



Mathematics-II

Mid Term Sallybus+ Sample Paper

BY

Huma IKhlaq

The Islamia University of Bahawalpur,

Department of Physics



2 CHAPTER 1 Linear Equations in Linear Algebra

seismic data for the equations are obtained from underwater shock waves created by explosions from air guns. The waves bounce off subsurface rocks and are measured by geophones attached to mile-long cables behind the ship.

- *Linear programming.* Many important management decisions today are made on the basis of linear programming models that utilize hundreds of variables. The airline industry, for instance,

employs linear programs that schedule flight crews, monitor the locations of aircraft, or plan the varied schedules of support services such as maintenance and terminal operations.

- *Electrical networks.* Engineers use simulation software to design electrical circuits and microchips involving millions of transistors. Such software relies on linear algebra techniques and systems of linear equations.

WEB

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

1.1 SYSTEMS OF LINEAR EQUATIONS

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

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \quad (1)$$

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

The equations

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

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

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

The equations

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

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

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

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



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

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

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

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

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

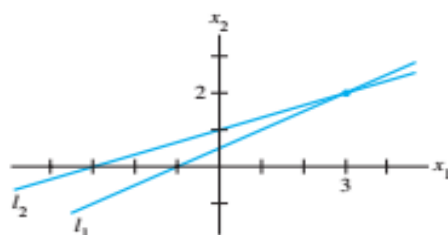


FIGURE 1 Exactly one solution.

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

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

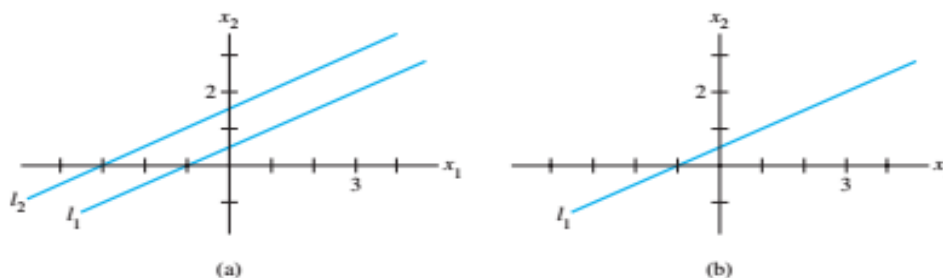


FIGURE 2 (a) No solution. (b) Infinitely many solutions.

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



A system of linear equations has

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

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

Matrix Notation

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

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

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

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

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

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

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

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

Solving a Linear System

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

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

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



EXAMPLE 1 Solve system (3).

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

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ 2x_2 - 8x_3 & = & 8 \\ -4x_1 + 5x_2 + 9x_3 & = & -9 \end{array} \quad \left[\begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array} \right]$$

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

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

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

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ 2x_2 - 8x_3 & = & 8 \\ -3x_2 + 13x_3 & = & -9 \end{array} \quad \left[\begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array} \right]$$

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

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ x_2 - 4x_3 & = & 4 \\ -3x_2 + 13x_3 & = & -9 \end{array} \quad \left[\begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{array} \right]$$

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

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

The new system has a *triangular* form:¹

$$\begin{array}{rcl} x_1 - 2x_2 + x_3 & = & 0 \\ x_2 - 4x_3 & = & 4 \\ x_3 & = & 3 \end{array} \quad \left[\begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

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

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

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

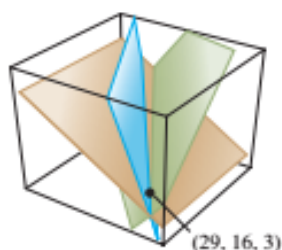


It is convenient to combine the results of these two operations:

$$\begin{array}{rcl} x_1 - 2x_2 & = & -3 \\ x_2 & = & 16 \\ x_3 & = & 3 \end{array} \quad \left[\begin{array}{cccc} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$

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

$$\begin{cases} x_1 & = & 29 \\ x_2 & = & 16 \\ x_3 & = & 3 \end{cases} \quad \left[\begin{array}{cccc} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array} \right]$$



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

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

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

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

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

ELEMENTARY ROW OPERATIONS

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

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

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

²A common paraphrase of row replacement is “Add to one row a multiple of another row.”



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

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

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

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

Existence and Uniqueness Questions

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

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

TWO FUNDAMENTAL QUESTIONS ABOUT A LINEAR SYSTEM

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

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

EXAMPLE 2 Determine if the following system is consistent:

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

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

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

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



only one possible value, and x_1 is therefore uniquely determined by equation 1. So the solution is unique.) ■

EXAMPLE 3 Determine if the following system is consistent:

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

SOLUTION The augmented matrix is

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

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

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

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

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

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

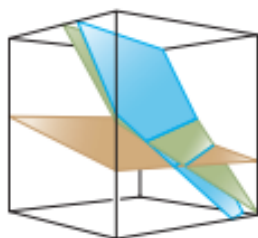
$$\begin{bmatrix} 2 & -3 & 2 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & 5/2 \end{bmatrix}\tag{7}$$

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

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

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

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



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



NUMERICAL NOTE

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

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

PRACTICE PROBLEMS

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

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

$$\begin{array}{ll} \text{a. } x_1 + 4x_2 - 2x_3 + 8x_4 = 12 & \text{b. } x_1 - 3x_2 + 5x_3 - 2x_4 = 0 \\ x_2 - 7x_3 + 2x_4 = -4 & x_2 + 8x_3 = -4 \\ 5x_3 - x_4 = 7 & 2x_3 = 3 \\ x_3 + 3x_4 = -5 & x_4 = 1 \end{array}$$

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

$$\left[\begin{array}{cccc} 1 & 5 & 2 & -6 \\ 0 & 4 & -7 & 2 \\ 0 & 0 & 5 & 0 \end{array} \right]$$

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

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

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

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

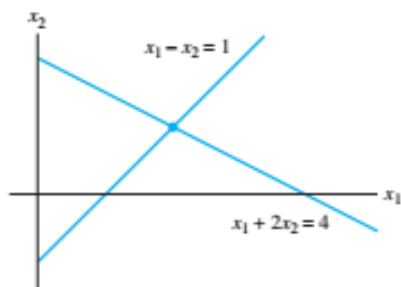


1.1 EXERCISES

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.

$$\begin{array}{ll} 1. & x_1 + 5x_2 = 7 \\ & -2x_1 - 7x_2 = -5 \end{array} \quad \begin{array}{ll} 2. & 3x_1 + 6x_2 = -3 \\ & 5x_1 + 7x_2 = 10 \end{array}$$

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.

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

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

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

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

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

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

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

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.

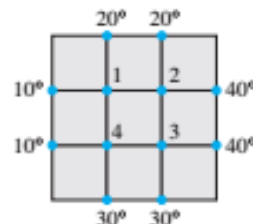
23. a. Every elementary row operation is reversible.
b. A 5×6 matrix has six rows.
c. The solution set of a linear system involving variables x_1, \dots, x_n is a list of numbers (s_1, \dots, s_n) that makes each equation in the system a true statement when the values s_1, \dots, s_n are substituted 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:
$$\left[\begin{array}{ccc|c} 1 & -4 & 7 & g \\ 0 & 3 & -5 & h \\ -2 & 5 & -9 & k \end{array} \right]$$
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.)



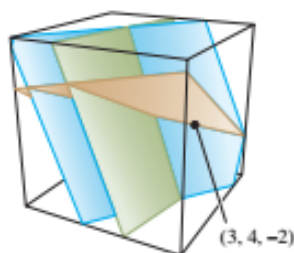
- b. The system is in triangular form. Further simplification begins with the x_4 in the fourth equation. Use the x_4 to eliminate all x_4 terms above it. The appropriate step now is to add 2 times equation 4 to equation 1. (After that, move to equation 3, multiply it by $1/2$, and then use the equation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

2. The system corresponding to the augmented matrix is

$$x_1 + 5x_2 + 2x_3 = -6$$

$$4x_2 - 7x_3 = 2$$

$$5x_3 = 0$$



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.

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.

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

$$5(3) - (4) + 2(-2) = 15 - 4 - 4 = 7$$

$$-2(3) + 6(4) + 9(-2) = -6 + 24 - 18 = 0$$

$$-7(3) + 5(4) - 3(-2) = -21 + 20 + 6 = 5$$

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

$$2x_1 - x_2 = h$$

$$0 = k + 3h$$

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).

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



DEFINITION

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

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

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

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

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

The “triangular” matrices of Section 1.1, such as

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

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

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

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

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

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

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

THEOREM 1

Uniqueness of the Reduced Echelon Form

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



14 CHAPTER 1 Linear Equations in Linear Algebra

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

Pivot Positions

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

DEFINITION

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

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

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

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

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

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

Pivot
Pivot column

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

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & -4 & -6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \quad (1)$$

Next pivot column



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

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

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

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

Pivot columns

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

$$A = \begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \quad \text{Pivot positions} \quad (3)$$

Pivot columns

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

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

The Row Reduction Algorithm

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

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

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

SOLUTION

STEP 1

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



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

↑ Pivot column

STEP 2

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

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

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

↑ Pivot

STEP 3

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

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

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

↑ Pivot

STEP 4

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

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

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

↑ New pivot column

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

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



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

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

Pivot

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

STEP 5

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

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

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

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

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

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

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

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

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

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

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

NUMERICAL NOTE

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



Solutions of Linear Systems

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

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

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

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

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

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

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

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

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

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

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

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

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

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



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

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

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

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

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



Parametric Descriptions of Solution Sets

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

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

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

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

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

Back-Substitution

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

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

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

Our matrix format for the backward phase of row reduction, which produces the reduced echelon form, has the same number of arithmetic operations as back-substitution. But the discipline of the matrix format substantially reduces the likelihood of errors



during hand computations. The best strategy is to use only the *reduced* echelon form to solve a system! The *Study Guide* that accompanies this text offers several helpful suggestions for performing row operations accurately and rapidly.

NUMERICAL NOTE

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

Existence and Uniqueness Questions

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

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

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

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

$$\left[\begin{array}{ccccc|c} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array} \right] \quad (8)$$

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

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

These remarks justify the following theorem.

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



THEOREM 2

Existence and Uniqueness Theorem

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

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

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

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

USING ROW REDUCTION TO SOLVE A LINEAR SYSTEM

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

PRACTICE PROBLEMS

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

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

2. Find the general solution of the system

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

1.2 EXERCISES

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

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

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

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

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



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

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

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

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

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

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

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

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

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

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

7. $\begin{bmatrix} 1 & 3 & 4 & 7 \\ 3 & 9 & 7 & 6 \end{bmatrix}$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

19. $\begin{aligned} x_1 + hx_2 &= 2 \\ 4x_1 + 8x_2 &= k \end{aligned}$

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

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

21. a. In some cases, a matrix may be row reduced to more than one matrix in reduced echelon form, using different sequences of row operations.
- b. The row reduction algorithm applies only to augmented matrices for a linear system.
- c. A basic variable in a linear system is a variable that corresponds to a pivot column in the coefficient matrix.
- d. Finding a parametric description of the solution set of a linear system is the same as *solving* the system.
- e. If one row in an echelon form of an augmented matrix is $[0 \ 0 \ 0 \ 5 \ 0]$, then the associated linear system is inconsistent.
22. a. The reduced echelon form of a matrix is unique.
- b. If every column of an augmented matrix contains a pivot, then the corresponding system is consistent.
- c. The pivot positions in a matrix depend on whether row interchanges are used in the row reduction process.
- d. A general solution of a system is an explicit description of all solutions of the system.
- e. Whenever a system has free variables, the solution set contains many solutions.

