Kingdom of Saudi Arabia
Ministry of Education
Taibah University
College of Science
Math Department

المملكة العربية السعودية
وزارة التعليم العالي
جامعة طيبة
كلية العلوم
قسم الرياضيات

Linear Algebra (1) MATH 243

.

Text Book: Howard Anton,Chris Rorres, Elementary Linear Algebra. Edition 11

Section	Topics
1.1 1.2	Introduction to Systems of Linear Equations Gaussian Elimination
1.3 1.4	Matrices and Matrix Operations Inverses: Algebraic Properties of Matrices
1.5 1.6	Elementary Matrices and a Method for finding A^{-1} More on Linear Systems and Invertible Matrices
1.7 2.1	Diagonal, Triangular, and Symmetric Matrices Determinants by Cofactor Expansion
2.2 2.3	Evaluating Determinants by Row Reduction Properties of Determinants; Cramer's Rule
4.1 4.2	Real Vector Spaces Subspaces
4.3 4.4	Linear Independence Coordinates and Basis
4.5	Dimensions
4.6	Change of Bases
4.7	Row Space, Column Space, and Null Space
4.8	Rank, Nullity, and the Fundamental Matrix Spaces
4.9	Matrix Transformations from R^n to R^m
4.10	Properties of Matrix transformations

CHAPTER 1

Systems of Linear Equations and Matrices

1.1 Introduction to Systems of Linear Equations

Linear Equations

Recall that in two dimensions a line in a rectangular xy-coordinate system can be represented by an equation of the form

$$ax + by = c \quad (a, b \text{ not both } 0)$$

and in three dimensions a plane in a rectangular xyz-coordinate system can be represented by an equation of the form

$$ax + by + cz = d \quad (a, b, c \text{ not all } 0)$$

These are examples of "linear equations," the first being a linear equation in the variables x and y and the second a linear equation in the variables x, y, and z. More generally, we define a ***linear equation*** in the n variables $x_1, x_2, \ldots, x_n$ to be one that can be expressed in the form

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

where $a_1, a_2, \ldots, a_n$ and b are constants, and the a's are not all zero. In the special cases where $n = 2$ or $n = 3$, we will often use variables without subscripts and write linear equations as

$$a_1x + a_2y = b \quad (a_1, a_2 \text{ not both } 0) \tag{2}$$

$$a_1x + a_2y + a_3z = b \quad (a_1, a_2, a_3 \text{ not all } 0) \tag{3}$$

In the special case where $b = 0$, Equation (1) has the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0 \tag{4}$$

which is called a ***homogeneous linear equation*** in the variables $x_1, x_2, \ldots, x_n$.

► EXAMPLE 1 Linear Equations

Observe that a linear equation does not involve any products or roots of variables. All variables occur only to the first power and do not appear, for example, as arguments of trigonometric, logarithmic, or exponential functions. The following are linear equations:

$$x + 3y = 7 \qquad x_1 - 2x_2 - 3x_3 + x_4 = 0$$

$$\tfrac{1}{2}x - y + 3z = -1 \qquad x_1 + x_2 + \cdots + x_n = 1$$

The following are not linear equations:

$$x + 3y^2 = 4 \qquad 3x + 2y - xy = 5$$

$$\sin x + y = 0 \qquad \sqrt{x_1} + 2x_2 + x_3 = 1$$ ◄

A finite set of linear equations is called a ***system of linear equations*** or, more briefly, a ***linear system***. The variables are called ***unknowns***. For example, system (5) that follows has unknowns x and y, and system (6) has unknowns x_1, x_2, and x_3.

$$\begin{aligned} 5x + y &= 3 \\ 2x - y &= 4 \end{aligned} \qquad \begin{aligned} 4x_1 - x_2 + 3x_3 &= -1 \\ 3x_1 + x_2 + 9x_3 &= -4 \end{aligned} \tag{5–6}$$

A general linear system of m equations in the n unknowns $x_1, x_2, \dots, x_n$ can be written as

$$\begin{array}{ccccccc} a_{11}x_1 & + a_{12}x_2 & + & \cdots & + a_{1n}x_n & = b_1 \\ a_{21}x_1 & + a_{22}x_2 & + & \cdots & + a_{2n}x_n & = b_2 \\ \vdots & \vdots & & & \vdots & \vdots \\ a_{m1}x_1 & + a_{m2}x_2 & + & \cdots & + a_{mn}x_n & = b_m \end{array} \tag{7}$$

A ***solution*** of a linear system in n unknowns $x_1, x_2, \dots, x_n$ is a sequence of n numbers $s_1, s_2, \dots, s_n$ for which the substitution

$$x_1 = s_1, \quad x_2 = s_2, \dots, \quad x_n = s_n$$

makes each equation a true statement. For example, the system in (5) has the solution

$$x = 1, \quad y = -2$$

and the system in (6) has the solution

$$x_1 = 1, \quad x_2 = 2, \quad x_3 = -1$$

These solutions can be written more succinctly as

$$(1, -2) \quad \text{and} \quad (1, 2, -1)$$

in which the names of the variables are omitted. This notation allows us to interpret these solutions geometrically as points in two-dimensional and three-dimensional space. More generally, a solution

$$x_1 = s_1, \quad x_2 = s_2, \dots, \quad x_n = s_n$$

of a linear system in n unknowns can be written as

$$(s_1, s_2, \dots, s_n)$$

which is called an ***ordered n-tuple***. With this notation it is understood that all variables appear in the same order in each equation. If $n = 2$, then the n-tuple is called an ***ordered pair***, and if $n = 3$, then it is called an ***ordered triple***.

Linear Systems in Two and Three Unknowns

Linear systems in two unknowns arise in connection with intersections of lines. For example, consider the linear system

$$a_1x + b_1y = c_1$$
$$a_2x + b_2y = c_2$$

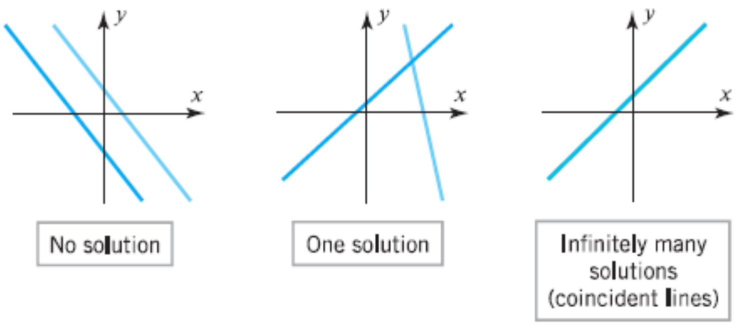
in which the graphs of the equations are lines in the xy-plane. Each solution (x, y) of this system corresponds to a point of intersection of the lines, so there are three possibilities (Figure 1.1.1):

1. The lines may be parallel and distinct, in which case there is no intersection and consequently no solution.
2. The lines may intersect at only one point, in which case the system has exactly one solution.
3. The lines may coincide, in which case there are infinitely many points of intersection (the points on the common line) and consequently infinitely many solutions.

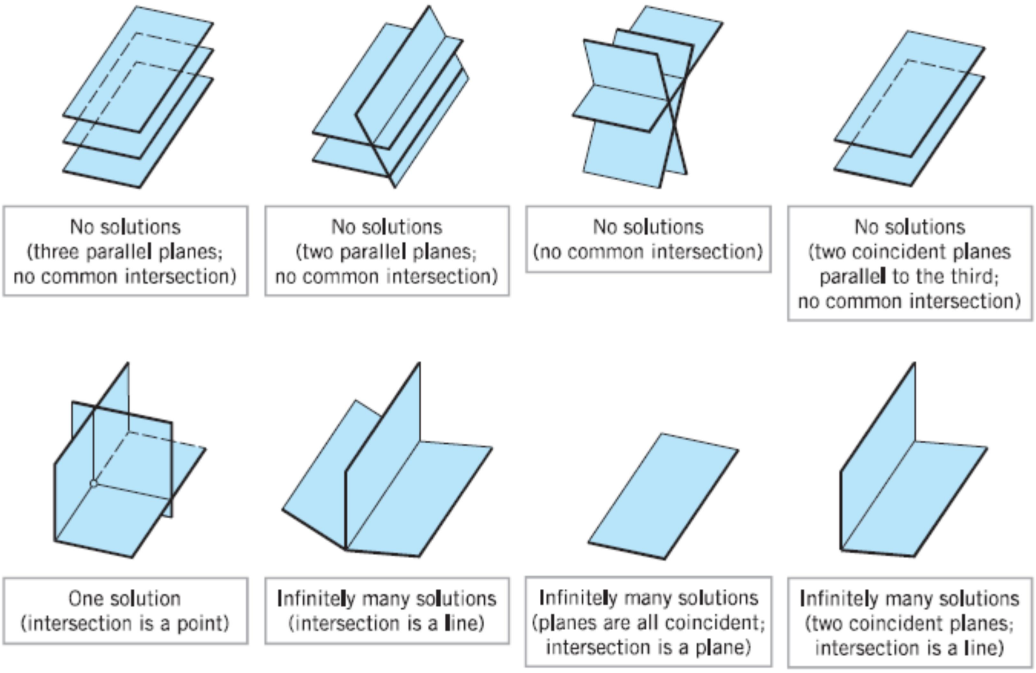
In general, we say that a linear system is ***consistent*** if it has at least one solution and ***inconsistent*** if it has no solutions. Thus, a *consistent* linear systemof two equations in two unknowns has either one solution or infinitely many solutions—there are no other possibilities. The same is true for a linear system of three equations in three unknowns

$$a_1x + b_1y + c_1z = d_1$$
$$a_2x + b_2y + c_2z = d_2$$
$$a_3x + b_3y + c_3z = d_3$$

in which the graphs of the equations are planes. The solutions of the system, if any, correspond to points where all three planes intersect, so again we see that there are only three possibilities—no solutions, one solution, or infinitely many solutions (Figure 1.1.2).



▶ Figure 1.1.1



▲ Figure 1.1.2

We will prove later that our observations about the number of solutions of linear systems of two equations in two unknowns and linear systems of three equations in three unknowns actually hold for *all* linear systems. That is:

Every system of linear equations has zero, one, or infinitely many solutions. There are no other possibilities.

▶ EXAMPLE 2 A Linear System with One Solution

Solve the linear system

$$\begin{aligned} x - y &= 1 \\ 2x + y &= 6 \end{aligned}$$

Solution We can eliminate x from the second equation by adding -2 times the first equation to the second. This yields the simplified system

$$\begin{aligned} x - y &= 1 \\ 3y &= 4 \end{aligned}$$

From the second equation we obtain $y = \frac{4}{3}$, and on substituting this value in the first equation we obtain $x = 1 + y = \frac{7}{3}$. Thus, the system has the unique solution

$$x = \tfrac{7}{3}, \quad y = \tfrac{4}{3}$$

Geometrically, this means that the lines represented by the equations in the system intersect at the single point $\left(\frac{7}{3}, \frac{4}{3}\right)$. We leave it for you to check this by graphing the lines.

▶ EXAMPLE 3 A Linear System with No Solutions

Solve the linear system

$$\begin{aligned} x + y &= 4 \\ 3x + 3y &= 6 \end{aligned}$$

Solution We can eliminate x from the second equation by adding -3 times the first equation to the second equation. This yields the simplified system

$$\begin{aligned} x + y &= 4 \\ 0 &= -6 \end{aligned}$$

The second equation is contradictory, so the given system has no solution. Geometrically, this means that the lines corresponding to the equations in the original system are parallel and distinct. We leave it for you to check this by graphing the lines or by showing that they have the same slope but different y-intercepts.

▶ EXAMPLE 4 A Linear System with Infinitely Many Solutions

Solve the linear system

$$\begin{aligned} 4x - 2y &= 1 \\ 16x - 8y &= 4 \end{aligned}$$

Solution We can eliminate x from the second equation by adding -4 times the first equation to the second. This yields the simplified system

$$\begin{aligned} 4x - 2y &= 1 \\ 0 &= 0 \end{aligned}$$

The second equation does not impose any restrictions on x and y and hence can be omitted. Thus, the solutions of the system are those values of x and y that satisfy the single equation

$$4x - 2y = 1 \tag{8}$$

Geometrically, this means the lines corresponding to the two equations in the original system coincide. One way to describe the solution set is to solve this equation for x in terms of y to obtain $x = \frac{1}{4} + \frac{1}{2}y$ and then assign an arbitrary value t (called a ***parameter***)

to y. This allows us to express the solution by the pair of equations (called ***parametric equations***)

$$x = \tfrac{1}{4} + \tfrac{1}{2}t, \quad y = t$$

We can obtain specific numerical solutions from these equations by substituting numerical values for the parameter t. For example, $t = 0$ yields the solution $\left(\frac{1}{4}, 0\right)$, $t = 1$ yields the solution $\left(\frac{3}{4}, 1\right)$, and $t = -1$ yields the solution $\left(-\frac{1}{4}, -1\right)$. You can confirm that these are solutions by substituting their coordinates into the given equations.

▶ EXAMPLE 5 A Linear System with Infinitely Many Solutions

Solve the linear system

$$\begin{aligned} x - y + 2z &= 5 \\ 2x - 2y + 4z &= 10 \\ 3x - 3y + 6z &= 15 \end{aligned}$$

Solution This system can be solved by inspection, since the second and third equations are multiples of the first. Geometrically, this means that the three planes coincide and that those values of x, y, and z that satisfy the equation

$$x - y + 2z = 5 \tag{9}$$

automatically satisfy all three equations. Thus, it suffices to find the solutions of (9). We can do this by first solving this equation for x in terms of y and z, then assigning arbitrary values r and s (parameters) to these two variables, and then expressing the solution by the three parametric equations

$$x = 5 + r - 2s, \quad y = r, \quad z = s$$

Specific solutions can be obtained by choosing numerical values for the parameters r and s. For example, taking $r = 1$ and $s = 0$ yields the solution $(6, 1, 0)$. ◀

Augmented Matrices and Elementary Row Operations

As the number of equations and unknowns in a linear system increases, so does the complexity of the algebra involved in finding solutions. The required computations can be made more manageable by simplifying notation and standardizing procedures. For example, by mentally keeping track of the location of the +'s, the x's, and the ='s in the linear system

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

we can abbreviate the system by writing only the rectangular array of numbers

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

This is called the ***augmented matrix*** for the system. For example, the augmented matrix for the system of equations

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 9 \\ 2x_1 + 4x_2 - 3x_3 &= 1 \\ 3x_1 + 6x_2 - 5x_3 &= 0 \end{aligned} \quad \text{is} \quad \begin{bmatrix} 1 & 1 & 2 & 9 \\ 2 & 4 & -3 & 1 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

The basic method for solving a linear system is to perform algebraic operations on the system that do not alter the solution set and that produce a succession of increasingly simpler systems, until a point is reached where it can be ascertained whether the system is consistent, and if so, what its solutions are. Typically, the algebraic operations are:

1. Multiply an equation through by a nonzero constant.
2. Interchange two equations.
3. Add a constant times one equation to another.

Since the rows (horizontal lines) of an augmented matrix correspond to the equations in the associated system, these three operations correspond to the following operations on the rows of the augmented matrix:

1. Multiply a row through by a nonzero constant.
2. Interchange two rows.
3. Add a constant times one row to another.

These are called ***elementary row operations*** on a matrix.

In the following example we will illustrate how to use elementary row operations and an augmented matrix to solve a linear system in three unknowns. Since a systematic procedure for solving linear systems will be developed in the next section, do not worry about how the steps in the example were chosen. Your objective here should be simply to understand the computations.

EXAMPLE 6 Using Elementary Row Operations

In the left column we solve a system of linear equations by operating on the equations in the system, and in the right column we solve the same system by operating on the rows of the augmented matrix.

$$\begin{aligned} x + y + 2z &= 9 \\ 2x + 4y - 3z &= 1 \\ 3x + 6y - 5z &= 0 \end{aligned}$$

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

Add −2 times the first equation to the second to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3x + 6y - 5z &= 0 \end{aligned}$$

Add −2 times the first row to the second to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 3 & 6 & -5 & 0 \end{bmatrix}$$

Add −3 times the first equation to the third to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ 2y - 7z &= -17 \\ 3y - 11z &= -27 \end{aligned}$$

Add −3 times the first row to the third to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 2 & -7 & -17 \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

Multiply the second equation by $\frac{1}{2}$ to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ 3y - 11z &= -27 \end{aligned}$$

Multiply the second row by $\frac{1}{2}$ to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 3 & -11 & -27 \end{bmatrix}$$

Add −3 times the second equation to the third to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ -\tfrac{1}{2}z &= -\tfrac{3}{2} \end{aligned}$$

Add −3 times the second row to the third to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$$

Multiply the third equation by −2 to obtain

$$\begin{aligned} x + y + 2z &= 9 \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ z &= 3 \end{aligned}$$

Multiply the third row by −2 to obtain

$$\begin{bmatrix} 1 & 1 & 2 & 9 \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Add −1 times the second equation to the first to obtain

$$\begin{aligned} x \quad\quad + \tfrac{11}{2}z &= \tfrac{35}{2} \\ y - \tfrac{7}{2}z &= -\tfrac{17}{2} \\ z &= 3 \end{aligned}$$

Add −1 times the second row to the first to obtain

$$\begin{bmatrix} 1 & 0 & \frac{11}{2} & \frac{35}{2} \\ 0 & 1 & -\frac{7}{2} & -\frac{17}{2} \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

Add $-\frac{11}{2}$ times the third equation to the first and $\frac{7}{2}$ times the third equation to the second to obtain

$$\begin{aligned} x \quad\quad &= 1 \\ y \quad &= 2 \\ z &= 3 \end{aligned}$$

Add $-\frac{11}{2}$ times the third row to the first and $\frac{7}{2}$ times the third row to the second to obtain

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

The solution $x = 1$, $y = 2$, $z = 3$ is now evident. ◀

Exercise Set 1.1

1. In each part, determine whether the equation is linear in x_1, x_2, and x_3.

(a) $x_1 + 5x_2 - \sqrt{2}\,x_3 = 1$ (b) $x_1 + 3x_2 + x_1x_3 = 2$

(c) $x_1 = -7x_2 + 3x_3$ (d) $x_1^{-2} + x_2 + 8x_3 = 5$

(e) $x_1^{3/5} - 2x_2 + x_3 = 4$ (f) $\pi x_1 - \sqrt{2}\,x_2 = 7^{1/3}$

2. In each part, determine whether the equation is linear in x and y.

(a) $2^{1/3}x + \sqrt{3}y = 1$ (b) $2x^{1/3} + 3\sqrt{y} = 1$

(c) $\cos\left(\frac{\pi}{7}\right)x - 4y = \log 3$ (d) $\frac{\pi}{7}\cos x - 4y = 0$

(e) $xy = 1$ (f) $y + 7 = x$

► In each part of Exercises 5–6, find a linear system in the unknowns $x_1, x_2, x_3, \ldots$, that corresponds to the given augmented matrix. ◄

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

6. (a) $\begin{bmatrix} 0 & 3 & -1 & -1 & -1 \\ 5 & 2 & 0 & -3 & -6 \end{bmatrix}$

(b) $\begin{bmatrix} 3 & 0 & 1 & -4 & 3 \\ -4 & 0 & 4 & 1 & -3 \\ -1 & 3 & 0 & -2 & -9 \\ 0 & 0 & 0 & -1 & -2 \end{bmatrix}$

► In each part of Exercises 7–8, find the augmented matrix for the linear system. ◄

7. (a)
$$\begin{aligned} -2x_1 &= 6 \\ 3x_1 &= 8 \\ 9x_1 &= -3 \end{aligned}$$

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

1.2 Gaussian Elimination

Echelon Forms

In Example 6 of the last section, we solved a linear system in the unknowns x, y, and z by reducing the augmented matrix to the form

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

from which the solution $x = 1$, $y = 2$, $z = 3$ became evident. This is an example of a matrix that is in ***reduced row echelon form***. To be of this form, a matrix must have the following properties:

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a ***leading 1***.
2. If there are any rows that consist entirely of zeros, then they are grouped together at the bottom of the matrix.
3. In any two successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.
4. Each column that contains a leading 1 has zeros everywhere else in that column.

A matrix that has the first three properties is said to be in ***row echelon form***. (Thus, a matrix in reduced row echelon form is of necessity in row echelon form, but not conversely.)

▶ EXAMPLE 1 Row Echelon and Reduced Row Echelon Form

The following matrices are in reduced row echelon form.

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

The following matrices are in row echelon form but not reduced row echelon form.

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

EXAMPLE 2 More on Row Echelon and Reduced Row Echelon Form

As Example 1 illustrates, a matrix in row echelon form has zeros below each leading 1, whereas a matrix in reduced row echelon form has zeros below *and above* each leading 1. Thus, with any real numbers substituted for the $*$'s, all matrices of the following types are in row echelon form:

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

All matrices of the following types are in reduced row echelon form:

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

EXAMPLE 3 Unique Solution

Suppose that the augmented matrix for a linear system in the unknowns x_1, x_2, x_3, and x_4 has been reduced by elementary row operations to

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

This matrix is in reduced row echelon form and corresponds to the equations

$$\begin{aligned}x_1 \qquad\qquad\qquad &= 3\\ x_2 \qquad\qquad &= -1\\ x_3 \qquad &= 0\\ x_4 &= 5\end{aligned}$$

Thus, the system has a unique solution, namely, $x_1 = 3$, $x_2 = -1$, $x_3 = 0$, $x_4 = 5$.

EXAMPLE 4 Linear Systems in Three Unknowns

In each part, suppose that the augmented matrix for a linear system in the unknowns x, y, and z has been reduced by elementary row operations to the given reduced row echelon form. Solve the system.

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

Solution (a) The equation that corresponds to the last row of the augmented matrix is

$$0x + 0y + 0z = 1$$

Since this equation is not satisfied by any values of x, y, and z, the system is inconsistent.

