p_3 = p_3 - 4q_1 - 6q_2$$

As a function, $q_3(t) = 12t^2 - 4 - 6(2t-1) = 12t^2 - 12t + 2$. The orthogonal basis for the subspace W is $\{q_1, q_2, q_3\}$. ■

PRACTICE PROBLEMS

Use the inner product axioms to verify the following statements.

1. $\langle \mathbf{v}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle = 0$.

2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$.

6.7 EXERCISES

1. Let $\mathbb{R}^2$ have the inner product of Example 1, and let $\mathbf{x} = (1, 1)$ and $\mathbf{y} = (5, -1)$.

a. Find $\|\mathbf{x}\|$, $\|\mathbf{y}\|$, and $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$.

b. Describe all vectors (z_1, z_2) that are orthogonal to $\mathbf{y}$.

2. Let $\mathbb{R}^2$ have the inner product of Example 1. Show that the Cauchy–Schwarz inequality holds for $\mathbf{x} = (3, -2)$ and $\mathbf{y} = (-2, 1)$. [*Suggestion:* Study $|\langle \mathbf{x}, \mathbf{y}\rangle|^2$.]

Exercises 3–8 refer to $\mathbb{P}_2$ with the inner product given by evaluation at -1, 0, and 1. (See Example 2.)

3. Compute $\langle p, q\rangle$, where $p(t) = 4 + t$, $q(t) = 5 - 4t^2$.

4. Compute $\langle p, q\rangle$, where $p(t) = 3t - t^2$, $q(t) = 3 + 2t^2$.

5. Compute $\|p\|$ and $\|q\|$, for p and q in Exercise 3.

6. Compute $\|p\|$ and $\|q\|$, for p and q in Exercise 4.

7. Compute the orthogonal projection of q onto the subspace spanned by p, for p and q in Exercise 3.

8. Compute the orthogonal projection of q onto the subspace spanned by p, for p and q in Exercise 4.

9. Let $\mathbb{P}_3$ have the inner product given by evaluation at -3, -1, 1, and 3. Let $p_0(t) = 1$, $p_1(t) = t$, and $p_2(t) = t^2$.

a. Compute the orthogonal projection of p_2 onto the subspace spanned by p_0 and p_1.

b. Find a polynomial q that is orthogonal to p_0 and p_1, such that $\{p_0, p_1, q\}$ is an orthogonal basis for $\text{Span}\{p_0, p_1, p_2\}$. Scale the polynomial q so that its vector of values at $(-3, -1, 1, 3)$ is $(1, -1, -1, 1)$.

10. Let $\mathbb{P}_3$ have the inner product as in Exercise 9, with p_0, p_1, and q the polynomials described there. Find the best approximation to $p(t) = t^3$ by polynomials in $\text{Span}\{p_0, p_1, q\}$.

11. Let p_0, p_1, and p_2 be the orthogonal polynomials described in Example 5, where the inner product on $\mathbb{P}_4$ is given by evaluation at -2, -1, 0, 1, and 2. Find the orthogonal projection of t^3 onto $\text{Span}\{p_0, p_1, p_2\}$.

12. Find a polynomial p_3 such that $\{p_0, p_1, p_2, p_3\}$ (see Exercise 11) is an orthogonal basis for the subspace $\mathbb{P}_3$ of $\mathbb{P}_4$. Scale the polynomial p_3 so that its vector of values is $(-1, 2, 0, -2, 1)$.

13. Let A be any invertible $n \times n$ matrix. Show that for $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^n$, the formula $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot (A\mathbf{v}) = (A\mathbf{u})^T(A\mathbf{v})$ defines an inner product on $\mathbb{R}^n$.

14. Let T be a one-to-one linear transformation from a vector space V into $\mathbb{R}^n$. Show that for $\mathbf{u}$, $\mathbf{v}$ in V, the formula $\langle \mathbf{u}, \mathbf{v} \rangle = T(\mathbf{u}) \cdot T(\mathbf{v})$ defines an inner product on V.

Use the inner product axioms and other results of this section to verify the statements in Exercises 15–18.

15. $\langle \mathbf{u}, c\mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$ for all scalars c.

16. If $\{\mathbf{u}, \mathbf{v}\}$ is an orthonormal set in V, then $\|\mathbf{u} - \mathbf{v}\| = \sqrt{2}$.

17. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$.

18. $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

19. Given $a \geq 0$ and $b \geq 0$, let $\mathbf{u} = \begin{bmatrix} \sqrt{a} \\ \sqrt{b} \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} \sqrt{b} \\ \sqrt{a} \end{bmatrix}$. Use the Cauchy–Schwarz inequality to compare the geometric mean $\sqrt{ab}$ with the arithmetic mean $(a + b)/2$.

20. Let $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Use the Cauchy–Schwarz inequality to show that

$$\left(\frac{a+b}{2}\right)^2 \leq \frac{a^2 + b^2}{2}$$

Exercises 21–24 refer to $V = C[0, 1]$, with the inner product given by an integral, as in Example 7.

21. Compute $\langle f, g \rangle$, where $f(t) = 1 - 3t^2$ and $g(t) = t - t^3$.

22. Compute $\langle f, g \rangle$, where $f(t) = 5t - 3$ and $g(t) = t^3 - t^2$.

23. Compute $\|f\|$ for f in Exercise 21.

24. Compute $\|g\|$ for g in Exercise 22.

25. Let V be the space $C[-1, 1]$ with the inner product of Example 7. Find an orthogonal basis for the subspace spanned by the polynomials 1, t, and t^2. The polynomials in this basis are called *Legendre polynomials*.

26. Let V be the space $C[-2, 2]$ with the inner product of Example 7. Find an orthogonal basis for the subspace spanned by the polynomials 1, t, and t^2.

27. [**M**] Let $\mathbb{P}_4$ have the inner product as in Example 5, and let p_0, p_1, p_2 be the orthogonal polynomials from that example. Using your matrix program, apply the Gram–Schmidt process to the set $\{p_0, p_1, p_2, t^3, t^4\}$ to create an orthogonal basis for $\mathbb{P}_4$.

28. [**M**] Let V be the space $C[0, 2\pi]$ with the inner product of Example 7. Use the Gram–Schmidt process to create an orthogonal basis for the subspace spanned by $\{1, \cos t, \cos^2 t, \cos^3 t\}$. Use a matrix program or computational program to compute the appropriate definite integrals.

SOLUTIONS TO PRACTICE PROBLEMS

1. By Axiom 1, $\langle \mathbf{v}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{v} \rangle$. Then $\langle \mathbf{0}, \mathbf{v} \rangle = \langle 0\mathbf{v}, \mathbf{v} \rangle = 0\langle \mathbf{v}, \mathbf{v} \rangle$, by Axiom 3, so $\langle \mathbf{0}, \mathbf{v} \rangle = 0$.

2. By Axioms 1, 2, and then 1 again, $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.

6.8 APPLICATIONS OF INNER PRODUCT SPACES

The examples in this section suggest how the inner product spaces defined in Section 6.7 arise in practical problems. The first example is connected with the massive least-squares problem of updating the North American Datum, described in the chapter's introductory example.

Weighted Least-Squares

Let $\mathbf{y}$ be a vector of n observations, $y_1, \dots, y_n$, and suppose we wish to approximate $\mathbf{y}$ by a vector $\hat{\mathbf{y}}$ that belongs to some specified subspace of $\mathbb{R}^n$. (In Section 6.5, $\hat{\mathbf{y}}$ was written as $A\mathbf{x}$ so that $\hat{\mathbf{y}}$ was in the column space of A.) Denote the entries in $\hat{\mathbf{y}}$ by $\hat{y}_1, \dots, \hat{y}_n$. Then the *sum of the squares for error*, or SS(E), in approximating $\mathbf{y}$ by $\hat{\mathbf{y}}$ is

$$\text{SS(E)} = (y_1 - \hat{y}_1)^2 + \cdots + (y_n - \hat{y}_n)^2 \tag{1}$$

This is simply $\|\mathbf{y} - \hat{\mathbf{y}}\|^2$, using the standard length in $\mathbb{R}^n$.

EXAMPLE 4 Find the nth-order Fourier approximation to the function $f(t) = t$ on the interval $[0, 2\pi]$.

SOLUTION Compute

$$\frac{a_0}{2} = \frac{1}{2}\cdot\frac{1}{\pi}\int_0^{2\pi} t\,dt = \frac{1}{2\pi}\left[\frac{1}{2}t^2\Big|_0^{2\pi}\right] = \pi$$

and for $k > 0$, using integration by parts,

$$a_k = \frac{1}{\pi}\int_0^{2\pi} t\cos kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\cos kt + \frac{t}{k}\sin kt\right]_0^{2\pi} = 0$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} t\sin kt\,dt = \frac{1}{\pi}\left[\frac{1}{k^2}\sin kt - \frac{t}{k}\cos kt\right]_0^{2\pi} = -\frac{2}{k}$$

Thus the nth-order Fourier approximation of $f(t) = t$ is

$$\pi - 2\sin t - \sin 2t - \frac{2}{3}\sin 3t - \cdots - \frac{2}{n}\sin nt$$

Figure 3 shows the third- and fourth-order Fourier approximations of f. ■

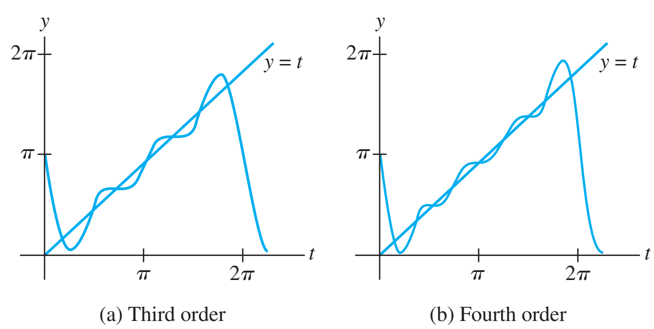


FIGURE 3 Fourier approximations of the function $f(t) = t$.