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

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

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



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

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

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,

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

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



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

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

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

SOLUTION

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

and

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

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

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} \neq [3 \ -1]$$

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

Geometric Descriptions of \mathbb{R}^2

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

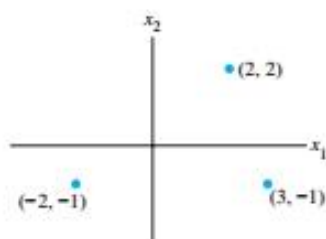


FIGURE 1 Vectors as points.

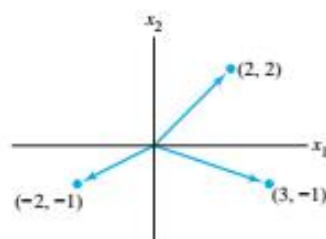


FIGURE 2 Vectors with arrows.

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

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

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



Parallelogram Rule for Addition

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

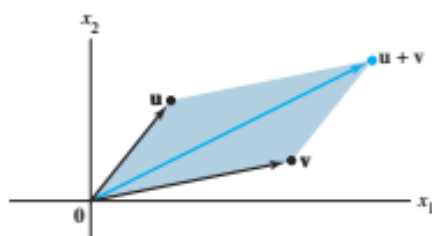


FIGURE 3 The parallelogram rule.

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

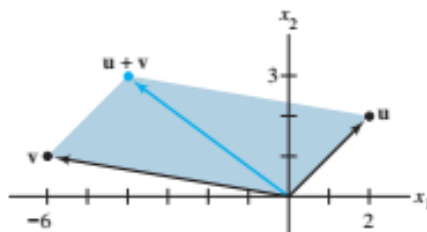
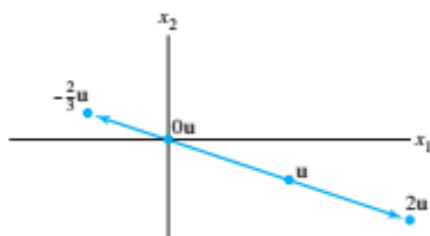


FIGURE 4

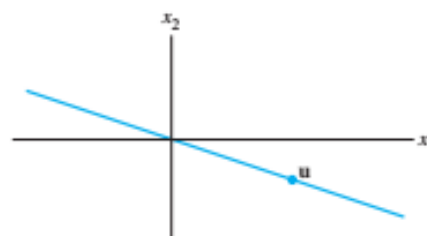
The next example illustrates the fact that the set of all scalar multiples of one fixed nonzero vector is a line through the origin, $(0, 0)$.

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

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



Typical multiples of \mathbf{u}



The set of all multiples of \mathbf{u}

FIGURE 5

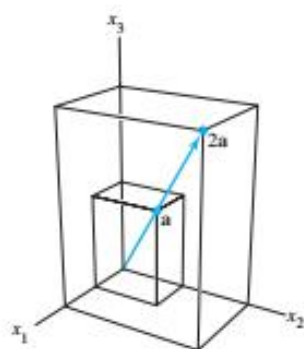


FIGURE 6
Scalar multiples .

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

Vectors in \mathbb{R}^3

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

Vectors in \mathbb{R}^n

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

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

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

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

Algebraic Properties of \mathbb{R}^n

For all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in \mathbb{R}^n and all scalars c and d :

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

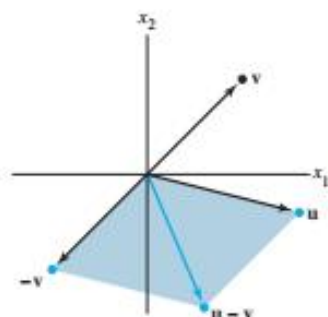


FIGURE 7
Vector subtraction.

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

Linear Combinations

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

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

is called a **linear combination** of $\mathbf{v}_1, \dots, \mathbf{v}_p$ with **weights** c_1, \dots, c_p . Property (ii) above permits us to omit parentheses when forming such a linear combination. The



weights in a linear combination can be any real numbers, including zero. For example, some linear combinations of vectors \mathbf{v}_1 and \mathbf{v}_2 are

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

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

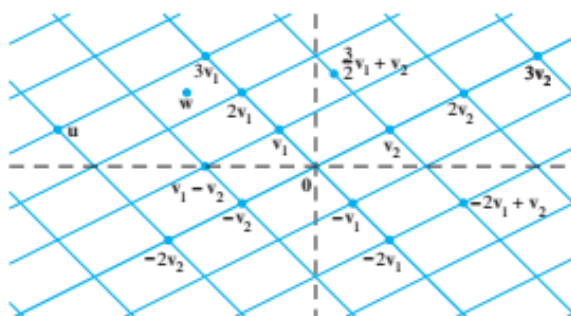


FIGURE 8 Linear combinations of \mathbf{v}_1 and \mathbf{v}_2 .

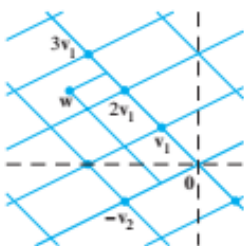


FIGURE 9

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

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

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

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

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

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

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

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

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

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

\uparrow \uparrow \uparrow
 \mathbf{a}_1 \mathbf{a}_2 \mathbf{b}



which is the same as

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

and

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

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

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

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

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

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

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

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

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

$\uparrow \quad \uparrow \quad \uparrow$
 $\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}$

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

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{b}] \quad (4)$$

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

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

A vector equation

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

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

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

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

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



One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{v_1, \dots, v_p\}$ of vectors.

DEFINITION

If v_1, \dots, v_p are in \mathbb{R}^n , then the set of all linear combinations of v_1, \dots, v_p is denoted by $\text{Span}\{v_1, \dots, v_p\}$ and is called the **subset of \mathbb{R}^n spanned (or generated) by v_1, \dots, v_p** . That is, $\text{Span}\{v_1, \dots, v_p\}$ is the collection of all vectors that can be written in the form

$$c_1 v_1 + c_2 v_2 + \dots + c_p v_p$$

with c_1, \dots, c_p scalars.

Asking whether a vector b is in $\text{Span}\{v_1, \dots, v_p\}$ amounts to asking whether the vector equation

$$x_1 v_1 + x_2 v_2 + \dots + x_p v_p = b$$

has a solution, or, equivalently, asking whether the linear system with augmented matrix $[v_1 \ \dots \ v_p \ b]$ has a solution.

Note that $\text{Span}\{v_1, \dots, v_p\}$ contains every scalar multiple of v_1 (for example), since $c v_1 = c v_1 + 0 v_2 + \dots + 0 v_p$. In particular, the zero vector must be in $\text{Span}\{v_1, \dots, v_p\}$.

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

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

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

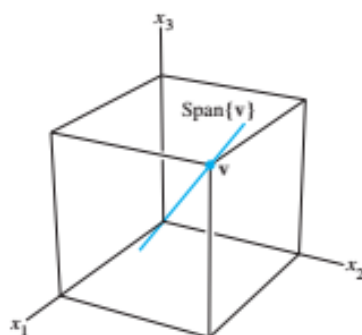


FIGURE 10 $\text{Span}\{v\}$ as a line through the origin.

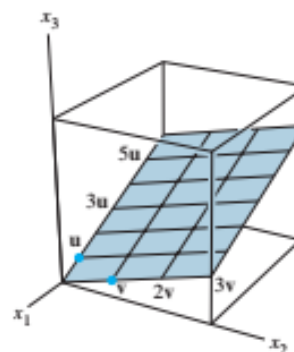


FIGURE 11 $\text{Span}\{u, v\}$ as a plane through the origin.

EXAMPLE 6 Let $a_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $a_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$, and $b = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Then

$\text{Span}\{a_1, a_2\}$ is a plane through the origin in \mathbb{R}^3 . Is b in that plane?



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

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

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 $\text{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$. ■

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:

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

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.

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

- 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.

- 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

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

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



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



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

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

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

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

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

In Exercises 17 and 18, list five vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. 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.

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

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

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

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

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

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 .



SOLUTIONS TO PRACTICE PROBLEMS

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

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \dots, u_n + v_n) \quad \text{Definition of vector addition}$$

$$= (v_1 + u_1, \dots, v_n + u_n) \quad \text{Commutativity of addition in } \mathbb{R}$$

$$= \mathbf{v} + \mathbf{u} \quad \text{Definition of vector addition}$$

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

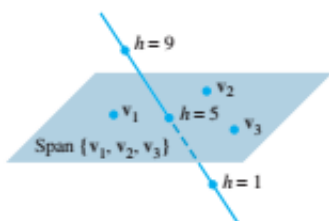
$$x_1 \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} 5 \\ -4 \\ -7 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}$$

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

$$\left[\begin{array}{cccc} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{array} \right] \sim \left[\begin{array}{cccc} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{array} \right] \sim \left[\begin{array}{cccc} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{array} \right]$$

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 $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if and only if $h = 5$.

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



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

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

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



DEFINITION

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

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

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

EXAMPLE 1

$$\begin{aligned} \text{a. } \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} &= 4 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -1 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 6 \\ -15 \end{bmatrix} + \begin{bmatrix} -7 \\ 21 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} \\ \text{b. } \begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} &= 4 \begin{bmatrix} 2 \\ 8 \\ -5 \end{bmatrix} + 7 \begin{bmatrix} -3 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 32 \\ -20 \end{bmatrix} + \begin{bmatrix} -21 \\ 0 \\ 14 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix} \quad \blacksquare \end{aligned}$$

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

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

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

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

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

is equivalent to

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

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

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

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

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

Here is the formal result.



THEOREM 3

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

$$A\mathbf{x} = \mathbf{b} \quad (4)$$

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{b} \quad (5)$$

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

$$\begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n & \mathbf{b} \end{bmatrix} \quad (6)$$

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

Existence of Solutions

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

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

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

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

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

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

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

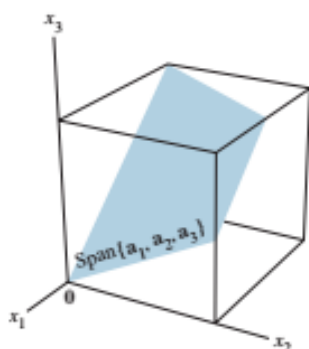


FIGURE 1
The columns of $A = [a_1 \ a_2 \ a_3]$ span a plane through $\mathbf{0}$.