Solution (b) The equation that corresponds to the last row of the augmented matrix is

$$0x + 0y + 0z = 0$$

This equation can be omitted since it imposes no restrictions on x, y, and z; hence, the linear system corresponding to the augmented matrix is

$$\begin{aligned} x \quad\; + 3z &= -1 \\ y - 4z &= 2 \end{aligned}$$

Since x and y correspond to the leading 1's in the augmented matrix, we call these the ***leading variables***. The remaining variables (in this case z) are called ***free variables***. Solving for the leading variables in terms of the free variables gives

$$\begin{aligned} x &= -1 - 3z \\ y &= 2 + 4z \end{aligned}$$

From these equations we see that the free variable z can be treated as a parameter and assigned an arbitrary value t, which then determines values for x and y. Thus, the solution set can be represented by the parametric equations

$$x = -1 - 3t, \quad y = 2 + 4t, \quad z = t$$

By substituting various values for t in these equations we can obtain various solutions of the system. For example, setting $t = 0$ yields the solution

$$x = -1, \quad y = 2, \quad z = 0$$

and setting $t = 1$ yields the solution

$$x = -4, \quad y = 6, \quad z = 1$$

Solution (c) As explained in part (b), we can omit the equations corresponding to the zero rows, in which case the linear system associated with the augmented matrix consists of the single equation

$$x - 5y + z = 4 \tag{1}$$

from which we see that the solution set is a plane in three-dimensional space. Although (1) is a valid form of the solution set, there are many applications in which it is preferable to express the solution set in parametric form. We can convert (1) to parametric form by solving for the leading variable x in terms of the free variables y and z to obtain

$$x = 4 + 5y - z$$

From this equation we see that the free variables can be assigned arbitrary values, say $y = s$ and $z = t$, which then determine the value of x. Thus, the solution set can be expressed parametrically as

$$x = 4 + 5s - t, \quad y = s, \quad z = t \;◀ \tag{2}$$

DEFINITION 1 If a linear system has infinitely many solutions, then a set of parametric equations from which all solutions can be obtained by assigning numerical values to the parameters is called a ***general solution*** of the system.

Elimination Methods

We have just seen how easy it is to solve a system of linear equations once its augmented matrix is in reduced row echelon form. Now we will give a step-by-step ***elimination procedure*** that can be used to reduce any matrix to reduced row echelon form. As we state each step in the procedure, we illustrate the idea by reducing the following matrix to reduced row echelon form.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Step 1. Locate the leftmost column that does not consist entirely of zeros.

$$\begin{bmatrix} 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -10 & 6 & 12 & 28 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

Leftmost nonzero column

Step 2. Interchange the top row with another row, if necessary, to bring a nonzero entry to the top of the column found in Step 1.

$$\begin{bmatrix} 2 & 4 & -10 & 6 & 12 & 28 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The first and second rows in the preceding matrix were interchanged.

Step 3. If the entry that is now at the top of the column found in Step 1 is a, multiply the first row by $1/a$ in order to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 2 & 4 & -5 & 6 & -5 & -1 \end{bmatrix}$$

The first row of the preceding matrix was multiplied by $\frac{1}{2}$.

Step 4. Add suitable multiples of the top row to the rows below so that all entries below the leading 1 become zeros.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

-2 times the first row of the preceding matrix was added to the third row.

Step 5. Now cover the top row in the matrix and begin again with Step 1 applied to the submatrix that remains. Continue in this way until the *entire* matrix is in row echelon form.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & -2 & 0 & 7 & 12 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

Leftmost nonzero column in the submatrix

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 5 & 0 & -17 & -29 \end{bmatrix}$$

The first row in the submatrix was multiplied by $-\frac{1}{2}$ to introduce a leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

← −5 times the first row of the submatrix was added to the second row of the submatrix to introduce a zero below the leading 1.

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

← The top row in the submatrix was covered, and we returned again to Step 1.

Leftmost nonzero column in the new submatrix

$$\begin{bmatrix} 1 & 2 & -5 & 3 & 6 & 14 \\ 0 & 0 & 1 & 0 & -\frac{7}{2} & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$$

← The first (and only) row in the new submatrix was multiplied by 2 to introduce a leading 1.

The *entire* matrix is now in row echelon form. To find the reduced row echelon form we need the following additional step.

Step 6. Beginning with the last nonzero row and working upward, add suitable multiples of each row to the rows above to introduce zeros above the leading 1's.

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

← $\frac{7}{2}$ times the third row of the preceding matrix was added to the second row.

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

← −6 times the third row was added to the first row.

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

← 5 times the second row was added to the first row.

The last matrix is in reduced row echelon form.

EXAMPLE 5 Gauss–Jordan Elimination

Solve by Gauss–Jordan elimination.

$$\begin{aligned}
x_1 + 3x_2 - 2x_3 \qquad + 2x_5 \qquad &= 0\\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= -1\\
5x_3 + 10x_4 \qquad + 15x_6 &= 5\\
2x_1 + 6x_2 \qquad + 8x_4 + 4x_5 + 18x_6 &= 6
\end{aligned}$$

Solution The augmented matrix for the system is

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0\\ 2 & 6 & -5 & -2 & 4 & -3 & -1\\ 0 & 0 & 5 & 10 & 0 & 15 & 5\\ 2 & 6 & 0 & 8 & 4 & 18 & 6 \end{bmatrix}$$

Adding −2 times the first row to the second and fourth rows gives

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0\\ 0 & 0 & -1 & -2 & 0 & -3 & -1\\ 0 & 0 & 5 & 10 & 0 & 15 & 5\\ 0 & 0 & 4 & 8 & 0 & 18 & 6 \end{bmatrix}$$

Multiplying the second row by −1 and then adding −5 times the new second row to the third row and −4 times the new second row to the fourth row gives

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

Interchanging the third and fourth rows and then multiplying the third row of the resulting matrix by $\frac{1}{6}$ gives the row echelon form

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

This completes the forward phase since there are zeros below the leading 1's.

Adding −3 times the third row to the second row and then adding 2 times the second row of the resulting matrix to the first row yields the reduced row echelon form

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

This completes the backward phase since there are zeros above the leading 1's.

The corresponding system of equations is

$$\begin{aligned}
x_1 + 3x_2 \qquad + 4x_4 + 2x_5 \qquad &= 0\\
x_3 + 2x_4 \qquad &= 0\\
x_6 &= \tfrac{1}{3}
\end{aligned} \tag{3}$$

Solving for the leading variables, we obtain

$$\begin{aligned}
x_1 &= -3x_2 - 4x_4 - 2x_5\\
x_3 &= -2x_4\\
x_6 &= \tfrac{1}{3}
\end{aligned}$$

Finally, we express the general solution of the system parametrically by assigning the free variables x_2, x_4, and x_5 arbitrary values r, s, and t, respectively. This yields

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = \tfrac{1}{3}$$ ◀

Homogeneous Linear Systems

A system of linear equations is said to be ***homogeneous*** if the constant terms are all zero; that is, the system has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Every homogeneous system of linear equations is consistent because all such systems have $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ as a solution. This solution is called the ***trivial solution***; if there are other solutions, they are called ***nontrivial solutions***.

Because a homogeneous linear system always has the trivial solution, there are only two possibilities for its solutions:

- The system has only the trivial solution.
- The system has infinitely many solutions in addition to the trivial solution.

▶ EXAMPLE 6 A Homogeneous System

Use Gauss–Jordan elimination to solve the homogeneous linear system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \qquad\quad + 2x_5 \qquad\quad &= 0 \\ 2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= 0 \\ 5x_3 + 10x_4 \qquad\quad + 15x_6 &= 0 \\ 2x_1 + 6x_2 \qquad\quad + 8x_4 + 4x_5 + 18x_6 &= 0 \end{aligned} \tag{4}$$

The augmented matrix for the given homogeneous system is

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 & 0 \\ 0 & 0 & 5 & 10 & 0 & 15 & 0 \\ 2 & 6 & 0 & 8 & 4 & 18 & 0 \end{bmatrix}$$

the reduced row echelon form of (5) is

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

The corresponding system of equations is

$$\begin{aligned} x_1 + 3x_2 \qquad + 4x_4 + 2x_5 \qquad &= 0 \\ x_3 + 2x_4 \qquad\qquad &= 0 \\ x_6 &= 0 \end{aligned}$$

Solving for the leading variables, we obtain

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= 0 \end{aligned} \tag{7}$$

If we now assign the free variables x_2, x_4, and x_5 arbitrary values r, s, and t, respectively, then we can express the solution set parametrically as

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = 0$$

Note that the trivial solution results when $r = s = t = 0$. ◀

THEOREM 1.2.2 *A homogeneous linear system with more unknowns than equations has infinitely many solutions.*

Gaussian Elimination and Back-Substitution

EXAMPLE 7 Example 5 Solved by Back-Substitution

From the computations in Example 5, a row echelon form of the augmented matrix is

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

To solve the corresponding system of equations

$$\begin{aligned} x_1 + 3x_2 - 2x_3 \quad\quad + 2x_5 \quad\quad &= 0 \\ x_3 + 2x_4 \quad\quad + 3x_6 &= 1 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

we proceed as follows:

Step 1. Solve the equations for the leading variables.

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 \\ x_3 &= 1 - 2x_4 - 3x_6 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Step 2. Beginning with the bottom equation and working upward, successively substitute each equation into all the equations above it.

Substituting $x_6 = \frac{1}{3}$ into the second equation yields

$$\begin{aligned} x_1 &= -3x_2 + 2x_3 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Substituting $x_3 = -2x_4$ into the first equation yields

$$\begin{aligned} x_1 &= -3x_2 - 4x_4 - 2x_5 \\ x_3 &= -2x_4 \\ x_6 &= \tfrac{1}{3} \end{aligned}$$

Step 3. Assign arbitrary values to the free variables, if any.

If we now assign x_2, x_4, and x_5 the arbitrary values r, s, and t, respectively, the general solution is given by the formulas

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = \tfrac{1}{3}$$

This agrees with the solution obtained in Example 5.

EXAMPLE 8

Suppose that the matrices below are augmented matrices for linear systems in the unknowns x_1, x_2, x_3, and x_4. These matrices are all in row echelon form but not reduced row echelon form. Discuss the existence and uniqueness of solutions to the corresponding linear systems

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

Solution (a) The last row corresponds to the equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 1$$

from which it is evident that the system is inconsistent.

Solution (b) The last row corresponds to the equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0$$

which has no effect on the solution set. In the remaining three equations the variables x_1, x_2, and x_3 correspond to leading 1's and hence are leading variables. The variable x_4 is a free variable. With a little algebra, the leading variables can be expressed in terms of the free variable, and the free variable can be assigned an arbitrary value. Thus, the system must have infinitely many solutions.

Solution (c) The last row corresponds to the equation

$$x_4 = 0$$

which gives us a numerical value for x_4. If we substitute this value into the third equation, namely,

$$x_3 + 6x_4 = 9$$

we obtain $x_3 = 9$. You should now be able to see that if we continue this process and substitute the known values of x_3 and x_4 into the equation corresponding to the second row, we will obtain a unique numerical value for x_2; and if, finally, we substitute the known values of x_4, x_3, and x_2 into the equation corresponding to the first row, we will produce a unique numerical value for x_1. Thus, the system has a unique solution. ◀

Exercise Set 1.2

▶ In Exercises 1–2, determine whether the matrix is in row echelon form, reduced row echelon form, both, or neither. ◀

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

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

(f) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$ (g) $\begin{bmatrix} 1 & -7 & 5 & 5 \\ 0 & 1 & 3 & 2 \end{bmatrix}$

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

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

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

▶ In Exercises 5–8, solve the linear system by Gaussian elimination. ◀

7.
$$\begin{aligned} x - y + 2z - w &= -1 \\ 2x + y - 2z - 2w &= -2 \\ -x + 2y - 4z + w &= 1 \\ 3x \qquad\qquad - 3w &= -3 \end{aligned}$$

▶ In Exercises 13–14, determine whether the homogeneous system has nontrivial solutions by inspection (without pencil and paper). ◀

13.
$$\begin{aligned} 2x_1 - 3x_2 + 4x_3 - x_4 &= 0 \\ 7x_1 + x_2 - 8x_3 + 9x_4 &= 0 \\ 2x_1 + 8x_2 + x_3 - x_4 &= 0 \end{aligned}$$

1.3 Matrices and Matrix Operations

DEFINITION 1 A ***matrix*** is a rectangular array of numbers. The numbers in the array are called the ***entries*** in the matrix.

EXAMPLE 1 Examples of Matrices

Some examples of matrices are

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{bmatrix}, \quad [2 \quad 1 \quad 0 \quad -3], \quad \begin{bmatrix} e & \pi & -\sqrt{2} \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad [4]$$

The ***size*** of a matrix is described in terms of the number of rows (horizontal lines) and columns (vertical lines) it contains. For example, the first matrix in Example 1 has three rows and two columns, so its size is 3 by 2 (written 3×2). In a size description, the first number always denotes the number of rows, and the second denotes the number of columns. The remaining matrices in Example 1 have sizes 1×4, 3×3, 2×1, and 1×1, respectively.

A matrix with only one row, such as the second in Example 1, is called a ***row vector*** (or a ***row matrix***), and a matrix with only one column, such as the fourth in that example, is called a ***column vector*** (or a ***column matrix***). The fifth matrix in that example is both a row vector and a column vector.

We will use capital letters to denote matrices and lowercase letters to denote numerical quantities; thus we might write

$$A = \begin{bmatrix} 2 & 1 & 7 \\ 3 & 4 & 2 \end{bmatrix} \quad \text{or} \quad C = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$$

When discussing matrices, it is common to refer to numerical quantities as ***scalars***. Unless stated otherwise, *scalars will be real numbers*; complex scalars will be considered later in the text.

The entry that occurs in row i and column j of a matrix A will be denoted by a_{ij}. Thus a general 3×4 matrix might be written as

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

and a general $m \times n$ matrix as

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

When a compact notation is desired, the preceding matrix can be written as

$$[a_{ij}]_{m\times n} \quad \text{or} \quad [a_{ij}]$$

the first notation being used when it is important in the discussion to know the size, and the second when the size need not be emphasized. Usually, we will match the letter denoting a matrix with the letter denoting its entries; thus, for a matrix B we would generally use b_{ij} for the entry in row i and column j, and for a matrix C we would use the notation c_{ij}.

The entry in row i and column j of a matrix A is also commonly denoted by the symbol $(A)_{ij}$. Thus, for matrix (1) above, we have

$$(A)_{ij} = a_{ij}$$

and for the matrix

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

we have $(A)_{11} = 2$, $(A)_{12} = -3$, $(A)_{21} = 7$, and $(A)_{22} = 0$.

Row and column vectors are of special importance, and it is common practice to denote them by boldface lowercase letters rather than capital letters. For such matrices, double subscripting of the entries is unnecessary. Thus a general $1 \times n$ row vector **a** and a general $m \times 1$ column vector **b** would be written as

$$\mathbf{a} = [a_1 \quad a_2 \quad \cdots \quad a_n] \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

A matrix A with n rows and n columns is called a ***square matrix of order n***, and the shaded entries $a_{11}, a_{22}, \ldots, a_{nn}$ in (2) are said to be on the ***main diagonal*** of A.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \tag{2}$$

Operations on Matrices

DEFINITION 2 Two matrices are defined to be ***equal*** if they have the same size and their corresponding entries are equal.

EXAMPLE 2 Equality of Matrices

Consider the matrices

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

If $x = 5$, then $A = B$, but for all other values of x the matrices A and B are not equal, since not all of their corresponding entries are equal. There is no value of x for which $A = C$ since A and C have different sizes. ◀

DEFINITION 3 If A and B are matrices of the same size, then the ***sum*** $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A, and the ***difference*** $A - B$ is the matrix obtained by subtracting the entries of B from the corresponding entries of A. Matrices of different sizes cannot be added or subtracted.

In matrix notation, if $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then

$$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij} \quad \text{and} \quad (A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}$$

EXAMPLE 3 Addition and Subtraction

Consider the matrices

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

Then

$$A + B = \begin{bmatrix} -2 & 4 & 5 & 4 \\ 1 & 2 & 2 & 3 \\ 7 & 0 & 3 & 5 \end{bmatrix} \quad \text{and} \quad A - B = \begin{bmatrix} 6 & -2 & -5 & 2 \\ -3 & -2 & 2 & 5 \\ 1 & -4 & 11 & -5 \end{bmatrix}$$

The expressions $A + C$, $B + C$, $A - C$, and $B - C$ are undefined. ◀

DEFINITION 4 If A is any matrix and c is any scalar, then the ***product*** cA is the matrix obtained by multiplying each entry of the matrix A by c. The matrix cA is said to be a ***scalar multiple*** of A.

In matrix notation, if $A = [a_{ij}]$, then

$$(cA)_{ij} = c(A)_{ij} = ca_{ij}$$

EXAMPLE 4 Scalar Multiples

For the matrices

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 2 & 7 \\ -1 & 3 & -5 \end{bmatrix}, \quad C = \begin{bmatrix} 9 & -6 & 3 \\ 3 & 0 & 12 \end{bmatrix}$$

we have

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

It is common practice to denote $(-1)B$ by $-B$. ◀

DEFINITION 5 If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then the ***product*** AB is the $m \times n$ matrix whose entries are determined as follows: To find the entry in row i and column j of AB, single out row i from the matrix A and column j from the matrix B. Multiply the corresponding entries from the row and column together, and then add up the resulting products.

▶ EXAMPLE 5 Multiplying Matrices

Consider the matrices

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

Since A is a 2×3 matrix and B is a 3×4 matrix, the product AB is a 2×4 matrix. To determine, for example, the entry in row 2 and column 3 of AB, we single out row 2 from A and column 3 from B. Then, as illustrated below, we multiply corresponding entries together and add up these products.

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \square \\ \square & \square & \boxed{26} & \square \end{bmatrix}$$

$$(2 \cdot 4) + (6 \cdot 3) + (0 \cdot 5) = 26$$

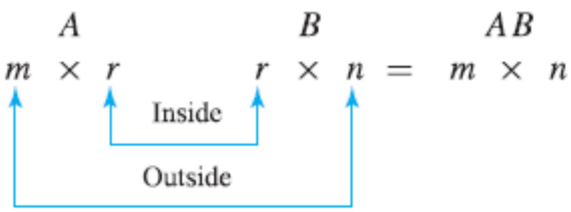
The entry in row 1 and column 4 of AB is computed as follows:

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} \square & \square & \square & \boxed{13} \\ \square & \square & \square & \square \end{bmatrix}$$

$$(1 \cdot 3) + (2 \cdot 1) + (4 \cdot 2) = 13$$

The computations for the remaining entries are

$$\begin{aligned} (1 \cdot 4) + (2 \cdot 0) + (4 \cdot 2) &= 12 \\ (1 \cdot 1) - (2 \cdot 1) + (4 \cdot 7) &= 27 \\ (1 \cdot 4) + (2 \cdot 3) + (4 \cdot 5) &= 30 \\ (2 \cdot 4) + (6 \cdot 0) + (0 \cdot 2) &= 8 \\ (2 \cdot 1) - (6 \cdot 1) + (0 \cdot 7) &= -4 \\ (2 \cdot 3) + (6 \cdot 1) + (0 \cdot 2) &= 12 \end{aligned} \qquad AB = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$ ◀



EXAMPLE 6 Determining Whether a Product Is Defined

Suppose that A, B, and C are matrices with the following sizes:

A	B	C
3×4	4×7	7×3

Then by (3), AB is defined and is a 3×7 matrix; BC is defined and is a 4×3 matrix; and CA is defined and is a 7×4 matrix. The products AC, CB, and BA are all undefined. ◀

In general, if $A = [a_{ij}]$ is an $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then, as illustrated by the shading in the following display,

$$AB = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ir} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{r1} & b_{r2} & \cdots & b_{rj} & \cdots & b_{rn} \end{bmatrix} \tag{4}$$

the entry $(AB)_{ij}$ in row i and column j of AB is given by

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj} \tag{5}$$

Formula (5) is called the ***row-column rule*** for matrix multiplication.

Matrix Form of a Linear System

Matrix multiplication has an important application to systems of linear equations. Consider a system of m linear equations in n unknowns:

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

Since two matrices are equal if and only if their corresponding entries are equal, we can replace the m equations in this system by the single matrix equation

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

The $m \times 1$ matrix on the left side of this equation can be written as a product to give

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

If we designate these matrices by A, $\mathbf{x}$, and $\mathbf{b}$, respectively, then we can replace the original system of m equations in n unknowns by the single matrix equation

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

The matrix A in this equation is called the ***coefficient matrix*** of the system. The augmented matrix for the system is obtained by adjoining $\mathbf{b}$ to A as the last column; thus the augmented matrix is

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Transpose of a Matrix

DEFINITION 7 If A is any $m \times n$ matrix, then the ***transpose of A***, denoted by A^T, is defined to be the $n \times m$ matrix that results by interchanging the rows and columns of A; that is, the first column of A^T is the first row of A, the second column of A^T is the second row of A, and so forth.

▶ EXAMPLE 11 Some Transposes

The following are some examples of matrices and their transposes.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 5 & 6 \end{bmatrix}, \quad C = [1 \quad 3 \quad 5], \quad D = [4]$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{bmatrix}, \quad B^T = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix}, \quad C^T = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}, \quad D^T = [4]$$ ◀

Observe that not only are the columns of A^T the rows of A, but the rows of A^T are the columns of A. Thus the entry in row i and column j of A^T is the entry in row j and column i of A; that is,

$$(A^T)_{ij} = (A)_{ji} \tag{14}$$

Note the reversal of the subscripts.

In the special case where A is a square matrix, the transpose of A can be obtained by interchanging entries that are symmetrically positioned about the main diagonal. In (15) we see that A^T can also be obtained by "reflecting" A about its main diagonal.

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 7 & 0 \\ -5 & 8 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -2 & 4 \\ 3 & 7 & 0 \\ -5 & 8 & 6 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 3 & -5 \\ -2 & 7 & 8 \\ 4 & 0 & 6 \end{bmatrix} \tag{15}$$

Interchange entries that are symmetrically positioned about the main diagonal.

Trace of a Matrix

> **DEFINITION 8** If A is a square matrix, then the ***trace of A***, denoted by $\text{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A. The trace of A is undefined if A is not a square matrix.