The norm of the difference between f and a Fourier approximation is called the **mean square error** in the approximation. (The term *mean* refers to the fact that the norm is determined by an integral.) It can be shown that the mean square error approaches zero as the order of the Fourier approximation increases. For this reason, it is common to write

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty}(a_m \cos mt + b_m \sin mt)$$

This expression for $f(t)$ is called the **Fourier series** for f on $[0, 2\pi]$. The term $a_m \cos mt$, for example, is the projection of f onto the one-dimensional subspace spanned by $\cos mt$.

PRACTICE PROBLEMS

1. Let $q_1(t) = 1$, $q_2(t) = t$, and $q_3(t) = 3t^2 - 4$. Verify that $\{q_1, q_2, q_3\}$ is an orthogonal set in $C[-2, 2]$ with the inner product of Example 7 in Section 6.7 (integration from -2 to 2).
2. Find the first-order and third-order Fourier approximations to

$$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$$

6.8 EXERCISES

1. Find the least-squares line $y = \beta_0 + \beta_1 x$ that best fits the data $(-2, 0)$, $(-1, 0)$, $(0, 2)$, $(1, 4)$, and $(2, 4)$, assuming that the first and last data points are less reliable. Weight them half as much as the three interior points.

2. Suppose 5 out of 25 data points in a weighted least-squares problem have a y-measurement that is less reliable than the others, and they are to be weighted half as much as the other 20 points. One method is to weight the 20 points by a factor of 1 and the other 5 by a factor of $\frac{1}{2}$. A second method is to weight the 20 points by a factor of 2 and the other 5 by a factor of 1. Do the two methods produce different results? Explain.

3. Fit a cubic trend function to the data in Example 2. The orthogonal cubic polynomial is $p_3(t) = \frac{5}{6}t^3 - \frac{17}{6}t$.

4. To make a trend analysis of six evenly spaced data points, one can use orthogonal polynomials with respect to evaluation at the points $t = -5, -3, -1, 1, 3$, and 5.

   a. Show that the first three orthogonal polynomials are

   $$p_0(t) = 1, \quad p_1(t) = t, \quad \text{and} \quad p_2(t) = \tfrac{3}{8}t^2 - \tfrac{35}{8}$$

   (The polynomial p_2 has been scaled so that its values at the evaluation points are small integers.)

   b. Fit a quadratic trend function to the data

   $$(-5, 1), (-3, 1), (-1, 4), (1, 4), (3, 6), (5, 8)$$

In Exercises 5–14, the space is $C[0, 2\pi]$ with the inner product (6).

5. Show that $\sin mt$ and $\sin nt$ are orthogonal when $m \neq n$.

6. Show that $\sin mt$ and $\cos nt$ are orthogonal for all positive integers m and n.

7. Show that $\|\cos kt\|^2 = \pi$ and $\|\sin kt\|^2 = \pi$ for $k > 0$.

8. Find the third-order Fourier approximation to $f(t) = t - 1$.

9. Find the third-order Fourier approximation to $f(t) = 2\pi - t$.

10. Find the third-order Fourier approximation to the *square wave function*, $f(t) = 1$ for $0 \leq t < \pi$ and $f(t) = -1$ for $\pi \leq t < 2\pi$.

11. Find the third-order Fourier approximation to $\sin^2 t$, without performing any integration calculations.

12. Find the third-order Fourier approximation to $\cos^3 t$, without performing any integration calculations.

13. Explain why a Fourier coefficient of the sum of two functions is the sum of the corresponding Fourier coefficients of the two functions.

14. Suppose the first few Fourier coefficients of some function f in $C[0, 2\pi]$ are a_0, a_1, a_2, and b_1, b_2, b_3. Which of the following trigonometric polynomials is closer to f? Defend your answer.

$$g(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \sin t$$

$$h(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \sin t + b_2 \sin 2t$$

15. **[M]** Refer to the data in Exercise 13 in Section 6.6, concerning the takeoff performance of an airplane. Suppose the possible measurement errors become greater as the speed of the airplane increases, and let W be the diagonal weighting matrix whose diagonal entries are 1, 1, 1, .9, .9, .8, .7, .6, .5, .4, .3, .2, and .1. Find the cubic curve that fits the data with minimum weighted least-squares error, and use it to estimate the velocity of the plane when $t = 4.5$ seconds.

16. **[M]** Let f_4 and f_5 be the fourth-order and fifth-order Fourier approximations in $C[0, 2\pi]$ to the square wave function in Exercise 10. Produce separate graphs of f_4 and f_5 on the interval $[0, 2\pi]$, and produce a graph of f_5 on $[-2\pi, 2\pi]$.

SG The Linearity of an Orthogonal Projection 6–25

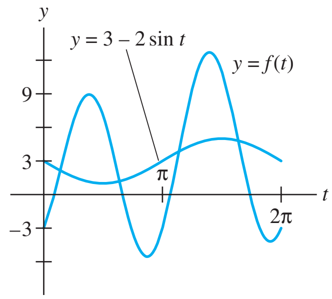
SOLUTIONS TO PRACTICE PROBLEMS

1. Compute

$$\langle q_1, q_2 \rangle = \int_{-2}^{2} 1 \cdot t \, dt = \frac{1}{2}t^2 \Big|_{-2}^{2} = 0$$

$$\langle q_1, q_3 \rangle = \int_{-2}^{2} 1 \cdot (3t^2 - 4) \, dt = (t^3 - 4t) \Big|_{-2}^{2} = 0$$

$$\langle q_2, q_3 \rangle = \int_{-2}^{2} t \cdot (3t^2 - 4) \, dt = \left(\frac{3}{4}t^4 - 2t^2 \right) \Big|_{-2}^{2} = 0$$



First- and third-order approximations to $f(t)$.

2. The third-order Fourier approximation to f is the best approximation in $C[0, 2\pi]$ to f by functions (vectors) in the subspace spanned by 1, $\cos t$, $\cos 2t$, $\cos 3t$, $\sin t$, $\sin 2t$, and $\sin 3t$. But f is obviously *in* this subspace, so f is its own best approximation:

$$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$$

For the first-order approximation, the closest function to f in the subspace $W = \text{Span}\{1, \cos t, \sin t\}$ is $3 - 2\sin t$. The other two terms in the formula for $f(t)$ are orthogonal to the functions in W, so they contribute nothing to the integrals that give the Fourier coefficients for a first-order approximation.

CHAPTER 6 SUPPLEMENTARY EXERCISES

1. The following statements refer to vectors in $\mathbb{R}^n$ (or $\mathbb{R}^m$) with the standard inner product. Mark each statement True or False. Justify each answer.

a. The length of every vector is a positive number.

b. A vector $\mathbf{v}$ and its negative $-\mathbf{v}$ have equal lengths.

c. The distance between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$.

d. If r is any scalar, then $\|r\mathbf{v}\| = r\|\mathbf{v}\|$.

e. If two vectors are orthogonal, they are linearly independent.

f. If $\mathbf{x}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, then $\mathbf{x}$ must be orthogonal to $\mathbf{u} - \mathbf{v}$.

g. If $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

h. If $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

i. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ is a scalar multiple of $\mathbf{y}$.

j. If a vector $\mathbf{y}$ coincides with its orthogonal projection onto a subspace W, then $\mathbf{y}$ is in W.

k. The set of all vectors in $\mathbb{R}^n$ orthogonal to one fixed vector is a subspace of $\mathbb{R}^n$.

l. If W is a subspace of $\mathbb{R}^n$, then W and $W^\perp$ have no vectors in common.

m. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set and if c_1, c_2, and c_3 are scalars, then $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, c_3\mathbf{v}_3\}$ is an orthogonal set.

n. If a matrix U has orthonormal columns, then $UU^T = I$.

o. A square matrix with orthogonal columns is an orthogonal matrix.

p. If a square matrix has orthonormal columns, then it also has orthonormal rows.

q. If W is a subspace, then $\|\text{proj}_W \mathbf{v}\|^2 + \|\mathbf{v} - \text{proj}_W \mathbf{v}\|^2 = \|\mathbf{v}\|^2$.

r. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is the vector $A\hat{\mathbf{x}}$ in Col A closest to $\mathbf{b}$, so that $\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$ for all $\mathbf{x}$.

s. The normal equations for a least-squares solution of $A\mathbf{x} = \mathbf{b}$ are given by $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

2. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be an orthonormal set. Verify the following equality by induction, beginning with $p = 2$. If $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, then

$$\|\mathbf{x}\|^2 = |c_1|^2 + \cdots + |c_p|^2$$

3. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be an orthonormal set in $\mathbb{R}^n$. Verify the following inequality, called *Bessel's inequality*, which is true for each $\mathbf{x}$ in $\mathbb{R}^n$:

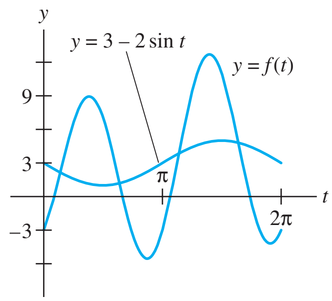
$$\|\mathbf{x}\|^2 \geq |\mathbf{x}\cdot\mathbf{v}_1|^2 + |\mathbf{x}\cdot\mathbf{v}_2|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_p|^2$$

4. Let U be an $n \times n$ orthogonal matrix. Show that if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then so is $\{U\mathbf{v}_1, \dots, U\mathbf{v}_n\}$.