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

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

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

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

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

THEOREM 4

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

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

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

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

Computation of $A\mathbf{x}$

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

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



SOLUTION From the definition,

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

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

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

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

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

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

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

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

EXAMPLE 5

- $\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 + 2 \cdot 3 + (-1) \cdot 7 \\ 0 \cdot 4 + (-5) \cdot 3 + 3 \cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$
- $\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \cdot 4 + (-3) \cdot 7 \\ 8 \cdot 4 + 0 \cdot 7 \\ (-5) \cdot 4 + 2 \cdot 7 \end{bmatrix} = \begin{bmatrix} -13 \\ 32 \\ -6 \end{bmatrix}$
- $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \begin{bmatrix} 1 \cdot r + 0 \cdot s + 0 \cdot t \\ 0 \cdot r + 1 \cdot s + 0 \cdot t \\ 0 \cdot r + 0 \cdot s + 1 \cdot t \end{bmatrix} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$ ■

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



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

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

THEOREM 5

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

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

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

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

Entries in $\mathbf{u} + \mathbf{v}$

Columns of A

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

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

NUMERICAL NOTE

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



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 $A\mathbf{u} + A\mathbf{v}$.



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 **b** as a linear combination of the columns of *A*.

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

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

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



1.5 SOLUTION SETS OF LINEAR SYSTEMS

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

Homogeneous Linear Systems

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

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

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

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

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

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

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

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

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

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

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

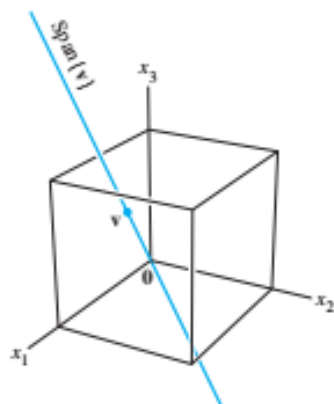


FIGURE 1



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

2.1 MATRIX OPERATIONS

If A is an $m \times n$ matrix—that is, a matrix with m rows and n columns—then the scalar entry in the i th row and j th column of A is denoted by a_{ij} and is called the (i, j) -entry of A . See Fig. 1. For instance, the $(3, 2)$ -entry is the number a_{32} in the third row, second column. Each column of A is a list of m real numbers, which identifies a vector in \mathbb{R}^m . Often, these columns are denoted by $\mathbf{a}_1, \dots, \mathbf{a}_n$, and the matrix A is written as

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

Observe that the number a_{ij} is the i th entry (from the top) of the j th column vector \mathbf{a}_j .

The **diagonal entries** in an $m \times n$ matrix $A = [a_{ij}]$ are $a_{11}, a_{22}, a_{33}, \dots$, and they form the **main diagonal** of A . A **diagonal matrix** is a square $n \times n$ matrix whose nondiagonal entries are zero. An example is the $n \times n$ identity matrix, I_n . An $m \times n$ matrix whose entries are all zero is a **zero matrix** and is written as 0 . The size of a zero matrix is usually clear from the context.



$$\begin{array}{c} \text{Column } j \\ \left[\begin{array}{cccc} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{array} \right] = A \\ \text{Row } i \\ \uparrow \quad \quad \quad \uparrow \quad \quad \quad \uparrow \\ a_1 \quad \quad \quad a_j \quad \quad \quad a_n \end{array}$$

FIGURE 1 Matrix notation.

Sums and Scalar Multiples

The arithmetic for vectors described earlier has a natural extension to matrices. We say that two matrices are **equal** if they have the same size (i.e., the same number of rows and the same number of columns) and if their corresponding columns are equal, which amounts to saying that their corresponding entries are equal. If A and B are $m \times n$ matrices, then the **sum** $A + B$ is the $m \times n$ matrix whose columns are the sums of the corresponding columns in A and B . Since vector addition of the columns is done entrywise, each entry in $A + B$ is the sum of the corresponding entries in A and B . The sum $A + B$ is defined only when A and B are the same size.

EXAMPLE 1 Let

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

Then

$$A + B = \begin{bmatrix} 5 & 1 & 6 \\ 2 & 8 & 9 \end{bmatrix}$$

but $A + C$ is not defined because A and C have different sizes. ■

If r is a scalar and A is a matrix, then the **scalar multiple** rA is the matrix whose columns are r times the corresponding columns in A . As with vectors, $-A$ stands for $(-1)A$, and $A - B$ is the same as $A + (-1)B$.

EXAMPLE 2 If A and B are the matrices in Example 1, then

$$\begin{aligned} 2B &= 2 \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} \\ A - 2B &= \begin{bmatrix} 4 & 0 & 5 \\ -1 & 3 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 2 & 2 \\ 6 & 10 & 14 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ -7 & -7 & -12 \end{bmatrix} \end{aligned}$$

It was unnecessary in Example 2 to compute $A - 2B$ as $A + (-1)2B$ because the usual rules of algebra apply to sums and scalar multiples of matrices, as the following theorem shows.

THEOREM 1

Let A , B , and C be matrices of the same size, and let r and s be scalars.

- | | |
|--------------------------------|-------------------------|
| a. $A + B = B + A$ | d. $r(A + B) = rA + rB$ |
| b. $(A + B) + C = A + (B + C)$ | e. $(r + s)A = rA + sA$ |
| c. $A + 0 = A$ | f. $r(sA) = (rs)A$ |



Each equality in Theorem 1 is verified by showing that the matrix on the left side has the same size as the matrix on the right and that corresponding columns are equal. Size is no problem because A , B , and C are equal in size. The equality of columns follows immediately from analogous properties of vectors. For instance, if the j th columns of A , B , and C are \mathbf{a}_j , \mathbf{b}_j , and \mathbf{c}_j , respectively, then the j th columns of $(A + B) + C$ and $A + (B + C)$ are

$$(\mathbf{a}_j + \mathbf{b}_j) + \mathbf{c}_j \quad \text{and} \quad \mathbf{a}_j + (\mathbf{b}_j + \mathbf{c}_j)$$

respectively. Since these two vector sums are equal for each j , property (b) is verified.

Because of the associative property of addition, we can simply write $A + B + C$ for the sum, which can be computed either as $(A + B) + C$ or as $A + (B + C)$. The same applies to sums of four or more matrices.

Matrix Multiplication

When a matrix B multiplies a vector \mathbf{x} , it transforms \mathbf{x} into the vector $B\mathbf{x}$. If this vector is then multiplied in turn by a matrix A , the resulting vector is $A(B\mathbf{x})$. See Fig. 2.



FIGURE 2 Multiplication by B and then A .

Thus $A(B\mathbf{x})$ is produced from \mathbf{x} by a *composition* of mappings—the linear transformations studied in Section 1.8. Our goal is to represent this composite mapping as multiplication by a single matrix, denoted by AB , so that

$$A(B\mathbf{x}) = (AB)\mathbf{x} \quad (1)$$

See Fig. 3.

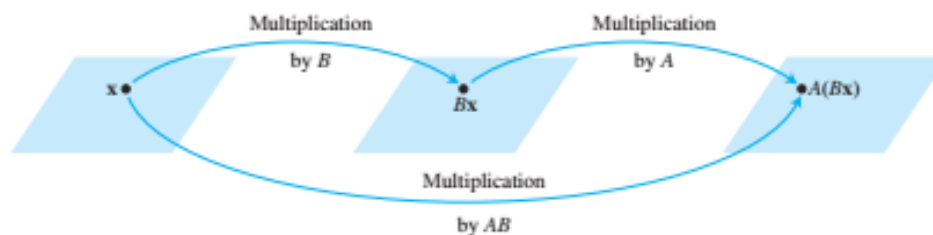


FIGURE 3 Multiplication by AB .

If A is $m \times n$, B is $n \times p$, and \mathbf{x} is in \mathbb{R}^p , denote the columns of B by $\mathbf{b}_1, \dots, \mathbf{b}_p$ and the entries in \mathbf{x} by x_1, \dots, x_p . Then

$$B\mathbf{x} = x_1\mathbf{b}_1 + \dots + x_p\mathbf{b}_p$$

By the linearity of multiplication by A ,

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1) + \dots + A(x_p\mathbf{b}_p) \\ &= x_1A\mathbf{b}_1 + \dots + x_pA\mathbf{b}_p \end{aligned}$$



The vector $A(B\mathbf{x})$ is a linear combination of the vectors $A\mathbf{b}_1, \dots, A\mathbf{b}_p$, using the entries in \mathbf{x} as weights. In matrix notation, this linear combination is written as

$$A(B\mathbf{x}) = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \dots \ A\mathbf{b}_p]\mathbf{x}$$

Thus multiplication by $[A\mathbf{b}_1 \ A\mathbf{b}_2 \ \dots \ A\mathbf{b}_p]$ transforms \mathbf{x} into $A(B\mathbf{x})$. We have found the matrix we sought!

DEFINITION

If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $\mathbf{b}_1, \dots, \mathbf{b}_p$, then the product AB is the $m \times p$ matrix whose columns are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$. That is,

$$AB = A[\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_p] = [A\mathbf{b}_1 \ A\mathbf{b}_2 \ \dots \ A\mathbf{b}_p]$$

This definition makes equation (1) true for all \mathbf{x} in \mathbb{R}^p . Equation (1) proves that the composite mapping in Fig. 3 is a linear transformation and that its standard matrix is AB . *Multiplication of matrices corresponds to composition of linear transformations.*

EXAMPLE 3 Compute AB , where $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

SOLUTION Write $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3]$, and compute:

$$\begin{aligned} A\mathbf{b}_1 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix}, & A\mathbf{b}_2 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix}, & A\mathbf{b}_3 &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 6 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 11 \\ -1 \end{bmatrix} & &= \begin{bmatrix} 0 \\ 13 \end{bmatrix} & &= \begin{bmatrix} 21 \\ -9 \end{bmatrix} \end{aligned}$$

Then

$$AB = A[\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}$$

$\uparrow \quad \uparrow \quad \uparrow$
 $A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad A\mathbf{b}_3$

Notice that since the first column of AB is $A\mathbf{b}_1$, this column is a linear combination of the columns of A using the entries in \mathbf{b}_1 as weights. A similar statement is true for each column of AB .

Each column of AB is a linear combination of the columns of A using weights from the corresponding column of B .

Obviously, the number of columns of A must match the number of rows in B in order for a linear combination such as $A\mathbf{b}_1$ to be defined. Also, the definition of AB shows that AB has the same number of rows as A and the same number of columns as B .

EXAMPLE 4 If A is a 3×5 matrix and B is a 5×2 matrix, what are the sizes of AB and BA , if they are defined?



SOLUTION Since A has 5 columns and B has 5 rows, the product AB is defined and is a 3×2 matrix:

$$\begin{array}{ccc} A & B & AB \\ \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} & \begin{bmatrix} * & * \\ * & * \\ * & * \\ * & * \\ * & * \end{bmatrix} & = \begin{bmatrix} * & * \\ * & * \\ * & * \end{bmatrix} \\ \begin{array}{c} 3 \times 5 \\ \uparrow \\ \text{Match} \\ \uparrow \\ 5 \times 2 \\ \text{Size of } AB \end{array} & & 3 \times 2 \end{array}$$

The product BA is *not* defined because the 2 columns of B do not match the 3 rows of A . ■

The definition of AB is important for theoretical work and applications, but the following rule provides a more efficient method for calculating the individual entries in AB when working small problems by hand.