EXAMPLE 12 Trace

The following are examples of matrices and their traces.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 2 & 7 & 0 \\ 3 & 5 & -8 & 4 \\ 1 & 2 & 7 & -3 \\ 4 & -2 & 1 & 0 \end{bmatrix}$$

$$\text{tr}(A) = a_{11} + a_{22} + a_{33} \qquad \text{tr}(B) = -1 + 5 + 7 + 0 = 11$$

Exercise Set 1.3

▶ In Exercises 1–2, suppose that A, B, C, D, and E are matrices with the following sizes:

A	B	C	D	E
(4×5)	(4×5)	(5×2)	(4×2)	(5×4)

In each part, determine whether the given matrix expression is defined. For those that are defined, give the size of the resulting matrix. ◀

2. (a) CD^T (b) DC (c) $BC - 3D$

(d) $D^T(BE)$ (e) $B^TD + ED$ (f) $BA^T + D$

▶ In Exercises 3–6, use the following matrices to compute the indicated expression if it is defined.

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

$$D = \begin{bmatrix} 1 & 5 & 2 \\ -1 & 0 & 1 \\ 3 & 2 & 4 \end{bmatrix}, \quad E = \begin{bmatrix} 6 & 1 & 3 \\ -1 & 1 & 2 \\ 4 & 1 & 3 \end{bmatrix}$$ ◀

3. (a) $D + E$ (b) $D - E$ (c) $5A$

(d) $-7C$ (e) $2B - C$ (f) $4E - 2D$

(g) $-3(D + 2E)$ (h) $A - A$ (i) $\text{tr}(D)$

(j) $\text{tr}(D - 3E)$ (k) $4\,\text{tr}(7B)$ (l) $\text{tr}(A)$

6. (a) $(2D^T - E)A$ (b) $(4B)C + 2B$

(c) $(-AC)^T + 5D^T$ (d) $(BA^T - 2C)^T$

(e) $B^T(CC^T - A^TA)$ (f) $D^TE^T - (ED)^T$

▶ In Exercises 23–24, solve the matrix equation for a, b, c, and d. ◀

23. $\begin{bmatrix} a & 3 \\ -1 & a + b \end{bmatrix} = \begin{bmatrix} 4 & d - 2c \\ d + 2c & -2 \end{bmatrix}$

1.4 Inverses; Algebraic Properties of Matrices

Properties of Matrix Addition and Scalar Multiplication

The following theorem lists the basic algebraic properties of the matrix operations.

THEOREM 1.4.1 Properties of Matrix Arithmetic

Assuming that the sizes of the matrices are such that the indicated operations can be performed, the following rules of matrix arithmetic are valid.

(*a*) $A + B = B + A$ **[Commutative law for matrix addition]**

(*b*) $A + (B + C) = (A + B) + C$ **[Associative law for matrix addition]**

(*c*) $A(BC) = (AB)C$ **[Associative law for matrix multiplication]**

(*d*) $A(B + C) = AB + AC$ **[Left distributive law]**

(*e*) $(B + C)A = BA + CA$ **[Right distributive law]**

(*f*) $A(B - C) = AB - AC$

(*g*) $(B - C)A = BA - CA$

(*h*) $a(B + C) = aB + aC$

(*i*) $a(B - C) = aB - aC$

(*j*) $(a + b)C = aC + bC$

(*k*) $(a - b)C = aC - bC$

(*l*) $a(bC) = (ab)C$

(*m*) $a(BC) = (aB)C = B(aC)$

▶ EXAMPLE 1 Associativity of Matrix Multiplication

As an illustration of the associative law for matrix multiplication, consider

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

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix} \quad \text{and} \quad BC = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix}$$

Thus

$$(AB)C = \begin{bmatrix} 8 & 5 \\ 20 & 13 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

and

$$A(BC) = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 10 & 9 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 18 & 15 \\ 46 & 39 \\ 4 & 3 \end{bmatrix}$$

so $(AB)C = A(BC)$, as guaranteed by Theorem 1.4.1(*c*). ◀

Properties of Matrix Multiplication

Do not let Theorem 1.4.1 lull you into believing that *all* laws of real arithmetic carry over to matrix arithmetic. For example, you know that in real arithmetic it is always true that $ab = ba$, which is called the *commutative law for multiplication*. In matrix arithmetic, however, the equality of AB and BA can fail for three possible reasons:

1. AB may be defined and BA may not (for example, if A is 2×3 and B is 3×4).
2. AB and BA may both be defined, but they may have different sizes (for example, if A is 2×3 and B is 3×2).
3. AB and BA may both be defined and have the same size, but the two products may be different (as illustrated in the next example).

EXAMPLE 2 Order Matters in Matrix Multiplication

Consider the matrices

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

Multiplying gives

$$AB = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$$

Thus, $AB \neq BA$. ◀

Zero Matrices

A matrix whose entries are all zero is called a ***zero matrix***. Some examples are

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad [0]$$

We will denote a zero matrix by 0 unless it is important to specify its size, in which case we will denote the $m \times n$ zero matrix by $0_{m \times n}$.

It should be evident that if A and 0 are matrices with the same size, then

$$A + 0 = 0 + A = A$$

Thus, 0 plays the same role in this matrix equation that the number 0 plays in the numerical equation $a + 0 = 0 + a = a$.

The following theorem lists the basic properties of zero matrices. Since the results should be self-evident, we will omit the formal proofs.

THEOREM 1.4.2 Properties of Zero Matrices

If c is a scalar, and if the sizes of the matrices are such that the operations can be perfomed, then:

(a) $A + 0 = 0 + A = A$

(b) $A - 0 = A$

(c) $A - A = A + (-A) = 0$

(d) $0A = 0$

(e) *If* $cA = 0$, *then* $c = 0$ *or* $A = 0$.

Since we know that the commutative law of real arithmetic is not valid in matrix arithmetic, it should not be surprising that there are other rules that fail as well. For example, consider the following two laws of real arithmetic:

- If $ab = ac$ and $a \neq 0$, then $b = c$. **[The cancellation law]**
- If $ab = 0$, then at least one of the factors on the left is 0.

The next two examples show that these laws are not true in matrix arithmetic.

▶ EXAMPLE 3 Failure of the Cancellation Law

Consider the matrices

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

We leave it for you to confirm that

$$AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Although $A \neq 0$, canceling A from both sides of the equation $AB = AC$ would lead to the incorrect conclusion that $B = C$. Thus, the cancellation law does not hold, in general, for matrix multiplication (though there may be particular cases where it is true).

▶ EXAMPLE 4 A Zero Product with Nonzero Factors

Here are two matrices for which $AB = 0$, but $A \neq 0$ and $B \neq 0$:

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

Identity Matrices

A square matrix with 1's on the main diagonal and zeros elsewhere is called an ***identity matrix***. Some examples are

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

An identity matrix is denoted by the letter I. If it is important to emphasize the size, we will write I_n for the $n \times n$ identity matrix.

To explain the role of identity matrices in matrix arithmetic, let us consider the effect of multiplying a general 2×3 matrix A on each side by an identity matrix. Multiplying on the right by the 3×3 identity matrix yields

$$AI_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

and multiplying on the left by the 2×2 identity matrix yields

$$I_2A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

The same result holds in general; that is, if A is any $m \times n$ matrix, then

$$AI_n = A \quad \text{and} \quad I_m A = A$$

Thus, the identity matrices play the same role in matrix arithmetic that the number 1 plays in the numerical equation $a \cdot 1 = 1 \cdot a = a$.

As the next theorem shows, identity matrices arise naturally in studying reduced row echelon forms of *square* matrices.

THEOREM 1.4.3 *If R is the reduced row echelon form of an $n \times n$ matrix A, then either R has a row of zeros or R is the identity matrix I_n.*

Inverse of a Matrix

DEFINITION 1 If A is a square matrix, and if a matrix B of the same size can be found such that $AB = BA = I$, then A is said to be ***invertible*** (or ***nonsingular***) and B is called an ***inverse*** of A. If no such matrix B can be found, then A is said to be ***singular***.

Remark The relationship $AB = BA = I$ is not changed by interchanging A and B, so if A is invertible and B is an inverse of A, then it is also true that B is invertible, and A is an inverse of B. Thus, when

$$AB = BA = I$$

we say that A and B are *inverses of one another*.

▶ EXAMPLE 5 An Invertible Matrix

Let

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

Then

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

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

Thus, A and B are invertible and each is an inverse of the other.

Properties of Inverses

THEOREM 1.4.4 *If B and C are both inverses of the matrix A, then $B = C$.*

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

THEOREM 1.4.5 *The matrix*

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

is invertible if and only if $ad - bc \neq 0$, *in which case the inverse is given by the formula*

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

▶ EXAMPLE 7 Calculating the Inverse of a 2 × 2 Matrix

In each part, determine whether the matrix is invertible. If so, find its inverse.

$$\text{(a) } A = \begin{bmatrix} 6 & 1 \\ 5 & 2 \end{bmatrix} \qquad \text{(b) } A = \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}$$

Solution (a) The determinant of A is $\det(A) = (6)(2) - (1)(5) = 7$, which is nonzero. Thus, A is invertible, and its inverse is

$$A^{-1} = \frac{1}{7}\begin{bmatrix} 2 & -1 \\ -5 & 6 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & -\frac{1}{7} \\ -\frac{5}{7} & \frac{6}{7} \end{bmatrix}$$

We leave it for you to confirm that $AA^{-1} = A^{-1}A = I$.

Solution (b) The matrix is not invertible since $\det(A) = (-1)(-6) - (2)(3) = 0$.

THEOREM 1.4.6 *If A and B are invertible matrices with the same size, then AB is invertible and*

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

Proof We can establish the invertibility and obtain the stated formula at the same time by showing that

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

But

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

and similarly, $(B^{-1}A^{-1})(AB) = I$. ◀

A product of any number of invertible matrices is invertible, and the inverse of the product is the product of the inverses in the reverse order.

▶ EXAMPLE 9 The Inverse of a Product

Consider the matrices

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$$

We leave it for you to show that

$$AB = \begin{bmatrix} 7 & 6 \\ 9 & 8 \end{bmatrix}, \quad (AB)^{-1} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

and also that

$$A^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}, \quad B^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}, \quad B^{-1}A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & \frac{3}{2} \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -3 \\ -\frac{9}{2} & \frac{7}{2} \end{bmatrix}$$

Thus, $(AB)^{-1} = B^{-1}A^{-1}$ as guaranteed by Theorem 1.4.6. ◀

Powers of a Matrix If A is a *square* matrix, then we define the nonnegative integer powers of A to be

$$A^0 = I \quad \text{and} \quad A^n = AA\cdots A \quad \text{[}n\text{ factors]}$$

and if A is invertible, then we define the negative integer powers of A to be

$$A^{-n} = (A^{-1})^n = A^{-1}A^{-1}\cdots A^{-1} \quad \text{[}n\text{ factors]}$$

$$A^r A^s = A^{r+s} \quad \text{and} \quad (A^r)^s = A^{rs}$$

THEOREM 1.4.7 *If A is invertible and n is a nonnegative integer, then:*

(*a*) *A^{-1} is invertible and $(A^{-1})^{-1} = A$.*

(*b*) *A^n is invertible and $(A^n)^{-1} = A^{-n} = (A^{-1})^n$.*

(*c*) *kA is invertible for any nonzero scalar k, and $(kA)^{-1} = k^{-1}A^{-1}$.*

▶ EXAMPLE 10 Properties of Exponents

Let A and A^{-1} be the matrices in Example 9; that is,

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

Then

$$A^{-3} = (A^{-1})^3 = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix}$$

Also,

$$A^3 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 30 \\ 15 & 41 \end{bmatrix}$$

so, as expected from Theorem 1.4.7(*b*),

$$(A^3)^{-1} = \frac{1}{(11)(41) - (30)(15)}\begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = \begin{bmatrix} 41 & -30 \\ -15 & 11 \end{bmatrix} = (A^{-1})^3$$

Matrix Polynomials If A is a square matrix, say $n \times n$, and if

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

is any polynomial, then we define the $n \times n$ matrix $p(A)$ to be

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_m A^m \tag{3}$$

where I is the $n \times n$ identity matrix; that is, $p(A)$ is obtained by substituting A for x and replacing the constant term a_0 by the matrix $a_0 I$. An expression of form (3) is called a ***matrix polynomial in A***.

▶ EXAMPLE 12 A Matrix Polynomial

Find $p(A)$ for

$$p(x) = x^2 - 2x - 3 \quad \text{and} \quad A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}$$

Solution

$$\begin{aligned} p(A) &= A^2 - 2A - 3I \\ &= \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}^2 - 2\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 4 \\ 0 & 9 \end{bmatrix} - \begin{bmatrix} -2 & 4 \\ 0 & 6 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

or more briefly, $p(A) = 0$. ◀

Properties of the Transpose

THEOREM 1.4.8 *If the sizes of the matrices are such that the stated operations can be performed, then:*

(a) $(A^T)^T = A$

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

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

(d) $(kA)^T = kA^T$

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

The transpose of a product of any number of matrices is the product of the transposes in the reverse order.

THEOREM 1.4.9 *If A is an invertible matrix, then A^T is also invertible and*

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

Proof We can establish the invertibility and obtain the formula at the same time by showing that

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

But from part (*e*) of Theorem 1.4.8 and the fact that $I^T = I$, we have

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

which completes the proof. ◀

Exercise Set 1.4

▶ In Exercises 5–8, use Theorem 1.4.5 to compute the inverse of the matrix. ◀

8. $D = \begin{bmatrix} 6 & 4 \\ -2 & -1 \end{bmatrix}$

10. Find the inverse of

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

▶ In Exercises 15–18, use the given information to find A. ◀

16. $(5A^T)^{-1} = \begin{bmatrix} -3 & -1 \\ 5 & 2 \end{bmatrix}$

17. $(I + 2A)^{-1} = \begin{bmatrix} -1 & 2 \\ 4 & 5 \end{bmatrix}$

▶ In Exercises 21–22, compute $p(A)$ for the given matrix A and the following polynomials.

(a) $p(x) = x - 2$

(b) $p(x) = 2x^2 - x + 1$

(c) $p(x) = x^3 - 2x + 1$ ◀

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

22. $A = \begin{bmatrix} 2 & 0 \\ 4 & 1 \end{bmatrix}$

1.5 Elementary Matrices and a Method for Finding A^{-1}

In Section 1.1 we defined three elementary row operations on a matrix A:

1. Multiply a row by a nonzero constant c.
2. Interchange two rows.
3. Add a constant c times one row to another.

It should be evident that if we let B be the matrix that results from A by performing one of the operations in this list, then the matrix A can be recovered from B by performing the corresponding operation in the following list:

1. Multiply the same row by $1/c$.
2. Interchange the same two rows.
3. If B resulted by adding c times row r_i of A to row r_j, then add $-c$ times r_j to r_i.

DEFINITION 1 Matrices A and B are said to be ***row equivalent*** if either (hence each) can be obtained from the other by a sequence of elementary row operations.

DEFINITION 2 A matrix E is called an ***elementary matrix*** if it can be obtained from an identity matrix by performing a *single* elementary row operation.

THEOREM 1.5.1 Row Operations by Matrix Multiplication

If the elementary matrix E results from performing a certain row operation on I_m and if A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.

THEOREM 1.5.2 *Every elementary matrix is invertible, and the inverse is also an elementary matrix.*

Equivalence Theorem

THEOREM 1.5.3 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent, that is, all true or all false.

(*a*) *A is invertible.*

(*b*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*c*) *The reduced row echelon form of A is I_n.*

(*d*) *A is expressible as a product of elementary matrices.*

A Method for Inverting Matrices

Inversion Algorithm To find the inverse of an invertible matrix A, find a sequence of elementary row operations that reduces A to the identity and then perform that same sequence of operations on I_n to obtain A^{-1}.

▶ EXAMPLE 4 Using Row Operations to Find A^{-1}

Find the inverse of

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

The computations are as follows:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

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

← We added −2 times the first row to the second and −1 times the first row to the third.

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

← We added 2 times the second row to the third.

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

← We multiplied the third row by −1.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & -14 & 6 & 3 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

← We added 3 times the third row to the second and −3 times the third row to the first.

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right]$$

← We added −2 times the second row to the first.

Thus,

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$ ◀

EXAMPLE 5 Showing That a Matrix Is Not Invertible

Consider the matrix

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

Applying the procedure of Example 4 yields

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

$$\left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 8 & 9 & 1 & 0 & 1 \end{array}\right]$$

← We added −2 times the first row to the second and added the first row to the third.

$$\left[\begin{array}{ccc|ccc} 1 & 6 & 4 & 1 & 0 & 0 \\ 0 & -8 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right]$$

← We added the second row to the third.

Since we have obtained a row of zeros on the left side, A is not invertible.

Exercise Set 1.5

▶ In Exercises 13–18, use the inversion algorithm to find the inverse of the matrix (if the inverse exists). ◀

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

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

▶ In Exercises 19–20, find the inverse of each of the following 4×4 matrices, where k_1, k_2, k_3, k_4, and k are all nonzero. ◀

19. (a) $\begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$

1.6 More on Linear Systems and Invertible Matrices

Number of Solutions of a Linear System

THEOREM 1.6.1 *A system of linear equations has zero, one, or infinitely many solutions. There are no other possibilities.*

Solving Linear Systems by Matrix Inversion

THEOREM 1.6.2 *If A is an invertible $n \times n$ matrix, then for each $n \times 1$ matrix $\mathbf{b}$, the system of equations $A\mathbf{x} = \mathbf{b}$ has exactly one solution, namely, $\mathbf{x} = A^{-1}\mathbf{b}$.*

Proof Since $A(A^{-1}\mathbf{b}) = \mathbf{b}$, it follows that $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution of $A\mathbf{x} = \mathbf{b}$. To show that this is the only solution, we will assume that $\mathbf{x}_0$ is an arbitrary solution and then show that $\mathbf{x}_0$ must be the solution $A^{-1}\mathbf{b}$.

If $\mathbf{x}_0$ is any solution of $A\mathbf{x} = \mathbf{b}$, then $A\mathbf{x}_0 = \mathbf{b}$. Multiplying both sides of this equation by A^{-1}, we obtain $\mathbf{x}_0 = A^{-1}\mathbf{b}$. ◀

▶ **EXAMPLE 1 Solution of a Linear System Using A^{-1}**

Consider the system of linear equations

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 5 \\ 2x_1 + 5x_2 + 3x_3 &= 3 \\ x_1 \qquad\quad + 8x_3 &= 17 \end{aligned}$$

In matrix form this system can be written as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix}$$

In Example 4 of the preceding section, we showed that A is invertible and

$$A^{-1} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix}$$

By Theorem 1.6.2, the solution of the system is

$$\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 17 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$$

or $x_1 = 1$, $x_2 = -1$, $x_3 = 2$. ◀

Linear Systems with a Common Coefficient Matrix

EXAMPLE 2 Solving Two Linear Systems at Once

Solve the systems

(a)
$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 4 \\ 2x_1 + 5x_2 + 3x_3 &= 5 \\ x_1 \qquad\quad + 8x_3 &= 9 \end{aligned}$$

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

Solution The two systems have the same coefficient matrix. If we augment this coefficient matrix with the columns of constants on the right sides of these systems, we obtain

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

Reducing this matrix to reduced row echelon form yields (verify)

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

It follows from the last two columns that the solution of system (a) is $x_1 = 1$, $x_2 = 0$, $x_3 = 1$ and the solution of system (b) is $x_1 = 2$, $x_2 = 1$, $x_3 = -1$. ◀

Properties of Invertible Matrices

THEOREM 1.6.3 *Let A be a square matrix.*

(a) If B is a square matrix satisfying $BA = I$, then $B = A^{-1}$.

(b) If B is a square matrix satisfying $AB = I$, then $B = A^{-1}$.

Equivalence Theorem

THEOREM 1.6.4 Equivalent Statements

If A is an $n \times n$ matrix, then the following are equivalent.

(a) A is invertible.

(b) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(c) The reduced row echelon form of A is I_n.

(d) A is expressible as a product of elementary matrices.

(e) $A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.

(f) $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.

Exercise Set 1.6

► In Exercises 1–8, solve the system by inverting the coefficient matrix and using Theorem 1.6.2. ◄

5.
$$\begin{aligned} x + y + z &= 5 \\ x + y - 4z &= 10 \\ -4x + y + z &= 0 \end{aligned}$$

8.
$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= b_1 \\ 2x_1 + 5x_2 + 5x_3 &= b_2 \\ 3x_1 + 5x_2 + 8x_3 &= b_3 \end{aligned}$$

1.7 Diagonal, Triangular, and Symmetric Matrices

Diagonal Matrices

A square matrix in which all the entries off the main diagonal are zero is called a ***diagonal matrix***. Here are some examples:

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

A general $n \times n$ diagonal matrix D can be written as

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} \tag{1}$$

A diagonal matrix is invertible if and only if all of its diagonal entries are nonzero; in this case the inverse of (1) is

$$D^{-1} = \begin{bmatrix} 1/d_1 & 0 & \cdots & 0 \\ 0 & 1/d_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1/d_n \end{bmatrix} \tag{2}$$

You can verify that this is so by multiplying (1) and (2).

Powers of diagonal matrices are easy to compute; we leave it for you to verify that if D is the diagonal matrix (1) and k is a positive integer, then

$$D^k = \begin{bmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & d_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_n^k \end{bmatrix} \tag{3}$$

▶ EXAMPLE 1 Inverses and Powers of Diagonal Matrices

If

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

then

$$A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad A^5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -243 & 0 \\ 0 & 0 & 32 \end{bmatrix}, \quad A^{-5} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{243} & 0 \\ 0 & 0 & \frac{1}{32} \end{bmatrix}$$

◀

Triangular Matrices

A square matrix in which all the entries above the main diagonal are zero is called ***lower triangular***, and a square matrix in which all the entries below the main diagonal are zero is called ***upper triangular***. A matrix that is either upper triangular or lower triangular is called ***triangular***.

EXAMPLE 2 Upper and Lower Triangular Matrices

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

A general 4×4 upper triangular matrix

A general 4×4 lower triangular matrix

Properties of Triangular Matrices

THEOREM 1.7.1

(*a*) *The transpose of a lower triangular matrix is upper triangular, and the transpose of an upper triangular matrix is lower triangular.*

(*b*) *The product of lower triangular matrices is lower triangular, and the product of upper triangular matrices is upper triangular.*

(*c*) *A triangular matrix is invertible if and only if its diagonal entries are all nonzero.*

(*d*) *The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.*

Symmetric Matrices

DEFINITION 1 A square matrix A is said to be ***symmetric*** if $A = A^T$.