5. Show that if an $n \times n$ matrix U satisfies $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, then U is an orthogonal matrix.

6. Show that if U is an orthogonal matrix, then any real eigenvalue of U must be ± 1.

7. A *Householder matrix*, or an *elementary reflector*, has the form $Q = I - 2\mathbf{u}\mathbf{u}^T$ where $\mathbf{u}$ is a unit vector. (See Exercise 13 in the Supplementary Exercises for Chapter 2.) Show that Q is an orthogonal matrix. (Elementary reflectors are often used in computer programs to produce a QR factorization of a matrix A. If A has linearly independent columns, then left-multiplication by a sequence of elementary reflectors can produce an upper triangular matrix.)



First- and third-order approximations to $f(t)$.

2. The third-order Fourier approximation to f is the best approximation in $C[0, 2\pi]$ to f by functions (vectors) in the subspace spanned by 1, $\cos t$, $\cos 2t$, $\cos 3t$, $\sin t$, $\sin 2t$, and $\sin 3t$. But f is obviously *in* this subspace, so f is its own best approximation:

$$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$$

For the first-order approximation, the closest function to f in the subspace $W = \text{Span}\{1, \cos t, \sin t\}$ is $3 - 2\sin t$. The other two terms in the formula for $f(t)$ are orthogonal to the functions in W, so they contribute nothing to the integrals that give the Fourier coefficients for a first-order approximation.

CHAPTER 6 SUPPLEMENTARY EXERCISES

1. The following statements refer to vectors in $\mathbb{R}^n$ (or $\mathbb{R}^m$) with the standard inner product. Mark each statement True or False. Justify each answer.

a. The length of every vector is a positive number.

b. A vector $\mathbf{v}$ and its negative $-\mathbf{v}$ have equal lengths.

c. The distance between $\mathbf{u}$ and $\mathbf{v}$ is $\|\mathbf{u} - \mathbf{v}\|$.

d. If r is any scalar, then $\|r\mathbf{v}\| = r\|\mathbf{v}\|$.

e. If two vectors are orthogonal, they are linearly independent.

f. If $\mathbf{x}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$, then $\mathbf{x}$ must be orthogonal to $\mathbf{u} - \mathbf{v}$.

g. If $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

h. If $\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$, then $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

i. The orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$ is a scalar multiple of $\mathbf{y}$.

j. If a vector $\mathbf{y}$ coincides with its orthogonal projection onto a subspace W, then $\mathbf{y}$ is in W.

k. The set of all vectors in $\mathbb{R}^n$ orthogonal to one fixed vector is a subspace of $\mathbb{R}^n$.

l. If W is a subspace of $\mathbb{R}^n$, then W and $W^\perp$ have no vectors in common.

m. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set and if c_1, c_2, and c_3 are scalars, then $\{c_1\mathbf{v}_1, c_2\mathbf{v}_2, c_3\mathbf{v}_3\}$ is an orthogonal set.

n. If a matrix U has orthonormal columns, then $UU^T = I$.

o. A square matrix with orthogonal columns is an orthogonal matrix.

p. If a square matrix has orthonormal columns, then it also has orthonormal rows.

q. If W is a subspace, then $\|\text{proj}_W \mathbf{v}\|^2 + \|\mathbf{v} - \text{proj}_W \mathbf{v}\|^2 = \|\mathbf{v}\|^2$.

r. A least-squares solution of $A\mathbf{x} = \mathbf{b}$ is the vector $A\hat{\mathbf{x}}$ in Col A closest to $\mathbf{b}$, so that $\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\|$ for all $\mathbf{x}$.

s. The normal equations for a least-squares solution of $A\mathbf{x} = \mathbf{b}$ are given by $\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

2. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be an orthonormal set. Verify the following equality by induction, beginning with $p = 2$. If $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_p\mathbf{v}_p$, then

$$\|\mathbf{x}\|^2 = |c_1|^2 + \cdots + |c_p|^2$$

3. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ be an orthonormal set in $\mathbb{R}^n$. Verify the following inequality, called *Bessel's inequality*, which is true for each $\mathbf{x}$ in $\mathbb{R}^n$:

$$\|\mathbf{x}\|^2 \ge |\mathbf{x}\cdot\mathbf{v}_1|^2 + |\mathbf{x}\cdot\mathbf{v}_2|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_p|^2$$

4. Let U be an $n \times n$ orthogonal matrix. Show that if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then so is $\{U\mathbf{v}_1, \dots, U\mathbf{v}_n\}$.

5. Show that if an $n \times n$ matrix U satisfies $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$, then U is an orthogonal matrix.

6. Show that if U is an orthogonal matrix, then any real eigenvalue of U must be ± 1.

7. A *Householder matrix*, or an *elementary reflector*, has the form $Q = I - 2\mathbf{u}\mathbf{u}^T$ where $\mathbf{u}$ is a unit vector. (See Exercise 13 in the Supplementary Exercises for Chapter 2.) Show that Q is an orthogonal matrix. (Elementary reflectors are often used in computer programs to produce a QR factorization of a matrix A. If A has linearly independent columns, then left-multiplication by a sequence of elementary reflectors can produce an upper triangular matrix.)

8. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation that preserves lengths; that is, $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

a. Show that T also preserves orthogonality; that is, $T(\mathbf{x})\cdot T(\mathbf{y}) = 0$ whenever $\mathbf{x}\cdot\mathbf{y} = 0$.

b. Show that the standard matrix of T is an orthogonal matrix.

9. Let $\mathbf{u}$ and $\mathbf{v}$ be linearly independent vectors in $\mathbb{R}^n$ that are *not* orthogonal. Describe how to find the best approximation to $\mathbf{z}$ in $\mathbb{R}^n$ by vectors of the form $x_1\mathbf{u} + x_2\mathbf{v}$ without first constructing an orthogonal basis for Span $\{\mathbf{u}, \mathbf{v}\}$.

10. Suppose the columns of A are linearly independent. Determine what happens to the least-squares solution $\hat{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$ when $\mathbf{b}$ is replaced by $c\mathbf{b}$ for some nonzero scalar c.

11. If a, b, and c are distinct numbers, then the following system is inconsistent because the graphs of the equations are parallel planes. Show that the set of all least-squares solutions of the system is precisely the plane whose equation is $x - 2y + 5z = (a + b + c)/3$.

$$\begin{aligned} x - 2y + 5z &= a \\ x - 2y + 5z &= b \\ x - 2y + 5z &= c \end{aligned}$$

12. Consider the problem of finding an eigenvalue of an $n \times n$ matrix A when an approximate eigenvector $\mathbf{v}$ is known. Since $\mathbf{v}$ is not exactly correct, the equation

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

will probably not have a solution. However, λ can be estimated by a least-squares solution when (1) is viewed properly. Think of $\mathbf{v}$ as an $n \times 1$ matrix V, think of λ as a vector in $\mathbb{R}^1$, and denote the vector $A\mathbf{v}$ by the symbol $\mathbf{b}$. Then (1) becomes $\mathbf{b} = \lambda V$, which may also be written as $V\lambda = \mathbf{b}$. Find the least-squares solution of this system of n equations in the one unknown λ, and write this solution using the original symbols. The resulting estimate for λ is called a *Rayleigh quotient*. See Exercises 11 and 12 in Section 5.8.

13. Use the steps below to prove the following relations among the four fundamental subspaces determined by an $m \times n$ matrix A.

$$\text{Row } A = (\text{Nul } A)^{\perp}, \quad \text{Col } A = (\text{Nul } A^T)^{\perp}$$

a. Show that Row A is contained in $(\text{Nul } A)^{\perp}$. (Show that if $\mathbf{x}$ is in Row A, then $\mathbf{x}$ is orthogonal to every $\mathbf{u}$ in Nul A.)

b. Suppose rank $A = r$. Find dim Nul A and dim $(\text{Nul } A)^{\perp}$, and then deduce from part (a) that Row $A = (\text{Nul } A)^{\perp}$. [*Hint:* Study the exercises for Section 6.3.]

c. Explain why Col $A = (\text{Nul } A^T)^{\perp}$.

14. Explain why an equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is orthogonal to all solutions of the equation $A^T\mathbf{x} = \mathbf{0}$.

Exercises 15 and 16 concern the (real) *Schur factorization* of an $n \times n$ matrix A in the form $A = URU^T$, where U is an orthogonal matrix and R is an $n \times n$ upper triangular matrix.[1]

15. Show that if A admits a (real) Schur factorization, $A = URU^T$, then A has n real eigenvalues, counting multiplicities.