ROW-COLUMN RULE FOR COMPUTING AB

If the product AB is defined, then the entry in row i and column j of AB is the sum of the products of corresponding entries from row i of A and column j of B . If $(AB)_{ij}$ denotes the (i, j) -entry in AB , and if A is an $m \times n$ matrix, then

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

To verify this rule, let $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$. Column j of AB is $A\mathbf{b}_j$, and we can compute $A\mathbf{b}_j$ by the row-vector rule for computing $A\mathbf{x}$ from Section 1.4. The i th entry in $A\mathbf{b}_j$ is the sum of the products of corresponding entries from row i of A and the vector \mathbf{b}_j , which is precisely the computation described in the rule for computing the (i, j) -entry of AB .

EXAMPLE 5 Use the row-column rule to compute two of the entries in AB for the matrices in Example 3. An inspection of the numbers involved will make it clear how the two methods for calculating AB produce the same matrix.

SOLUTION To find the entry in row 1 and column 3 of AB , consider row 1 of A and column 3 of B . Multiply corresponding entries and add the results, as shown below:

$$AB = \begin{array}{c} \rightarrow \end{array} \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{array}{c} \downarrow \\ \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} \end{array} = \begin{bmatrix} \square & \square & 2(6) + 3(3) \\ \square & \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & \square & \square \end{bmatrix}$$

For the entry in row 2 and column 2 of AB , use row 2 of A and column 2 of B :

$$\begin{array}{c} \rightarrow \end{array} \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{array}{c} \downarrow \\ \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} \end{array} = \begin{bmatrix} \square & \square & 21 \\ \square & 1(3) + -5(-2) & \square \end{bmatrix} = \begin{bmatrix} \square & \square & 21 \\ \square & 13 & \square \end{bmatrix}$$



EXAMPLE 6 Find the entries in the second row of AB , where

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

SOLUTION By the row-column rule, the entries of the second row of AB come from row 2 of A (and the columns of B):

$$\begin{aligned} & \begin{bmatrix} 2 & -5 & 0 \\ -1 & 3 & -4 \\ 6 & -8 & -7 \\ -3 & 0 & 9 \end{bmatrix} \begin{bmatrix} 4 & -6 \\ 7 & 1 \\ 3 & 2 \end{bmatrix} \\ &= \begin{bmatrix} \square & \square \\ -4 & 21 - 12 & 6 + 3 - 8 \\ \square & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} \square & \square \\ 5 & 1 \\ \square & \square \\ \square & \square \end{bmatrix} \end{aligned}$$

Notice that since Example 6 requested only the second row of AB , we could have written just the second row of A to the left of B and computed

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

This observation about rows of AB is true in general and follows from the row-column rule. Let $\text{row}_i(A)$ denote the i th row of a matrix A . Then

$$\text{row}_i(AB) = \text{row}_i(A) \cdot B \quad (2)$$

Properties of Matrix Multiplication

The following theorem lists the standard properties of matrix multiplication. Recall that I_m represents the $m \times m$ identity matrix and $I_m \mathbf{x} = \mathbf{x}$ for all \mathbf{x} in \mathbb{R}^m .

THEOREM 2

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

- $A(BC) = (AB)C$ (associative law of multiplication)
- $A(B + C) = AB + AC$ (left distributive law)
- $(B + C)A = BA + CA$ (right distributive law)
- $r(AB) = (rA)B = A(rB)$ for any scalar r
- $I_m A = A = A I_n$ (identity for matrix multiplication)

PROOF Properties (b)–(e) are considered in the exercises. Property (a) follows from the fact that matrix multiplication corresponds to composition of linear transformations (which are functions), and it is known (or easy to check) that the composition of functions is associative. Here is another proof of (a) that rests on the “column definition” of the product of two matrices. Let

$$C = [\mathbf{c}_1 \ \cdots \ \mathbf{c}_p]$$



By the definition of matrix multiplication,

$$BC = [Bc_1 \ \cdots \ Bc_p]$$

$$A(BC) = [A(Bc_1) \ \cdots \ A(Bc_p)]$$

Recall from equation (1) that the definition of AB makes $A(Bx) = (AB)x$ for all x , so

$$A(BC) = [(AB)c_1 \ \cdots \ (AB)c_p] = (AB)C \quad \blacksquare$$

The associative and distributive laws in Theorems 1 and 2 say essentially that pairs of parentheses in matrix expressions can be inserted and deleted in the same way as in the algebra of real numbers. In particular, we can write ABC for the product, which can be computed either as $A(BC)$ or as $(AB)C$.¹ Similarly, a product $ABCD$ of four matrices can be computed as $A(BCD)$ or $(ABC)D$ or $A(BC)D$, and so on. It does not matter how we group the matrices when computing the product, so long as the left-to-right order of the matrices is preserved.

The left-to-right order in products is critical because AB and BA are usually not the same. This is not surprising, because the columns of AB are linear combinations of the columns of A , whereas the columns of BA are constructed from the columns of B . The position of the factors in the product AB is emphasized by saying that A is *right-multiplied* by B or that B is *left-multiplied* by A . If $AB = BA$, we say that A and B **commute** with one another.

EXAMPLE 7 Let $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$. Show that these matrices do not commute. That is, verify that $AB \neq BA$.

SOLUTION

$$AB = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$$

$$BA = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 29 & -2 \end{bmatrix} \quad \blacksquare$$

Example 7 illustrates the first of the following list of important differences between matrix algebra and the ordinary algebra of real numbers. See Exercises 9–12 for examples of these situations.

WARNINGS:

1. In general, $AB \neq BA$.
2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)
3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

Powers of a Matrix

WEB

If A is an $n \times n$ matrix and if k is a positive integer, then A^k denotes the product of k

¹When B is square and C has fewer columns than A has rows, it is more efficient to compute $A(BC)$ than $(AB)C$.



copies of A :

$$A^k = \underbrace{A \cdots A}_k$$

If A is nonzero and if \mathbf{x} is in \mathbb{R}^n , then $A^k \mathbf{x}$ is the result of left-multiplying \mathbf{x} by A repeatedly k times. If $k = 0$, then $A^0 \mathbf{x}$ should be \mathbf{x} itself. Thus A^0 is interpreted as the identity matrix. Matrix powers are useful in both theory and applications (Sections 2.6, 4.9, and later in the text).

The Transpose of a Matrix

Given an $m \times n$ matrix A , the **transpose** of A is the $n \times m$ matrix, denoted by A^T , whose columns are formed from the corresponding rows of A .

EXAMPLE 8 Let

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

Then

$$A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \quad B^T = \begin{bmatrix} -5 & 1 & 0 \\ 2 & -3 & 4 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 & -3 \\ 1 & 5 \\ 1 & -2 \\ 1 & 7 \end{bmatrix} \quad \blacksquare$$

THEOREM 3

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- For any scalar r , $(rA)^T = rA^T$
- $(AB)^T = B^T A^T$

Proofs of (a)–(c) are straightforward and are omitted. For (d), see Exercise 33. Usually, $(AB)^T$ is not equal to $A^T B^T$, even when A and B have sizes such that the product $A^T B^T$ is defined.

The generalization of Theorem 3(d) to products of more than two factors can be stated in words as follows:

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

The exercises contain numerical examples that illustrate properties of transposes.



NUMERICAL NOTES

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

PRACTICE PROBLEMS

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

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

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

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

2.1 EXERCISES

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

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

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

1. $-2A$, $B - 2A$, AC , CD

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

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

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

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 $B \neq C$.



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,

$$\mathbf{x}\mathbf{x}^T = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \begin{bmatrix} 5 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 15 \\ 15 & 9 \end{bmatrix}$$

$$\mathbf{x}^T \mathbf{x} = \begin{bmatrix} 5 & 3 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = [25 + 9] = 34$$

A 1×1 matrix such as $\mathbf{x}^T \mathbf{x}$ is usually written without the brackets. Finally, $A^T \mathbf{x}^T$ is not defined, because \mathbf{x}^T does not have two rows to match the two columns of A^T .

2. The fastest way to compute $A^2 \mathbf{x}$ is to compute $A(A\mathbf{x})$. The product $A\mathbf{x}$ requires 16 multiplications, 4 for each entry, and $A(A\mathbf{x})$ requires 16 more. In contrast, the product A^2 requires 64 multiplications, 4 for each of the 16 entries in A^2 . After that, $A^2 \mathbf{x}$ takes 16 more multiplications, for a total of 80.

2.2 THE INVERSE OF A MATRIX

Matrix algebra provides tools for manipulating matrix equations and creating various useful formulas in ways similar to doing ordinary algebra with real numbers. This section investigates the matrix analogue of the reciprocal, or multiplicative inverse, of a nonzero number.



Recall that the multiplicative inverse of a number such as 5 is $1/5$ or 5^{-1} . This inverse satisfies the equations

$$5^{-1} \cdot 5 = 1 \quad \text{and} \quad 5 \cdot 5^{-1} = 1$$

The matrix generalization requires *both* equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative. Furthermore, a full generalization is possible only if the matrices involved are square.¹

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A . In fact, C is uniquely determined by A , because if B were another inverse of A , then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1} , so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

EXAMPLE 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and}$$

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

Thus $C = A^{-1}$. ■

Here is a simple formula for the inverse of a 2×2 matrix, along with a test to tell if the inverse exists.

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A , and we write

$$\det A = ad - bc$$

Theorem 4 says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

¹One could say that an $m \times n$ matrix A is invertible if there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. However, these equations imply that A is square and $C = D$. Thus A is invertible as defined above. See Exercises 23–25 in Section 2.1.



Recall that the multiplicative inverse of a number such as 5 is $1/5$ or 5^{-1} . This inverse satisfies the equations

$$5^{-1} \cdot 5 = 1 \quad \text{and} \quad 5 \cdot 5^{-1} = 1$$

The matrix generalization requires *both* equations and avoids the slanted-line notation (for division) because matrix multiplication is not commutative. Furthermore, a full generalization is possible only if the matrices involved are square.¹

An $n \times n$ matrix A is said to be **invertible** if there is an $n \times n$ matrix C such that

$$CA = I \quad \text{and} \quad AC = I$$

where $I = I_n$, the $n \times n$ identity matrix. In this case, C is an **inverse** of A . In fact, C is uniquely determined by A , because if B were another inverse of A , then $B = BI = B(AC) = (BA)C = IC = C$. This unique inverse is denoted by A^{-1} , so that

$$A^{-1}A = I \quad \text{and} \quad AA^{-1} = I$$

A matrix that is *not* invertible is sometimes called a **singular matrix**, and an invertible matrix is called a **nonsingular matrix**.

EXAMPLE 1 If $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$ and $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$, then

$$AC = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \\ CA = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus $C = A^{-1}$. ■

Here is a simple formula for the inverse of a 2×2 matrix, along with a test to tell if the inverse exists.

THEOREM 4

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

If $ad - bc = 0$, then A is not invertible.

The simple proof of Theorem 4 is outlined in Exercises 25 and 26. The quantity $ad - bc$ is called the **determinant** of A , and we write

$$\det A = ad - bc$$

Theorem 4 says that a 2×2 matrix A is invertible if and only if $\det A \neq 0$.

¹One could say that an $m \times n$ matrix A is invertible if there exist $n \times m$ matrices C and D such that $CA = I_n$ and $AD = I_m$. However, these equations imply that A is square and $C = D$. Thus A is invertible as defined above. See Exercises 23–25 in Section 2.1.



