# 1 Matrices and Determinants

Beginning with your first algebra course you have encountered problems such as the following:

> *A boat traveling at a constant speed in a river with a constant current speed can travel 48 miles downstream in 4 hours. The same trip upstream takes 6 hours. What is the speed of the boat in still water and what is the speed of the current?*

Here is one way we can solve this problem: Let $x$ be the speed of the boat in still water and $y$ be the speed of the current. Since the speed of the boat going downstream is $x + y$, we have

$$4(x + y) = 48 \quad \text{or} \quad x + y = 12.$$

Since the speed of the boat going upstream is $x - y$,

$$6(x - y) = 48 \quad \text{or} \quad x - y = 8.$$

Thus we can determine the speed of the boat in still water and the speed of the current by solving the system of equations:

$$\begin{aligned} x + y &= 12 \\ x - y &= 8. \end{aligned}$$

Doing so (try it), we find the speed of the boat in still water is $x = 10$ miles per hour and the speed of the current is $y = 2$ miles per hour.

The system of equations we have just considered is an example of a system of linear equations, and you have encountered many such linear systems over the years. Of course, you probably have come to realize that the larger the system—that is, the more variables and/or equations in the system—the more difficult it often is to solve the system. For instance, suppose we needed to find the partial fraction decomposition of

$$\frac{1}{(x^2+1)(x^2+4)},$$

which, as you saw in calculus, is used to integrate this expression. (In our study of the Laplace transform in Chapter 7, we will see another place where finding partial fraction decompositions of expressions such as this arises.) This partial fraction decomposition has the form

$$\frac{1}{(x^2+1)(x^2+4)} = \frac{Ax+B}{x^2+1} + \frac{Cx+D}{x^2+4},$$

and finding it involves solving a system of four linear equations in the four unknowns $A$, $B$, $C$, and $D$, which takes more time and effort to solve than the problem with the boat. There is no limit to the size of linear systems that arise in practice. It is not unheard of to encounter systems of linear equations with tens, hundreds, or even thousands of unknowns and equations.

The larger the linear system, the easier it is to get lost in your work if you are not careful. Because of this, we are going to begin this chapter by showing you a systematic way of solving linear systems of equations so that, if you follow this approach, you will always be led to the correct solutions of a given linear system. Our approach will involve representing linear systems of equations by a type of expression called a matrix. After you have seen this particular use of matrices (it will be just one of many more to come) in Section 1.1, we will go on to study matrices in their own right in the rest of this chapter. We begin with a discussion of some of the basics.

## 1.1 Systems of Linear Equations

A **linear equation** in the variables or unknowns $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$\boxed{a_1x_1 + a_2x_2 + \cdots + a_nx_n = b}$$

where $a_1, a_2, \ldots, a_n, b$ are constants. For instance,

$$2x - 3y = 1$$

is a linear equation in the variables $x$ and $y$,

$$3x - y + 2z = 8$$

is a linear equation in the variables $x$, $y$, and $z$, and

$$-x_1 + 5x_2 - \pi x_3 + \sqrt{2}x_4 - 9x_5 = e^2$$

is a linear equation in the variables $x_1$, $x_2$, $x_3$, $x_4$, and $x_5$. The graph of a linear equation in two variables such as $2x - 3y = 1$ is a line in the $xy$-plane, and the graph of a linear equation in three variables such as $3x - y + 2z = 8$ is a plane in 3-space.

When considered together, a collection of linear equations

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

is called a **system of linear equations.** For instance,

$$\begin{aligned} x - y + z &= 0 \\ 2x - 3y + 4z &= -2 \\ -2x - y + z &= 7 \end{aligned}$$

is a system of three linear equations in three variables.

A **solution** to a system of equations with variables $x_1, x_2, \ldots, x_n$ consists of values of $x_1, x_2, \ldots, x_n$ that satisfy each equation in the system. From your first algebra course you should recall that the solutions to a system of two linear equations in $x$ and $y$,

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2, \end{aligned}$$

are the points at which the graphs of the lines given by these two equations intersect. Consequently, such a system will have exactly one solution if the graphs intersect in a single point, will have infinitely many solutions if the graphs are the same line, and will have no solution if the graphs are parallel. As we shall see, this in fact holds for all systems of linear equations; that is, a linear system either has exactly one solution, infinitely many solutions, or no solutions.

The main purpose of this section is to present the **Gauss-Jordan elimination method,**[1] a systematic way for solving systems of linear equations that will always lead us to solutions of the system. The Gauss-Jordan method involves the repeated use of three basic transformations on a system. We shall call the following transformations **elementary operations.**

1. Interchange two equations in the system.
2. Multiply an equation by a nonzero number.
3. Replace an equation by itself plus a multiple of another equation.

Two systems of equations are said to be **equivalent** if they have the same solutions. It is not difficult to see that applying an elementary operation to a system produces an equivalent system.