16. Let A be an $n \times n$ matrix with n real eigenvalues, counting multiplicities, denoted by $\lambda_1, \dots, \lambda_n$. It can be shown that A admits a (real) Schur factorization. Parts (a) and (b) show the key ideas in the proof. The rest of the proof amounts to repeating (a) and (b) for successively smaller matrices, and then piecing together the results.

a. Let $\mathbf{u}_1$ be a unit eigenvector corresponding to λ_1, let $\mathbf{u}_2, \dots, \mathbf{u}_n$ be any other vectors such that $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis for $\mathbb{R}^n$, and then let $U = [\,\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \cdots \;\; \mathbf{u}_n\,]$. Show that the first column of U^TAU is $\lambda_1\mathbf{e}_1$, where $\mathbf{e}_1$ is the first column of the $n \times n$ identity matrix.

b. Part (a) implies that U^TAU has the form shown below. Explain why the eigenvalues of A_1 are $\lambda_2, \dots, \lambda_n$. [*Hint:* See the Supplementary Exercises for Chapter 5.]

$$U^TAU = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & & & & \\ \vdots & & A_1 & & \\ 0 & & & & \end{bmatrix}$$

[**M**] When the right side of an equation $A\mathbf{x} = \mathbf{b}$ is changed slightly—say, to $A\mathbf{x} = \mathbf{b} + \Delta\mathbf{b}$ for some vector $\Delta\mathbf{b}$—the solution changes from $\mathbf{x}$ to $\mathbf{x} + \Delta\mathbf{x}$, where $\Delta\mathbf{x}$ satisfies $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$. The quotient $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ is called the **relative change** in $\mathbf{b}$ (or the **relative error** in $\mathbf{b}$ when $\Delta\mathbf{b}$ represents possible error in the entries of $\mathbf{b}$). The relative change in the solution is $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$. When A is invertible, the **condition number** of A, written as cond(A), produces a bound on how large the relative change in $\mathbf{x}$ can be:

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \text{cond}(A)\cdot\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|} \tag{2}$$

In Exercises 17–20, solve $A\mathbf{x} = \mathbf{b}$ and $A(\Delta\mathbf{x}) = \Delta\mathbf{b}$, and show that the inequality (2) holds in each case. (See the discussion of *ill-conditioned* matrices in Exercises 41–43 in Section 2.3.)

17. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}$, $\Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

18. $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .500 \\ -1.407 \end{bmatrix}$, $\Delta\mathbf{b} = \begin{bmatrix} .001 \\ -.003 \end{bmatrix}$

[1] If complex numbers are allowed, *every* $n \times n$ matrix A admits a (complex) Schur factorization, $A = URU^{-1}$, where R is upper triangular and U^{-1} is the *conjugate* transpose of U. This very useful fact is discussed in *Matrix Analysis*, by Roger A. Horn and Charles R. Johnson (Cambridge: Cambridge University Press, 1985), pp. 79–100.

19. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} .100 \\ 2.888 \\ -1.404 \\ 1.462 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .49 \\ -1.28 \\ 5.78 \\ 8.04 \end{bmatrix}$

20. $A = \begin{bmatrix} 7 & -6 & -4 & 1 \\ -5 & 1 & 0 & -2 \\ 10 & 11 & 7 & -3 \\ 19 & 9 & 7 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 4.230 \\ -11.043 \\ 49.991 \\ 69.536 \end{bmatrix}$,

$\Delta\mathbf{b} = 10^{-4}\begin{bmatrix} .27 \\ 7.76 \\ -3.77 \\ 3.93 \end{bmatrix}$

NUMERICAL NOTE

When A is symmetric and not too large, modern high-performance computer algorithms calculate eigenvalues and eigenvectors with great precision. They apply a sequence of similarity transformations to A involving orthogonal matrices. The diagonal entries of the transformed matrices converge rapidly to the eigenvalues of A. (See the Numerical Notes in Section 5.2.) Using orthogonal matrices generally prevents numerical errors from accumulating during the process. When A is symmetric, the sequence of orthogonal matrices combines to form an orthogonal matrix whose columns are eigenvectors of A.

A nonsymmetric matrix cannot have a full set of orthogonal eigenvectors, but the algorithm still produces fairly accurate eigenvalues. After that, nonorthogonal techniques are needed to calculate eigenvectors.

PRACTICE PROBLEMS

1. Show that if A is a symmetric matrix, then A^2 is symmetric.
2. Show that if A is orthogonally diagonalizable, then so is A^2.

7.1 EXERCISES

Determine which of the matrices in Exercises 1–6 are symmetric.

1. $\begin{bmatrix} 3 & 5 \\ 5 & -7 \end{bmatrix}$ **2.** $\begin{bmatrix} -3 & 5 \\ -5 & 3 \end{bmatrix}$

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

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

Determine which of the matrices in Exercises 7–12 are orthogonal. If orthogonal, find the inverse.

7. $\begin{bmatrix} .6 & .8 \\ .8 & -.6 \end{bmatrix}$ **8.** $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

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

11. $\begin{bmatrix} 2/3 & 2/3 & 1/3 \\ 0 & 1/\sqrt{5} & -2/\sqrt{5} \\ \sqrt{5}/3 & -4/\sqrt{45} & -2/\sqrt{45} \end{bmatrix}$

12. $\begin{bmatrix} .5 & .5 & -.5 & -.5 \\ -.5 & .5 & -.5 & .5 \\ .5 & .5 & .5 & .5 \\ -.5 & .5 & .5 & -.5 \end{bmatrix}$

Orthogonally diagonalize the matrices in Exercises 13–22, giving an orthogonal matrix P and a diagonal matrix D. To save you time, the eigenvalues in Exercises 17–22 are: (17) 5, 2, −2; (18) 25, 3, −50; (19) 7, −2; (20) 13, 7, 1; (21) 9, 5, 1; (22) 2, 0.

13. $\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ **14.** $\begin{bmatrix} 1 & 5 \\ 5 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 16 & -4 \\ -4 & 1 \end{bmatrix}$ **16.** $\begin{bmatrix} -7 & 24 \\ 24 & 7 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ **18.** $\begin{bmatrix} -2 & -36 & 0 \\ -36 & -23 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

19. $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ **20.** $\begin{bmatrix} 7 & -4 & 4 \\ -4 & 5 & 0 \\ 4 & 0 & 9 \end{bmatrix}$

21. $\begin{bmatrix} 4 & 1 & 3 & 1 \\ 1 & 4 & 1 & 3 \\ 3 & 1 & 4 & 1 \\ 1 & 3 & 1 & 4 \end{bmatrix}$ **22.** $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

23. Let $A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Verify that 2 is an eigenvalue of A and $\mathbf{v}$ is an eigenvector. Then orthogonally diagonalize A.

24. Let $A = \begin{bmatrix} 5 & -4 & -2 \\ -4 & 5 & 2 \\ -2 & 2 & 2 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. Verify that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of A. Then orthogonally diagonalize A.

In Exercises 25 and 26, mark each statement True or False. Justify each answer.

25. a. An $n \times n$ matrix that is orthogonally diagonalizable must be symmetric.

b. If $A^T = A$ and if vectors $\mathbf{u}$ and $\mathbf{v}$ satisfy $A\mathbf{u} = 3\mathbf{u}$ and $A\mathbf{v} = 4\mathbf{v}$, then $\mathbf{u} \cdot \mathbf{v} = 0$.

c. An $n \times n$ symmetric matrix has n distinct real eigenvalues.

d. For a nonzero $\mathbf{v}$ in $\mathbb{R}^n$, the matrix $\mathbf{v}\mathbf{v}^T$ is called a projection matrix.

26. a. Every symmetric matrix is orthogonally diagonalizable.

b. If $B = PDP^T$, where $P^T = P^{-1}$ and D is a diagonal matrix, then B is a symmetric matrix.

c. An orthogonal matrix is orthogonally diagonalizable.

d. The dimension of an eigenspace of a symmetric matrix equals the multiplicity of the corresponding eigenvalue.

27. Suppose A is a symmetric $n \times n$ matrix and B is any $n \times m$ matrix. Show that B^TAB, B^TB, and BB^T are symmetric matrices.

28. Show that if A is an $n \times n$ symmetric matrix, then $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$ for all $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$.

29. Suppose A is invertible and orthogonally diagonalizable. Explain why A^{-1} is also orthogonally diagonalizable.

30. Suppose A and B are both orthogonally diagonalizable and $AB = BA$. Explain why AB is also orthogonally diagonalizable.

31. Let $A = PDP^{-1}$, where P is orthogonal and D is diagonal, and let λ be an eigenvalue of A of multiplicity k. Then λ appears k times on the diagonal of D. Explain why the dimension of the eigenspace for λ is k.

32. Suppose $A = PRP^{-1}$, where P is orthogonal and R is upper triangular. Show that if A is symmetric, then R is symmetric and hence is actually a diagonal matrix.

33. Construct a spectral decomposition of A from Example 2.

34. Construct a spectral decomposition of A from Example 3.

35. Let $\mathbf{u}$ be a unit vector in $\mathbb{R}^n$, and let $B = \mathbf{u}\mathbf{u}^T$.

a. Given any $\mathbf{x}$ in $\mathbb{R}^n$, compute $B\mathbf{x}$ and show that $B\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}$, as described in Section 6.2.

b. Show that B is a symmetric matrix and $B^2 = B$.

c. Show that $\mathbf{u}$ is an eigenvector of B. What is the corresponding eigenvalue?

36. Let B be an $n \times n$ symmetric matrix such that $B^2 = B$. Any such matrix is called a **projection matrix** (or an **orthogonal projection matrix**). Given any $\mathbf{y}$ in $\mathbb{R}^n$, let $\hat{\mathbf{y}} = B\mathbf{y}$ and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

a. Show that $\mathbf{z}$ is orthogonal to $\hat{\mathbf{y}}$.

b. Let W be the column space of B. Show that $\mathbf{y}$ is the sum of a vector in W and a vector in $W^\perp$. Why does this prove that $B\mathbf{y}$ is the orthogonal projection of $\mathbf{y}$ onto the column space of B?

[**M**] Orthogonally diagonalize the matrices in Exercises 37–40. To practice the methods of this section, do not use an eigenvector routine from your matrix program. Instead, use the program to find the eigenvalues, and, for each eigenvalue λ, find an orthonormal basis for Nul$(A - \lambda I)$, as in Examples 2 and 3.