▶ EXAMPLE 4 Symmetric Matrices

The following matrices are symmetric, since each is equal to its own transpose (verify).

$$\begin{bmatrix} 7 & -3 \\ -3 & 5 \end{bmatrix}, \quad \begin{bmatrix} 1 & 4 & 5 \\ 4 & -3 & 0 \\ 5 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} d_1 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 \\ 0 & 0 & d_3 & 0 \\ 0 & 0 & 0 & d_4 \end{bmatrix} \blacktriangleleft$$

THEOREM 1.7.2 *If A and B are symmetric matrices with the same size, and if k is any scalar, then:*

(*a*) A^T *is symmetric.*

(*b*) $A + B$ *and* $A - B$ *are symmetric.*

(*c*) kA *is symmetric.*

It is not true, in general, that the product of symmetric matrices is symmetric. To see why this is so, let A and B be symmetric matrices with the same size. Then it follows from part (*e*) of Theorem 1.4.8 and the symmetry of A and B that

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

Thus, $(AB)^T = AB$ if and only if $AB = BA$, that is, if and only if A and B commute. In summary, we have the following result.

THEOREM 1.7.3 *The product of two symmetric matrices is symmetric if and only if the matrices commute.*

▶ EXAMPLE 5 Products of Symmetric Matrices

The first of the following equations shows a product of symmetric matrices that *is not* symmetric, and the second shows a product of symmetric matrices that *is* symmetric. We conclude that the factors in the first equation do not commute, but those in the second equation do. We leave it for you to verify that this is so.

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

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

Invertibility of Symmetric Matrices

THEOREM 1.7.4 *If A is an invertible symmetric matrix, then A^{-1} is symmetric.*

Proof Assume that A is symmetric and invertible. From Theorem 1.4.9 and the fact that $A = A^T$, we have

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

which proves that A^{-1} is symmetric. ◀

THEOREM 1.7.4 *If A is an invertible symmetric matrix, then A^{-1} is symmetric.*

Proof Assume that A is symmetric and invertible. From Theorem 1.4.9 and the fact that $A = A^T$, we have

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

which proves that A^{-1} is symmetric. ◀

Products AA^T and A^TA are Symmetric

$$(AA^T)^T = (A^T)^T A^T = AA^T \quad \text{and} \quad (A^TA)^T = A^T(A^T)^T = A^TA$$

▶ **EXAMPLE 6 The Product of a Matrix and Its Transpose Is Symmetric**

Let A be the 2×3 matrix

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

Then

$$A^TA = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 10 & -2 & -11 \\ -2 & 4 & -8 \\ -11 & -8 & 41 \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} 21 & -17 \\ -17 & 34 \end{bmatrix}$$

Observe that A^TA and AA^T are symmetric as expected. ◀

THEOREM 1.7.5 *If A is an invertible matrix, then AA^T and A^TA are also invertible.*

Exercise Set 1.7

▶ In Exercises 7–10, find A^2, A^{-2}, and A^{-k} (where k is any integer) by inspection. ◀

8. $A = \begin{bmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

▶ In Exercises 13–14, compute the indicated quantity. ◀

14. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{1000}$

▶ In Exercises 19–22, determine by inspection whether the matrix is invertible. ◀

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

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

CHAPTER 2

Determinants

2.1 Determinants by Cofactor Expansion

Recall from Theorem 1.4.5 that the 2×2 matrix

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

is invertible if and only if $ad - bc \neq 0$ and that the expression $ad - bc$ is called the ***determinant*** of the matrix A. Recall also that this determinant is denoted by writing

$$\det(A) = ad - bc \quad \text{or} \quad \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{1}$$

and that the inverse of A can be expressed in terms of the determinant as

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \tag{2}$$

Minors and Cofactors

DEFINITION 1 If A is a square matrix, then the ***minor of entry*** $\boldsymbol{a_{ij}}$ is denoted by M_{ij} and is defined to be the determinant of the submatrix that remains after the ith row and jth column are deleted from A. The number $(-1)^{i+j}M_{ij}$ is denoted by C_{ij} and is called the ***cofactor of entry*** $\boldsymbol{a_{ij}}$.

EXAMPLE 1 Finding Minors and Cofactors

Let

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

The minor of entry a_{11} is

$$M_{11} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 5 & 6 \\ 4 & 8 \end{vmatrix} = 16$$

The cofactor of a_{11} is

$$C_{11} = (-1)^{1+1}M_{11} = M_{11} = 16$$

Similarly, the minor of entry a_{32} is

$$M_{32} = \begin{vmatrix} 3 & 1 & -4 \\ 2 & 5 & 6 \\ 1 & 4 & 8 \end{vmatrix} = \begin{vmatrix} 3 & -4 \\ 2 & 6 \end{vmatrix} = 26$$

The cofactor of a_{32} is

$$C_{32} = (-1)^{3+2} M_{32} = -M_{32} = -26$$

Remark Note that a minor M_{ij} and its corresponding cofactor C_{ij} are either the same or negatives of each other and that the relating sign $(-1)^{i+j}$ is either $+1$ or -1 in accordance with the pattern in the "checkerboard" array

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

For example,

$$C_{11} = M_{11}, \quad C_{21} = -M_{21}, \quad C_{22} = M_{22}$$

EXAMPLE 2 Cofactor Expansions of a 2 × 2 Matrix

The checkerboard pattern for a 2×2 matrix $A = [a_{ij}]$ is

$$\begin{bmatrix} + & - \\ - & + \end{bmatrix}$$

so that

$$\begin{aligned} C_{11} &= M_{11} = a_{22} & \qquad C_{12} &= -M_{12} = -a_{21} \\ C_{21} &= -M_{21} = -a_{12} & \qquad C_{22} &= M_{22} = a_{11} \end{aligned}$$

We leave it for you to use Formula (3) to verify that $\det(A)$ can be expressed in terms of cofactors in the following four ways:

$$\begin{aligned} \det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}C_{11} + a_{12}C_{12} \\ &= a_{21}C_{21} + a_{22}C_{22} \\ &= a_{11}C_{11} + a_{21}C_{21} \\ &= a_{12}C_{12} + a_{22}C_{22} \end{aligned} \tag{6}$$

Each of the last four equations is called a *cofactor expansion* of $\det(A)$. In each cofactor expansion the entries and cofactors all come from the same row or same column of A.

For example, in the first equation the entries and cofactors all come from the first row of A, in the second they all come from the second row of A, in the third they all come from the first column of A, and in the fourth they all come from the second column of A.

Definition of a General Determinant

THEOREM 2.1.1 *If A is an $n \times n$ matrix, then regardless of which row or column of A is chosen, the number obtained by multiplying the entries in that row or column by the corresponding cofactors and adding the resulting products is always the same.*

DEFINITION 2 If A is an $n \times n$ matrix, then the number obtained by multiplying the entries in any row or column of A by the corresponding cofactors and adding the resulting products is called the ***determinant of A***, and the sums themselves are called ***cofactor expansions of A***. That is,

$$\det(A) = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{7}$$

[cofactor expansion along the *j*th column]

and

$$\det(A) = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \tag{8}$$

[cofactor expansion along the *i*th row]

▶ EXAMPLE 3 Cofactor Expansion Along the First Row

Find the determinant of the matrix

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

by cofactor expansion along the first row.

Solution

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

$$= 3(-4) - (1)(-11) + 0 = -1$$

▶ EXAMPLE 4 Cofactor Expansion Along the First Column

Let A be the matrix in Example 3, and evaluate $\det(A)$ by cofactor expansion along the first column of A.

Solution

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

$$= 3(-4) - (-2)(-2) + 5(3) = -1$$

This agrees with the result obtained in Example 3.

EXAMPLE 5 Smart Choice of Row or Column

If A is the 4×4 matrix

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

then to find $\det(A)$ it will be easiest to use cofactor expansion along the second column, since it has the most zeros:

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

For the 3×3 determinant, it will be easiest to use cofactor expansion along its second column, since it has the most zeros:

$$\begin{aligned} \det(A) &= 1 \cdot -2 \cdot \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} \\ &= -2(1+2) \\ &= -6 \end{aligned}$$

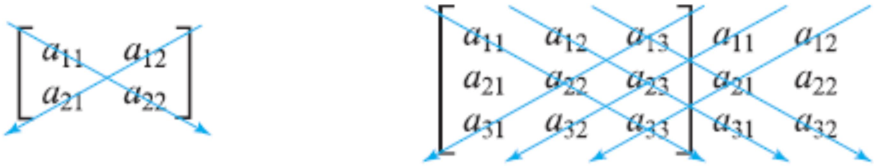
EXAMPLE 6 Determinant of a Lower Triangular Matrix

The following computation shows that the determinant of a 4×4 lower triangular matrix is the product of its diagonal entries. Each part of the computation uses a cofactor expansion along the first row.

$$\begin{aligned} \begin{vmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & 0 & 0 \\ a_{32} & a_{33} & 0 \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \\ &= a_{11}a_{22} \begin{vmatrix} a_{33} & 0 \\ a_{43} & a_{44} \end{vmatrix} \\ &= a_{11}a_{22}a_{33}|a_{44}| = a_{11}a_{22}a_{33}a_{44} \end{aligned}$$

THEOREM 2.1.2 *If A is an $n \times n$ triangular matrix (upper triangular, lower triangular, or diagonal), then $\det(A)$ is the product of the entries on the main diagonal of the matrix; that is, $\det(A) = a_{11}a_{22}\cdots a_{nn}$.*

A Useful Technique for Evaluating 2×2 and 3×3 Determinants



EXAMPLE 7 A Technique for Evaluating 2 × 2 and 3 × 3 Determinants

$$\begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 4 & -2 \end{vmatrix} = (3)(-2) - (1)(4) = -10$$

$$\begin{vmatrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{vmatrix} = \left|\begin{matrix} 1 & 2 & 3 \\ -4 & 5 & 6 \\ 7 & -8 & 9 \end{matrix}\right| \begin{matrix} 1 & 2 \\ -4 & 5 \\ 7 & -8 \end{matrix}$$

$$= [45 + 84 + 96] - [105 - 48 - 72] = 240$$ ◀

Exercise Set 2.1

3. Let

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

Find

(a) M_{13} and C_{13}.

(b) M_{23} and C_{23}.

(c) M_{22} and C_{22}.

(d) M_{21} and C_{21}.

▶ In Exercises 9–14, use the arrow technique to evaluate the determinant. ◀

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

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

▶ In Exercises 21–26, evaluate $\det(A)$ by a cofactor expansion along a row or column of your choice. ◀

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

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

2.2 Evaluating Determinants by Row Reduction

THEOREM 2.2.1 *Let A be a square matrix. If A has a row of zeros or a column of zeros, then* $\det(A) = 0$.

Proof Since the determinant of A can be found by a cofactor expansion along any row or column, we can use the row or column of zeros. Thus, if we let $C_1, C_2, \ldots, C_n$ denote the cofactors of A along that row or column, then it follows from Formula (7) or (8) in Section 2.1 that

$$\det(A) = 0 \cdot C_1 + 0 \cdot C_2 + \cdots + 0 \cdot C_n = 0 \blacktriangleleft$$

THEOREM 2.2.2 *Let A be a square matrix. Then* $\det(A) = \det(A^T)$.

Elementary Row Operations

Relationship	Operation
$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = k \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = k\det(A)$	In the matrix B the first row of A was multiplied by k.
$\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = -\det(A)$	In the matrix B the first and second rows of A were interchanged.
$\begin{vmatrix} a_{11}+ka_{21} & a_{12}+ka_{22} & a_{13}+ka_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ $\det(B) = \det(A)$	In the matrix B a multiple of the second row of A was added to the first row.

THEOREM 2.2.3 *Let A be an $n \times n$ matrix.*

(*a*) *If B is the matrix that results when a single row or single column of A is multiplied by a scalar k, then* $\det(B) = k\det(A)$.

(*b*) *If B is the matrix that results when two rows or two columns of A are interchanged, then* $\det(B) = -\det(A)$.

(*c*) *If B is the matrix that results when a multiple of one row of A is added to another or when a multiple of one column is added to another, then* $\det(B) = \det(A)$.

THEOREM 2.2.5 *If A is a square matrix with two proportional rows or two proportional columns, then* $\det(A) = 0$.

EXAMPLE 2 Proportional Rows or Columns

Each of the following matrices has two proportional rows or columns; thus, each has a determinant of zero.

$$\begin{bmatrix} -1 & 4 \\ -2 & 8 \end{bmatrix}, \quad \begin{bmatrix} 1 & -2 & 7 \\ -4 & 8 & 5 \\ 2 & -4 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 & 4 & -5 \\ 6 & -2 & 5 & 2 \\ 5 & 8 & 1 & 4 \\ -9 & 3 & -12 & 15 \end{bmatrix}$$

Evaluating Determinants by Row Reduction

EXAMPLE 3 Using Row Reduction to Evaluate a Determinant

Evaluate $\det(A)$ where

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

Solution We will reduce A to row echelon form (which is upper triangular) and then apply Theorem 2.1.2.

$$\det(A) = \begin{vmatrix} 0 & 1 & 5 \\ 3 & -6 & 9 \\ 2 & 6 & 1 \end{vmatrix} = - \begin{vmatrix} 3 & -6 & 9 \\ 0 & 1 & 5 \\ 2 & 6 & 1 \end{vmatrix}$$

← The first and second rows of A were interchanged.

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

← A common factor of 3 from the first row was taken through the determinant sign.

$$= -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 10 & -5 \end{vmatrix}$$

← −2 times the first row was added to the third row.

$$= -3 \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & -55 \end{vmatrix}$$

← −10 times the second row was added to the third row.

$$= (-3)(-55) \begin{vmatrix} 1 & -2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{vmatrix}$$

← A common factor of −55 from the last row was taken through the determinant sign.

$$= (-3)(-55)(1) = 165$$

▶ EXAMPLE 4 Using Column Operations to Evaluate a Determinant

Compute the determinant of

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

Solution This determinant could be computed as above by using elementary row operations to reduce A to row echelon form, but we can put A in lower triangular form in one step by adding -3 times the first column to the fourth to obtain

$$\det(A) = \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 7 & 0 & 0 \\ 0 & 6 & 3 & 0 \\ 7 & 3 & 1 & -26 \end{bmatrix} = (1)(7)(3)(-26) = -546 \blacktriangleleft$$

Cofactor expansion and row or column operations can sometimes be used in combination to provide an effective method for evaluating determinants. The following example illustrates this idea.

▶ EXAMPLE 5 Row Operations and Cofactor Expansion

Evaluate $\det(A)$ where

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

Solution By adding suitable multiples of the second row to the remaining rows, we obtain

$$\det(A) = \begin{vmatrix} 0 & -1 & 1 & 3 \\ 1 & 2 & -1 & 1 \\ 0 & 0 & 3 & 3 \\ 0 & 1 & 8 & 0 \end{vmatrix}$$

$$= -\begin{vmatrix} -1 & 1 & 3 \\ 0 & 3 & 3 \\ 1 & 8 & 0 \end{vmatrix} \quad \leftarrow \text{Cofactor expansion along the first column}$$

$$= -\begin{vmatrix} -1 & 1 & 3 \\ 0 & 3 & 3 \\ 0 & 9 & 3 \end{vmatrix} \quad \leftarrow \text{We added the first row to the third row.}$$

$$= -(-1)\begin{vmatrix} 3 & 3 \\ 9 & 3 \end{vmatrix} \quad \leftarrow \text{Cofactor expansion along the first column}$$

$$= -18 \blacktriangleleft$$

Exercise Set 2.2

▶ In Exercises 9–14, evaluate the determinant of the matrix by first reducing the matrix to row echelon form and then using some combination of row operations and cofactor expansion. ◀

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

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

In Exercises 15–22, evaluate the determinant, given that

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = -6$$ ◀

15. $\begin{vmatrix} d & e & f \\ g & h & i \\ a & b & c \end{vmatrix}$

18. $\begin{vmatrix} a+d & b+e & c+f \\ -d & -e & -f \\ g & h & i \end{vmatrix}$

19. $\begin{vmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{vmatrix}$

20. $\begin{vmatrix} a & b & c \\ 2d & 2e & 2f \\ g+3a & h+3b & i+3c \end{vmatrix}$

21. $\begin{vmatrix} -3a & -3b & -3c \\ d & e & f \\ g-4d & h-4e & i-4f \end{vmatrix}$

▶ In Exercises 29–30, show that $\det(A) = 0$ without directly evaluating the determinant. ◀

30. $A = \begin{bmatrix} -4 & 1 & 1 & 1 & 1 \\ 1 & -4 & 1 & 1 & 1 \\ 1 & 1 & -4 & 1 & 1 \\ 1 & 1 & 1 & -4 & 1 \\ 1 & 1 & 1 & 1 & -4 \end{bmatrix}$

2.3 Properties of Determinants; Cramer's Rule

Basic Properties of Determinants

Suppose that A and B are $n \times n$ matrices and k is any scalar. We begin by considering possible relationships among $\det(A)$, $\det(B)$, and

$$\det(kA), \quad \det(A+B), \quad \text{and} \quad \det(AB)$$

$$\det(kA) = k^n \det(A) \tag{1}$$

For example,

$$\begin{vmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{vmatrix} = k^3 \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

Unfortunately, no simple relationship exists among $\det(A)$, $\det(B)$, and $\det(A+B)$. In particular, $\det(A+B)$ will usually *not* be equal to $\det(A) + \det(B)$. The following example illustrates this fact.

EXAMPLE 1 $\det(A+B) \neq \det(A) + \det(B)$

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}, \quad A + B = \begin{bmatrix} 4 & 3 \\ 3 & 8 \end{bmatrix}$$

We have $\det(A) = 1$, $\det(B) = 8$, and $\det(A+B) = 23$; thus

$$\det(A+B) \neq \det(A) + \det(B)$$

Determinant Test for Invertibility

THEOREM 2.3.3 *A square matrix A is invertible if and only if $\det(A) \neq 0$.*

EXAMPLE 3 Determinant Test for Invertibility

Since the first and third rows of

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

are proportional, $\det(A) = 0$. Thus A is not invertible.

We are now ready for the main result concerning products of matrices.

THEOREM 2.3.4 *If A and B are square matrices of the same size, then*

$$\det(AB) = \det(A)\det(B)$$

EXAMPLE 4 Verifying that $\det(AB) = \det(A)\det(B)$

Consider the matrices

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 3 \\ 5 & 8 \end{bmatrix}, \quad AB = \begin{bmatrix} 2 & 17 \\ 3 & 14 \end{bmatrix}$$

We leave it for you to verify that

$$\det(A) = 1, \quad \det(B) = -23, \quad \text{and} \quad \det(AB) = -23$$

Thus $\det(AB) = \det(A)\det(B)$, as guaranteed by Theorem 2.3.4.

THEOREM 2.3.5 *If A is invertible, then*

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Proof Since $A^{-1}A = I$, it follows that $\det(A^{-1}A) = \det(I)$. Therefore, we must have $\det(A^{-1})\det(A) = 1$. Since $\det(A) \neq 0$, the proof can be completed by dividing through by $\det(A)$. ◀

Adjoint of a Matrix

▶ **EXAMPLE 5 Entries and Cofactors from Different Rows**

Let

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

We leave it for you to verify that the cofactors of A are

$$\begin{array}{lll} C_{11} = 12 & C_{12} = 6 & C_{13} = -16 \\ C_{21} = 4 & C_{22} = 2 & C_{23} = 16 \\ C_{31} = 12 & C_{32} = -10 & C_{33} = 16 \end{array}$$

so, for example, the cofactor expansion of $\det(A)$ along the first row is

$$\det(A) = 3C_{11} + 2C_{12} + (-1)C_{13} = 36 + 12 + 16 = 64$$

and along the first column is

$$\det(A) = 3C_{11} + C_{21} + 2C_{31} = 36 + 4 + 24 = 64$$

Suppose, however, we multiply the entries in the first row by the corresponding cofactors from the *second row* and add the resulting products. The result is

$$3C_{21} + 2C_{22} + (-1)C_{23} = 12 + 4 - 16 = 0$$

Or suppose we multiply the entries in the first column by the corresponding cofactors from the *second column* and add the resulting products. The result is again zero since

$$3C_{12} + 1C_{22} + 2C_{32} = 18 + 2 - 20 = 0 \ ◀$$

DEFINITION 1 If A is any $n \times n$ matrix and C_{ij} is the cofactor of a_{ij}, then the matrix

$$\begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}$$

is called the ***matrix of cofactors from A***. The transpose of this matrix is called the ***adjoint of A*** and is denoted by $\text{adj}(A)$.

▶ EXAMPLE 6 Adjoint of a 3 × 3 Matrix

Let

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

As noted in Example 5, the cofactors of A are

$$\begin{array}{lll} C_{11} = 12 & C_{12} = 6 & C_{13} = -16 \\ C_{21} = 4 & C_{22} = 2 & C_{23} = 16 \\ C_{31} = 12 & C_{32} = -10 & C_{33} = 16 \end{array}$$

so the matrix of cofactors is

$$\begin{bmatrix} 12 & 6 & -16 \\ 4 & 2 & 16 \\ 12 & -10 & 16 \end{bmatrix}$$

and the adjoint of A is

$$\text{adj}(A) = \begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix}$$ ◀

THEOREM 2.3.6 Inverse of a Matrix Using Its Adjoint

If A is an invertible matrix, then

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) \tag{6}$$

▶ EXAMPLE 7 Using the Adjoint to Find an Inverse Matrix

Use Formula (6) to find the inverse of the matrix A in Example 6.