EXAMPLE 2 Find the inverse of $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$.

SOLUTION Since $\det A = 3(6) - 4(5) = -2 \neq 0$, A is invertible, and

$$A^{-1} = \frac{1}{-2} \begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} 6/(-2) & -4/(-2) \\ -5/(-2) & 3/(-2) \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix} \quad \blacksquare$$

Invertible matrices are indispensable in linear algebra—mainly for algebraic calculations and formula derivations, as in the next theorem. There are also occasions when an inverse matrix provides insight into a mathematical model of a real-life situation, as in Example 3, below.

THEOREM 5

If A is an invertible $n \times n$ matrix, then for each \mathbf{b} in \mathbb{R}^n , the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

PROOF Take any \mathbf{b} in \mathbb{R}^n . A solution exists because if $A^{-1}\mathbf{b}$ is substituted for \mathbf{x} , then $A\mathbf{x} = A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. So $A^{-1}\mathbf{b}$ is a solution. To prove that the solution is unique, show that if \mathbf{u} is any solution, then \mathbf{u} , in fact, must be $A^{-1}\mathbf{b}$. Indeed, if $A\mathbf{u} = \mathbf{b}$, we can multiply both sides by A^{-1} and obtain

$$A^{-1}A\mathbf{u} = A^{-1}\mathbf{b}, \quad I\mathbf{u} = A^{-1}\mathbf{b}, \quad \text{and} \quad \mathbf{u} = A^{-1}\mathbf{b} \quad \blacksquare$$



THEOREM 6

a. If A is an invertible matrix, then A^{-1} is invertible and

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

b. If A and B are $n \times n$ invertible matrices, then so is AB , and the inverse of AB is the product of the inverses of A and B in the reverse order. That is,

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

c. If A is an invertible matrix, then so is A^T , and the inverse of A^T is the transpose of A^{-1} . That is,

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

PROOF To verify statement (a), find a matrix C such that

$$A^{-1}C = I \quad \text{and} \quad CA^{-1} = I$$

In fact, these equations are satisfied with A in place of C . Hence A^{-1} is invertible, and A is its inverse. Next, to prove statement (b), compute:

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

A similar calculation shows that $(B^{-1}A^{-1})(AB) = I$. For statement (c), use Theorem 3(d), read from right to left, $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$. Similarly, $A^T (A^{-1})^T = I^T = I$. Hence A^T is invertible, and its inverse is $(A^{-1})^T$. ■

The following generalization of Theorem 6(b) is needed later.

The product of $n \times n$ invertible matrices is invertible, and the inverse is the product of their inverses in the reverse order.



There is an important connection between invertible matrices and row operations that leads to a method for computing inverses. As we shall see, an invertible matrix A is row equivalent to an identity matrix, and we can find A^{-1} by *watching the row reduction of A to I* .

Elementary Matrices

An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix. The next example illustrates the three kinds of elementary matrices.

EXAMPLE 5 Let

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix},$$

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

Compute E_1A , E_2A , and E_3A , and describe how these products can be obtained by elementary row operations on A .

SOLUTION Verify that

$$E_1A = \begin{bmatrix} a & b & c \\ d & e & f \\ g - 4a & h - 4b & i - 4c \end{bmatrix}, \quad E_2A = \begin{bmatrix} d & e & f \\ a & b & c \\ g & h & i \end{bmatrix},$$

$$E_3A = \begin{bmatrix} a & b & c \\ d & e & f \\ 5g & 5h & 5i \end{bmatrix}.$$

Addition of -4 times row 1 of A to row 3 produces E_1A . (This is a row replacement operation.) An interchange of rows 1 and 2 of A produces E_2A , and multiplication of row 3 of A by 5 produces E_3A . ■

Left-multiplication (that is, multiplication on the left) by E_1 in Example 5 has the same effect on any $3 \times n$ matrix. It adds -4 times row 1 to row 3. In particular, since $E_1 \cdot I = E_1$, we see that E_1 itself is produced by this same row operation on the identity. Thus Example 5 illustrates the following general fact about elementary matrices. See Exercises 27 and 28.

If an elementary row operation is performed on an $m \times n$ matrix A , the resulting matrix can be written as EA , where the $m \times m$ matrix E is created by performing the same row operation on I_m .

Since row operations are reversible, as shown in Section 1.1, elementary matrices are invertible, for if E is produced by a row operation on I , then there is another row operation of the same type that changes E back into I . Hence there is an elementary matrix F such that $FE = I$. Since E and F correspond to reverse operations, $EF = I$, too.



Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I .

EXAMPLE 6 Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

SOLUTION To transform E_1 into I , add +4 times row 1 to row 3. The elementary matrix that does this is

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ +4 & 0 & 1 \end{bmatrix}$$

The following theorem provides the best way to “visualize” an invertible matrix, and the theorem leads immediately to a method for finding the inverse of a matrix.

THEOREM 7

An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n , and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1} .

PROOF Suppose that A is invertible. Then, since the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each \mathbf{b} (Theorem 5), A has a pivot position in every row (Theorem 4 in Section 1.4). Because A is square, the n pivot positions must be on the diagonal, which implies that the reduced echelon form of A is I_n . That is, $A \sim I_n$.

Now suppose, conversely, that $A \sim I_n$. Then, since each step of the row reduction of A corresponds to left-multiplication by an elementary matrix, there exist elementary matrices E_1, \dots, E_p such that

$$A \sim E_1 A \sim E_2 (E_1 A) \sim \dots \sim E_p (E_{p-1} \dots E_1 A) = I_n$$

That is,

$$E_p \dots E_1 A = I_n \quad (1)$$

Since the product $E_p \dots E_1$ of invertible matrices is invertible, (1) leads to

$$\begin{aligned} (E_p \dots E_1)^{-1} (E_p \dots E_1) A &= (E_p \dots E_1)^{-1} I_n \\ A &= (E_p \dots E_1)^{-1} \end{aligned}$$

Thus A is invertible, as it is the inverse of an invertible matrix (Theorem 6). Also,

$$A^{-1} = [(E_p \dots E_1)^{-1}]^{-1} = E_p \dots E_1$$

Then $A^{-1} = E_p \dots E_1 \cdot I_n$, which says that A^{-1} results from applying E_1, \dots, E_p successively to I_n . This is the same sequence in (1) that reduced A to I_n . ■

An Algorithm for Finding A^{-1}

If we place A and I side-by-side to form an augmented matrix $[A \ I]$, then row operations on this matrix produce identical operations on A and on I . By Theorem 7, either there are row operations that transform A to I_n and I_n to A^{-1} or else A is not invertible.



ALGORITHM FOR FINDING A^{-1}

Row reduce the augmented matrix $[A \ I]$. If A is row equivalent to I , then $[A \ I]$ is row equivalent to $[I \ A^{-1}]$. Otherwise, A does not have an inverse.

EXAMPLE 7 Find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$, if it exists.

SOLUTION

$$\begin{aligned} [A \ I] &= \begin{bmatrix} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{bmatrix} \end{aligned}$$

Theorem 7 shows, since $A \sim I$, that A is invertible, and

$$A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

It is a good idea to check the final answer:

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

It is not necessary to check that $A^{-1}A = I$ since A is invertible. ■



WEB

NUMERICAL NOTE

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.

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.

$$\begin{aligned} 2. [A \ I] &\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ -1 & 5 & 6 & 0 & 1 & 0 \\ 5 & -4 & 5 & 0 & 0 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 6 & 10 & -5 & 0 & 1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -2 & -1 & 1 & 0 & 0 \\ 0 & 3 & 5 & 1 & 1 & 0 \\ 0 & 0 & 0 & -7 & -2 & 1 \end{bmatrix} \end{aligned}$$

So $[A \ I]$ is row equivalent to a matrix of the form $[B \ D]$, 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.



2.4 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

$$A = \left[\begin{array}{ccc|cc|c} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ -8 & -6 & 3 & 1 & 7 & -4 \end{array} \right]$$

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*)

$$\begin{aligned} A_{11} &= \begin{bmatrix} 3 & 0 & -1 \\ -5 & 2 & 4 \end{bmatrix}, & A_{12} &= \begin{bmatrix} 5 & 9 \\ 0 & -3 \end{bmatrix}, & A_{13} &= \begin{bmatrix} -2 \\ 1 \end{bmatrix} \\ A_{21} &= \begin{bmatrix} -8 & -6 & 3 \end{bmatrix}, & A_{22} &= \begin{bmatrix} 1 & 7 \end{bmatrix}, & A_{23} &= \begin{bmatrix} -4 \end{bmatrix} \end{aligned}$$



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

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

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.



Addition and Scalar Multiplication

If matrices A and B are the same size and are partitioned in exactly the same way, then it is natural to make the same partition of the ordinary matrix sum $A + B$. In this case, each block of $A + B$ is the (matrix) sum of the corresponding blocks of A and B . Multiplication of a partitioned matrix by a scalar is also computed block by block.

Multiplication of Partitioned Matrices

Partitioned matrices can be multiplied by the usual row-column rule as if the block entries were scalars, provided that for a product AB , the column partition of A matches the row partition of B .

EXAMPLE 3 Let

$$A = \left[\begin{array}{ccc|cc} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ 0 & -4 & -2 & 7 & -1 \end{array} \right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \left[\begin{array}{cc} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ -1 & 3 \\ 5 & 2 \end{array} \right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

The 5 columns of A are partitioned into a set of 3 columns and then a set of 2 columns. The 5 rows of B are partitioned in the same way—into a set of 3 rows and then a set of 2 rows. We say that the partitions of A and B are **conformable** for **block multiplication**. It can be shown that the ordinary product AB can be written as

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \\ 2 & 1 \end{bmatrix}$$

It is important for each smaller product in the expression for AB to be written with the submatrix from A on the left, since matrix multiplication is not commutative. For instance,

$$A_{11}B_1 = \begin{bmatrix} 2 & -3 & 1 \\ 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ -2 & 1 \\ -3 & 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix}$$

$$A_{12}B_2 = \begin{bmatrix} 0 & -4 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix}$$

Hence the top block in AB is

$$A_{11}B_1 + A_{12}B_2 = \begin{bmatrix} 15 & 12 \\ 2 & -5 \end{bmatrix} + \begin{bmatrix} -20 & -8 \\ -8 & 7 \end{bmatrix} = \begin{bmatrix} -5 & 4 \\ -6 & 2 \end{bmatrix} \quad \blacksquare$$

The row-column rule for multiplication of block matrices provides the most general way to regard the product of two matrices. Each of the following views of a product has already been described using simple partitions of matrices: (1) the definition of Ax using the columns of A , (2) the column definition of AB , (3) the row-column rule for computing AB , and (4) the rows of AB as products of the rows of A and the matrix B . A fifth view of AB , again using partitions, follows in Theorem 10 below.

The calculations in the next example prepare the way for Theorem 10. Here $\text{col}_k(A)$ is the k th column of A , and $\text{row}_k(B)$ is the k th row of B .