37. $\begin{bmatrix} 5 & 2 & 9 & -6 \\ 2 & 5 & -6 & 9 \\ 9 & -6 & 5 & 2 \\ -6 & 9 & 2 & 5 \end{bmatrix}$

38. $\begin{bmatrix} .38 & -.18 & -.06 & -.04 \\ -.18 & .59 & -.04 & .12 \\ -.06 & -.04 & .47 & -.12 \\ -.04 & .12 & -.12 & .41 \end{bmatrix}$

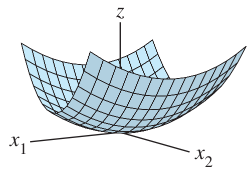
39. $\begin{bmatrix} .31 & .58 & .08 & .44 \\ .58 & -.56 & .44 & -.58 \\ .08 & .44 & .19 & -.08 \\ .44 & -.58 & -.08 & .31 \end{bmatrix}$

40. $\begin{bmatrix} 10 & 2 & 2 & -6 & 9 \\ 2 & 10 & 2 & -6 & 9 \\ 2 & 2 & 10 & -6 & 9 \\ -6 & -6 & -6 & 26 & 9 \\ 9 & 9 & 9 & 9 & -19 \end{bmatrix}$

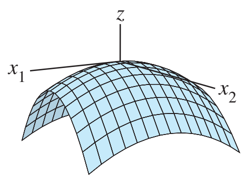
SOLUTIONS TO PRACTICE PROBLEMS

1. $(A^2)^T = (AA)^T = A^TA^T$, by a property of transposes. By hypothesis, $A^T = A$. So $(A^2)^T = AA = A^2$, which shows that A^2 is symmetric.

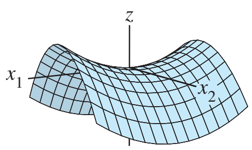
2. If A is orthogonally diagonalizable, then A is symmetric, by Theorem 2. By Practice Problem 1, A^2 is symmetric and hence is orthogonally diagonalizable (Theorem 2).



Positive definite



Negative definite



Indefinite

PROOF By the Principal Axes Theorem, there exists an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$ such that

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \quad (4)$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A. Since P is invertible, there is a one-to-one correspondence between all nonzero $\mathbf{x}$ and all nonzero $\mathbf{y}$. Thus the values of $Q(\mathbf{x})$ for $\mathbf{x} \neq \mathbf{0}$ coincide with the values of the expression on the right side of (4), which is obviously controlled by the signs of the eigenvalues $\lambda_1, \ldots, \lambda_n$, in the three ways described in the theorem. ■

EXAMPLE 6 Is $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$ positive definite?

SOLUTION Because of all the plus signs, this form "looks" positive definite. But the matrix of the form is

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

and the eigenvalues of A turn out to be 5, 2, and -1. So Q is an indefinite quadratic form, not positive definite. ■

WEB

The classification of a quadratic form is often carried over to the matrix of the form. Thus a **positive definite matrix** A is a *symmetric* matrix for which the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite. Other terms, such as **positive semidefinite matrix**, are defined analogously.

NUMERICAL NOTE

A fast way to determine whether a symmetric matrix A is positive definite is to attempt to factor A in the form $A = R^T R$, where R is upper triangular with positive diagonal entries. (A slightly modified algorithm for an LU factorization is one approach.) Such a *Cholesky factorization* is possible if and only if A is positive definite. See Supplementary Exercise 7 at the end of Chapter 7.

PRACTICE PROBLEM

Describe a positive semidefinite matrix A in terms of its eigenvalues.

WEB

7.2 EXERCISES

1. Compute the quadratic form $\mathbf{x}^T A\mathbf{x}$, when $A = \begin{bmatrix} 5 & 1/3 \\ 1/3 & 1 \end{bmatrix}$ and

a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

2. Compute the quadratic form $\mathbf{x}^T A\mathbf{x}$, for $A = \begin{bmatrix} 4 & 3 & 0 \\ 3 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ and

a. $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ b. $\mathbf{x} = \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$ c. $\mathbf{x} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$

3. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^2$.

a. $10x_1^2 - 6x_1x_2 - 3x_2^2$ b. $5x_1^2 + 3x_1x_2$

4. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^2$.

a. $20x_1^2 + 15x_1x_2 - 10x_2^2$ b. x_1x_2

5. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^3$.

a. $8x_1^2 + 7x_2^2 - 3x_3^2 - 6x_1x_2 + 4x_1x_3 - 2x_2x_3$

b. $4x_1x_2 + 6x_1x_3 - 8x_2x_3$

6. Find the matrix of the quadratic form. Assume $\mathbf{x}$ is in $\mathbb{R}^3$.

a. $5x_1^2 - x_2^2 + 7x_3^2 + 5x_1x_2 - 3x_1x_3$

b. $x_3^2 - 4x_1x_2 + 4x_2x_3$

7. Make a change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $x_1^2 + 10x_1x_2 + x_2^2$ into a quadratic form with no cross-product term. Give P and the new quadratic form.

8. Let A be the matrix of the quadratic form

$$9x_1^2 + 7x_2^2 + 11x_3^2 - 8x_1x_2 + 8x_1x_3$$

It can be shown that the eigenvalues of A are 3, 9, and 15. Find an orthogonal matrix P such that the change of variable $\mathbf{x} = P\mathbf{y}$ transforms $\mathbf{x}^TA\mathbf{x}$ into a quadratic form with no cross-product term. Give P and the new quadratic form.

Classify the quadratic forms in Exercises 9–18. Then make a change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form into one with no cross-product term. Write the new quadratic form. Construct P using the methods of Section 7.1.

9. $3x_1^2 - 4x_1x_2 + 6x_2^2$

10. $9x_1^2 - 8x_1x_2 + 3x_2^2$

11. $2x_1^2 + 10x_1x_2 + 2x_2^2$

12. $-5x_1^2 + 4x_1x_2 - 2x_2^2$

13. $x_1^2 - 6x_1x_2 + 9x_2^2$

14. $8x_1^2 + 6x_1x_2$

15. **[M]** $-2x_1^2 - 6x_2^2 - 9x_3^2 - 9x_4^2 + 4x_1x_2 + 4x_1x_3 + 4x_1x_4 + 6x_3x_4$

16. **[M]** $4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_4^2 + 3x_1x_2 + 3x_3x_4 - 4x_1x_4 + 4x_2x_3$

17. **[M]** $x_1^2 + x_2^2 + x_3^2 + x_4^2 + 9x_1x_2 - 12x_1x_4 + 12x_2x_3 + 9x_3x_4$

18. **[M]** $11x_1^2 - x_2^2 - 12x_1x_2 - 12x_1x_3 - 12x_1x_4 - 2x_3x_4$

19. What is the largest possible value of the quadratic form $5x_1^2 + 8x_2^2$ if $\mathbf{x} = (x_1, x_2)$ and $\mathbf{x}^T\mathbf{x} = 1$, that is, if $x_1^2 + x_2^2 = 1$? (Try some examples of $\mathbf{x}$.)

20. What is the largest value of the quadratic form $5x_1^2 - 3x_2^2$ if $\mathbf{x}^T\mathbf{x} = 1$?

In Exercises 21 and 22, matrices are $n \times n$ and vectors are in $\mathbb{R}^n$. Mark each statement True or False. Justify each answer.

21. a. The matrix of a quadratic form is a symmetric matrix.

b. A quadratic form has no cross-product terms if and only if the matrix of the quadratic form is a diagonal matrix.

c. The principal axes of a quadratic form $\mathbf{x}^TA\mathbf{x}$ are eigenvectors of A.

d. A positive definite quadratic form Q satisfies $Q(\mathbf{x}) > 0$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

e. If the eigenvalues of a symmetric matrix A are all positive, then the quadratic form $\mathbf{x}^TA\mathbf{x}$ is positive definite.

f. A Cholesky factorization of a symmetric matrix A has the form $A = R^TR$, for an upper triangular matrix R with positive diagonal entries.

22. a. The expression $\|\mathbf{x}\|^2$ is a quadratic form.

b. If A is symmetric and P is an orthogonal matrix, then the change of variable $\mathbf{x} = P\mathbf{y}$ transforms $\mathbf{x}^TA\mathbf{x}$ into a quadratic form with no cross-product term.

c. If A is a 2×2 symmetric matrix, then the set of $\mathbf{x}$ such that $\mathbf{x}^TA\mathbf{x} = c$ (for a constant c) corresponds to either a circle, an ellipse, or a hyperbola.

d. An indefinite quadratic form is either positive semidefinite or negative semidefinite.

e. If A is symmetric and the quadratic form $\mathbf{x}^TA\mathbf{x}$ has only negative values for $\mathbf{x} \neq \mathbf{0}$, then the eigenvalues of A are all negative.

Exercises 23 and 24 show how to classify a quadratic form $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, when $A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$ and $\det A \neq 0$, without finding the eigenvalues of A.

23. If λ_1 and λ_2 are the eigenvalues of A, then the characteristic polynomial of A can be written in two ways: $\det(A - \lambda I)$ and $(\lambda - \lambda_1)(\lambda - \lambda_2)$. Use this fact to show that $\lambda_1 + \lambda_2 = a + d$ (the diagonal entries of A) and $\lambda_1\lambda_2 = \det A$.

24. Verify the following statements.

a. Q is positive definite if $\det A > 0$ and $a > 0$.

b. Q is negative definite if $\det A > 0$ and $a < 0$.

c. Q is indefinite if $\det A < 0$.

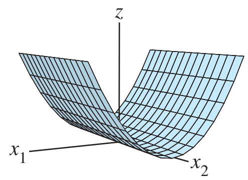
25. Show that if B is $m \times n$, then B^TB is positive semidefinite; and if B is $n \times n$ and invertible, then B^TB is positive definite.