Solution We showed in Example 5 that $\det(A) = 64$. Thus,

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) = \frac{1}{64}\begin{bmatrix} 12 & 4 & 12 \\ 6 & 2 & -10 \\ -16 & 16 & 16 \end{bmatrix} = \begin{bmatrix} \frac{12}{64} & \frac{4}{64} & \frac{12}{64} \\ \frac{6}{64} & \frac{2}{64} & -\frac{10}{64} \\ -\frac{16}{64} & \frac{16}{64} & \frac{16}{64} \end{bmatrix}$$ ◀

Cramer's Rule

THEOREM 2.3.7 Cramer's Rule

If $A\mathbf{x} = \mathbf{b}$ is a system of n linear equations in n unknowns such that $\det(A) \neq 0$, *then the system has a unique solution. This solution is*

$$x_1 = \frac{\det(A_1)}{\det(A)}, \quad x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, \quad x_n = \frac{\det(A_n)}{\det(A)}$$

where A_j is the matrix obtained by replacing the entries in the jth column of A by the entries in the matrix

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

EXAMPLE 8 Using Cramer's Rule to Solve a Linear System

Use Cramer's rule to solve

$$\begin{aligned} x_1 + \quad\quad + 2x_3 &= 6 \\ -3x_1 + 4x_2 + 6x_3 &= 30 \\ -x_1 - 2x_2 + 3x_3 &= 8 \end{aligned}$$

Solution

$$A = \begin{bmatrix} 1 & 0 & 2 \\ -3 & 4 & 6 \\ -1 & -2 & 3 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 6 & 0 & 2 \\ 30 & 4 & 6 \\ 8 & -2 & 3 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} 1 & 6 & 2 \\ -3 & 30 & 6 \\ -1 & 8 & 3 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 & 6 \\ -3 & 4 & 30 \\ -1 & -2 & 8 \end{bmatrix}$$

Therefore,

$$x_1 = \frac{\det(A_1)}{\det(A)} = \frac{-40}{44} = \frac{-10}{11}, \quad x_2 = \frac{\det(A_2)}{\det(A)} = \frac{72}{44} = \frac{18}{11},$$

$$x_3 = \frac{\det(A_3)}{\det(A)} = \frac{152}{44} = \frac{38}{11}$$

THEOREM 2.3.8 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(*a*) *A is invertible.*

(*b*) *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*

(*c*) *The reduced row echelon form of A is I_n.*

(*d*) *A can be expressed as a product of elementary matrices.*

(*e*) *$A\mathbf{x} = \mathbf{b}$ is consistent for every $n \times 1$ matrix $\mathbf{b}$.*

(*f*) *$A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{b}$.*

(*g*) $\det(A) \neq 0$.

Exercise Set 2.3

▶ In Exercises 1–4, verify that $\det(kA) = k^n \det(A)$. ◀

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

▶ In Exercises 7–14, use determinants to decide whether the given matrix is invertible. ◀

12. $A = \begin{bmatrix} 1 & 0 & -1 \\ 9 & -1 & 4 \\ 8 & 9 & -1 \end{bmatrix}$

▶ In Exercises 15–18, find the values of k for which the matrix A is invertible. ◀

16. $A = \begin{bmatrix} k & 2 \\ 2 & k \end{bmatrix}$

▶ In Exercises 19–23, decide whether the matrix is invertible, and if so, use the adjoint method to find its inverse. ◀

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

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

▶ In Exercises 24–29, solve by Cramer's rule, where it applies. ◀

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

29. $\begin{aligned} 3x_1 - x_2 + x_3 &= 4 \\ -x_1 + 7x_2 - 2x_3 &= 1 \\ 2x_1 + 6x_2 - x_3 &= 5 \end{aligned}$

CHAPTER 4

General Vector Spaces

4.1 Real Vector Spaces

Vector Space Axioms

DEFINITION 1 Let V be an arbitrary nonempty set of objects on which two operations are defined: addition, and multiplication by numbers called ***scalars***. By ***addition*** we mean a rule for associating with each pair of objects $\mathbf{u}$ and $\mathbf{v}$ in V an object $\mathbf{u}+\mathbf{v}$, called the ***sum*** of $\mathbf{u}$ and $\mathbf{v}$; by ***scalar multiplication*** we mean a rule for associating with each scalar k and each object $\mathbf{u}$ in V an object $k\mathbf{u}$, called the ***scalar multiple*** of $\mathbf{u}$ by k. If the following axioms are satisfied by all objects $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in V and all scalars k and m, then we call V a ***vector space*** and we call the objects in V ***vectors***.

1. If $\mathbf{u}$ and $\mathbf{v}$ are objects in V, then $\mathbf{u}+\mathbf{v}$ 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 an object $\mathbf{0}$ in V, called a ***zero vector*** for V, such that $\mathbf{0}+\mathbf{u}=\mathbf{u}+\mathbf{0}=\mathbf{u}$ for all $\mathbf{u}$ in V.
5. For each $\mathbf{u}$ in V, there is an object $-\mathbf{u}$ in V, called a ***negative*** of $\mathbf{u}$, such that $\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}$.
6. If k is any scalar and $\mathbf{u}$ is any object in V, then $k\mathbf{u}$ is in V.
7. $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$
8. $(k+m)\mathbf{u}=k\mathbf{u}+m\mathbf{u}$
9. $k(m\mathbf{u})=(km)(\mathbf{u})$
10. $1\mathbf{u}=\mathbf{u}$

To Show That a Set with Two Operations Is a Vector Space

Step 1. Identify the set V of objects that will become vectors.

Step 2. Identify the addition and scalar multiplication operations on V.

Step 3. Verify Axioms 1 and 6; that is, adding two vectors in V produces a vector in V, and multiplying a vector in V by a scalar also produces a vector in V. Axiom 1 is called ***closure under addition***, and Axiom 6 is called ***closure under scalar multiplication***.

Step 4. Confirm that Axioms 2, 3, 4, 5, 7, 8, 9, and 10 hold.

▶ EXAMPLE 1 The Zero Vector Space

Let V consist of a single object, which we denote by $\mathbf{0}$, and define

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{and} \quad k\mathbf{0} = \mathbf{0}$$

for all scalars k. It is easy to check that all the vector space axioms are satisfied. We call this the ***zero vector space***. ◀

Our second example is one of the most important of all vector spaces—the familiar space R^n. It should not be surprising that the operations on R^n satisfy the vector space axioms because those axioms were based on known properties of operations on R^n.

▶ EXAMPLE 2 R^n Is a Vector Space

Let $V = R^n$, and define the vector space operations on V to be the usual operations of addition and scalar multiplication of n-tuples; that is,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1, u_2, \ldots, u_n) + (v_1, v_2, \ldots, v_n) = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n) \\ k\mathbf{u} &= (ku_1, ku_2, \ldots, ku_n) \end{aligned}$$

The set $V = R^n$ is closed under addition and scalar multiplication because the foregoing operations produce n-tuples as their end result, and these operations satisfy Axioms 2, 3, 4, 5, 7, 8, 9, and 10 by virtue of Theorem 3.1.1. ◀

▶ EXAMPLE 4 The Vector Space of 2 × 2 Matrices

Let V be the set of 2×2 matrices with real entries, and take the vector space operations on V to be the usual operations of matrix addition and scalar multiplication; that is,

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} u_{11} + v_{11} & u_{12} + v_{12} \\ u_{21} + v_{21} & u_{22} + v_{22} \end{bmatrix} \tag{1}$$

$$k\mathbf{u} = k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix}$$

The set V is closed under addition and scalar multiplication because the foregoing operations produce 2×2 matrices as the end result. Thus, it remains to confirm that Axioms 2, 3, 4, 5, 7, 8, 9, and 10 hold. Some of these are standard properties of matrix operations. For example, Axiom 2 follows from Theorem 1.4.1(a) since

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} + \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} + \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} = \mathbf{v} + \mathbf{u}$$

Similarly, Axioms 3, 7, 8, and 9 follow from parts (b), (h), (j), and (e), respectively, of that theorem (verify). This leaves Axioms 4, 5, and 10 that remain to be verified.

To confirm that Axiom 4 is satisfied, we must find a 2×2 matrix $\mathbf{0}$ in V for which $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u}$ for all 2×2 matrices in V. We can do this by taking

$$\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

With this definition,

$$\mathbf{0}+\mathbf{u}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}+\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}=\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}=\mathbf{u}$$

and similarly $\mathbf{u}+\mathbf{0}=\mathbf{u}$. To verify that Axiom 5 holds we must show that each object $\mathbf{u}$ in V has a negative $-\mathbf{u}$ in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$ and $(-\mathbf{u})+\mathbf{u}=\mathbf{0}$. This can be done by defining the negative of $\mathbf{u}$ to be

$$-\mathbf{u}=\begin{bmatrix}-u_{11} & -u_{12}\\ -u_{21} & -u_{22}\end{bmatrix}$$

With this definition,

$$\mathbf{u}+(-\mathbf{u})=\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}+\begin{bmatrix}-u_{11} & -u_{12}\\ -u_{21} & -u_{22}\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=\mathbf{0}$$

and similarly $(-\mathbf{u})+\mathbf{u}=\mathbf{0}$. Finally, Axiom 10 holds because

$$1\mathbf{u}=1\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}=\begin{bmatrix}u_{11} & u_{12}\\ u_{21} & u_{22}\end{bmatrix}=\mathbf{u}$$

▶ EXAMPLE 5 The Vector Space of $m \times n$ Matrices

Example 4 is a special case of a more general class of vector spaces. You should have no trouble adapting the argument used in that example to show that the set V of all $m \times n$ matrices with the usual matrix operations of addition and scalar multiplication is a vector space. We will denote this vector space by the symbol M_{mn}. Thus, for example, the vector space in Example 4 is denoted as M_{22}.

EXAMPLE 6 The Vector Space of Real-Valued Functions

Let V be the set of real-valued functions that are defined at each x in the interval $(-\infty, \infty)$. If $\mathbf{f} = f(x)$ and $\mathbf{g} = g(x)$ are two functions in V and if k is any scalar, then define the operations of addition and scalar multiplication by

$$(\mathbf{f} + \mathbf{g})(x) = f(x) + g(x) \tag{2}$$

$$(k\mathbf{f})(x) = kf(x) \tag{3}$$

One way to think about these operations is to view the numbers $f(x)$ and $g(x)$ as "components" of $\mathbf{f}$ and $\mathbf{g}$ at the point x, in which case Equations (2) and (3) state that two functions are added by adding corresponding components, and a function is multiplied by a scalar by multiplying each component by that scalar—exactly as in R^n and R^∞. This idea is illustrated in parts (*a*) and (*b*) of Figure 4.1.2. The set V with these operations is denoted by the symbol $F(-\infty, \infty)$. We can prove that this is a vector space as follows:

Axioms 1 and 6: These closure axioms require that if we add two functions that are defined at each x in the interval $(-\infty, \infty)$, then sums and scalar multiples of those functions must also be defined at each x in the interval $(-\infty, \infty)$. This follows from Formulas (2) and (3).

Axiom 4: This axiom requires that there exists a function $\mathbf{0}$ in $F(-\infty, \infty)$, which when added to any other function $\mathbf{f}$ in $F(-\infty, \infty)$ produces $\mathbf{f}$ back again as the result. The function whose value at every point x in the interval $(-\infty, \infty)$ is zero has this property. Geometrically, the graph of the function $\mathbf{0}$ is the line that coincides with the x-axis.

Axiom 5: This axiom requires that for each function $\mathbf{f}$ in $F(-\infty, \infty)$ there exists a function $-\mathbf{f}$ in $F(-\infty, \infty)$, which when added to $\mathbf{f}$ produces the function $\mathbf{0}$. The function defined by $-\mathbf{f}(x) = -f(x)$ has this property. The graph of $-\mathbf{f}$ can be obtained by reflecting the graph of $\mathbf{f}$ about the x-axis (Figure 4.1.2*c*).

Axioms 2, 3, 7, 8, 9, 10: The validity of each of these axioms follows from properties of real numbers. For example, if $\mathbf{f}$ and $\mathbf{g}$ are functions in $F(-\infty, \infty)$, then Axiom 2 requires that $\mathbf{f} + \mathbf{g} = \mathbf{g} + \mathbf{f}$. This follows from the computation

$$(\mathbf{f} + \mathbf{g})(x) = \mathbf{f}(x) + \mathbf{g}(x) = \mathbf{g}(x) + \mathbf{f}(x) = (\mathbf{g} + \mathbf{f})(x)$$

in which the first and last equalities follow from (2), and the middle equality is a property of real numbers. We will leave the proofs of the remaining parts as exercises. ◀

▶ EXAMPLE 7 A Set That Is Not a Vector Space

Let $V = R^2$ and define addition and scalar multiplication operations as follows: If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, then define

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2)$$

and if k is any real number, then define

$$k\mathbf{u} = (ku_1, 0)$$

For example, if $\mathbf{u} = (2, 4)$, $\mathbf{v} = (-3, 5)$, and $k = 7$, then

$$\mathbf{u} + \mathbf{v} = (2 + (-3), 4 + 5) = (-1, 9)$$
$$k\mathbf{u} = 7\mathbf{u} = (7 \cdot 2, 0) = (14, 0)$$

The addition operation is the standard one from R^2, but the scalar multiplication is not. In the exercises we will ask you to show that the first nine vector space axioms are satisfied. However, Axiom 10 fails to hold for certain vectors. For example, if $\mathbf{u} = (u_1, u_2)$ is such that $u_2 \neq 0$, then

$$1\mathbf{u} = 1(u_1, u_2) = (1 \cdot u_1, 0) = (u_1, 0) \neq \mathbf{u}$$

Thus, V is not a vector space with the stated operations. ◀

Our final example will be an unusual vector space that we have included to illustrate how varied vector spaces can be. Since the vectors in this space will be real numbers, it will be important for you to keep track of which operations are intended as vector operations and which ones as ordinary operations on real numbers.

▶ EXAMPLE 8 An Unusual Vector Space

Let V be the set of positive real numbers, let $\mathbf{u} = u$ and $\mathbf{v} = v$ be any vectors (i.e., positive real numbers) in V, and let k be any scalar. Define the operations on V to be

$$u + v = uv \quad \text{[Vector addition is numerical multiplication.]}$$
$$ku = u^k \quad \text{[Scalar multiplication is numerical exponentiation.]}$$

Thus, for example, $1 + 1 = 1$ and $(2)(1) = 1^2 = 1$—strange indeed, but nevertheless the set V with these operations satisfies the ten vector space axioms and hence is a vector space. We will confirm Axioms 4, 5, and 7, and leave the others as exercises.

- Axiom 4—The zero vector in this space is the number 1 (i.e., $\mathbf{0} = 1$) since

$$u + 1 = u \cdot 1 = u$$

- Axiom 5—The negative of a vector u is its reciprocal (i.e., $-u = 1/u$) since

$$u + \frac{1}{u} = u\left(\frac{1}{u}\right) = 1\ (= \mathbf{0})$$

- Axiom 7—$k(u + v) = (uv)^k = u^k v^k = (ku) + (kv)$. ◀

Some Properties of Vectors

THEOREM 4.1.1 *Let V be a vector space,* $\mathbf{u}$ *a vector in V, and k a scalar; then:*

(*a*) $0\mathbf{u} = \mathbf{0}$

(*b*) $k\mathbf{0} = \mathbf{0}$

(*c*) $(-1)\mathbf{u} = -\mathbf{u}$

(*d*) *If* $k\mathbf{u} = \mathbf{0}$, *then* $k = 0$ *or* $\mathbf{u} = \mathbf{0}$.

Proof (*a*) We can write

$$\begin{aligned} 0\mathbf{u} + 0\mathbf{u} &= (0 + 0)\mathbf{u} && \text{[Axiom 8]} \\ &= 0\mathbf{u} && \text{[Property of the number 0]} \end{aligned}$$

By Axiom 5 the vector $0\mathbf{u}$ has a negative, $-0\mathbf{u}$. Adding this negative to both sides above yields

$$[0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u}) = 0\mathbf{u} + (-0\mathbf{u})$$

or

$$\begin{aligned} 0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})] &= 0\mathbf{u} + (-0\mathbf{u}) && \text{[Axiom 3]} \\ 0\mathbf{u} + \mathbf{0} &= \mathbf{0} && \text{[Axiom 5]} \\ 0\mathbf{u} &= \mathbf{0} && \text{[Axiom 4]} \end{aligned}$$

Proof (*c*) To prove that $(-1)\mathbf{u} = -\mathbf{u}$, we must show that $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. The proof is as follows:

$$\begin{aligned} \mathbf{u} + (-1)\mathbf{u} &= 1\mathbf{u} + (-1)\mathbf{u} && \text{[Axiom 10]} \\ &= (1 + (-1))\mathbf{u} && \text{[Axiom 8]} \\ &= 0\mathbf{u} && \text{[Property of numbers]} \\ &= \mathbf{0} && \text{[Part (a) of this theorem]} \end{aligned}$$ ◀

Exercise Set 4.1

1. Let V be the set of all ordered pairs of real numbers, and consider the following addition and scalar multiplication operations on $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2), \quad k\mathbf{u} = (0, ku_2)$$

(a) Compute $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for $\mathbf{u} = (-1, 2)$, $\mathbf{v} = (3, 4)$, and $k = 3$.

(b) In words, explain why V is closed under addition and scalar multiplication.

(c) Since addition on V is the standard addition operation on R^2, certain vector space axioms hold for V because they are known to hold for R^2. Which axioms are they?

(d) Show that Axioms 7, 8, and 9 hold.

(e) Show that Axiom 10 fails and hence that V is not a vector space under the given operations.

2. Let V be the set of all ordered pairs of real numbers, and consider the following addition and scalar multiplication operations on $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1 + 1, u_2 + v_2 + 1), \quad k\mathbf{u} = (ku_1, ku_2)$$

(a) Compute $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for $\mathbf{u} = (0, 4)$, $\mathbf{v} = (1, -3)$, and $k = 2$.

(b) Show that $(0, 0) \neq \mathbf{0}$.

(c) Show that $(-1, -1) = \mathbf{0}$.

(d) Show that Axiom 5 holds by producing an ordered pair $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ for $\mathbf{u} = (u_1, u_2)$.

(e) Find two vector space axioms that fail to hold.

▶ In Exercises 3–12, determine whether each set equipped with the given operations is a vector space. For those that are not vector spaces identify the vector space axioms that fail. ◀

11. The set of all pairs of real numbers of the form $(1, x)$ with the operations

$$(1, y) + (1, y') = (1, y + y') \quad \text{and} \quad k(1, y) = (1, ky)$$

4.2 Subspaces

DEFINITION 1 A subset W of a vector space V is called a ***subspace*** of V if W is itself a vector space under the addition and scalar multiplication defined on V.

THEOREM 4.2.1 *If W is a set of one or more vectors in a vector space V, then W is a subspace of V if and only if the following conditions are satisfied.*

(*a*) *If $\mathbf{u}$ and $\mathbf{v}$ are vectors in W, then $\mathbf{u}+\mathbf{v}$ is in W.*

(*b*) *If k is a scalar and $\mathbf{u}$ is a vector in W, then $k\mathbf{u}$ is in W.*

▶ EXAMPLE 1 The Zero Subspace

If V is any vector space, and if $W = \{\mathbf{0}\}$ is the subset of V that consists of the zero vector only, then W is closed under addition and scalar multiplication since

$$\mathbf{0}+\mathbf{0}=\mathbf{0} \quad \text{and} \quad k\mathbf{0}=\mathbf{0}$$

for any scalar k. We call W the ***zero subspace*** of V.

▶ EXAMPLE 4 A Subset of R^2 That Is Not a Subspace

Let W be the set of all points (x, y) in R^2 for which $x \geq 0$ and $y \geq 0$ (the shaded region in Figure 4.2.4). This set is not a subspace of R^2 because it is not closed under scalar multiplication. For example, $\mathbf{v} = (1, 1)$ is a vector in W, but $(-1)\mathbf{v} = (-1, -1)$ is not.

▶ EXAMPLE 5 Subspaces of M_{nn}

We know from Theorem 1.7.2 that the sum of two symmetric $n \times n$ matrices is symmetric and that a scalar multiple of a symmetric $n \times n$ matrix is symmetric. Thus, the set of symmetric $n \times n$ matrices is closed under addition and scalar multiplication and hence is a subspace of M_{nn}. Similarly, the sets of upper triangular matrices, lower triangular matrices, and diagonal matrices are subspaces of M_{nn}.

▶ EXAMPLE 6 A Subset of M_{nn} That Is Not a Subspace

The set W of invertible $n \times n$ matrices is not a subspace of M_{nn}, failing on two counts—it is not closed under addition and not closed under scalar multiplication. We will illustrate this with an example in M_{22} that you can readily adapt to M_{nn}. Consider the matrices

$$U = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -1 & 2 \\ -2 & 5 \end{bmatrix}$$

The matrix $0U$ is the 2×2 zero matrix and hence is not invertible, and the matrix $U + V$ has a column of zeros so it also is not invertible.

▶ EXAMPLE 9 The Subspace of All Polynomials

Recall that a ***polynomial*** is a function that can be expressed in the form

$$p(x) = a_0 + a_1x + \cdots + a_nx^n \tag{1}$$

where $a_0, a_1, \ldots, a_n$ are constants. It is evident that the sum of two polynomials is a polynomial and that a constant times a polynomial is a polynomial. Thus, the set W of all polynomials is closed under addition and scalar multiplication and hence is a subspace of $F(-\infty, \infty)$. We will denote this space by P_∞.

▶ EXAMPLE 10 The Subspace of Polynomials of Degree $\leq n$

Recall that the ***degree*** of a polynomial is the highest power of the variable that occurs with a nonzero coefficient. Thus, for example, if $a_n \neq 0$ in Formula (1), then that polynomial has degree n. It is *not* true that the set W of polynomials with positive degree n is a subspace of $F(-\infty, \infty)$ because that set is not closed under addition. For example, the polynomials

$$1 + 2x + 3x^2 \quad \text{and} \quad 5 + 7x - 3x^2$$

both have degree 2, but their sum has degree 1. What *is* true, however, is that for each nonnegative integer n the polynomials of degree n *or less* form a subspace of $F(-\infty, \infty)$. We will denote this space by P_n. ◀

THEOREM 4.2.2 *If $W_1, W_2, \ldots, W_r$ are subspaces of a vector space V, then the intersection of these subspaces is also a subspace of V.*

Proof Let W be the intersection of the subspaces $W_1, W_2, \ldots, W_r$. This set is not empty because each of these subspaces contains the zero vector of V, and hence so does their intersection. Thus, it remains to show that W is closed under addition and scalar multiplication.