[1] Named in honor of Karl Friedrich Gauss (1777–1855), who is one of the greatest mathematicians of all time and is often referred to as the "prince of mathematics," and Wilhelm Jordan (1842–1899), a German engineer.

To illustrate the Gauss-Jordan elimination method, consider the system:

$$\begin{aligned} x - y + z &= 0 \\ 2x - 3y + 4z &= -2 \\ -2x - y + z &= 7. \end{aligned}$$

We are going to use elementary operations to transform this system to one of the form

$$\begin{array}{lllll} x & & & = & * \\ & y & & = & * \\ & & z & = & * \end{array}$$

where each $*$ is a constant from which we have the solution. To this end, let us first replace the second equation by itself plus $-2$ times the first equation (or subtracting 2 times the first equation from the second) and replace the third equation by itself plus 2 times the first equation to eliminate $x$ from the second and third equations. (We are doing two elementary operations simultaneously here.) This gives us the system

$$\begin{aligned} x - y + z &= 0 \\ -y + 2z &= -2 \\ -3y + 3z &= 7. \end{aligned}$$

Next, let us use the second equation to eliminate $y$ in the first and third equations by replacing the first equation by itself minus the second equation and replacing the third equation by itself plus $-3$ times the second equation, obtaining

$$\begin{aligned} x \qquad - z &= 2 \\ -y + 2z &= -2 \\ -3z &= 13. \end{aligned}$$

Now we are going to use the third equation to eliminate $z$ from the first two equations by multiplying the first equation by 3 and then subtracting the third equation from it (we actually are doing two elementary operations here) and multiplying the second equation by 3 and then adding 2 times the third equation to it (here too we are doing two elementary operations). This gives us the system:

$$\begin{array}{lllll} 3x & & & = & -7 \\ & -3y & & = & 20 \\ & & 3z & = & -13. \end{array}$$

Finally, dividing the first equation by 3 (or multiplying it by $1/3$), dividing the second equation by $-3$, and dividing the third equation by 3, we have our system in the promised form as

$$\begin{array}{lllll} x & & & = & -\dfrac{7}{3} \\ & y & & = & -\dfrac{20}{3} \\ & & z & = & -\dfrac{13}{3}, \end{array}$$

which tells us the solution.

You might notice that we only really need to keep track of the coefficients as we transform our system. To keep track of them, we will indicate a system such as

$$\begin{aligned} x - y + z &= 0 \\ 2x - 3y + 4z &= -2 \\ -2x - y + z &= 7 \end{aligned}$$

by the following array of numbers:

$$\left[\begin{array}{rrr:r} 1 & -1 & 1 & 0 \\ 2 & -3 & 4 & -2 \\ -2 & -1 & 1 & 7 \end{array}\right].$$

This array is called the **augmented matrix** for the system. The entries appearing to the left of the dashed vertical line are the coefficients of the variables as they appear in the system. This part of the augmented matrix is called the **coefficient matrix** of the system. The numbers to the right of the dashed vertical line are the constants on the right-hand side of the system as they appear in the system. In general, the augmented matrix for the system

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

is

$$\boxed{\left[\begin{array}{cccc:c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].}$$

The portion of the augmented matrix to the left of the dashed line with entries $a_{ij}$ is the coefficient matrix of the system.

Corresponding to the elementary operations for systems of equations are elementary row operations that we perform on the augmented matrix for a linear system. These are as follows.

1. Interchange two rows.[2]
2. Multiply a row by a nonzero number.
3. Replace a row by itself plus a multiple of another row.

---

[2] A line of numbers going across the matrix from left to right is called a **row**; a line of numbers going down the matrix is called a **column.**

As our first formal example of this section, we are going to redo the work we did in solving the system

$$\begin{aligned} x - y + z &= 0 \\ 2x - 3y + 4z &= -2 \\ -2x - y + z &= 7 \end{aligned}$$

with augmented matrices.

**EXAMPLE 1** Solve the system:

$$\begin{aligned} x - y + z &= 0 \\ 2x - 3y + 4z &= -2 \\ -2x - y + z &= 7. \end{aligned}$$

*Solution* Our work will consist of four steps. In the first step, we shall use the first row and row operations to make all other entries in the first column zero. In the second step, we shall use the second row to make all other entries in the second column zero. In the third step, we shall use the third row to make all other entries in the third column zero. In the fourth step, we shall make the nonzero entries in the coefficient matrix 1 at which point we will be able to read off our solution. To aid you in following the steps, an expression such as $R_2 - 2R_1$ next to the second row indicates that we are replacing the second row by itself plus $-2$ times the first row; an expression such as $R_1/3$ next to the first row indicates we are dividing this row by 3. Arrows are used to indicate the progression of our steps.