26. Show that if an $n \times n$ matrix A is positive definite, then there exists a positive definite matrix B such that $A = B^TB$. [*Hint:* Write $A = PDP^T$, with $P^T = P^{-1}$. Produce a diagonal matrix C such that $D = C^TC$, and let $B = PCP^T$. Show that B works.]

27. Let A and B be symmetric $n \times n$ matrices whose eigenvalues are all positive. Show that the eigenvalues of $A + B$ are all positive. [*Hint:* Consider quadratic forms.]

28. Let A be an $n \times n$ invertible symmetric matrix. Show that if the quadratic form $\mathbf{x}^TA\mathbf{x}$ is positive definite, then so is the quadratic form $\mathbf{x}^TA^{-1}\mathbf{x}$. [*Hint:* Consider eigenvalues.]

SG Mastering: Diagonalization and Quadratic Forms 7–7



Positive semidefinite

SOLUTION TO PRACTICE PROBLEM

Make an orthogonal change of variable $\mathbf{x} = P\mathbf{y}$, and write

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

as in equation (4). If an eigenvalue—say, λ_i—were negative, then $\mathbf{x}^T A\mathbf{x}$ would be negative for the $\mathbf{x}$ corresponding to $\mathbf{y} = \mathbf{e}_i$ (the ith column of I_n). So the eigenvalues of a positive semidefinite quadratic form must all be nonnegative. Conversely, if the eigenvalues are nonnegative, the expansion above shows that $\mathbf{x}^T A\mathbf{x}$ must be positive semidefinite.

7.3 CONSTRAINED OPTIMIZATION

Engineers, economists, scientists, and mathematicians often need to find the maximum or minimum value of a quadratic form $Q(\mathbf{x})$ for $\mathbf{x}$ in some specified set. Typically, the problem can be arranged so that $\mathbf{x}$ varies over the set of unit vectors. This *constrained optimization problem* has an interesting and elegant solution. Example 6 below and the discussion in Section 7.5 will illustrate how such problems arise in practice.

The requirement that a vector $\mathbf{x}$ in $\mathbb{R}^n$ be a unit vector can be stated in several equivalent ways:

$$\|\mathbf{x}\| = 1, \qquad \|\mathbf{x}\|^2 = 1, \qquad \mathbf{x}^T\mathbf{x} = 1$$

and

$$x_1^2 + x_2^2 + \cdots + x_n^2 = 1 \tag{1}$$

The expanded version (1) of $\mathbf{x}^T\mathbf{x} = 1$ is commonly used in applications.

When a quadratic form Q has no cross-product terms, it is easy to find the maximum and minimum of $Q(\mathbf{x})$ for $\mathbf{x}^T\mathbf{x} = 1$.

EXAMPLE 1 Find the maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$.

SOLUTION Since x_2^2 and x_3^2 are nonnegative, note that

$$4x_2^2 \le 9x_2^2 \qquad \text{and} \qquad 3x_3^2 \le 9x_3^2$$

and hence

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 3x_3^2 \\ &\le 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

whenever $x_1^2 + x_2^2 + x_3^2 = 1$. So the maximum value of $Q(\mathbf{x})$ cannot exceed 9 when $\mathbf{x}$ is a unit vector. Furthermore, $Q(\mathbf{x}) = 9$ when $\mathbf{x} = (1, 0, 0)$. Thus 9 is the maximum value of $Q(\mathbf{x})$ for $\mathbf{x}^T\mathbf{x} = 1$.

To find the minimum value of $Q(\mathbf{x})$, observe that

$$9x_1^2 \ge 3x_1^2, \qquad 4x_2^2 \ge 3x_2^2$$

and hence

$$Q(\mathbf{x}) \ge 3x_1^2 + 3x_2^2 + 3x_3^2 = 3(x_1^2 + x_2^2 + x_3^2) = 3$$

whenever $x_1^2 + x_2^2 + x_3^2 = 1$. Also, $Q(\mathbf{x}) = 3$ when $x_1 = 0$, $x_2 = 0$, and $x_3 = 1$. So 3 is the minimum value of $Q(\mathbf{x})$ when $\mathbf{x}^T\mathbf{x} = 1$. ■

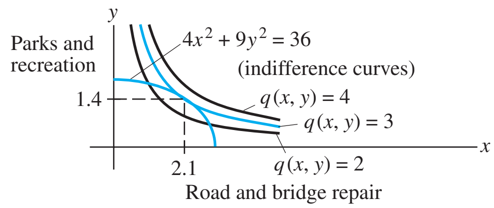


FIGURE 4 The optimum public works schedule is (2.1, 1.4).

and define

$$x_1 = \frac{x}{3}, \quad x_2 = \frac{y}{2}, \qquad \text{that is,} \quad x = 3x_1 \quad \text{and} \quad y = 2x_2$$

Then the constraint equation becomes

$$x_1^2 + x_2^2 = 1$$

and the utility function becomes $q(3x_1, 2x_2) = (3x_1)(2x_2) = 6x_1x_2$. Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then the problem is to maximize $Q(\mathbf{x}) = 6x_1x_2$ subject to $\mathbf{x}^T\mathbf{x} = 1$. Note that $Q(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$, where

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

The eigenvalues of A are ± 3, with eigenvectors $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ for $\lambda = 3$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ for $\lambda = -3$. Thus the maximum value of $Q(\mathbf{x}) = q(x_1, x_2)$ is 3, attained when $x_1 = 1/\sqrt{2}$ and $x_2 = 1/\sqrt{2}$.

In terms of the original variables, the optimum public works schedule is $x = 3x_1 = 3/\sqrt{2} \approx 2.1$ hundred miles of roads and bridges and $y = 2x_2 = \sqrt{2} \approx 1.4$ hundred acres of parks and recreational areas. The optimum public works schedule is the point where the constraint curve and the indifference curve $q(x, y) = 3$ just meet. Points (x, y) with a higher utility lie on indifference curves that do not touch the constraint curve. See Fig. 4. ■

PRACTICE PROBLEMS

1. Let $Q(\mathbf{x}) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Find a change of variable that transforms Q into a quadratic form with no cross-product term, and give the new quadratic form.
2. With Q as in Problem 1, find the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, and find a unit vector at which the maximum is attained.

7.3 EXERCISES

In Exercises 1 and 2, find the change of variable $\mathbf{x} = P\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^TA\mathbf{x}$ into $\mathbf{y}^TD\mathbf{y}$ as shown.

1. $5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3 = 9y_1^2 + 6y_2^2 + 3y_3^2$
2. $3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3 = 5y_1^2 + 2y_2^2$ [*Hint:* $\mathbf{x}$ and $\mathbf{y}$ must have the same number of coordinates, so the quadratic form shown here must have a coefficient of zero for y_3^2.]

In Exercises 3–6, find (a) the maximum value of $Q(\mathbf{x})$ subject to the constraint $\mathbf{x}^T\mathbf{x} = 1$, (b) a unit vector $\mathbf{u}$ where this maximum is attained, and (c) the maximum of $Q(\mathbf{x})$ subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u} = 0$.

3. $Q(\mathbf{x}) = 5x_1^2 + 6x_2^2 + 7x_3^2 + 4x_1x_2 - 4x_2x_3$ (See Exercise 1.)

4. $Q(\mathbf{x}) = 3x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$ (See Exercise 2.)

5. $Q(\mathbf{x}) = 5x_1^2 + 5x_2^2 - 4x_1x_2$

6. $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 + 3x_1x_2$

7. Let $Q(\mathbf{x}) = -2x_1^2 - x_2^2 + 4x_1x_2 + 4x_2x_3$. Find a unit vector $\mathbf{x}$ in $\mathbb{R}^3$ at which $Q(\mathbf{x})$ is maximized, subject to $\mathbf{x}^T\mathbf{x} = 1$. [*Hint:* The eigenvalues of the matrix of the quadratic form Q are 2, -1, and -4.]

8. Let $Q(\mathbf{x}) = 7x_1^2 + x_2^2 + 7x_3^2 - 8x_1x_2 - 4x_1x_3 - 8x_2x_3$. Find a unit vector $\mathbf{x}$ in $\mathbb{R}^3$ at which $Q(\mathbf{x})$ is maximized, subject to $\mathbf{x}^T\mathbf{x} = 1$. [*Hint:* The eigenvalues of the matrix of the quadratic form Q are 9 and -3.]

9. Find the maximum value of $Q(\mathbf{x}) = 7x_1^2 + 3x_2^2 - 2x_1x_2$, subject to the constraint $x_1^2 + x_2^2 = 1$. (Do not go on to find a vector where the maximum is attained.)

10. Find the maximum value of $Q(\mathbf{x}) = -3x_1^2 + 5x_2^2 - 2x_1x_2$, subject to the constraint $x_1^2 + x_2^2 = 1$. (Do not go on to find a vector where the maximum is attained.)

11. Suppose $\mathbf{x}$ is a unit eigenvector of a matrix A corresponding to an eigenvalue 3. What is the value of $\mathbf{x}^TA\mathbf{x}$?

12. Let λ be any eigenvalue of a symmetric matrix A. Justify the statement made in this section that $m \le \lambda \le M$, where m and M are defined as in (2). [*Hint:* Find an $\mathbf{x}$ such that $\lambda = \mathbf{x}^TA\mathbf{x}$.]