EXAMPLE 4 Let $A = \begin{bmatrix} -3 & 1 & 2 \\ 1 & -4 & 5 \end{bmatrix}$ and $B = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}$. Verify that

$$AB = \text{col}_1(A) \text{row}_1(B) + \text{col}_2(A) \text{row}_2(B) + \text{col}_3(A) \text{row}_3(B)$$

SOLUTION Each term above is an *outer product*. (See Exercises 27 and 28 in Section 2.1.) By the row-column rule for computing a matrix product,

$$\text{col}_1(A) \text{row}_1(B) = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} -3a & -3b \\ a & b \end{bmatrix}$$

$$\text{col}_2(A) \text{row}_2(B) = \begin{bmatrix} 1 \\ -4 \end{bmatrix} \begin{bmatrix} c & d \end{bmatrix} = \begin{bmatrix} c & d \\ -4c & -4d \end{bmatrix}$$

$$\text{col}_3(A) \text{row}_3(B) = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \begin{bmatrix} e & f \end{bmatrix} = \begin{bmatrix} 2e & 2f \\ 5e & 5f \end{bmatrix}$$

Thus

$$\sum_{k=1}^3 \text{col}_k(A) \text{row}_k(B) = \begin{bmatrix} -3a + c + 2e & -3b + d + 2f \\ a - 4c + 5e & b - 4d + 5f \end{bmatrix}$$

This matrix is obviously AB . Notice that the $(1, 1)$ -entry in AB is the sum of the $(1, 1)$ -entries in the three outer products, the $(1, 2)$ -entry in AB is the sum of the $(1, 2)$ -entries in the three outer products, and so on. ■

THEOREM 10

Column-Row Expansion of AB

If A is $m \times n$ and B is $n \times p$, then

$$AB = [\text{col}_1(A) \quad \text{col}_2(A) \quad \cdots \quad \text{col}_n(A)] \begin{bmatrix} \text{row}_1(B) \\ \text{row}_2(B) \\ \vdots \\ \text{row}_n(B) \end{bmatrix} \quad (1)$$

$$= \text{col}_1(A) \text{row}_1(B) + \cdots + \text{col}_n(A) \text{row}_n(B)$$

PROOF For each row index i and column index j , the (i, j) -entry in $\text{col}_k(A) \text{row}_k(B)$ is the product of a_{ik} from $\text{col}_k(A)$ and b_{kj} from $\text{row}_k(B)$. Hence the (i, j) -entry in the sum shown in equation (1) is

$$\underbrace{a_{i1}b_{1j}}_{(k=1)} + \underbrace{a_{i2}b_{2j}}_{(k=2)} + \cdots + \underbrace{a_{in}b_{nj}}_{(k=n)}$$

This sum is also the (i, j) -entry in AB , by the row-column rule. ■

Inverses of Partitioned Matrices

The next example illustrates calculations involving inverses and partitioned matrices.

EXAMPLE 5 A matrix of the form

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

is said to be *block upper triangular*. Assume that A_{11} is $p \times p$, A_{22} is $q \times q$, and A is invertible. Find a formula for A^{-1} .



SOLUTION Denote A^{-1} by B and partition B so that

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_q \end{bmatrix} \quad (2)$$

This matrix equation provides four equations that will lead to the unknown blocks B_{11}, \dots, B_{22} . Compute the product on the left side of equation (2), and equate each entry with the corresponding block in the identity matrix on the right. That is, set

$$A_{11}B_{11} + A_{12}B_{21} = I_p \quad (3)$$

$$A_{11}B_{12} + A_{12}B_{22} = 0 \quad (4)$$

$$A_{22}B_{21} = 0 \quad (5)$$

$$A_{22}B_{22} = I_q \quad (6)$$

By itself, equation (6) does not show that A_{22} is invertible. However, since A_{22} is square, the Invertible Matrix Theorem and (6) together show that A_{22} is invertible and $B_{22} = A_{22}^{-1}$. Next, left-multiply both sides of (5) by A_{22}^{-1} and obtain

$$B_{21} = A_{22}^{-1}0 = 0$$

so that (3) simplifies to

$$A_{11}B_{11} + 0 = I_p$$

Since A_{11} is square, this shows that A_{11} is invertible and $B_{11} = A_{11}^{-1}$. Finally, use these results with (4) to find that

$$A_{11}B_{12} = -A_{12}B_{22} = -A_{12}A_{22}^{-1} \quad \text{and} \quad B_{12} = -A_{11}^{-1}A_{12}A_{22}^{-1}$$

Thus

$$A^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix} \quad \blacksquare$$

A **block diagonal matrix** is a partitioned matrix with zero blocks off the main diagonal (of blocks). Such a matrix is invertible if and only if each block on the diagonal is invertible. See Exercises 13 and 14.

NUMERICAL NOTES

1. When matrices are too large to fit in a computer's high-speed memory, partitioning permits the computer to work with only two or three submatrices at a time. For instance, one linear programming research team simplified a problem by partitioning the matrix into 837 rows and 51 columns. The problem's solution took about 4 minutes on a Cray supercomputer.¹
2. Some high-speed computers, particularly those with vector pipeline architecture, perform matrix calculations more efficiently when the algorithms use partitioned matrices.²
3. Professional software for high-performance numerical linear algebra, such as LAPACK, makes intensive use of partitioned matrix calculations.

¹The solution time doesn't sound too impressive until you learn that each of the 51 block columns contained about 250,000 individual columns. The original problem had 837 equations and over 12,750,000 variables! Nearly 100 million of the more than 10 billion entries in the matrix were nonzero. See Robert E. Bixby et al., "Very Large-Scale Linear Programming: A Case Study in Combining Interior Point and Simplex Methods," *Operations Research*, 40, no. 5 (1992): 885–897.

²The importance of block matrix algorithms for computer calculations is described in *Matrix Computations*, 3rd ed., by Gene H. Golub and Charles F. van Loan (Baltimore: Johns Hopkins University Press, 1996).



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 $X^T X$, 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. $\begin{bmatrix} X & 0 & 0 \\ Y & 0 & I \end{bmatrix} \begin{bmatrix} A & Z \\ 0 & 0 \\ B & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$

In Exercises 11 and 12, mark each statement True or False. Justify each answer.

11. a. If $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$ and $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, with A_1 and A_2 the same sizes as B_1 and B_2 , respectively, then $A + B = \begin{bmatrix} A_1 + B_1 & A_2 + B_2 \end{bmatrix}$.
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 $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, 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 $A^T = \begin{bmatrix} P^T & R^T \\ Q^T & S^T \end{bmatrix}$.
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.



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. $X^T X = \begin{bmatrix} X_1^T \\ X_2^T \end{bmatrix} \begin{bmatrix} X_1 & X_2 \end{bmatrix} = \begin{bmatrix} X_1^T X_1 & X_1^T X_2 \\ X_2^T X_1 & X_2^T X_2 \end{bmatrix}$. 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 $X^T X$ is used in several computer algorithms for matrix computations.



3.1 INTRODUCTION TO DETERMINANTS

Recall from Section 2.2 that a 2×2 matrix is invertible if and only if its determinant is nonzero. To extend this useful fact to larger matrices, we need a definition for the determinant of an $n \times n$ matrix. We can discover the definition for the 3×3 case by watching what happens when an invertible 3×3 matrix A is row reduced.

Consider $A = [a_{ij}]$ with $a_{11} \neq 0$. If we multiply the second and third rows of A by a_{11} and then subtract appropriate multiples of the first row from the other two rows, we find that A is row equivalent to the following two matrices:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11}a_{21} & a_{11}a_{22} & a_{11}a_{23} \\ a_{11}a_{31} & a_{11}a_{32} & a_{11}a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{bmatrix} \quad (1)$$

Since A is invertible, either the $(2, 2)$ -entry or the $(3, 2)$ -entry on the right in (1) is nonzero. Let us suppose that the $(2, 2)$ -entry is nonzero. (Otherwise, we can make a row interchange before proceeding.) Multiply row 3 by $a_{11}a_{22} - a_{12}a_{21}$, and then to the new row 3 add $-(a_{11}a_{32} - a_{12}a_{31})$ times row 2. This will show that

$$A \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & 0 & a_{11}\Delta \end{bmatrix}$$

where

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \quad (2)$$

Since A is invertible, Δ must be nonzero. The converse is true, too, as we will see in Section 3.2. We call Δ in (2) the **determinant** of the 3×3 matrix A .

Recall that the determinant of a 2×2 matrix, $A = [a_{ij}]$, is the number

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

For a 1×1 matrix—say, $A = [a_{11}]$ —we define $\det A = a_{11}$. To generalize the definition of the determinant to larger matrices, we'll use 2×2 determinants to rewrite the 3×3 determinant Δ described above. Since the terms in Δ can be grouped as $(a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31})$,

$$\Delta = a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

For brevity, write

$$\Delta = a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13} \quad (3)$$

where A_{11} , A_{12} , and A_{13} are obtained from A by deleting the first row and one of the three columns. For any square matrix A , let A_{ij} denote the submatrix formed by deleting



the i th row and j th column of A . For instance, if

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

then A_{32} is obtained by crossing out row 3 and column 2,

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

so that

$$A_{32} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$$

We can now give a *recursive* definition of a determinant. When $n = 3$, $\det A$ is defined using determinants of the 2×2 submatrices A_{ij} , as in (3) above. When $n = 4$, $\det A$ uses determinants of the 3×3 submatrices A_{ij} . In general, an $n \times n$ determinant is defined by determinants of $(n - 1) \times (n - 1)$ submatrices.

DEFINITION

For $n \geq 2$, the **determinant** of an $n \times n$ matrix $A = [a_{ij}]$ is the sum of n terms of the form $\pm a_{1j} \det A_{1j}$, with plus and minus signs alternating, where the entries $a_{11}, a_{12}, \dots, a_{1n}$ are from the first row of A . In symbols,

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned}$$

EXAMPLE 1 Compute the determinant of

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

SOLUTION Compute $\det A = a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13}$:

$$\begin{aligned} \det A &= 1 \cdot \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \cdot \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1(0 - 2) - 5(0 - 0) + 0(-4 - 0) = -2 \end{aligned}$$

Another common notation for the determinant of a matrix uses a pair of vertical lines in place of brackets. Thus the calculation in Example 1 can be written as

$$\det A = 1 \begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5 \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} = \cdots = -2$$

To state the next theorem, it is convenient to write the definition of $\det A$ in a slightly different form. Given $A = [a_{ij}]$, the (i, j) -**cofactor** of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij} \quad (4)$$

Then

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}$$



This formula is called a **cofactor expansion across the first row** of A . We omit the proof of the following fundamental theorem to avoid a lengthy digression.

THEOREM 1

The determinant of an $n \times n$ matrix A can be computed by a cofactor expansion across any row or down any column. The expansion across the i th row using the cofactors in (4) is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

The cofactor expansion down the j th column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

The plus or minus sign in the (i, j) -cofactor depends on the position of a_{ij} in the matrix, regardless of the sign of a_{ij} itself. The factor $(-1)^{i+j}$ determines the following checkerboard pattern of signs:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

EXAMPLE 2 Use a cofactor expansion across the third row to compute $\det A$, where

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

SOLUTION Compute

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= (-1)^{3+1}a_{31}\det A_{31} + (-1)^{3+2}a_{32}\det A_{32} + (-1)^{3+3}a_{33}\det A_{33} \\ &= 0 \begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2) \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0 \begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= 0 + 2(-1) + 0 = -2 \end{aligned}$$

Theorem 1 is helpful for computing the determinant of a matrix that contains many zeros. For example, if a row is mostly zeros, then the cofactor expansion across that row has many terms that are zero, and the cofactors in those terms need not be calculated. The same approach works with a column that contains many zeros.