$$\left[\begin{array}{rrr:r} 1 & -1 & 1 & 0 \\ 2 & -3 & 4 & -2 \\ -2 & -1 & 1 & 7 \end{array}\right] \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 + 2R_1 \end{array}$$

$$\to \left[\begin{array}{rrr:r} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & -3 & 3 & 7 \end{array}\right] \begin{array}{l} R_1 - R_2 \\ \\ R_3 - 3R_2 \end{array}$$

$$\to \left[\begin{array}{rrr:r} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & -3 & 13 \end{array}\right] \begin{array}{l} 3R_1 - R_3 \\ 3R_2 + 2R_3 \\ \\ \end{array} \to \left[\begin{array}{rrr:r} 3 & 0 & 0 & -7 \\ 0 & -3 & 0 & 20 \\ 0 & 0 & -3 & 13 \end{array}\right] \begin{array}{l} R_1/3 \\ -R_2/3 \\ -R_3/3 \end{array}$$

$$\to \left[\begin{array}{rrr:r} 1 & 0 & 0 & -7/3 \\ 0 & 1 & 0 & -20/3 \\ 0 & 0 & 1 & -13/3 \end{array}\right]$$

The solution is then $x = -7/3$, $y = -20/3$, $z = -13/3$. ●

In Gauss-Jordan elimination, we use elementary row operations on the augmented matrix of the system to transform it so that the final coefficient matrix has a form called **reduced row-echelon form** with the following properties.

1. Any rows of zeros (called **zero rows**) appear at the bottom.
2. The first nonzero entry of a nonzero row is 1 (called a **leading 1**).
3. The leading 1 of a nonzero row appears to the right of the leading 1 of any preceding row.
4. All the other entries of a column containing a leading 1 are zero.

Looking back at Example 1, you will see that the coefficient matrix in our final augmented matrix is in reduced row-echelon form. Once we have the coefficient matrix in reduced row-echelon form, the solutions to the system are easily determined.

Let us do some more examples.

**EXAMPLE 2** Solve the system:

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

*Solution* We try to proceed as we did in Example 1. Notice, however, that we will have to modify our approach here. The symbol $R_2 \leftrightarrow R_3$ after the first step is used to indicate that we are interchanging the second and third rows.

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 1 \\ 1 & 1 & 0 & 1 & 2 \\ 1 & 2 & -4 & 0 & 1 \\ 2 & 1 & 2 & 5 & 1 \end{array}\right] \begin{array}{l} \\ R_2 - R_1 \\ R_3 - R_1 \\ R_4 - 2R_1 \end{array}$$

$$\to \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & -3 & -2 & 0 \\ 0 & -1 & 4 & 1 & -1 \end{array}\right] \begin{array}{l} \\ R_2 \leftrightarrow R_3 \\ \\ \\ \end{array}$$

$$\to \left[\begin{array}{cccc|c} 1 & 1 & -1 & 2 & 1 \\ 0 & 1 & -3 & -2 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & -1 & 4 & 1 & -1 \end{array}\right] \begin{array}{l} R_1 - R_2 \\ \\ \\ R_4 + R_2 \end{array}$$

$$\rightarrow \left[\begin{array}{cccc:c} 1 & 0 & 2 & 4 & 1 \\ 0 & 1 & -3 & -2 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 & -1 \end{array}\right] \begin{array}{l} R_1 - 2R_3 \\ R_2 + 3R_3 \\ \\ R_4 - R_3 \end{array} \rightarrow \left[\begin{array}{cccc:c} 1 & 0 & 0 & 6 & -1 \\ 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{array}\right]$$

We now have the coefficient matrix in reduced row-echelon form. Our final augmented matrix represents the system

$$\begin{aligned} x_1 + 6x_4 &= -1 \\ x_2 - 5x_4 &= 3 \\ x_3 - x_4 &= 1 \\ 0 &= -2, \end{aligned}$$

which is equivalent to our original system. Since this last system contains the false equation $0 = -2$, it has no solutions. Hence our original system has no solutions. ●

**EXAMPLE 3** Solve the system:

$$\begin{aligned} 2x + 3y - z &= 3 \\ -x - y + 3z &= 0 \\ x + 2y + 2z &= 3 \\ y + 5z &= 3. \end{aligned}$$

*Solution* We first reduce the augmented matrix for this system so that its coefficient matrix is in reduced row-echelon form.

$$\left[\begin{array}{ccc:c} 2 & 3 & -1 & 3 \\ -1 & -1 & 3 & 0 \\ 1 & 2 & 2 & 3 \\ 0 & 1 & 5 & 3 \end{array}\right] \begin{array}{l} \\ 2R_2 + R_1 \\ 2R_3 - R_1 \\ \\ \end{array}$$

$$\rightarrow \left[\begin{array}{ccc:c} 2 & 3 & -1 & 3 \\ 0 & 1 & 5 & 3 \\ 0 & 1 & 5 & 3 \\ 0 & 1 & 5 & 3 \end{array}\right] \begin{array}{l} R_1 - 3R_2 \\ \\ R_3 - R_2 \\ R_4 - R_2 \end{array}$$

$$\rightarrow \left[\begin{array}{ccc:c} 2 & 0 & -16 & -6 \\ 0 & 1 & 5 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} R_1/2 \\ \\ \\ \\ \end{array} \rightarrow \left[\begin{array}{ccc:c} 1 & 0 & -8 & -3 \\ 0 & 1 & 5 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