To prove closure under addition, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in W. Since W is the intersection of $W_1, W_2, \ldots, W_r$, it follows that $\mathbf{u}$ and $\mathbf{v}$ also lie in each of these subspaces. Moreover, since these subspaces are closed under addition and scalar multiplication, they also all contain the vectors $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for every scalar k, and hence so does their intersection W. This proves that W is closed under addition and scalar multiplication. ◀

DEFINITION 2 If $\mathbf{w}$ is a vector in a vector space V, then $\mathbf{w}$ is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in V if $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \tag{2}$$

where $k_1, k_2, \ldots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

THEOREM 4.2.3 *If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then:*

(*a*) *The set W of all possible linear combinations of the vectors in S is a subspace of V.*

(*b*) *The set W in part (a) is the "smallest" subspace of V that contains all of the vectors in S in the sense that any other subspace that contains those vectors contains W.*

Proof (*a*) Let W be the set of all possible linear combinations of the vectors in S. We must show that W is closed under addition and scalar multiplication. To prove closure under addition, let

$$\mathbf{u} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_r\mathbf{w}_r \quad \text{and} \quad \mathbf{v} = k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_r\mathbf{w}_r$$

be two vectors in W. It follows that their sum can be written as

$$\mathbf{u} + \mathbf{v} = (c_1 + k_1)\mathbf{w}_1 + (c_2 + k_2)\mathbf{w}_2 + \cdots + (c_r + k_r)\mathbf{w}_r$$

which is a linear combination of the vectors in S. Thus, W is closed under addition. We leave it for you to prove that W is also closed under scalar multiplication and hence is a subspace of V.

Proof (*b*) Let W' be any subspace of V that contains all of the vectors in S. Since W' is closed under addition and scalar multiplication, it contains all linear combinations of the vectors in S and hence contains W. ◀

The following definition gives some important notation and terminology related to Theorem 4.2.3.

DEFINITION 3 If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space V, then the subspace W of V that consists of all possible linear combinations of the vectors in S is called the subspace of V ***generated*** by S, and we say that the vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ ***span*** W. We denote this subspace as

$$W = \text{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\} \quad \text{or} \quad W = \text{span}(S)$$

▶ EXAMPLE 11 The Standard Unit Vectors Span R^n

Recall that the standard unit vectors in R^n are

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

These vectors span R^n since every vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in R^n can be expressed as

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n$$

which is a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Thus, for example, the vectors

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

span R^3 since every vector $\mathbf{v} = (a, b, c)$ in this space can be expressed as

$$\mathbf{v} = (a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

▶ EXAMPLE 13 A Spanning Set for P_n

The polynomials $1, x, x^2, \ldots, x^n$ span the vector space P_n defined in Example 10 since each polynomial $\mathbf{p}$ in P_n can be written as

$$\mathbf{p} = a_0 + a_1 x + \cdots + a_n x^n$$

which is a linear combination of $1, x, x^2, \ldots, x^n$. We can denote this by writing

$$P_n = \text{span}\{1, x, x^2, \ldots, x^n\}$$ ◀

▶ EXAMPLE 14 Linear Combinations

Consider the vectors $\mathbf{u} = (1, 2, -1)$ and $\mathbf{v} = (6, 4, 2)$ in R^3. Show that $\mathbf{w} = (9, 2, 7)$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$ and that $\mathbf{w}' = (4, -1, 8)$ is *not* a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

Solution In order for $\mathbf{w}$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w} = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(9, 2, 7) = k_1(1, 2, -1) + k_2(6, 4, 2) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 9 \\ 2k_1 + 4k_2 &= 2 \\ -k_1 + 2k_2 &= 7 \end{aligned}$$

Solving this system using Gaussian elimination yields $k_1 = -3$, $k_2 = 2$, so

$$\mathbf{w} = -3\mathbf{u} + 2\mathbf{v}$$

Similarly, for $\mathbf{w}'$ to be a linear combination of $\mathbf{u}$ and $\mathbf{v}$, there must be scalars k_1 and k_2 such that $\mathbf{w}' = k_1\mathbf{u} + k_2\mathbf{v}$; that is,

$$(4, -1, 8) = k_1(1, 2, -1) + k_2(6, 4, 2) = (k_1 + 6k_2, 2k_1 + 4k_2, -k_1 + 2k_2)$$

Equating corresponding components gives

$$\begin{aligned} k_1 + 6k_2 &= 4 \\ 2k_1 + 4k_2 &= -1 \\ -k_1 + 2k_2 &= 8 \end{aligned}$$

This system of equations is inconsistent (verify), so no such scalars k_1 and k_2 exist. Consequently, $\mathbf{w}'$ is not a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

EXAMPLE 15 Testing for Spanning

Determine whether the vectors $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 0, 1)$, and $\mathbf{v}_3 = (2, 1, 3)$ span the vector space R^3.

Solution We must determine whether an arbitrary vector $\mathbf{b} = (b_1, b_2, b_3)$ in R^3 can be expressed as a linear combination

$$\mathbf{b} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3$$

of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Expressing this equation in terms of components gives

$$(b_1, b_2, b_3) = k_1(1, 1, 2) + k_2(1, 0, 1) + k_3(2, 1, 3)$$

or

$$(b_1, b_2, b_3) = (k_1 + k_2 + 2k_3,\ k_1 + k_3,\ 2k_1 + k_2 + 3k_3)$$

or

$$\begin{aligned} k_1 + k_2 + 2k_3 &= b_1 \\ k_1 \phantom{{}+k_2} + k_3 &= b_2 \\ 2k_1 + k_2 + 3k_3 &= b_3 \end{aligned}$$

Thus, our problem reduces to ascertaining whether this system is consistent for all values of b_1, b_2, and b_3. One way of doing this is to use parts (e) and (g) of Theorem 2.3.8, which state that the system is consistent if and only if its coefficient matrix

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

has a nonzero determinant. But this is *not* the case here since $\det(A) = 0$ (verify), so $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ do not span R^3. ◀

Solution Spaces of Homogeneous Systems

THEOREM 4.2.4 *The solution set of a homogeneous linear system $A\mathbf{x} = \mathbf{0}$ of m equations in n unknowns is a subspace of R^n.*

Proof Let W be the solution set of the system. The set W is not empty because it contains at least the trivial solution $\mathbf{x} = \mathbf{0}$.

To show that W is a subspace of R^n, we must show that it is closed under addition and scalar multiplication. To do this, let $\mathbf{x}_1$ and $\mathbf{x}_2$ be vectors in W. Since these vectors are solutions of $A\mathbf{x} = \mathbf{0}$, we have

$$A\mathbf{x}_1 = \mathbf{0} \quad \text{and} \quad A\mathbf{x}_2 = \mathbf{0}$$

It follows from these equations and the distributive property of matrix multiplication that

$$A(\mathbf{x}_1 + \mathbf{x}_2) = A\mathbf{x}_1 + A\mathbf{x}_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

so W is closed under addition. Similarly, if k is any scalar then

$$A(k\mathbf{x}_1) = kA\mathbf{x}_1 = k\mathbf{0} = \mathbf{0}$$

so W is also closed under scalar multiplication. ◀

► EXAMPLE 16 Solution Spaces of Homogeneous Systems

In each part, solve the system by any method and then give a geometric description of the solution set.

(a) $\begin{bmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ -2 & 4 & -6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -2 & 3 \\ -3 & 7 & -8 \\ 4 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Solution

(a) The solutions are

$$x = 2s - 3t, \quad y = s, \quad z = t$$

from which it follows that

$$x = 2y - 3z \quad \text{or} \quad x - 2y + 3z = 0$$

This is the equation of a plane through the origin that has $\mathbf{n} = (1, -2, 3)$ as a normal.

(b) The solutions are

$$x = -5t, \quad y = -t, \quad z = t$$

which are parametric equations for the line through the origin that is parallel to the vector $\mathbf{v} = (-5, -1, 1)$.

(c) The only solution is $x = 0$, $y = 0$, $z = 0$, so the solution space consists of the single point $\{\mathbf{0}\}$.

(d) This linear system is satisfied by all real values of x, y, and z, so the solution space is all of R^3. ◄

Exercise Set 4.2

1. Use Theorem 4.2.1 to determine which of the following are subspaces of R^3.
   (a) All vectors of the form $(a, 0, 0)$.
   (b) All vectors of the form $(a, 1, 1)$.
   (c) All vectors of the form (a, b, c), where $b = a + c$.
   (d) All vectors of the form (a, b, c), where $b = a + c + 1$.
   (e) All vectors of the form $(a, b, 0)$.

2. Use Theorem 4.2.1 to determine which of the following are subspaces of M_{nn}.
   (a) The set of all diagonal $n \times n$ matrices.
   (b) The set of all $n \times n$ matrices A such that $\det(A) = 0$.
   (c) The set of all $n \times n$ matrices A such that $\operatorname{tr}(A) = 0$.
   (d) The set of all symmetric $n \times n$ matrices.
   (e) The set of all $n \times n$ matrices A such that $A^T = -A$.
   (f) The set of all $n \times n$ matrices A for which $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.
   (g) The set of all $n \times n$ matrices A such that $AB = BA$ for some fixed $n \times n$ matrix B.

4. Which of the following are subspaces of $F(-\infty, \infty)$?
   (a) All functions f in $F(-\infty, \infty)$ for which $f(0) = 0$.
   (b) All functions f in $F(-\infty, \infty)$ for which $f(0) = 1$.
   (c) All functions f in $F(-\infty, \infty)$ for which $f(-x) = f(x)$.
   (d) All polynomials of degree 2.

7. Which of the following are linear combinations of $\mathbf{u} = (0, -2, 2)$ and $\mathbf{v} = (1, 3, -1)$?
   (a) $(2, 2, 2)$ (b) $(0, 4, 5)$ (c) $(0, 0, 0)$

8. Express the following as linear combinations of $\mathbf{u} = (2, 1, 4)$, $\mathbf{v} = (1, -1, 3)$, and $\mathbf{w} = (3, 2, 5)$.
   (a) $(-9, -7, -15)$ (b) $(6, 11, 6)$ (c) $(0, 0, 0)$

11. In each part, determine whether the vectors span R^3.

(a) $\mathbf{v}_1 = (2, 2, 2)$, $\mathbf{v}_2 = (0, 0, 3)$, $\mathbf{v}_3 = (0, 1, 1)$

(b) $\mathbf{v}_1 = (2, -1, 3)$, $\mathbf{v}_2 = (4, 1, 2)$, $\mathbf{v}_3 = (8, -1, 8)$

13. Determine whether the following polynomials span P_2.

$$\mathbf{p}_1 = 1 - x + 2x^2, \quad \mathbf{p}_2 = 3 + x,$$
$$\mathbf{p}_3 = 5 - x + 4x^2, \quad \mathbf{p}_4 = -2 - 2x + 2x^2$$

4.3 Linear Independence

Linear Independence and Dependence

DEFINITION 1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a set of two or more vectors in a vector space V, then S is said to be a ***linearly independent set*** if no vector in S can be expressed as a linear combination of the others. A set that is not linearly independent is said to be ***linearly dependent***.

THEOREM 4.3.1 *A nonempty set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ in a vector space V is linearly independent if and only if the only coefficients satisfying the vector equation*

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}$$

are $k_1 = 0, k_2 = 0, \ldots, k_r = 0$.

▶ EXAMPLE 1 Linear Independence of the Standard Unit Vectors in R^n

The most basic linearly independent set in R^n is the set of standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

To illustrate this in R^3, consider the standard unit vectors

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

To prove linear independence we must show that the only coefficients satisfying the vector equation

$$k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k} = \mathbf{0}$$

are $k_1 = 0, k_2 = 0, k_3 = 0$. But this becomes evident by writing this equation in its component form

$$(k_1, k_2, k_3) = (0, 0, 0)$$

You should have no trouble adapting this argument to establish the linear independence of the standard unit vectors in R^n.

▶ EXAMPLE 2 Linear Independence in R^3

Determine whether the vectors

$$\mathbf{v}_1 = (1, -2, 3), \quad \mathbf{v}_2 = (5, 6, -1), \quad \mathbf{v}_3 = (3, 2, 1) \tag{2}$$

are linearly independent or linearly dependent in R^3.

Solution The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0} \tag{3}$$

can be satisfied with coefficients that are not all zero. To see whether this is so, let us rewrite (3) in the component form

$$k_1(1, -2, 3) + k_2(5, 6, -1) + k_3(3, 2, 1) = (0, 0, 0)$$

Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0 \end{aligned} \tag{4}$$

Thus, our problem reduces to determining whether this system has nontrivial solutions. There are various ways to do this; one possibility is to simply solve the system, which yields

$$k_1 = -\tfrac{1}{2}t, \quad k_2 = -\tfrac{1}{2}t, \quad k_3 = t$$

(we omit the details). This shows that the system has nontrivial solutions and hence that the vectors are linearly dependent. A second method for establishing the linear dependence is to take advantage of the fact that the coefficient matrix

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

is square and compute its determinant. We leave it for you to show that $\det(A) = 0$ from which it follows that (4) has nontrivial solutions by parts (b) and (g) of Theorem 2.3.8.

Because we have established that the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ in (2) are linearly dependent, we know that at least one of them is a linear combination of the others. We leave it for you to confirm, for example, that

$$\mathbf{v}_3 = \tfrac{1}{2}\mathbf{v}_1 + \tfrac{1}{2}\mathbf{v}_2$$

▶ EXAMPLE 3 Linear Independence in R^4

Determine whether the vectors

$$\mathbf{v}_1 = (1, 2, 2, -1), \quad \mathbf{v}_2 = (4, 9, 9, -4), \quad \mathbf{v}_3 = (5, 8, 9, -5)$$

in R^4 are linearly dependent or linearly independent.

Solution The linear independence or linear dependence of these vectors is determined by whether there exist nontrivial solutions of the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0}$$

or, equivalently, of

$$k_1(1, 2, 2, -1) + k_2(4, 9, 9, -4) + k_3(5, 8, 9, -5) = (0, 0, 0, 0)$$

Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{aligned} k_1 + 4k_2 + 5k_3 &= 0 \\ 2k_1 + 9k_2 + 8k_3 &= 0 \\ 2k_1 + 9k_2 + 9k_3 &= 0 \\ -k_1 - 4k_2 - 5k_3 &= 0 \end{aligned}$$

We leave it for you to show that this system has only the trivial solution

$$k_1 = 0, \quad k_2 = 0, \quad k_3 = 0$$

from which you can conclude that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are linearly independent.

▶ EXAMPLE 4 An Important Linearly Independent Set in P_n

Show that the polynomials

$$1, \quad x, \quad x^2, \ldots, \quad x^n$$

form a linearly independent set in P_n.

Solution For convenience, let us denote the polynomials as

$$\mathbf{p}_0 = 1, \quad \mathbf{p}_1 = x, \quad \mathbf{p}_2 = x^2, \ldots, \quad \mathbf{p}_n = x^n$$

We must show that the only coefficients satisfying the vector equation

$$a_0\mathbf{p}_0 + a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \cdots + a_n\mathbf{p}_n = \mathbf{0} \tag{5}$$

are

$$a_0 = a_1 = a_2 = \cdots = a_n = 0$$

But (5) is equivalent to the statement that

$$a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = 0 \tag{6}$$

for all x in $(-\infty, \infty)$, so we must show that this is true if and only if each coefficient in (6) is zero. To see that this is so, recall from algebra that a nonzero polynomial of degree n has at most n distinct roots. That being the case, each coefficient in (6) must be zero, for otherwise the left side of the equation would be a nonzero polynomial with infinitely many roots. Thus, (5) has only the trivial solution. ◀

▶ EXAMPLE 5 Linear Independence of Polynomials

Determine whether the polynomials

$$\mathbf{p}_1 = 1 - x, \quad \mathbf{p}_2 = 5 + 3x - 2x^2, \quad \mathbf{p}_3 = 1 + 3x - x^2$$

are linearly dependent or linearly independent in P_2.

Solution The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1\mathbf{p}_1 + k_2\mathbf{p}_2 + k_3\mathbf{p}_3 = \mathbf{0} \tag{7}$$

can be satisfied with coefficients that are not all zero. To see whether this is so, let us rewrite (7) in its polynomial form

$$k_1(1 - x) + k_2(5 + 3x - 2x^2) + k_3(1 + 3x - x^2) = 0 \tag{8}$$

or, equivalently, as

$$(k_1 + 5k_2 + k_3) + (-k_1 + 3k_2 + 3k_3)x + (-2k_2 - k_3)x^2 = 0$$

Since this equation must be satisfied by all x in $(-\infty, \infty)$, each coefficient must be zero (as explained in the previous example). Thus, the linear dependence or independence of the given polynomials hinges on whether the following linear system has a nontrivial solution:

$$\begin{aligned} k_1 + 5k_2 + k_3 &= 0 \\ -k_1 + 3k_2 + 3k_3 &= 0 \\ -2k_2 - k_3 &= 0 \end{aligned} \tag{9}$$

We leave it for you to show that this linear system has nontrivial solutions either by solving it directly or by showing that the coefficient matrix has determinant zero. Thus, the set $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is linearly dependent. ◀

Sets with One or Two Vectors

THEOREM 4.3.2

(*a*) *A finite set that contains* $\mathbf{0}$ *is linearly dependent.*

(*b*) *A set with exactly one vector is linearly independent if and only if that vector is not* $\mathbf{0}$.

(*c*) *A set with exactly two vectors is linearly independent if and only if neither vector is a scalar multiple of the other.*

THEOREM 4.3.3 *Let* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ *be a set of vectors in* R^n. *If* $r > n$, *then* S *is linearly dependent.*

Exercise Set 4.3

1. Explain why the following form linearly dependent sets of vectors. (Solve this problem by inspection.)

   (a) $\mathbf{u}_1 = (-1, 2, 4)$ and $\mathbf{u}_2 = (5, -10, -20)$ in R^3

   (b) $\mathbf{u}_1 = (3, -1),\ \mathbf{u}_2 = (4, 5),\ \mathbf{u}_3 = (-4, 7)$ in R^2

   (c) $\mathbf{p}_1 = 3 - 2x + x^2$ and $\mathbf{p}_2 = 6 - 4x + 2x^2$ in P_2

   (d) $A = \begin{bmatrix} -3 & 4 \\ 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -4 \\ -2 & 0 \end{bmatrix}$ in M_{22}

2. In each part, determine whether the vectors are linearly independent or are linearly dependent in R^3.

   (a) $(-3, 0, 4),\ (5, -1, 2),\ (1, 1, 3)$

4. In each part, determine whether the vectors are linearly independent or are linearly dependent in P_2.

   (a) $2 - x + 4x^2,\ 3 + 6x + 2x^2,\ 2 + 10x - 4x^2$

   (b) $1 + 3x + 3x^2,\ x + 4x^2,\ 5 + 6x + 3x^2,\ 7 + 2x - x^2$

5. In each part, determine whether the matrices are linearly independent or dependent.

   (a) $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ in M_{22}

   (b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ in M_{23}

10. (a) Show that the vectors $\mathbf{v}_1 = (1, 2, 3, 4)$, $\mathbf{v}_2 = (0, 1, 0, -1)$, and $\mathbf{v}_3 = (1, 3, 3, 3)$ form a linearly dependent set in R^4.

    (b) Express each vector in part (a) as a linear combination of the other two.

4.4 Coordinates and Basis

Coordinate Systems in Linear Algebra

DEFINITION 1 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a set of vectors in a finite-dimensional vector space V, then S is called a ***basis*** for V if:

(a) S spans V.

(b) S is linearly independent.

EXAMPLE 1 The Standard Basis for R^n

Recall from Example 11 of Section 4.2 that the standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

span R^n and from Example 1 of Section 4.3 that they are linearly independent. Thus, they form a basis for R^n that we call the ***standard basis for*** $\boldsymbol{R^n}$. In particular,

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

is the standard basis for R^3.

EXAMPLE 2 The Standard Basis for P_n

Show that $S = \{1, x, x^2, \ldots, x^n\}$ is a basis for the vector space P_n of polynomials of degree n or less.

Solution We must show that the polynomials in S are linearly independent and span P_n. Let us denote these polynomials by

$$\mathbf{p}_0 = 1, \quad \mathbf{p}_1 = x, \quad \mathbf{p}_2 = x^2, \ldots, \quad \mathbf{p}_n = x^n$$

We showed in Example 13 of Section 4.2 that these vectors span P_n and in Example 4 of Section 4.3 that they are linearly independent. Thus, they form a basis for P_n that we call the ***standard basis for*** $\boldsymbol{P_n}$.

EXAMPLE 3 Another Basis for R^3

Show that the vectors $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 9, 0)$, and $\mathbf{v}_3 = (3, 3, 4)$ form a basis for R^3.

Solution We must show that these vectors are linearly independent and span R^3. To prove linear independence we must show that the vector equation

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{0} \tag{1}$$

has only the trivial solution; and to prove that the vectors span R^3 we must show that every vector $\mathbf{b} = (b_1, b_2, b_3)$ in R^3 can be expressed as

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \mathbf{b} \tag{2}$$

By equating corresponding components on the two sides, these two equations can be expressed as the linear systems

$$\begin{aligned} c_1 + 2c_2 + 3c_3 &= 0 \\ 2c_1 + 9c_2 + 3c_3 &= 0 \\ c_1 \qquad + 4c_3 &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} c_1 + 2c_2 + 3c_3 &= b_1 \\ 2c_1 + 9c_2 + 3c_3 &= b_2 \\ c_1 \qquad + 4c_3 &= b_3 \end{aligned} \tag{3}$$

▶ EXAMPLE 4 The Standard Basis for M_{mn}

Show that the matrices

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis for the vector space M_{22} of 2×2 matrices.

Solution We must show that the matrices are linearly independent and span M_{22}. To prove linear independence we must show that the equation

$$c_1M_1 + c_2M_2 + c_3M_3 + c_4M_4 = \mathbf{0} \tag{4}$$

has only the trivial solution, where $\mathbf{0}$ is the 2×2 zero matrix; and to prove that the matrices span M_{22} we must show that every 2×2 matrix

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

can be expressed as

$$c_1M_1 + c_2M_2 + c_3M_3 + c_4M_4 = B \tag{5}$$

The matrix forms of Equations (4) and (5) are

$$c_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$c_1\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c_2\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c_3\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + c_4\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

which can be rewritten as

$$\begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Since the first equation has only the trivial solution

$$c_1 = c_2 = c_3 = c_4 = 0$$

the matrices are linearly independent, and since the second equation has the solution

$$c_1 = a, \quad c_2 = b, \quad c_3 = c, \quad c_4 = d$$

the matrices span M_{22}. This proves that the matrices M_1, M_2, M_3, M_4 form a basis for M_{22}. More generally, the mn different matrices whose entries are zero except for a single entry of 1 form a basis for M_{mn} called the ***standard basis for*** $\boldsymbol{M_{mn}}$. ◀

▶ EXAMPLE 5 An Infinite-Dimensional Vector Space

Show that the vector space of P_∞ of all polynomials with real coefficients is infinite-dimensional by showing that it has no finite spanning set.