13. Let A be an $n \times n$ symmetric matrix, let M and m denote the maximum and minimum values of the quadratic form $\mathbf{x}^TA\mathbf{x}$, and denote corresponding unit eigenvectors by $\mathbf{u}_1$ and $\mathbf{u}_n$. The following calculations show that given any number t between M and m, there is a unit vector $\mathbf{x}$ such that $t = \mathbf{x}^TA\mathbf{x}$. Verify that $t = (1-\alpha)m + \alpha M$ for some number α between 0 and 1. Then let $\mathbf{x} = \sqrt{1-\alpha}\,\mathbf{u}_n + \sqrt{\alpha}\,\mathbf{u}_1$, and show that $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^TA\mathbf{x} = t$.

[M] In Exercises 14–17, follow the instructions given for Exercises 3–6.

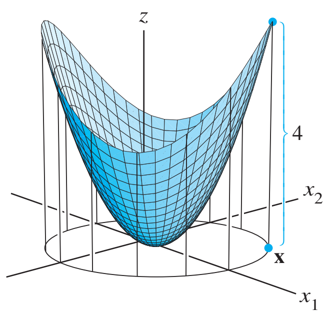
14. $x_1x_2 + 3x_1x_3 + 30x_1x_4 + 30x_2x_3 + 3x_2x_4 + x_3x_4$

15. $3x_1x_2 + 5x_1x_3 + 7x_1x_4 + 7x_2x_3 + 5x_2x_4 + 3x_3x_4$

16. $4x_1^2 - 6x_1x_2 - 10x_1x_3 - 10x_1x_4 - 6x_2x_3 - 6x_2x_4 - 2x_3x_4$

17. $-6x_1^2 - 10x_2^2 - 13x_3^2 - 13x_4^2 - 4x_1x_2 - 4x_1x_3 - 4x_1x_4 + 6x_3x_4$

SOLUTIONS TO PRACTICE PROBLEMS



The maximum value of $Q(\mathbf{x})$ subject to $\mathbf{x}^T\mathbf{x} = 1$ is 4.

1. The matrix of the quadratic form is $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. It is easy to find the eigenvalues, 4 and 2, and corresponding unit eigenvectors, $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ and $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. So the desired change of variable is $\mathbf{x} = P\mathbf{y}$, where $P = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$. (A common error here is to forget to normalize the eigenvectors.) The new quadratic form is $\mathbf{y}^TD\mathbf{y} = 4y_1^2 + 2y_2^2$.

2. The maximum of $Q(\mathbf{x})$ for $\mathbf{x}$ a unit vector is 4, and the maximum is attained at the unit eigenvector $\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. [A common incorrect answer is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. This vector maximizes the quadratic form $\mathbf{y}^TD\mathbf{y}$ instead of $Q(\mathbf{x})$.]

7.4 THE SINGULAR VALUE DECOMPOSITION

The diagonalization theorems in Sections 5.3 and 7.1 play a part in many interesting applications. Unfortunately, as we know, not all matrices can be factored as $A = PDP^{-1}$ with D diagonal. However, a factorization $A = QDP^{-1}$ *is* possible for any $m \times n$ matrix A! A special factorization of this type, called the *singular value decomposition*, is one of the most useful matrix factorizations in applied linear algebra.

The singular value decomposition is based on the following property of the ordinary diagonalization that can be imitated for rectangular matrices: The absolute values of the eigenvalues of a symmetric matrix A measure the amounts that A stretches or shrinks

Moler, C. B., and D. Morrison, "Singular Value Analysis of Cryptograms." *Amer. Math. Monthly* **90** (1983), pp. 78–87.

Strang, Gilbert, *Linear Algebra and Its Applications,* 4th ed. (Belmont, CA: Brooks/Cole, 2005).

Watkins, David S., *Fundamentals of Matrix Computations* (New York: Wiley, 1991), pp. 390–398, 409–421.

PRACTICE PROBLEM

WEB

Given a singular value decomposition, $A = U\Sigma V^T$, find an SVD of A^T. How are the singular values of A and A^T related?

7.4 EXERCISES

Find the singular values of the matrices in Exercises 1–4.

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

3. $\begin{bmatrix} \sqrt{6} & 1 \\ 0 & \sqrt{6} \end{bmatrix}$ **4.** $\begin{bmatrix} \sqrt{3} & 2 \\ 0 & \sqrt{3} \end{bmatrix}$

Find an SVD of each matrix in Exercises 5–12. [*Hint:* In Exercise 11, one choice for U is $\begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$. In Exercise 12, one column of U can be $\begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$.]

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

7. $\begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$ **8.** $\begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$

9. $\begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$ **10.** $\begin{bmatrix} 4 & -2 \\ 2 & -1 \\ 0 & 0 \end{bmatrix}$

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

13. Find the SVD of $A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$ [*Hint:* Work with A^T.]

14. In Exercise 7, find a unit vector $\mathbf{x}$ at which $A\mathbf{x}$ has maximum length.

15. Suppose the factorization below is an SVD of a matrix A, with the entries in U and V rounded to two decimal places.

$$A = \begin{bmatrix} .40 & -.78 & .47 \\ .37 & -.33 & -.87 \\ -.84 & -.52 & -.16 \end{bmatrix} \begin{bmatrix} 7.10 & 0 & 0 \\ 0 & 3.10 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .30 & -.51 & -.81 \\ .76 & .64 & -.12 \\ .58 & -.58 & .58 \end{bmatrix}$$

a. What is the rank of A?

b. Use this decomposition of A, with no calculations, to write a basis for Col A and a basis for Nul A. [*Hint:* First write the columns of V.]

16. Repeat Exercise 15 for the following SVD of a 3×4 matrix A:

$$A = \begin{bmatrix} -.86 & -.11 & -.50 \\ .31 & .68 & -.67 \\ .41 & -.73 & -.55 \end{bmatrix} \begin{bmatrix} 12.48 & 0 & 0 & 0 \\ 0 & 6.34 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} .66 & -.03 & -.35 & .66 \\ -.13 & -.90 & -.39 & -.13 \\ .65 & .08 & -.16 & -.73 \\ -.34 & .42 & -.84 & -.08 \end{bmatrix}$$

In Exercises 17–24, A is an $m \times n$ matrix with a singular value decomposition $A = U\Sigma V^T$, where U is an $m \times m$ orthogonal matrix, Σ is an $m \times n$ "diagonal" matrix with r positive entries and no negative entries, and V is an $n \times n$ orthogonal matrix. Justify each answer.

17. Suppose A is square and invertible. Find a singular value decomposition of A^{-1}.

18. Show that if A is square, then $|\det A|$ is the product of the singular values of A.

19. Show that the columns of V are eigenvectors of A^TA, the columns of U are eigenvectors of AA^T, and the diagonal entries of Σ are the singular values of A. [*Hint:* Use the SVD to compute A^TA and AA^T.]

20. Show that if A is an $n \times n$ positive definite matrix, then an orthogonal diagonalization $A = PDP^T$ is a singular value decomposition of A.

21. Show that if P is an orthogonal $m \times m$ matrix, then PA has the same singular values as A.

22. Justify the statement in Example 2 that the second singular value of a matrix A is the maximum of $\|A\mathbf{x}\|$ as $\mathbf{x}$ varies over all unit vectors orthogonal to $\mathbf{v}_1$, with $\mathbf{v}_1$ a right singular vector corresponding to the first singular value of A. [*Hint:* Use Theorem 7 in Section 7.3.]

23. Let $U = [\,\mathbf{u}_1 \;\; \cdots \;\; \mathbf{u}_m\,]$ and $V = [\,\mathbf{v}_1 \;\; \cdots \;\; \mathbf{v}_n\,]$, where the $\mathbf{u}_i$ and $\mathbf{v}_i$ are as in Theorem 10. Show that

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \sigma_2\mathbf{u}_2\mathbf{v}_2^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T.$$

24. Using the notation of Exercise 23, show that $A^T\mathbf{u}_j = \sigma_j\mathbf{v}_j$ for $1 \le j \le r = \text{rank}\,A$.

25. Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Describe how to find a basis $\mathcal{B}$ for $\mathbb{R}^n$ and a basis $\mathcal{C}$ for $\mathbb{R}^m$ such that the matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$ is an $m \times n$ "diagonal" matrix.

[**M**] Compute an SVD of each matrix in Exercises 26 and 27. Report the final matrix entries accurate to two decimal places. Use the method of Examples 3 and 4.

26. $A = \begin{bmatrix} -18 & 13 & -4 & 4 \\ 2 & 19 & -4 & 12 \\ -14 & 11 & -12 & 8 \\ -2 & 21 & 4 & 8 \end{bmatrix}$

27. $A = \begin{bmatrix} 6 & -8 & -4 & 5 & -4 \\ 2 & 7 & -5 & -6 & 4 \\ 0 & -1 & -8 & 2 & 2 \\ -1 & -2 & 4 & 4 & -8 \end{bmatrix}$

28. [**M**] Compute the singular values of the 4×4 matrix in Exercise 9 in Section 2.3, and compute the condition number σ_1/σ_4.

29. [**M**] Compute the singular values of the 5×5 matrix in Exercise 10 in Section 2.3, and compute the condition number σ_1/σ_5.

SOLUTION TO PRACTICE PROBLEM

If $A = U\Sigma V^T$, where Σ is $m \times n$, then $A^T = (V^T)^T\Sigma^T U^T = V\Sigma^T U^T$. This is an SVD of A^T because V and U are orthogonal matrices and Σ^T is an $n \times m$ "diagonal" matrix. Since Σ and Σ^T have the same nonzero diagonal entries, A and A^T have the same nonzero singular values. [*Note:* If A is $2 \times n$, then AA^T is only 2×2 and its eigenvalues may be easier to compute (by hand) than the eigenvalues of A^TA.]