This final augmented matrix represents the equivalent system:

$$\begin{aligned} x - 8z &= -3 \\ y + 5z &= 3 \\ 0 &= 0 \\ 0 &= 0. \end{aligned}$$

Solving the first two equations for $x$ and $y$ in terms of $z$, we can say that our solutions have the form

$$x = -3 + 8z, \qquad y = 3 - 5z$$

where $z$ is any real number. In particular, we have infinitely many solutions in this example. (Any choice of $z$ gives us a solution. If $z = 0$, we have $x = -3$, $y = 3$, $z = 0$ as a solution; if $z = 1$, we have $x = 5$, $y = -2$, $z = 1$ as a solution; if $z = \sqrt{17}$, we have $x = -3 + 8\sqrt{17}$, $y = 3 - 5\sqrt{17}$, $z = \sqrt{17}$ as a solution; and so on.) In a case such as this, we refer to $z$ as the **free variable** and $x$ and $y$ as the **dependent** variables in our solutions. When specifying our solutions to systems like this, we will follow the convention of using variables that correspond to leading ones as dependent variables and those that do not as free variables. It is not necessary to specify our solutions this way, however. For instance, in this example we could solve for $z$ in terms of $x$, obtaining

$$z = \frac{x}{8} + \frac{3}{8}$$

and

$$y = 3 - 5z = 3 - 5\left(\frac{x}{8} + \frac{3}{8}\right) = -\frac{5x}{8} + \frac{9}{8},$$

giving us the solutions with $x$ as the free variable and $y$ and $z$ as the dependent variables. ●

**EXAMPLE 4** Solve the system:

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

*Solution* We again begin by reducing the augmented matrix to the point where its coefficient matrix is in reduced row-echelon form:

$$\left[\begin{array}{ccccc|c} 4 & -8 & -1 & 1 & 3 & 0 \\ 5 & -10 & -1 & 2 & 3 & 0 \\ 3 & -6 & -1 & 1 & 2 & 0 \end{array}\right] \begin{array}{l} \\ 4R_2 - 5R_1 \\ 4R_3 - 3R_1 \end{array}$$

$$\to \left[\begin{array}{rrrrr|r} 4 & -8 & -1 & 1 & 3 & 0 \\ 0 & 0 & 1 & 3 & -3 & 0 \\ 0 & 0 & -1 & 1 & -1 & 0 \end{array}\right] \begin{array}{l} R_1 + R_2 \\ \\ R_3 + R_2 \end{array}$$

$$\to \left[\begin{array}{rrrrr|r} 4 & -8 & 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 3 & -3 & 0 \\ 0 & 0 & 0 & 4 & -4 & 0 \end{array}\right] \begin{array}{l} R_1/4 \\ \\ R_3/4 \end{array}$$

$$\to \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & -3 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right] \begin{array}{l} R_1 - R_3 \\ R_2 - 3R_3 \\ \\ \end{array}$$

$$\to \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right].$$

We now have arrived at the equivalent system

$$\begin{aligned} x_1 - 2x_2 + x_5 &= 0 \\ x_3 &= 0 \\ x_4 - x_5 &= 0, \end{aligned}$$

which has solutions

$$x_1 = 2x_2 - x_5, \qquad x_3 = 0, \qquad x_4 = x_5$$

with $x_2$ and $x_5$ as the free variables and $x_1$ and $x_4$ as the dependent variables. ●

Systems of equations that have solutions such as those in Examples 1, 3, and 4 are called **consistent systems**; those that do not have solutions as occurred in Example 2 are called **inconsistent systems.** Notice that an inconsistent system is easily recognized once the coefficient matrix of its augmented matrix is put in reduced row-echelon form: There will be a row with zeros in the coefficient matrix with nonzero entry in the right-hand entry of this row. If we do not have this, the system is consistent. Consistent systems break down into two types. Once the coefficient matrix of the augmented matrix is put in reduced row-echelon form, the number of nonzero rows in the coefficient matrix is always less than or equal to the number of columns of the coefficient matrix. (That is, there will never be more nonzero rows than columns when the coefficient matrix is in reduced row-echelon form. Why is this the case?) If there are fewer nonzero rows than columns, as we had in Examples 3 and 4, the system will have infinitely many solutions. If we have as many nonzero rows as columns, as occurred in Example 1, we have exactly one solution. Recall that it was mentioned at the beginning of this section that every system of linear equations either has exactly one solution, infinitely many solutions, or no solutions. Now we can see why this is true.