Solution If there were a finite spanning set, say $S = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_r\}$, then the degrees of the polynomials in S would have a maximum value, say n; and this in turn would imply that any linear combination of the polynomials in S would have degree at most n. Thus, there would be no way to express the polynomial x^{n+1} as a linear combination of the polynomials in S, contradicting the fact that the vectors in S span P_∞.

Coordinates Relative to a Basis

THEOREM 4.4.1 Uniqueness of Basis Representation

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for a vector space V, then every vector $\mathbf{v}$ in V can be expressed in the form $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$ in exactly one way.

Proof Since S spans V, it follows from the definition of a spanning set that every vector in V is expressible as a linear combination of the vectors in S. To see that there is only *one* way to express a vector as a linear combination of the vectors in S, suppose that some vector $\mathbf{v}$ can be written as

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

and also as

$$\mathbf{v} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n$$

Subtracting the second equation from the first gives

$$\mathbf{0} = (c_1 - k_1)\mathbf{v}_1 + (c_2 - k_2)\mathbf{v}_2 + \cdots + (c_n - k_n)\mathbf{v}_n$$

Since the right side of this equation is a linear combination of vectors in S, the linear independence of S implies that

$$c_1 - k_1 = 0, \quad c_2 - k_2 = 0, \dots, \quad c_n - k_n = 0$$

that is,

$$c_1 = k_1, \quad c_2 = k_2, \dots, \quad c_n = k_n$$

Thus, the two expressions for $\mathbf{v}$ are the same. ◀

DEFINITION 2 If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis for a vector space V, and

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

is the expression for a vector $\mathbf{v}$ in terms of the basis S, then the scalars $c_1, c_2, \dots, c_n$ are called the ***coordinates*** of $\mathbf{v}$ relative to the basis S. The vector $(c_1, c_2, \dots, c_n)$ in R^n constructed from these coordinates is called the ***coordinate vector of* $\mathbf{v}$ *relative to S***; it is denoted by

$$(\mathbf{v})_S = (c_1, c_2, \dots, c_n) \tag{6}$$

▶ EXAMPLE 7 Coordinates Relative to the Standard Basis for R^n

In the special case where $V = R^n$ and S is the *standard basis*, the coordinate vector $(\mathbf{v})_S$ and the vector $\mathbf{v}$ are the same; that is,

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

For example, in R^3 the representation of a vector $\mathbf{v} = (a, b, c)$ as a linear combination of the vectors in the standard basis $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is

$$\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

so the coordinate vector relative to this basis is $(\mathbf{v})_S = (a, b, c)$, which is the same as the vector $\mathbf{v}$.

▶ EXAMPLE 8 Coordinate Vectors Relative to Standard Bases

(a) Find the coordinate vector for the polynomial

$$\mathbf{p}(x) = c_0 + c_1x + c_2x^2 + \cdots + c_nx^n$$

relative to the standard basis for the vector space P_n.

(b) Find the coordinate vector of

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

relative to the standard basis for M_{22}.

Solution (a) The given formula for $\mathbf{p}(x)$ expresses this polynomial as a linear combination of the standard basis vectors $S = \{1, x, x^2, \ldots, x^n\}$. Thus, the coordinate vector for $\mathbf{p}$ relative to S is

$$(\mathbf{p})_S = (c_0, c_1, c_2, \ldots, c_n)$$

Solution (b) We showed in Example 4 that the representation of a vector

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

as a linear combination of the standard basis vectors is

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

so the coordinate vector of B relative to S is

$$(B)_S = (a, b, c, d)$$

▶ EXAMPLE 9 Coordinates in R^3

(a) We showed in Example 3 that the vectors

$$\mathbf{v}_1 = (1, 2, 1), \quad \mathbf{v}_2 = (2, 9, 0), \quad \mathbf{v}_3 = (3, 3, 4)$$

form a basis for R^3. Find the coordinate vector of $\mathbf{v} = (5, -1, 9)$ relative to the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(b) Find the vector $\mathbf{v}$ in R^3 whose coordinate vector relative to S is $(\mathbf{v})_S = (-1, 3, 2)$.

Solution (a) To find $(\mathbf{v})_S$ we must first express $\mathbf{v}$ as a linear combination of the vectors in S; that is, we must find values of c_1, c_2, and c_3 such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$$

or, in terms of components,

$$(5, -1, 9) = c_1(1, 2, 1) + c_2(2, 9, 0) + c_3(3, 3, 4)$$

Equating corresponding components gives

$$\begin{aligned} c_1 + 2c_2 + 3c_3 &= 5 \\ 2c_1 + 9c_2 + 3c_3 &= -1 \\ c_1 \qquad\;\; + 4c_3 &= 9 \end{aligned}$$

Solving this system we obtain $c_1 = 1$, $c_2 = -1$, $c_3 = 2$ (verify). Therefore,

$$(\mathbf{v})_S = (1, -1, 2)$$

Solution (b) Using the definition of $(\mathbf{v})_S$, we obtain

$$\begin{aligned} \mathbf{v} &= (-1)\mathbf{v}_1 + 3\mathbf{v}_2 + 2\mathbf{v}_3 \\ &= (-1)(1, 2, 1) + 3(2, 9, 0) + 2(3, 3, 4) = (11, 31, 7) \end{aligned}$$ ◀

Exercise Set 4.4

1. Use the method of Example 3 to show that the following set of vectors forms a basis for R^2.

$$\{(2, 1), (3, 0)\}$$

2. Use the method of Example 3 to show that the following set of vectors forms a basis for R^3.

$$\{(3, 1, -4), (2, 5, 6), (1, 4, 8)\}$$

3. Show that the following polynomials form a basis for P_2.

$$x^2 + 1, \quad x^2 - 1, \quad 2x - 1$$

6. Show that the following matrices form a basis for M_{22}.

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

11. Find the coordinate vector of $\mathbf{w}$ relative to the basis $S = \{\mathbf{u}_1, \mathbf{u}_2\}$ for R^2.

(a) $\mathbf{u}_1 = (2, -4),\ \mathbf{u}_2 = (3, 8);\ \mathbf{w} = (1, 1)$

(b) $\mathbf{u}_1 = (1, 1),\ \mathbf{u}_2 = (0, 2);\ \mathbf{w} = (a, b)$

14. Find the coordinate vector of $\mathbf{p}$ relative to the basis $S = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ for P_2.

(a) $\mathbf{p} = 4 - 3x + x^2;\ \mathbf{p}_1 = 1,\ \mathbf{p}_2 = x,\ \mathbf{p}_3 = x^2$

(b) $\mathbf{p} = 2 - x + x^2;\ \mathbf{p}_1 = 1 + x,\ \mathbf{p}_2 = 1 + x^2,\ \mathbf{p}_3 = x + x^2$

4.5 Dimension

Number of Vectors in a Basis

THEOREM 4.5.1 *All bases for a finite-dimensional vector space have the same number of vectors.*

THEOREM 4.5.2 *Let V be an n-dimensional vector space, and let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be any basis.*

(*a*) *If a set in V has more than n vectors, then it is linearly dependent.*

(*b*) *If a set in V has fewer than n vectors, then it does not span V.*

DEFINITION 1 The ***dimension*** of a finite-dimensional vector space V is denoted by $\dim(V)$ and is defined to be the number of vectors in a basis for V. In addition, the zero vector space is defined to have dimension zero.

▶ EXAMPLE 1 Dimensions of Some Familiar Vector Spaces

$$\dim(R^n) = n \qquad \text{[The standard basis has } n \text{ vectors.]}$$

$$\dim(P_n) = n + 1 \qquad \text{[The standard basis has } n+1 \text{ vectors.]}$$

$$\dim(M_{mn}) = mn \qquad \text{[The standard basis has } mn \text{ vectors.]}$$

▶ EXAMPLE 2 Dimension of Span(S)

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ then every vector in span(S) is expressible as a linear combination of the vectors in S. Thus, if the vectors in S are *linearly independent*, they automatically form a basis for span(S), from which we can conclude that

$$\dim[\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}] = r$$

In words, the dimension of the space spanned by a linearly independent set of vectors is equal to the number of vectors in that set.

▶ EXAMPLE 3 Dimension of a Solution Space

Find a basis for and the dimension of the solution space of the homogeneous system

$$\begin{aligned}
x_1 + 3x_2 - 2x_3 \phantom{{}- 2x_4} + 2x_5 \phantom{{}- 3x_6} &= 0 \\
2x_1 + 6x_2 - 5x_3 - 2x_4 + 4x_5 - 3x_6 &= 0 \\
5x_3 + 10x_4 \phantom{{}+ 4x_5} + 15x_6 &= 0 \\
2x_1 + 6x_2 \phantom{{}- 5x_3} + 8x_4 + 4x_5 + 18x_6 &= 0
\end{aligned}$$

Solution In Example 6 of Section 1.2 we found the solution of this system to be

$$x_1 = -3r - 4s - 2t, \quad x_2 = r, \quad x_3 = -2s, \quad x_4 = s, \quad x_5 = t, \quad x_6 = 0$$

which can be written in vector form as

$$(x_1, x_2, x_3, x_4, x_5, x_6) = (-3r - 4s - 2t, r, -2s, s, t, 0)$$

or, alternatively, as

$$(x_1, x_2, x_3, x_4, x_5, x_6) = r(-3, 1, 0, 0, 0, 0) + s(-4, 0, -2, 1, 0, 0) + t(-2, 0, 0, 0, 1, 0)$$

This shows that the vectors

$$\mathbf{v}_1 = (-3, 1, 0, 0, 0, 0), \quad \mathbf{v}_2 = (-4, 0, -2, 1, 0, 0), \quad \mathbf{v}_3 = (-2, 0, 0, 0, 1, 0)$$

span the solution space. We leave it for you to check that these vectors are linearly independent by showing that none of them is a linear combination of the other two (but see the remark that follows). Thus, the solution space has dimension 3. ◀

THEOREM 4.5.3 Plus/Minus Theorem

Let S be a nonempty set of vectors in a vector space V.

(*a*) *If S is a linearly independent set, and if $\mathbf{v}$ is a vector in V that is outside of span(S), then the set $S \cup \{\mathbf{v}\}$ that results by inserting $\mathbf{v}$ into S is still linearly independent.*

(*b*) *If $\mathbf{v}$ is a vector in S that is expressible as a linear combination of other vectors in S, and if $S - \{\mathbf{v}\}$ denotes the set obtained by removing $\mathbf{v}$ from S, then S and $S - \{\mathbf{v}\}$ span the same space; that is,*

$$\text{span}(S) = \text{span}(S - \{\mathbf{v}\})$$

▶ EXAMPLE 4 Applying the Plus/Minus Theorem

Show that $\mathbf{p}_1 = 1 - x^2$, $\mathbf{p}_2 = 2 - x^2$, and $\mathbf{p}_3 = x^3$ are linearly independent vectors.

Solution The set $S = \{\mathbf{p}_1, \mathbf{p}_2\}$ is linearly independent since neither vector in S is a scalar multiple of the other. Since the vector $\mathbf{p}_3$ cannot be expressed as a linear combination of the vectors in S (why?), it can be adjoined to S to produce a linearly independent set $S \cup \{\mathbf{p}_3\} = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$. ◀

THEOREM 4.5.4 *Let V be an n-dimensional vector space, and let S be a set in V with exactly n vectors. Then S is a basis for V if and only if S spans V or S is linearly independent.*

▶ EXAMPLE 5 Bases by Inspection

(a) Explain why the vectors $\mathbf{v}_1 = (-3, 7)$ and $\mathbf{v}_2 = (5, 5)$ form a basis for R^2.

(b) Explain why the vectors $\mathbf{v}_1 = (2, 0, -1)$, $\mathbf{v}_2 = (4, 0, 7)$, and $\mathbf{v}_3 = (-1, 1, 4)$ form a basis for R^3.

Solution (a) Since neither vector is a scalar multiple of the other, the two vectors form a linearly independent set in the two-dimensional space R^2, and hence they form a basis by Theorem 4.5.4.

Solution (b) The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a linearly independent set in the xz-plane (why?). The vector $\mathbf{v}_3$ is outside of the xz-plane, so the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also linearly independent. Since R^3 is three-dimensional, Theorem 4.5.4 implies that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for the vector space R^3. ◀

THEOREM 4.5.5 *Let S be a finite set of vectors in a finite-dimensional vector space V.*

(*a*) *If S spans V but is not a basis for V, then S can be reduced to a basis for V by removing appropriate vectors from S.*

(*b*) *If S is a linearly independent set that is not already a basis for V, then S can be enlarged to a basis for V by inserting appropriate vectors into S.*

THEOREM 4.5.6 *If W is a subspace of a finite-dimensional vector space V, then:*

(*a*) *W is finite-dimensional.*

(*b*) $\dim(W) \leq \dim(V)$.

(*c*) $W = V$ *if and only if* $\dim(W) = \dim(V)$.

Exercise Set 4.5

In Exercises 1–6, find a basis for the solution space of the homogeneous linear system, and find the dimension of that space.

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

7. In each part, find a basis for the given subspace of R^3, and state its dimension.

(a) The plane $3x - 2y + 5z = 0$.

(b) The plane $x - y = 0$.

(c) The line $x = 2t,\ y = -t,\ z = 4t$.

(d) All vectors of the form (a, b, c), where $b = a + c$.

11. (a) Show that the set W of all polynomials in P_2 such that $p(1) = 0$ is a subspace of P_2.

(b) Make a conjecture about the dimension of W.

(c) Confirm your conjecture by finding a basis for W.

14. Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a basis for a vector space V. Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is also a basis, where $\mathbf{u}_1 = \mathbf{v}_1$, $\mathbf{u}_2 = \mathbf{v}_1 + \mathbf{v}_2$, and $\mathbf{u}_3 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$.

18. Find a basis for the subspace of R^4 that is spanned by the vectors

$$\mathbf{v}_1 = (1, 1, 1, 1), \quad \mathbf{v}_2 = (2, 2, 2, 0), \quad \mathbf{v}_3 = (0, 0, 0, 3),$$
$$\mathbf{v}_4 = (3, 3, 3, 4)$$

4.6 Change of Basis

Coordinate Maps

If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a finite-dimensional vector space V, and if

$$(\mathbf{v})_S = (c_1, c_2, \ldots, c_n)$$

is the coordinate vector of $\mathbf{v}$ relative to S, then, as illustrated in Figure 4.4.6, the mapping

$$\mathbf{v} \rightarrow (\mathbf{v})_S \tag{1}$$

creates a connection (a one-to-one correspondence) between vectors in the *general* vector space V and vectors in the *Euclidean* vector space R^n. We call (1) the ***coordinate map relative to S*** from V to R^n. In this section we will find it convenient to express coordinate vectors in the matrix form

$$[\mathbf{v}]_S = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{2}$$

where the square brackets emphasize the matrix notation (Figure 4.6.1).

Change of Basis

> **The Change-of-Basis Problem** If $\mathbf{v}$ is a vector in a finite-dimensional vector space V, and if we change the basis for V from a basis B to a basis B', how are the coordinate vectors $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ related?

For simplicity, we will solve this problem for two-dimensional spaces. The solution for n-dimensional spaces is similar. Let

$$B = \{\mathbf{u}_1, \mathbf{u}_2\} \quad \text{and} \quad B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$$

be the old and new bases, respectively. We will need the coordinate vectors for the new basis vectors relative to the old basis. Suppose they are

$$[\mathbf{u}'_1]_B = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad [\mathbf{u}'_2]_B = \begin{bmatrix} c \\ d \end{bmatrix} \tag{3}$$

That is,

$$\begin{aligned} \mathbf{u}'_1 &= a\mathbf{u}_1 + b\mathbf{u}_2 \\ \mathbf{u}'_2 &= c\mathbf{u}_1 + d\mathbf{u}_2 \end{aligned} \tag{4}$$

Now let $\mathbf{v}$ be any vector in V, and let

$$[\mathbf{v}]_{B'} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \tag{5}$$

be the new coordinate vector, so that

$$\mathbf{v} = k_1\mathbf{u}'_1 + k_2\mathbf{u}'_2 \tag{6}$$

In order to find the old coordinates of $\mathbf{v}$, we must express $\mathbf{v}$ in terms of the old basis B. To do this, we substitute (4) into (6). This yields

$$\mathbf{v} = k_1(a\mathbf{u}_1 + b\mathbf{u}_2) + k_2(c\mathbf{u}_1 + d\mathbf{u}_2)$$

or

$$\mathbf{v} = (k_1a + k_2c)\mathbf{u}_1 + (k_1b + k_2d)\mathbf{u}_2$$

Thus, the old coordinate vector for $\mathbf{v}$ is

$$[\mathbf{v}]_B = \begin{bmatrix} k_1a + k_2c \\ k_1b + k_2d \end{bmatrix}$$

which, by using (5), can be written as

$$[\mathbf{v}]_B = \begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}[\mathbf{v}]_{B'}$$

This equation states that the old coordinate vector $[\mathbf{v}]_B$ results when we multiply the new coordinate vector $[\mathbf{v}]_{B'}$ on the left by the matrix

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

Since the columns of this matrix are the coordinates of the new basis vectors relative to the old basis [see (3)], we have the following solution of the change-of-basis problem.

Solution of the Change-of-Basis Problem If we change the basis for a vector space V from an old basis $B = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ to a new basis $B' = \{\mathbf{u}'_1, \mathbf{u}'_2, \dots, \mathbf{u}'_n\}$, then for each vector $\mathbf{v}$ in V, the old coordinate vector $[\mathbf{v}]_B$ is related to the new coordinate vector $[\mathbf{v}]_{B'}$ by the equation

$$[\mathbf{v}]_B = P[\mathbf{v}]_{B'} \tag{7}$$

where the columns of P are the coordinate vectors of the new basis vectors relative to the old basis; that is, the column vectors of P are

$$[\mathbf{u}'_1]_B, \quad [\mathbf{u}'_2]_B, \dots, \quad [\mathbf{u}'_n]_B \tag{8}$$

Transition Matrices

The matrix P in Equation (7) is called the ***transition matrix*** from B' to B. For emphasis, we will often denote it by $P_{B' \to B}$. It follows from (8) that this matrix can be expressed in terms of its column vectors as

$$P_{B' \to B} = \left[[\mathbf{u}'_1]_B \mid [\mathbf{u}'_2]_B \mid \cdots \mid [\mathbf{u}'_n]_B\right] \tag{9}$$

Similarly, the transition matrix from B to B' can be expressed in terms of its column vectors as

$$P_{B \to B'} = \left[[\mathbf{u}_1]_{B'} \mid [\mathbf{u}_2]_{B'} \mid \cdots \mid [\mathbf{u}_n]_{B'}\right] \tag{10}$$

The columns of the transition matrix from an old basis to a new basis are the coordinate vectors of the old basis relative to the new basis.

► EXAMPLE 1 Finding Transition Matrices

Consider the bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$ for R^2, where

$$\mathbf{u}_1 = (1, 0), \quad \mathbf{u}_2 = (0, 1), \quad \mathbf{u}'_1 = (1, 1), \quad \mathbf{u}'_2 = (2, 1)$$

(a) Find the transition matrix $P_{B' \to B}$ from B' to B.

(b) Find the transition matrix $P_{B \to B'}$ from B to B'.

Solution (a) Here the old basis vectors are $\mathbf{u}'_1$ and $\mathbf{u}'_2$ and the new basis vectors are $\mathbf{u}_1$ and $\mathbf{u}_2$. We want to find the coordinate matrices of the old basis vectors $\mathbf{u}'_1$ and $\mathbf{u}'_2$ relative to the new basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. To do this, observe that

$$\begin{aligned} \mathbf{u}'_1 &= \mathbf{u}_1 + \mathbf{u}_2 \\ \mathbf{u}'_2 &= 2\mathbf{u}_1 + \mathbf{u}_2 \end{aligned}$$

from which it follows that

$$[\mathbf{u}'_1]_B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{u}'_2]_B = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

and hence that

$$P_{B' \to B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Solution (b) Here the old basis vectors are $\mathbf{u}_1$ and $\mathbf{u}_2$ and the new basis vectors are $\mathbf{u}'_1$ and $\mathbf{u}'_2$. As in part (a), we want to find the coordinate matrices of the old basis vectors $\mathbf{u}'_1$ and $\mathbf{u}'_2$ relative to the new basis vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. To do this, observe that

$$\begin{aligned} \mathbf{u}_1 &= -\mathbf{u}'_1 + \mathbf{u}'_2 \\ \mathbf{u}_2 &= 2\mathbf{u}'_1 - \mathbf{u}'_2 \end{aligned}$$

from which it follows that

$$[\mathbf{u}_1]_{B'} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{and} \quad [\mathbf{u}_2]_{B'} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

and hence that

$$P_{B \to B'} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \quad ◄$$

Suppose now that B and B' are bases for a finite-dimensional vector space V. Since multiplication by $P_{B'\to B}$ maps coordinate vectors relative to the basis B' into coordinate vectors relative to a basis B, and $P_{B\to B'}$ maps coordinate vectors relative to B into coordinate vectors relative to B', it follows that for every vector $\mathbf{v}$ in V we have

$$[\mathbf{v}]_B = P_{B'\to B}[\mathbf{v}]_{B'} \tag{11}$$

$$[\mathbf{v}]_{B'} = P_{B\to B'}[\mathbf{v}]_B \tag{12}$$

▶ EXAMPLE 2 Computing Coordinate Vectors

Let B and B' be the bases in Example 1. Use an appropriate formula to find $[\mathbf{v}]_B$ given that

$$[\mathbf{v}]_{B'} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

Solution To find $[\mathbf{v}]_B$ we need to make the transition from B' to B. It follows from Formula (11) and part (a) of Example 1 that

$$[\mathbf{v}]_B = P_{B'\to B}[\mathbf{v}]_{B'} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix} \blacktriangleleft$$

Invertibility of Transition Matrices

If B and B' are bases for a finite-dimensional vector space V, then

$$(P_{B'\to B})(P_{B\to B'}) = P_{B\to B}$$

because multiplication by the product $(P_{B'\to B})(P_{B\to B'})$ first maps the B-coordinates of a vector into its B'-coordinates, and then maps those B'-coordinates back into the original B-coordinates. Since the net effect of the two operations is to leave each coordinate vector unchanged, we are led to conclude that $P_{B\to B}$ must be the identity matrix, that is,

$$(P_{B'\to B})(P_{B\to B'}) = I \tag{13}$$

THEOREM 4.6.1 *If P is the transition matrix from a basis B' to a basis B for a finite-dimensional vector space V, then P is invertible and P^{-1} is the transition matrix from B to B'.*

An Efficient Method for Computing Transition Matrices for R^n

A Procedure for Computing $P_{B\to B'}$

Step 1. Form the matrix $[B' \mid B]$.