EXAMPLE 3 Compute $\det A$, where

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

SOLUTION The cofactor expansion down the first column of A has all terms equal to zero except the first. Thus

$$\det A = 3 \cdot \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} - 0 \cdot C_{21} + 0 \cdot C_{31} - 0 \cdot C_{41} + 0 \cdot C_{51}$$



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



168 CHAPTER 3 Determinants

$$7. \begin{vmatrix} 4 & 3 & 0 \\ 6 & 5 & 2 \\ 9 & 7 & 3 \end{vmatrix}$$

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

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.

$$9. \begin{vmatrix} 6 & 0 & 0 & 5 \\ 1 & 7 & 2 & -5 \\ 2 & 0 & 0 & 0 \\ 8 & 3 & 1 & 8 \end{vmatrix}$$

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

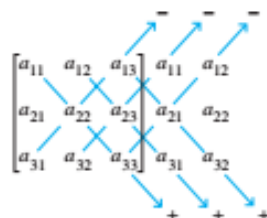
$$11. \begin{vmatrix} 3 & 5 & -8 & 4 \\ 0 & -2 & 3 & -7 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 2 \end{vmatrix}$$

$$12. \begin{vmatrix} 4 & 0 & 0 & 0 \\ 7 & -1 & 0 & 0 \\ 2 & 6 & 3 & 0 \\ 5 & -8 & 4 & -3 \end{vmatrix}$$

$$13. \begin{vmatrix} 4 & 0 & -7 & 3 & -5 \\ 0 & 0 & 2 & 0 & 0 \\ 7 & 3 & -6 & 4 & -8 \\ 5 & 0 & 5 & 2 & -3 \\ 0 & 0 & 9 & -1 & 2 \end{vmatrix}$$

$$14. \begin{vmatrix} 6 & 3 & 2 & 4 & 0 \\ 9 & 0 & -4 & 1 & 0 \\ 8 & -5 & 6 & 7 & 1 \\ 3 & 0 & 0 & 0 & 0 \\ 4 & 2 & 3 & 2 & 0 \end{vmatrix}$$

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 i th row and j th 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 $\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}$. 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 $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ 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 $\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{v} & \mathbf{u} \end{bmatrix}$. 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.

- If a multiple of one row of A is added to another row to produce a matrix B , then $\det B = \det A$.
- If two rows of A are interchanged to produce B , then $\det B = -\det A$.
- 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.



EXAMPLE 1 Compute $\det A$, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

SOLUTION The strategy is to reduce A to echelon form and then to use the fact that the determinant of a triangular matrix is the product of the diagonal entries. The first two row replacements in column 1 do not change the determinant:

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

An interchange of rows 2 and 3 reverses the sign of the determinant, so

$$\det A = - \begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -(1)(3)(-5) = 15 \quad \blacksquare$$

A common use of Theorem 3(c) in hand calculations is to *factor out a common multiple of one row* of a matrix. For instance,

$$\begin{vmatrix} * & * & * \\ 5k & -2k & 3k \\ * & * & * \end{vmatrix} = k \begin{vmatrix} * & * & * \\ 5 & -2 & 3 \\ * & * & * \end{vmatrix}$$

where the starred entries are unchanged. We use this step in the next example.

EXAMPLE 2 Compute $\det A$, where $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

SOLUTION To simplify the arithmetic, we want a 1 in the upper-left corner. We could interchange rows 1 and 4. Instead, we factor out 2 from the top row, and then proceed with row replacements in the first column:

$$\det A = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Next, we could factor out another 2 from row 3 or use the 3 in the second column as a pivot. We choose the latter operation, adding 4 times row 2 to row 3:

$$\det A = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

Finally, adding $-1/2$ times row 3 to row 4, and computing the “triangular” determinant, we find that

$$\det A = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 2 \cdot (1)(3)(-6)(1) = -36 \quad \blacksquare$$



$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix}$$

$\det U \neq 0$

$$U = \begin{bmatrix} \blacksquare & * & * & * \\ 0 & \blacksquare & * & * \\ 0 & 0 & 0 & \blacksquare \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$\det U = 0$

FIGURE 1
Typical echelon forms of square matrices.

Suppose a square matrix A has been reduced to an echelon form U by row replacements and row interchanges. (This is always possible. See the row reduction algorithm in Section 1.2.) If there are r interchanges, then Theorem 3 shows that

$$\det A = (-1)^r \det U$$

Since U is in echelon form, it is triangular, and so $\det U$ is the product of the diagonal entries u_{11}, \dots, u_{nn} . If A is invertible, the entries u_{ii} are all pivots (because $A \sim I_n$ and the u_{ii} have not been scaled to 1's). Otherwise, at least u_{nn} is zero, and the product $u_{11} \cdots u_{nn}$ is zero. See Fig. 1. Thus

$$\det A = \begin{cases} (-1)^r \cdot \left(\text{product of pivots in } U \right) & \text{when } A \text{ is invertible} \\ 0 & \text{when } A \text{ is not invertible} \end{cases} \quad (1)$$

It is interesting to note that although the echelon form U described above is not unique (because it is not completely row reduced), and the pivots are not unique, the *product* of the pivots *is* unique, except for a possible minus sign.

Formula (1) not only gives a concrete interpretation of $\det A$ but also proves the main theorem of this section:

THEOREM 4

A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 4 adds the statement " $\det A \neq 0$ " to the Invertible Matrix Theorem. A useful corollary is that $\det A = 0$ when the columns of A are linearly dependent. Also, $\det A = 0$ when the rows of A are linearly dependent. (Rows of A are columns of A^T , and linearly dependent columns of A^T make A^T singular. When A^T is singular, so is A , by the Invertible Matrix Theorem.) In practice, linear dependence is obvious when two columns or two rows are the same or a column or a row is zero.

EXAMPLE 3 Compute $\det A$, where $A = \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ -6 & 7 & -7 & 4 \\ -5 & -8 & 0 & 9 \end{bmatrix}$.

SOLUTION Add 2 times row 1 to row 3 to obtain

$$\det A = \det \begin{bmatrix} 3 & -1 & 2 & -5 \\ 0 & 5 & -3 & -6 \\ 0 & 5 & -3 & -6 \\ -5 & -8 & 0 & 9 \end{bmatrix} = 0$$

because the second and third rows of the second matrix are equal. ■

NUMERICAL NOTES

1. Most computer programs that compute $\det A$ for a general matrix A use the method of formula (1) above.
2. It can be shown that evaluation of an $n \times n$ determinant using row operations requires about $2n^3/3$ arithmetic operations. Any modern microcomputer can calculate a 25×25 determinant in a fraction of a second, since only about 10,000 operations are required.

WEB



Computers can also handle large “sparse” matrices, with special routines that take advantage of the presence of many zeros. Of course, zero entries can speed hand computations, too. The calculations in the next example combine the power of row operations with the strategy from Section 3.1 of using zero entries in cofactor expansions.

EXAMPLE 4 Compute $\det A$, where $A = \begin{bmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{bmatrix}$.

SOLUTION A good way to begin is to use the 2 in column 1 as a pivot, eliminating the -2 below it. Then use a cofactor expansion to reduce the size of the determinant, followed by another row replacement operation. Thus

$$\det A = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix} = -2 \begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix}$$

An interchange of rows 2 and 3 would produce a “triangular determinant.” Another approach is to make a cofactor expansion down the first column:

$$\det A = (-2)(1) \begin{vmatrix} 0 & 5 \\ -3 & 1 \end{vmatrix} = -2 \cdot (15) = -30$$

Column Operations

We can perform operations on the columns of a matrix in a way that is analogous to the row operations we have considered. The next theorem shows that column operations have the same effects on determinants as row operations.

THEOREM 5

If A is an $n \times n$ matrix, then $\det A^T = \det A$.

PROOF The theorem is obvious for $n = 1$. Suppose the theorem is true for $k \times k$ determinants and let $n = k + 1$. Then the cofactor of a_{ij} in A equals the cofactor of a_{ji} in A^T , because the cofactors involve $k \times k$ determinants. Hence the cofactor expansion of $\det A$ along the first *row* equals the cofactor expansion of $\det A^T$ down the first *column*. That is, A and A^T have equal determinants. Thus the theorem is true for $n = 1$, 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 is true for all $n \geq 1$.

Because of Theorem 5, each statement in Theorem 3 is true when the word *row* is replaced everywhere by *column*. To verify this property, one merely applies the original Theorem 3 to A^T . A row operation on A^T amounts to a column operation on A .

Column operations are useful for both theoretical purposes and hand computations. However, for simplicity we’ll perform only row operations in numerical calculations.

Determinants and Matrix Products

The proof of the following useful theorem is at the end of the section. Applications are in the exercises.



THEOREM 6

Multiplicative Property

If A and B are $n \times n$ matrices, then $\det AB = (\det A)(\det B)$.

EXAMPLE 5 Verify Theorem 6 for $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$.

SOLUTION

$$AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}$$

and

$$\det AB = 25 \cdot 13 - 20 \cdot 14 = 325 - 280 = 45$$

Since $\det A = 9$ and $\det B = 5$,

$$(\det A)(\det B) = 9 \cdot 5 = 45 = \det AB$$

Warning: A common misconception is that Theorem 6 has an analogue for *sums* of matrices. However, $\det(A + B)$ is *not* equal to $\det A + \det B$, in general.

A Linearity Property of the Determinant Function

For an $n \times n$ matrix A , we can consider $\det A$ as a function of the n column vectors in A . We will show that if all columns except one are held fixed, then $\det A$ is a *linear function* of that one (vector) variable.

Suppose that the j th column of A is allowed to vary, and write

$$A = [a_1 \quad \cdots \quad a_{j-1} \quad x \quad a_{j+1} \quad \cdots \quad a_n]$$

Define a transformation T from \mathbb{R}^n to \mathbb{R} by

$$T(x) = \det [a_1 \quad \cdots \quad a_{j-1} \quad x \quad a_{j+1} \quad \cdots \quad a_n]$$

Then,

$$T(cx) = cT(x) \quad \text{for all scalars } c \text{ and all } x \text{ in } \mathbb{R}^n \quad (2)$$

$$T(u + v) = T(u) + T(v) \quad \text{for all } u, v \text{ in } \mathbb{R}^n \quad (3)$$

Property (2) is Theorem 3(c) applied to the columns of A . A proof of property (3) follows from a cofactor expansion of $\det A$ down the j th column. (See Exercise 43.) This (multi-) linearity property of the determinant turns out to have many useful consequences that are studied in more advanced courses.