A system of linear equations that can be written in the form

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

is called a **homogeneous system.** The system of equations in Example 4 is homogeneous. Notice that

$$x_1 = 0, \qquad x_2 = 0, \qquad \ldots, \qquad x_n = 0$$

is a solution to the homogeneous system in Equations (1). This is called the **trivial solution** of the homogeneous system. Because homogeneous systems always have a trivial solution, they are never inconsistent systems. Homogeneous systems will occur frequently in our future work and we will often be interested in whether such a system has solutions other than the trivial one, which we naturally call **nontrivial solutions.** The system in Example 4 has nontrivial solutions. For instance, we would obtain one (among the infinitely many such nontrivial solutions) by letting $x_2 = 1$ and $x_5 = 2$, in which case we have the nontrivial solution $x_1 = 0$, $x_2 = 1$, $x_3 = 0$, $x_4 = 2$, $x_5 = 2$. Actually, we can tell ahead of time that the system in Example 4 has nontrivial solutions. Because this system has fewer equations than variables, the reduced row-echelon form of the coefficient matrix will have fewer nonzero rows than columns and hence must have infinitely many solutions (only one of which is the trivial solution) and consequently must have infinitely many nontrivial solutions. This reasoning applies to any homogeneous system with fewer equations than variables, and hence we have the following theorem.

**THEOREM 1.1** A homogeneous system of $m$ linear equations in $n$ variables with $m < n$ has infinitely many nontrivial solutions.

Of course, if a homogeneous system has at least as many equations as variables such as the systems

$$\begin{aligned} x + y + z &= 0 \\ x - y - z &= 0 \\ 2x + y + z &= 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} 2x + y + z &= 0 \\ x - 2y - z &= 0 \\ 3x - y &= 0 \\ 4x - 3y - z &= 0 \end{aligned}$$

we would have to do some work toward solving these systems before we would be able to see whether they have nontrivial solutions. We shall do this for the second system a bit later.

**Gaussian elimination,** which is another systematic approach for solving linear systems, is similar to the approach we have been using but does not require that all the other entries of the column containing a leading 1 be zero. That is, it uses row operations to transform the augmented matrix so that the coefficient matrix has the following form:

**1.** Any zero rows appear at the bottom.

2. The first nonzero entry of a nonzero row is 1.
3. The leading 1 of a nonzero row appears to the right of the leading 1 of any preceding row.

Such a form is called a **row-echelon form** for the coefficient matrix. In essence, we do not eliminate (make zero) entries above the leading 1s in Gaussian elimination. Here is how this approach can be applied to the system in Example 1.

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 2 & -3 & 4 & -2 \\ -2 & -1 & 1 & 7 \end{array}\right] \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 + 2R_1 \end{array}$$

$$\to \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & -3 & 3 & 7 \end{array}\right] \begin{array}{l} \\ \\ R_3 - 3R_2 \end{array}$$

$$\to \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & -3 & 13 \end{array}\right] \begin{array}{l} \\ -R_2 \\ -R_3/3 \end{array} \to \left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & -13/3 \end{array}\right]$$

We now have the coefficient matrix in a row-echelon form and use this result to find the solutions. The third row tells us

$$z = -\frac{13}{3}.$$

The values of the remaining variables are found by a process called **back substitution.** From the second row, we have the equation

$$y - 2z = 2$$

from which we can find $y$:

$$y + \frac{26}{3} = 2$$
$$y = -\frac{20}{3}.$$

Finally, the first row represents the equation

$$x - y + z = 0$$

from which we can find $x$:

$$x + \frac{20}{3} - \frac{13}{3} = 0$$
$$x = -\frac{7}{3}.$$

On the plus side, Gaussian elimination requires fewer row operations. But on the minus side, the work is sometimes messy when doing the back substitutions. Often, we find ourselves having to deal with fractions even if our original system involves only integers. The back substitutions are also cumbersome to do when dealing with systems that have infinitely many solutions. Try the Gaussian elimination procedure in Example 3 or 4 if you would like to see how it goes.

As a rule we will tend to use Gauss-Jordan elimination when we have to find the solutions to a linear system in this text. Sometimes, however, we will not have to completely solve a system and will use Gaussian elimination since it will involve less work. The next example illustrates an instance of this. In fact, in this example we will not even have to bother completing Gaussian elimination by making the leading entries one.

**EXAMPLE 5** Determine the values of $a$, $b$, and $c$ so that the system

$$\begin{aligned} x - y + 2z &= a \\ 2x + y - z &= b \\ x + 2y - 3z &= c \end{aligned}$$

has solutions.