Step 2. Use elementary row operations to reduce the matrix in Step 1 to reduced row echelon form.

Step 3. The resulting matrix will be $[I \mid P_{B\to B'}]$.

Step 4. Extract the matrix $P_{B\to B'}$ from the right side of the matrix in Step 3.

This procedure is captured in the following diagram.

$$[\text{new basis} \mid \text{old basis}] \xrightarrow{\text{row operations}} [I \mid \text{transition from old to new}] \tag{14}$$

▶ EXAMPLE 3 Example 1 Revisited

In Example 1 we considered the bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$ for R^2, where

$$\mathbf{u}_1 = (1, 0), \quad \mathbf{u}_2 = (0, 1), \quad \mathbf{u}'_1 = (1, 1), \quad \mathbf{u}'_2 = (2, 1)$$

(a) Use Formula (14) to find the transition matrix from B' to B.

(b) Use Formula (14) to find the transition matrix from B to B'.

Solution (a) Here B' is the old basis and B is the new basis, so

$$[\text{new basis} \mid \text{old basis}] = \left[\begin{array}{cc|cc} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right]$$

Since the left side is already the identity matrix, no reduction is needed. We see by inspection that the transition matrix is

$$P_{B'\to B} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$$

which agrees with the result in Example 1.

Solution (b) Here B is the old basis and B' is the new basis, so

$$[\text{new basis} \mid \text{old basis}] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right]$$

By reducing this matrix, so the left side becomes the identity, we obtain (verify)

$$[I \mid \text{transition from old to new}] = \left[\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right]$$

so the transition matrix is

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

which also agrees with the result in Example 1. ◀

Exercise Set 4.6

1. Consider the bases $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2\}$ for R^2, where

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \quad \mathbf{u}'_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \mathbf{u}'_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

(a) Find the transition matrix from B' to B.

(b) Find the transition matrix from B to B'.

(c) Compute the coordinate vector $[\mathbf{w}]_B$, where

$$\mathbf{w} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$$

and use (12) to compute $[\mathbf{w}]_{B'}$.

(d) Check your work by computing $[\mathbf{w}]_{B'}$ directly.

3. Consider the bases $B = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ and $B' = \{\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3\}$ for R^3, where

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{u}'_1 = \begin{bmatrix} 3 \\ 1 \\ -5 \end{bmatrix}, \quad \mathbf{u}'_2 = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}, \quad \mathbf{u}'_3 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

(a) Find the transition matrix B to B'.

(b) Compute the coordinate vector $[\mathbf{w}]_B$, where

$$\mathbf{w} = \begin{bmatrix} -5 \\ 8 \\ -5 \end{bmatrix}$$

and use (12) to compute $[\mathbf{w}]_{B'}$.

(c) Check your work by computing $[\mathbf{w}]_{B'}$ directly.

6. Consider the bases $B = \{\mathbf{p}_1, \mathbf{p}_2\}$ and $B' = \{\mathbf{q}_1, \mathbf{q}_2\}$ for P_1, where

$$\mathbf{p}_1 = 6 + 3x, \quad \mathbf{p}_2 = 10 + 2x, \quad \mathbf{q}_1 = 2, \quad \mathbf{q}_2 = 3 + 2x$$

(a) Find the transition matrix from B' to B.

(b) Find the transition matrix from B to B'.

(c) Compute the coordinate vector $[\mathbf{p}]_B$, where $\mathbf{p} = -4 + x$, and use (12) to compute $[\mathbf{p}]_{B'}$.

(d) Check your work by computing $[\mathbf{p}]_{B'}$ directly.

9. Let S be the standard basis for R^3, and let $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be the basis in which $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 5, 0)$, and $\mathbf{v}_3 = (3, 3, 8)$.

(a) Find the transition matrix $P_{B \to S}$ by inspection.

(b) Use Formula (14) to find the transition matrix $P_{S \to B}$.

(c) Confirm that $P_{B \to S}$ and $P_{S \to B}$ are inverses of one another.

(d) Let $\mathbf{w} = (5, -3, 1)$. Find $[\mathbf{w}]_B$ and then use Formula (11) to compute $[\mathbf{w}]_S$.

(e) Let $\mathbf{w} = (3, -5, 0)$. Find $[\mathbf{w}]_S$ and then use Formula (12) to compute $[\mathbf{w}]_B$.

4.7 Row Space, Column Space, and Null Space

Row Space, Column Space, and Null Space

DEFINITION 1 For an $m \times n$ matrix

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

the vectors

$$\begin{aligned} \mathbf{r}_1 &= [a_{11} \quad a_{12} \quad \cdots \quad a_{1n}] \\ \mathbf{r}_2 &= [a_{21} \quad a_{22} \quad \cdots \quad a_{2n}] \\ &\vdots \\ \mathbf{r}_m &= [a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}] \end{aligned}$$

in R^n that are formed from the rows of A are called the ***row vectors*** of A, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, \quad \mathbf{c}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

in R^m formed from the columns of A are called the ***column vectors*** of A.

EXAMPLE 1 Row and Column Vectors of a 2 × 3 Matrix

Let

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

The row vectors of A are

$$\mathbf{r}_1 = [2 \quad 1 \quad 0] \quad \text{and} \quad \mathbf{r}_2 = [3 \quad -1 \quad 4]$$

and the column vectors of A are

$$\mathbf{c}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \text{and} \quad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$$ ◀

DEFINITION 2 If A is an $m \times n$ matrix, then the subspace of R^n spanned by the row vectors of A is called the ***row space*** of A, and the subspace of R^m spanned by the column vectors of A is called the ***column space*** of A. The solution space of the homogeneous system of equations $A\mathbf{x} = \mathbf{0}$, which is a subspace of R^n, is called the ***null space*** of A.

THEOREM 4.7.1 *A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.*

THEOREM 4.7.2 *If $\mathbf{x}_0$ is any solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$, and if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for the null space of A, then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed in the form*

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \tag{3}$$

Conversely, for all choices of scalars $c_1, c_2, \ldots, c_k$, the vector $\mathbf{x}$ in this formula is a solution of $A\mathbf{x} = \mathbf{b}$.

The vector $\mathbf{x}_0$ in Formula (3) is called a ***particular solution of*** $A\mathbf{x} = \mathbf{b}$, and the remaining part of the formula is called the ***general solution of*** $A\mathbf{x} = \mathbf{0}$. With this terminology Theorem 4.7.2 can be rephrased as:

The general solution of a consistent linear system can be expressed as the sum of a particular solution of that system and the general solution of the corresponding homogeneous system.

EXAMPLE 3 General Solution of a Linear System $A\mathbf{x} = \mathbf{b}$

In the concluding subsection of Section 3.4 we compared solutions of the linear systems

$$\begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 5 \\ 6 \end{bmatrix}$$

and deduced that the general solution $\mathbf{x}$ of the nonhomogeneous system and the general solution $\mathbf{x}_h$ of the corresponding homogeneous system (when written in column-vector form) are related by

$$\underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}}_{\mathbf{x}} = \begin{bmatrix} -3r - 4s - 2t \\ r \\ -2s \\ s \\ t \\ \frac{1}{3} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \end{bmatrix}}_{\mathbf{x}_0} + \underbrace{r\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -4 \\ 0 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{x}_h}$$

Bases for Row Spaces, Column Spaces, and Null Spaces

THEOREM 4.7.3 *Elementary row operations do not change the null space of a matrix.*

THEOREM 4.7.4 *Elementary row operations do not change the row space of a matrix.*

EXAMPLE 4 Finding a Basis for the Null Space of a Matrix

Find a basis for the null space of the matrix

$$A = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 8 & 4 & 18 \end{bmatrix}$$

Solution The null space of A is the solution space of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, which, as shown in Example 3, has the basis

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

THEOREM 4.7.5 *If a matrix R is in row echelon form, then the row vectors with the leading 1's (the nonzero row vectors) form a basis for the row space of R, and the column vectors with the leading 1's of the row vectors form a basis for the column space of R.*

EXAMPLE 5 Bases for the Row and Column Spaces of a Matrix in Row Echelon Form

Find bases for the row and column spaces of the matrix

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

Solution Since the matrix R is in row echelon form, it follows from Theorem 4.7.5 that the vectors

$$\mathbf{r}_1 = [1 \quad -2 \quad 5 \quad 0 \quad 3]$$
$$\mathbf{r}_2 = [0 \quad 1 \quad 3 \quad 0 \quad 0]$$
$$\mathbf{r}_3 = [0 \quad 0 \quad 0 \quad 1 \quad 0]$$

form a basis for the row space of R, and the vectors

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R.

EXAMPLE 6 Basis for a Row Space by Row Reduction

Find a basis for the row space of the matrix

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

Solution Since elementary row operations do not change the row space of a matrix, we can find a basis for the row space of A by finding a basis for the row space of any row echelon form of A. Reducing A to row echelon form, we obtain (verify)

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

By Theorem 4.7.5, the nonzero row vectors of R form a basis for the row space of R and hence form a basis for the row space of A. These basis vectors are

$$\begin{aligned} \mathbf{r}_1 &= [1 \quad -3 \quad 4 \quad -2 \quad 5 \quad 4] \\ \mathbf{r}_2 &= [0 \quad 0 \quad 1 \quad 3 \quad -2 \quad -6] \\ \mathbf{r}_3 &= [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 5] \end{aligned}$$

Basis for the Column Space of a Matrix

THEOREM 4.7.6 *If A and B are row equivalent matrices, then:*

(*a*) *A given set of column vectors of A is linearly independent if and only if the corresponding column vectors of B are linearly independent.*

(*b*) *A given set of column vectors of A forms a basis for the column space of A if and only if the corresponding column vectors of B form a basis for the column space of B.*

EXAMPLE 7 Basis for a Column Space by Row Reduction

Find a basis for the column space of the matrix

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

that consists of column vectors of A.

Solution We observed in Example 6 that the matrix

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

is a row echelon form of A. Keeping in mind that A and R can have different column spaces, we cannot find a basis for the column space of A directly from the column vectors of R. However, it follows from Theorem 4.7.6(*b*) that if we can find a set of column vectors of R that forms a basis for the column space of R, then the *corresponding* column vectors of A will form a basis for the column space of A.

Since the first, third, and fifth columns of R contain the leading 1's of the row vectors the vectors

$$\mathbf{c}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_3' = \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_5' = \begin{bmatrix} 5 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

form a basis for the column space of R. Thus, the corresponding column vectors of A, which are

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} 4 \\ 9 \\ 9 \\ -4 \end{bmatrix}, \quad \mathbf{c}_5 = \begin{bmatrix} 5 \\ 8 \\ 9 \\ -5 \end{bmatrix}$$

form a basis for the column space of A. ◀

▶ EXAMPLE 8 Basis for the Space Spanned by a Set of Vectors

The following vectors span a subspace of R^4. Find a subset of these vectors that forms a basis of this subspace.

$$\mathbf{v}_1 = (1, 2, 2, -1), \quad \mathbf{v}_2 = (-3, -6, -6, 3),$$
$$\mathbf{v}_3 = (4, 9, 9, -4), \quad \mathbf{v}_4 = (-2, -1, -1, 2),$$
$$\mathbf{v}_5 = (5, 8, 9, -5), \quad \mathbf{v}_6 = (4, 2, 7, -4)$$

Solution If we rewrite these vectors in column form and construct the matrix that has those vectors as its successive columns, then we obtain the matrix A in Example 7 (verify). Thus,

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\} = \text{col}(A)$$

Proceeding as in that example (and adjusting the notation appropriately), we see that the vectors $\mathbf{v}_1$, $\mathbf{v}_3$, and $\mathbf{v}_5$ form a basis for

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5, \mathbf{v}_6\} \quad ◀$$

> **Problem** Given a set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in R^n, find a subset of these vectors that forms a basis for span(S), and express each vector that is not in that basis as a linear combination of the basis vectors.

▶ EXAMPLE 10 Basis and Linear Combinations

(a) Find a subset of the vectors

$$\mathbf{v}_1 = (1, -2, 0, 3), \quad \mathbf{v}_2 = (2, -5, -3, 6),$$
$$\mathbf{v}_3 = (0, 1, 3, 0), \quad \mathbf{v}_4 = (2, -1, 4, -7), \quad \mathbf{v}_5 = (5, -8, 1, 2)$$

that forms a basis for the subspace of R^4 spanned by these vectors.

(b) Express each vector not in the basis as a linear combination of the basis vectors.

Solution (a) We begin by constructing a matrix that has $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_5$ as its column vectors:

$$\begin{bmatrix} 1 & 2 & 0 & 2 & 5 \\ -2 & -5 & 1 & -1 & -8 \\ 0 & -3 & 3 & 4 & 1 \\ 3 & 6 & 0 & -7 & 2 \end{bmatrix} \tag{5}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 & \mathbf{v}_5 \end{matrix}$$

The first part of our problem can be solved by finding a basis for the column space of this matrix. Reducing the matrix to *reduced* row echelon form and denoting the column vectors of the resulting matrix by $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4$, and $\mathbf{w}_5$ yields

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{6}$$
$$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 & \mathbf{w}_4 & \mathbf{w}_5 \end{matrix}$$

The leading 1's occur in columns 1, 2, and 4, so by Theorem 4.7.5,

$$\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4\}$$

is a basis for the column space of (6), and consequently,

$$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$$

is a basis for the column space of (5).

Solution (b) We will start by expressing $\mathbf{w}_3$ and $\mathbf{w}_5$ as linear combinations of the basis vectors $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_4$. The simplest way of doing this is to express $\mathbf{w}_3$ and $\mathbf{w}_5$ in terms of basis vectors with smaller subscripts. Accordingly, we will express $\mathbf{w}_3$ as a linear combination of $\mathbf{w}_1$ and $\mathbf{w}_2$, and we will express $\mathbf{w}_5$ as a linear combination of $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_4$. By inspection of (6), these linear combinations are

$$\mathbf{w}_3 = 2\mathbf{w}_1 - \mathbf{w}_2$$
$$\mathbf{w}_5 = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_4$$

We call these the ***dependency equations***. The corresponding relationships in (5) are

$$\mathbf{v}_3 = 2\mathbf{v}_1 - \mathbf{v}_2$$
$$\mathbf{v}_5 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_4 \quad \blacktriangleleft$$

The following is a summary of the steps that we followed in our last example to solve the problem posed above.

Basis for the Space Spanned by a Set of Vectors

Step 1. Form the matrix A whose columns are the vectors in the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

Step 2. Reduce the matrix A to reduced row echelon form R.

Step 3. Denote the column vectors of R by $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$.

Step 4. Identify the columns of R that contain the leading 1's. The corresponding column vectors of A form a basis for span(S).

This completes the first part of the problem.

Step 5. Obtain a set of dependency equations for the column vectors $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k$ of R by successively expressing each $\mathbf{w}_i$ that does not contain a leading 1 of R as a linear combination of predecessors that do.

Step 6. In each dependency equation obtained in Step 5, replace the vector $\mathbf{w}_i$ by the vector $\mathbf{v}_i$ for $i = 1, 2, \ldots, k$.

This completes the second part of the problem.

Exercise Set 4.7

▶ In Exercises 7–8, find the vector form of the general solution of the linear system $A\mathbf{x} = \mathbf{b}$, and then use that result to find the vector form of the general solution of $A\mathbf{x} = \mathbf{0}$. ◀

(b)
$$\begin{aligned} x_1 + 2x_2 - 3x_3 + x_4 &= 4 \\ -2x_1 + x_2 + 2x_3 + x_4 &= -1 \\ -x_1 + 3x_2 - x_3 + 2x_4 &= 3 \\ 4x_1 - 7x_2 \qquad\quad - 5x_4 &= -5 \end{aligned}$$

13. (a) Use the methods of Examples 6 and 7 to find bases for the row space and column space of the matrix

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

(b) Use the method of Example 9 to find a basis for the row space of A that consists entirely of row vectors of A.

▶ In Exercises 14–15, find a basis for the subspace of R^4 that is spanned by the given vectors. ◀

14. $(1, 1, -4, -3),\ (2, 0, 2, -2),\ (2, -1, 3, 2)$

▶ In Exericses 16–17, find a subset of the given vectors that forms a basis for the space spanned by those vectors, and then express each vector that is not in the basis as a linear combination of the basis vectors. ◀

16. $\mathbf{v}_1 = (1, 0, 1, 1),\ \mathbf{v}_2 = (-3, 3, 7, 1),$
$\mathbf{v}_3 = (-1, 3, 9, 3),\ \mathbf{v}_4 = (-5, 3, 5, -1)$

4.8 Rank, Nullity,

Row and Column Spaces Have Equal Dimensions

In Examples 6 and 7 of Section 4.7 we found that the row and column spaces of the matrix

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

both have three basis vectors and hence are both three-dimensional. The fact that these spaces have the same dimension is not accidental, but rather a consequence of the following theorem.

THEOREM 4.8.1 *The row space and the column space of a matrix A have the same dimension.*

Rank and Nullity

The dimensions of the row space, column space, and null space of a matrix are such important numbers that there is some notation and terminology associated with them.

DEFINITION 1 The common dimension of the row space and column space of a matrix A is called the ***rank*** of A and is denoted by $\text{rank}(A)$; the dimension of the null space of A is called the ***nullity*** of A and is denoted by $\text{nullity}(A)$.

EXAMPLE 1 Rank and Nullity of a 4 × 6 Matrix

Find the rank and nullity of the matrix

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

Solution The reduced row echelon form of A is

$$\begin{bmatrix} 1 & 0 & -4 & -28 & -37 & 13 \\ 0 & 1 & -2 & -12 & -16 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{1}$$

(verify). Since this matrix has two leading 1's, its row and column spaces are two-dimensional and $\text{rank}(A) = 2$. To find the nullity of A, we must find the dimension of the solution space of the linear system $A\mathbf{x} = \mathbf{0}$. This system can be solved by reducing its augmented matrix to reduced row echelon form. The resulting matrix will be identical to (1), except that it will have an additional last column of zeros, and hence the corresponding system of equations will be

$$\begin{aligned} x_1 - 4x_3 - 28x_4 - 37x_5 + 13x_6 &= 0 \\ x_2 - 2x_3 - 12x_4 - 16x_5 + 5x_6 &= 0 \end{aligned}$$

Solving these equations for the leading variables yields

$$\begin{aligned} x_1 &= 4x_3 + 28x_4 + 37x_5 - 13x_6 \\ x_2 &= 2x_3 + 12x_4 + 16x_5 - 5x_6 \end{aligned} \tag{2}$$

from which we obtain the general solution

$$\begin{aligned} x_1 &= 4r + 28s + 37t - 13u \\ x_2 &= 2r + 12s + 16t - 5u \\ x_3 &= r \\ x_4 &= s \\ x_5 &= t \\ x_6 &= u \end{aligned}$$

or in column vector form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = r\begin{bmatrix} 4 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 28 \\ 12 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 37 \\ 16 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} -13 \\ -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

Because the four vectors on the right side of (3) form a basis for the solution space, $\text{nullity}(A) = 4$.

▶ EXAMPLE 2 Maximum Value for Rank

What is the maximum possible rank of an $m \times n$ matrix A that is not square?

Solution Since the row vectors of A lie in R^n and the column vectors in R^m, the row space of A is at most n-dimensional and the column space is at most m-dimensional. Since the rank of A is the common dimension of its row and column space, it follows that the rank is at most the smaller of m and n. We denote this by writing

$$\text{rank}(A) \leq \min(m, n)$$

in which $\min(m, n)$ is the minimum of m and n. ◀

THEOREM 4.8.2 Dimension Theorem for Matrices

If A is a matrix with n columns, then

$$\textit{rank}(A) + \textit{nullity}(A) = n \tag{4}$$

▶ EXAMPLE 3 The Sum of Rank and Nullity

The matrix

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

has 6 columns, so

$$\text{rank}(A) + \text{nullity}(A) = 6$$

This is consistent with Example 1, where we showed that

$$\text{rank}(A) = 2 \quad \text{and} \quad \text{nullity}(A) = 4$$ ◀

THEOREM 4.8.8 Equivalent Statements

If A is an $n \times n$ matrix, then the following statements are equivalent.

(*a*) A *is invertible.*

(*b*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*c*) *The reduced row echelon form of* A *is* I_n.

(*d*) A *is expressible as a product of elementary matrices.*

(*e*) $A\mathbf{x} = \mathbf{b}$ *is consistent for every* $n \times 1$ *matrix* $\mathbf{b}$.

(*f*) $A\mathbf{x} = \mathbf{b}$ *has exactly one solution for every* $n \times 1$ *matrix* $\mathbf{b}$.

(*g*) $\det(A) \neq 0$.

(*h*) *The column vectors of* A *are linearly independent.*

(*i*) *The row vectors of* A *are linearly independent.*

(*j*) *The column vectors of* A *span* R^n.

(*k*) *The row vectors of* A *span* R^n.

(*l*) *The column vectors of* A *form a basis for* R^n.

(*m*) *The row vectors of* A *form a basis for* R^n.

(*n*) A *has rank* n.

(*o*) A *has nullity* 0.

In Exercises 1–2, find the rank and nullity of the matrix A by reducing it to row echelon form.

1. (a) $A = \begin{bmatrix} 1 & 2 & -1 & 1 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -3 & 3 \\ 4 & 8 & -4 & 4 \end{bmatrix}$

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

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

(b) $A = \begin{bmatrix} 1 & 3 & 1 & 3 \\ 0 & 1 & 1 & 0 \\ -3 & 0 & 6 & -1 \\ 3 & 4 & -2 & 1 \\ 2 & 0 & -4 & -2 \end{bmatrix}$