7.5 APPLICATIONS TO IMAGE PROCESSING AND STATISTICS

The satellite photographs in this chapter's introduction provide an example of multidimensional, or *multivariate*, data—information organized so that each datum in the data set is identified with a point (vector) in $\mathbb{R}^n$. The main goal of this section is to explain a technique, called *principal component analysis*, used to analyze such multivariate data. The calculations will illustrate the use of orthogonal diagonalization and the singular value decomposition.

Principal component analysis can be applied to any data that consist of lists of measurements made on a collection of objects or individuals. For instance, consider a chemical process that produces a plastic material. To monitor the process, 300 samples are taken of the material produced, and each sample is subjected to a battery of eight tests, such as melting point, density, ductility, tensile strength, and so on. The laboratory report for each sample is a vector in $\mathbb{R}^8$, and the set of such vectors forms an 8×300 matrix, called the **matrix of observations**.

Loosely speaking, we can say that the process control data are eight-dimensional. The next two examples describe data that can be visualized graphically.

EXAMPLE 1 An example of two-dimensional data is given by a set of weights and heights of N college students. Let $\mathbf{X}_j$ denote the **observation vector** in $\mathbb{R}^2$ that lists the weight and height of the jth student. If w denotes weight and h height, then the matrix in the sense that the sum of the squares of the *orthogonal* distances to the line is minimized. In fact, principal component analysis is equivalent to what is termed *orthogonal regression*, but that is a story for another day. Perhaps we'll meet again.

CHAPTER 7 SUPPLEMENTARY EXERCISES

1. Mark each statement True or False. Justify each answer. In each part, A represents an $n \times n$ matrix.

a. If A is orthogonally diagonalizable, then A is symmetric.

b. If A is an orthogonal matrix, then A is symmetric.

c. If A is an orthogonal matrix, then $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x}$ in $\mathbb{R}^n$.

d. The principal axes of a quadratic form $\mathbf{x}^T A\mathbf{x}$ can be the columns of any matrix P that diagonalizes A.

e. If P is an $n \times n$ matrix with orthogonal columns, then $P^T = P^{-1}$.

f. If every coefficient in a quadratic form is positive, then the quadratic form is positive definite.

g. If $\mathbf{x}^T A\mathbf{x} > 0$ for some $\mathbf{x}$, then the quadratic form $\mathbf{x}^T A\mathbf{x}$ is positive definite.

h. By a suitable change of variable, any quadratic form can be changed into one with no cross-product term.

i. The largest value of a quadratic form $\mathbf{x}^T A\mathbf{x}$, for $\|\mathbf{x}\| = 1$, is the largest entry on the diagonal of A.

j. The maximum value of a positive definite quadratic form $\mathbf{x}^T A\mathbf{x}$ is the greatest eigenvalue of A.

k. A positive definite quadratic form can be changed into a negative definite form by a suitable change of variable $\mathbf{x} = P\mathbf{u}$, for some orthogonal matrix P.

l. An indefinite quadratic form is one whose eigenvalues are not definite.

m. If P is an $n \times n$ orthogonal matrix, then the change of variable $\mathbf{x} = P\mathbf{u}$ transforms $\mathbf{x}^T A\mathbf{x}$ into a quadratic form whose matrix is $P^{-1}AP$.

n. If U is $m \times n$ with orthogonal columns, then $UU^T\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto Col U.

o. If B is $m \times n$ and $\mathbf{x}$ is a unit vector in $\mathbb{R}^n$, then $\|B\mathbf{x}\| \le \sigma_1$, where σ_1 is the first singular value of B.

p. A singular value decomposition of an $m \times n$ matrix B can be written as $B = P\Sigma Q$, where P is an $m \times m$ orthogonal matrix, Q is an $n \times n$ orthogonal matrix, and Σ is an $m \times n$ "diagonal" matrix.

q. If A is $n \times n$, then A and $A^T A$ have the same singular values.

2. Let $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ be an orthonormal basis for $\mathbb{R}^n$, and let $\lambda_1, \dots, \lambda_n$ be any real scalars. Define

$$A = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \cdots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T$$

a. Show that A is symmetric.

b. Show that $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A.

3. Let A be an $n \times n$ symmetric matrix of rank r. Explain why the spectral decomposition of A represents A as the sum of r rank 1 matrices.

4. Let A be an $n \times n$ symmetric matrix.

a. Show that $(\text{Col } A)^\perp = \text{Nul } A$. [*Hint:* See Section 6.1.]

b. Show that each $\mathbf{y}$ in $\mathbb{R}^n$ can be written in the form $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, with $\hat{\mathbf{y}}$ in Col A and $\mathbf{z}$ in Nul A.

5. Show that if $\mathbf{v}$ is an eigenvector of an $n \times n$ matrix A and $\mathbf{v}$ corresponds to a nonzero eigenvalue of A, then $\mathbf{v}$ is in Col A. [*Hint:* Use the definition of an eigenvector.]

6. Let A be an $n \times n$ symmetric matrix. Use Exercise 5 and an eigenvector basis for $\mathbb{R}^n$ to give a second proof of the decomposition in Exercise 4(b).

7. Prove that an $n \times n$ matrix A is positive definite if and only if A admits a *Cholesky factorization*, namely, $A = R^T R$ for some invertible upper triangular matrix R whose diagonal entries are all positive. [*Hint:* Use a QR factorization and Exercise 26 in Section 7.2.]

8. Use Exercise 7 to show that if A is positive definite, then A has an LU factorization, $A = LU$, where U has positive pivots on its diagonal. (The converse is true, too.)

If A is $m \times n$, then the matrix $G = A^T A$ is called the *Gram matrix of* A. In this case, the entries of G are the inner products of the columns of A. (See Exercises 9 and 10.)

9. Show that the Gram matrix of any matrix A is positive semidefinite, with the same rank as A. (See the Exercises in Section 6.5.)

10. Show that if an $n \times n$ matrix G is positive semidefinite and has rank r, then G is the Gram matrix of some $r \times n$ matrix A. This is called a *rank-revealing factorization* of G. [*Hint:* Consider the spectral decomposition of G, and first write G as BB^T for an $n \times r$ matrix B.]

11. Prove that any $n \times n$ matrix A admits a *polar decomposition* of the form $A = PQ$, where P is an $n \times n$ positive semidefinite matrix with the same rank as A and where Q is an $n \times n$ orthogonal matrix. [*Hint:* Use a singular value decomposition, $A = U\Sigma V^T$, and observe that $A = (U\Sigma U^T)(UV^T)$.] This decomposition is used, for instance, in mechanical engineering to model the deformation of a material. The matrix P describes the stretching or compression of the material (in the directions of the eigenvectors of P), and Q describes the rotation of the material in space.

Exercises 12–14 concern an $m \times n$ matrix A with a reduced singular value decomposition, $A = U_r D V_r^T$, and the pseudoinverse $A^+ = V_r D^{-1} U_r^T$.

12. Verify the properties of A^+:

a. For each $\mathbf{y}$ in $\mathbb{R}^m$, $AA^+\mathbf{y}$ is the orthogonal projection of $\mathbf{y}$ onto Col A.

b. For each $\mathbf{x}$ in $\mathbb{R}^n$, $A^+A\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto Row A.

c. $AA^+A = A$ and $A^+AA^+ = A^+$.

13. Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent, and let $\mathbf{x}^+ = A^+\mathbf{b}$. By Exercise 23 in Section 6.3, there is exactly one vector $\mathbf{p}$ in Row A such that $A\mathbf{p} = \mathbf{b}$. The following steps prove that $\mathbf{x}^+ = \mathbf{p}$ and $\mathbf{x}^+$ is the *minimum length solution* of $A\mathbf{x} = \mathbf{b}$.

a. Show that $\mathbf{x}^+$ is in Row A. [*Hint:* Write $\mathbf{b}$ as $A\mathbf{x}$ for some $\mathbf{x}$, and use Exercise 12.]

b. Show that $\mathbf{x}^+$ is a solution of $A\mathbf{x} = \mathbf{b}$.

c. Show that if $\mathbf{u}$ is any solution of $A\mathbf{x} = \mathbf{b}$, then $\|\mathbf{x}^+\| \leq \|\mathbf{u}\|$, with equality only if $\mathbf{u} = \mathbf{x}^+$.

14. Given any $\mathbf{b}$ in $\mathbb{R}^m$, adapt Exercise 13 to show that $A^+\mathbf{b}$ is the *least-squares solution of minimum length*. [*Hint:* Consider the equation $A\mathbf{x} = \hat{\mathbf{b}}$, where $\hat{\mathbf{b}}$ is the orthogonal projection of $\mathbf{b}$ onto Col A.]

[**M**] In Exercises 15 and 16, construct the pseudoinverse of A. Begin by using a matrix program to produce the SVD of A, or, if that is not available, begin with an orthogonal diagonalization of A^TA. Use the pseudoinverse to solve $A\mathbf{x} = \mathbf{b}$, for $\mathbf{b} = (6, -1, -4, 6)$, and let $\hat{\mathbf{x}}$ be the solution. Make a calculation to verify that $\hat{\mathbf{x}}$ is in Row A. Find a nonzero vector $\mathbf{u}$ in Nul A, and verify that $\|\hat{\mathbf{x}}\| < \|\hat{\mathbf{x}} + \mathbf{u}\|$, which must be true by Exercise 13(c).

15. $A = \begin{bmatrix} -3 & -3 & -6 & 6 & 1 \\ -1 & -1 & -1 & 1 & -2 \\ 0 & 0 & -1 & 1 & -1 \\ 0 & 0 & -1 & 1 & -1 \end{bmatrix}$

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