***Solution*** We begin doing row operations as follows.

$$\left[\begin{array}{rrr|c} 1 & -1 & 2 & a \\ 2 & 1 & -1 & b \\ 1 & 2 & -3 & c \end{array}\right] \begin{array}{l} \\ R_2 - 2R_1 \\ R_3 - R_1 \end{array} \rightarrow \left[\begin{array}{rrr|c} 1 & -1 & 2 & a \\ 0 & 3 & -5 & b-2a \\ 0 & 3 & -5 & c-a \end{array}\right] R_3 - R_2$$

$$\rightarrow \left[\begin{array}{rrr|c} 1 & -1 & 2 & a \\ 0 & 3 & -5 & b-2a \\ 0 & 0 & 0 & a-b+c \end{array}\right]$$

Now we can see that this system has solutions if and only if $a$, $b$, and $c$ satisfy the equation

$$a - b + c = 0.$$

Another place where we will sometimes use an abbreviated version of Gaussian elimination is when we are trying to see if a homogeneous system has nontrivial solutions.

**EXAMPLE 6** Determine if the system

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

has nontrivial solutions.

*Solution* Perform row operations:

$$\left[\begin{array}{rrr|r} 2 & 1 & 1 & 0 \\ 1 & -2 & -1 & 0 \\ 3 & -1 & 0 & 0 \\ 4 & -3 & -1 & 0 \end{array}\right] \begin{array}{c} \\ 2R_2 - R_1 \\ 2R_3 - 3R_1 \\ R_4 - 2R_1 \end{array} \to \left[\begin{array}{rrr|r} 2 & 1 & 1 & 0 \\ 0 & -5 & -3 & 0 \\ 0 & -5 & -3 & 0 \\ 0 & -5 & -3 & 0 \end{array}\right] \begin{array}{c} \\ \\ R_3 - R_2 \\ R_4 - R_2 \end{array}$$

$$\to \left[\begin{array}{rrr|r} 2 & 1 & 1 & 0 \\ 0 & -5 & -3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

It is now apparent that this system has nontrivial solutions. In fact, you should be able to see this after the first set of row operations. ●

It is not difficult to write computer programs for solving systems of linear equations using the Gauss-Jordan or Gaussian elimination methods. Thus it is not surprising that there are computer software packages for solving systems of linear systems.[3] Maple is one among several available mathematical software packages that can be used to find the solutions of linear systems of equations.

In the preface we mentioned that we will use Maple as our accompanying software package within this text. The use of Maple is at the discretion of your instructor. Some may use it, others may prefer to use a different software package, and yet others may choose to not use any such package (and give an excellent and complete course). For those intructors who wish to use Maple—or for students who are independently interested in gaining some knowledge of its capabilities—we will include occasional remarks about how to use it when we deem it appropriate. On many other occasions we will not include any remarks and will simply provide some exercises asking you to use indicated Maple commands. In these cases, you are expected to look up the command in the Help menu under Topic Search to see how to use it. This is one place where we will include a few remarks to get you started. For those who wish to use the software packages Mathematica or MATLAB, the accompanying *Technology Resource Manual* contains corresponding commands for these software packages.

Here we explain how to use Maple to find solutions to linear systems. One way to do this is to use the *linsolve* command. To use this command in a Maple worksheet, you will first have to load Maple's linear algebra package by typing

```
with(linalg);
```

at the command prompt > and then hitting the enter key. After doing this, you will get a list of Maple's linear algebra commands. To solve the system in Example 1, first enter the coefficient matrix of the system by typing

```
A:= matrix([[1,-1,1],[2,-3,4],[-2,-1,1]]);
```

[3] Often these packages employ methods that are more efficient than Gauss-Jordan or Gaussian elimination, but we will not concern ourselves with these issues in this text.

at the command prompt and then hitting the enter key. (The symbol := is used in Maple for indicating that we are defining A to be the coefficient matrix we type on the right.) The constants on the right-hand side of the system are typed and entered as

```
b:=vector([0,-2,7]);
```

at the command prompt. Finally, type and enter

```
linsolve(A,b);
```

at the commmand prompt and Maple will give us the solution as

$$\left[\frac{-7}{3}, \frac{-20}{3}, \frac{-13}{3}\right]^4.$$

Doing the same set of steps for the system in Example 2 results in no output, indicating there is no solution. Doing them in Example 3 yields the output

$$[-3 + 8\_t_1, 3 - 5\_t_1, \_t_1],$$

which is Maple's way of indicating our solutions in Example 3 with $t_1$ in place of $z$. In Example 4, these steps yield

$$[2\_t_1 - \_t_2, \_t_1, 0, \_t_2, \_t_2].$$

## EXERCISES 1.1

Solve the systems of equations in Exercises 1–16.