Proofs of Theorems 3 and 6

It is convenient to prove Theorem 3 when it is stated in terms of the elementary matrices discussed in Section 2.2. We call an elementary matrix E a *row replacement (matrix)* if E is obtained from the identity I by adding a multiple of one row to another row; E is an *interchange* if E is obtained by interchanging two rows of I ; and E is a *scale by r* if E is obtained by multiplying a row of I by a nonzero scalar r . With this terminology, Theorem 3 can be reformulated as follows:



If A is an $n \times n$ matrix and E is an $n \times n$ elementary matrix, then

$$\det EA = (\det E)(\det A)$$

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, \dots, 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 v_1, v_2, v_3 are linearly independent, when

$$v_1 = \begin{bmatrix} 5 \\ -7 \\ 9 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 3 \\ -5 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 2 \\ -7 \\ 5 \end{bmatrix}$$

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 $\det A^{-1} = \frac{1}{\det A}$.
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 $U^T U = I$. 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. $\det A^T A$ 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 $\det A^T A$ and $\det A A^T$ for several random 4×5 matrices and several random 5×6 matrices. What can you say about $A^T A$ and $A A^T$ 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.

$$\begin{vmatrix} 1 & -3 & 1 & -2 \\ 2 & -5 & -1 & -2 \\ 0 & -4 & 5 & 1 \\ -3 & 10 & -6 & 8 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 1 & -3 & 2 \end{vmatrix} = \begin{vmatrix} 1 & -3 & 1 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & -4 & 5 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = 0$$



3.3 Cramer's Rule, Volume, and Linear Transformations 177

$$\begin{aligned}
 2. \det[v_1 \ v_2 \ v_3] &= \begin{vmatrix} 5 & -3 & 2 \\ -7 & 3 & -7 \\ 9 & -5 & 5 \end{vmatrix} = \begin{vmatrix} 5 & -3 & 2 \\ -2 & 0 & -5 \\ 9 & -5 & 5 \end{vmatrix} && \text{Row 1 added to row 2} \\
 &= -(-3) \begin{vmatrix} -2 & -5 \\ 9 & 5 \end{vmatrix} - (-5) \begin{vmatrix} 5 & 2 \\ -2 & -5 \end{vmatrix} && \text{Cofactors of column 2} \\
 &= 3 \cdot (35) + 5 \cdot (-21) = 0
 \end{aligned}$$

By Theorem 4, the matrix $[v_1 \ v_2 \ v_3]$ 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 $Ax = b$ is affected by changes in the entries of 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 b in \mathbb{R}^n , let $A_i(b)$ be the matrix obtained from A by replacing column i by the vector b .

$$A_i(b) = [a_1 \ \cdots \ \underset{\substack{\uparrow \\ \text{col } i}}{b} \ \cdots \ a_n]$$

THEOREM 7

Cramer's Rule

Let A be an invertible $n \times n$ matrix. For any b in \mathbb{R}^n , the unique solution x of $Ax = b$ has entries given by

$$x_i = \frac{\det A_i(b)}{\det A}, \quad i = 1, 2, \dots, n \quad (1)$$

PROOF Denote the columns of A by a_1, \dots, a_n and the columns of the $n \times n$ identity matrix I by e_1, \dots, e_n . If $Ax = b$, the definition of matrix multiplication shows that

$$\begin{aligned}
 A \cdot I_i(x) &= A [e_1 \ \cdots \ x \ \cdots \ e_n] = [Ae_1 \ \cdots \ Ax \ \cdots \ Ae_n] \\
 &= [a_1 \ \cdots \ b \ \cdots \ a_n] = A_i(b)
 \end{aligned}$$

By the multiplicative property of determinants,

$$(\det A)(\det I_i(x)) = \det A_i(b)$$

The second determinant on the left is simply x_i . (Make a cofactor expansion along the i th row.) Hence $(\det A) \cdot x_i = \det A_i(b)$. 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 View the system as $A\mathbf{x} = \mathbf{b}$. Using the notation introduced above,

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \quad A_1(\mathbf{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \quad A_2(\mathbf{b}) = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

Since $\det A = 2$, the system has a unique solution. By Cramer's rule,

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{24 + 16}{2} = 20$$

$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{24 + 30}{2} = 27$$

Application to Engineering

A number of important engineering problems, particularly in electrical engineering and control theory, can be analyzed by *Laplace transforms*. This approach converts an appropriate system of linear differential equations into a system of linear algebraic equations whose coefficients involve a parameter. The next example illustrates the type of algebraic system that may arise.

EXAMPLE 2 Consider the following system in which s is an unspecified parameter. Determine the values of s for which the system has a unique solution, and use Cramer's rule to describe the solution.

$$\begin{aligned} 3sx_1 - 2x_2 &= 4 \\ -6x_1 + sx_2 &= 1 \end{aligned}$$

SOLUTION View the system as $A\mathbf{x} = \mathbf{b}$. Then

$$A = \begin{bmatrix} 3s & -2 \\ -6 & s \end{bmatrix}, \quad A_1(\mathbf{b}) = \begin{bmatrix} 4 & -2 \\ 1 & s \end{bmatrix}, \quad A_2(\mathbf{b}) = \begin{bmatrix} 3s & 4 \\ -6 & 1 \end{bmatrix}$$

Since

$$\det A = 3s^2 - 12 = 3(s + 2)(s - 2)$$

the system has a unique solution precisely when $s \neq \pm 2$. For such an s , the solution is (x_1, x_2) , where

$$x_1 = \frac{\det A_1(\mathbf{b})}{\det A} = \frac{4s + 2}{3(s + 2)(s - 2)}$$

$$x_2 = \frac{\det A_2(\mathbf{b})}{\det A} = \frac{3s + 24}{3(s + 2)(s - 2)} = \frac{s + 8}{(s + 2)(s - 2)}$$

A Formula for A^{-1}

Cramer's rule leads easily to a general formula for the inverse of an $n \times n$ matrix A . The j th column of A^{-1} is a vector \mathbf{x} that satisfies

$$A\mathbf{x} = \mathbf{e}_j$$

where \mathbf{e}_j is the j th column of the identity matrix, and the i th entry of \mathbf{x} is the (i, j) -entry of A^{-1} . By Cramer's rule,

$$\{(i, j)\text{-entry of } A^{-1}\} = x_i = \frac{\det A_i(\mathbf{e}_j)}{\det A} \quad (2)$$



Recall that A_{ji} denotes the submatrix of A formed by deleting row j and column i . A cofactor expansion down column i of $A_i(\mathbf{e}_j)$ shows that

$$\det A_i(\mathbf{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji} \quad (3)$$

where C_{ji} is a cofactor of A . By (2), the (i, j) -entry of A^{-1} is the cofactor C_{ji} divided by $\det A$. [Note that the subscripts on C_{ji} are the reverse of (i, j) .] Thus

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} \quad (4)$$

The matrix of cofactors on the right side of (4) is called the **adjugate** (or **classical adjoint**) of A , denoted by $\text{adj } A$. (The term *adjoint* also has another meaning in advanced texts on linear transformations.) The next theorem simply restates (4).

THEOREM 8

An Inverse Formula

Let A be an invertible $n \times n$ matrix. Then

$$A^{-1} = \frac{1}{\det A} \text{adj } A$$

EXAMPLE 3 Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

SOLUTION The nine cofactors are

$$\begin{aligned} C_{11} &= + \begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, & C_{12} &= - \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, & C_{13} &= + \begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5 \\ C_{21} &= - \begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, & C_{22} &= + \begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, & C_{23} &= - \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7 \\ C_{31} &= + \begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, & C_{32} &= - \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, & C_{33} &= + \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3 \end{aligned}$$

The adjugate matrix is the *transpose* of the matrix of cofactors. [For instance, C_{12} goes in the $(2, 1)$ position.] Thus

$$\text{adj } A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}$$

We could compute $\det A$ directly, but the following computation provides a check on the calculations above *and* produces $\det A$:

$$(\text{adj } A) \cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = 14I$$

Since $(\text{adj } A)A = 14I$, Theorem 8 shows that $\det A = 14$ and

$$A^{-1} = \frac{1}{14} \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix} \quad \blacksquare$$



3.3 EXERCISES

Use Cramer's rule to compute the solutions of the systems in Exercises 1–6.

1. $5x_1 + 7x_2 = 3$
 $2x_1 + 4x_2 = 1$
2. $4x_1 + x_2 = 6$
 $5x_1 + 2x_2 = 7$
3. $3x_1 - 2x_2 = 7$
 $-5x_1 + 6x_2 = -5$
4. $-5x_1 + 3x_2 = 9$
 $3x_1 - x_2 = -5$
5. $2x_1 + x_2 = 7$
 $-3x_1 + x_3 = -8$
 $x_2 + 2x_3 = -3$
6. $2x_1 + x_2 + x_3 = 4$
 $-x_1 + 2x_3 = 2$
 $3x_1 + x_2 + 3x_3 = -2$

In Exercises 7–10, determine the values of the parameter s for which the system has a unique solution, and describe the solution.

7. $6sx_1 + 4x_2 = 5$
 $9x_1 + 2sx_2 = -2$
8. $3sx_1 - 5x_2 = 3$
 $9x_1 + 5sx_2 = 2$
9. $sx_1 - 2sx_2 = -1$
 $3x_1 + 6sx_2 = 4$
10. $2sx_1 + x_2 = 1$
 $3sx_1 + 6sx_2 = 2$

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

$$15. \begin{bmatrix} 3 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 3 & 2 \end{bmatrix} \quad 16. \begin{bmatrix} 1 & 2 & 4 \\ 0 & -3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

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 $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear transformation, and let \mathbf{p} be a vector and S a set in \mathbb{R}^n . Show that the image of $\mathbf{p} + S$ under T is the translated set $T(\mathbf{p}) + T(S)$ in \mathbb{R}^n .



4.1 VECTOR SPACES AND SUBSPACES

Much of the theory in Chapters 1 and 2 rested on certain simple and obvious algebraic properties of \mathbb{R}^n , listed in Section 1.3. In fact, many other mathematical systems have the same properties. The specific properties of interest are listed in the following definition.

DEFINITION

A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms (or rules) listed below.¹ The axioms must hold for all vectors \mathbf{u} , \mathbf{v} , and \mathbf{w} in V and for all scalars c and d .

1. The sum of \mathbf{u} and \mathbf{v} , denoted by $\mathbf{u} + \mathbf{v}$, is in V .
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. There is a **zero** vector $\mathbf{0}$ in V such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. For each \mathbf{u} in V , there is a vector $-\mathbf{u}$ in V such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. The scalar multiple of \mathbf{u} by c , denoted by $c\mathbf{u}$, is in V .
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$.
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$.
10. $1\mathbf{u} = \mathbf{u}$.

Subspaces

In many problems, a vector space consists of an appropriate subset of vectors from some larger vector space. In this case, only three of the ten vector space axioms need to be checked; the rest are automatically satisfied.

DEFINITION

A **subspace** of a vector space V is a subset H of V that has three properties:

- a. The zero vector of V is in H .²
- b. H is closed under vector addition. That is, for each \mathbf{u} and \mathbf{v} in H , the sum $\mathbf{u} + \mathbf{v}$ is in H .
- c. H is closed under multiplication by scalars. That is, for each \mathbf{u} in H and each scalar c , the vector $c\mathbf{u}$ is in H .