**1.** $\begin{aligned} x + y - z &= 0 \\ 2x + 3y - 2z &= 6 \\ x + 2y + 2z &= 10 \end{aligned}$

**2.** $\begin{aligned} 2x + y - 2z &= 0 \\ 2x - y - 2z &= 0 \\ x + 2y - 4z &= 0 \end{aligned}$

**3.** $\begin{aligned} 2x + 3y - 4z &= 3 \\ 2x + 3y - 2z &= 3 \\ 4x + 6y - 2z &= 7 \end{aligned}$

**4.** $\begin{aligned} 3x + y - 2z &= 3 \\ x - 8y - 14z &= -14 \\ x + 2y + z &= 2 \end{aligned}$

**5.** $\begin{aligned} x + 3z &= 0 \\ 2x + y - z &= 0 \\ 4x + y + 5z &= 0 \end{aligned}$

**6.** $\begin{aligned} 2x + 3y + z &= 4 \\ x + 9y - 4z &= 2 \\ x - y + 2z &= 3 \end{aligned}$

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

**8.** $\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 1 \\ x_1 + x_2 - x_3 - x_4 &= -1 \\ x_1 + 2x_2 + x_3 + 2x_4 &= -1 \\ 2x_1 + 2x_2 + x_3 + x_4 &= 2 \end{aligned}$

**9.** $\begin{aligned} x_1 + 2x_2 - 3x_3 + 4x_4 &= 2 \\ 2x_1 - 4x_2 + 6x_3 - 5x_4 &= 10 \\ x_1 - 6x_2 + 9x_3 - 9x_4 &= 8 \\ 3x_1 - 2x_2 + 4x_3 - x_4 &= 12 \end{aligned}$

---

[4] Software packages such as Maple often will have several ways of doing things. This is the case for solving systems of linear equations. One variant is to enter $b$ as a matrix with one column by typing and entering

```
b:=matrix([[0],[-2],[7]]);
```

When we then type and enter

```
linsolve(A,b);
```

our solution is given in column form. Another way is to use Maple's *solve* command for solving equations and systems of equations. (With this approach it is not necessary to load Maple's linear algebra package.) To do it this way for the system in Example 1, we type and enter

```
solve({x-y+z=0,2*x-3*y+4*z=-2,-2*x-y+z=7},{x,y,z});
```

**10.**
$$\begin{aligned} x_1 - x_2 + x_3 + x_4 - x_5 &= 0 \\ 2x_1 - x_2 + 2x_3 - x_4 + 3x_5 &= 0 \\ 2x_1 - x_2 - 2x_4 + x_5 &= 0 \\ x_1 + x_2 - x_3 - x_4 + 2x_5 &= 0 \\ 2x_1 + 4x_3 + x_4 + 3x_5 &= 0 \end{aligned}$$

**11.**
$$\begin{aligned} x + 2y + z &= -2 \\ 2x + 2y - 2z &= 3 \end{aligned}$$

**12.**
$$\begin{aligned} 2x - 4y + 6z &= 2 \\ -3x + 6y - 9z &= 3 \end{aligned}$$

**13.**
$$\begin{aligned} x - 2y &= 2 \\ x + 8y &= -4 \\ 2x + y &= 1 \end{aligned}$$

**14.**
$$\begin{aligned} 2x + 3y &= 5 \\ 2x + y &= 2 \\ x - 2y &= 1 \end{aligned}$$

**15.**
$$\begin{aligned} 2x_1 - x_2 - x_3 + x_4 + x_5 &= 0 \\ x_1 - x_2 + x_3 + 2x_4 - 3x_5 &= 0 \\ 3x_1 - 2x_2 - x_3 - x_4 + 2x_5 &= 0 \end{aligned}$$

**16.**
$$\begin{aligned} x_1 - 3x_2 + x_3 - x_4 - x_5 &= 1 \\ 2x_1 + x_2 - x_3 + 2x_4 + x_5 &= 2 \\ -x_1 + 3x_2 - x_3 - 2x_4 - x_5 &= 3 \\ 2x_1 + x_2 - x_3 - x_4 - x_5 &= 6 \end{aligned}$$

Determine conditions on $a$, $b$, and $c$ so that the systems of equations in Exercises 17 and 18 have solutions.

**17.**
$$\begin{aligned} 2x - y + 3z &= a \\ x - 3y + 2z &= b \\ x + 2y + z &= c \end{aligned}$$

**18.**
$$\begin{aligned} x + 2y - z &= a \\ x + y - 2z &= b \\ 2x + y - 3z &= c \end{aligned}$$

Determine conditions on $a$, $b$, $c$, and $d$ so that the systems of equations in Exercises 19 and 20 have solutions.

**19.**
$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= a \\ x_1 - x_2 - x_3 + x_4 &= b \\ x_1 + x_2 + x_3 + x_4 &= c \\ x_1 - x_2 + x_3 + x_4 &= d \end{aligned}$$

**20.**
$$\begin{aligned} x_1 - x_2 + x_3 + x_4 &= a \\ x_1 + x_2 - 2x_3 + 3x_4 &= b \\ 3x_1 - 2x_2 + 3x_3 - 2x_4 &= c \\ 2x_2 - 3x_3 + 2x_4 &= d \end{aligned}$$

Determine if the homogeneous systems of linear equations in Exercises 21–24 have nontrivial solutions. You do not have to solve the systems.

**21.**
$$\begin{aligned} 9x - 2y + 17z &= 0 \\ 13x + 81y - 27z &= 0 \end{aligned}$$

**22.**
$$\begin{aligned} 99x_1 + \pi x_2 - \sqrt{5}x_3 &= 0 \\ 2x_1 + (\sin 1)x_2 + 2x_4 &= 0 \\ 3.38x_1 - ex_3 + (\ln 2)x_4 &= 0 \end{aligned}$$

**23.**
$$\begin{aligned} x - y + z &= 0 \\ 2x + y + 2z &= 0 \\ 3x - 5y + 3z &= 0 \end{aligned}$$

**24.**
$$\begin{aligned} x + y + 2z &= 0 \\ 3x - y - 2z &= 0 \\ 2x - 2y - 4z &= 0 \\ x + 3y + 6z &= 0 \end{aligned}$$

**25.** We have seen that homogeneous linear systems with fewer equations than variables always have infinitely many solutions. What possibilities can arise for nonhomogeneous linear systems with fewer equations than variables? Explain your answer.

**26.** Give an example of a system of linear equations with more equations than variables that illustrates each of the following possibilities: Has exactly one solution, has infinitely many solutions, and has no solution.

**27.** Describe graphically the possible solutions to a system of two linear equations in $x$, $y$, and $z$.

**28.** Describe graphically the possible solutions to a system of three linear equations in $x$, $y$, and $z$.

Use Maple or another appropriate software package to solve the systems of equations in Exercises 29–32. If you are using Mathematica or MATLAB, see the *Technology Resource Manual* for appropriate commands. (To become more comfortable with the software package you are using, you may wish to practice using it to solve some of the smaller systems in Exercises 1–16 before doing these.)

**29.**
$$\begin{aligned} 7x_1 - 3x_2 + 5x_3 - 8x_4 + 2x_5 &= 13 \\ 12x_1 + 4x_2 - 16x_3 - 9x_4 + 7x_5 &= 21 \\ -22x_1 - 8x_2 + 25x_3 - 16x_4 - 8x_5 &= 47 \\ -52x_1 - 40x_2 + 118x_3 - 37x_4 - 29x_5 &= 62 \end{aligned}$$

**30.**
$$\begin{aligned} 46x_1 + 82x_2 - 26x_3 + 44x_4 &= 122 \\ 69x_1 + 101x_2 + 43x_3 + 30x_4 &= 261 \\ -437x_1 - 735x_2 + 335x_3 + 437x_4 &= -406 \\ 299x_1 + 379x_2 - 631x_3 - 2501x_4 &= -4146 \\ 1863x_1 + 2804x_2 + 62x_3 - 1983x_4 &= 4857 \\ 1748x_1 + 2291x_2 - 461x_3 - 9863x_4 &= 4166 \end{aligned}$$

**31.** $62x_1 + 82x_2 + 26x_3 - 4x_4$
$+ 32x_5 + 34x_6 - 2x_7 - 4x_8 = 0$
$93x_1 + 123x_2 + 67x_3 - 36x_4$
$+ 106x_5 + 51x_6 + 31x_7 - 188x_8 = 0$
$-589x_1 - 779x_2 - 303x_3 + 647x_4$
$- 330x_5 - 323x_6 - 256x_7 - 246x_8 = 0$
$403x_1 + 533x_2 + 365x_3 - 2493x_4$
$+ 263x_5 + 50x_6 + 981x_7 + 1345x_8 = 0$
$2511x_1 + 3321x_2 + 1711x_3 - 2636x_4$
$+ 2358x_5 + 1357x_6 + 1457x_7 - 2323x_8 = 0$
$2356x_1 + 3116x_2 + 2038x_3 - 6828x_4$
$+ 2418x_5 + 1936x_6 + 3596x_7 - 357x_8 = 0$

**32.** $3.3x_1 + 3.3x_2 + 12.1x_3 + 2.2x_4$
$+ 45.1x_5 + 7.7x_6 + 12.1x_7$
$+ 35.2x_8 + 1.1x_9 = -3.3$
$3x_1 + 3x_2 + 15.8x_3 - 4x_4$
$+ 61.4x_5 + 82x_6 + 5x_7$
$+ 21.2x_8 + 5.8x_9 = -0.6$
*(continued)*

$-3.3x_1 - 3.3x_2 - 16.1x_3 + 1.8x_4$
$- 61.1x_5 - 9.7x_6 - 10.1x_7$
$- 28.2x_8 - 4.2x_9 = 7.3$
$3x_1 + 3x_2 + 15x_3$
$+ 56.3x_5 + 8.4x_6 + 13.7x_7$
$+ 30.3x_8 + 9.8x_9 = -9.9$
$3x_1 + 3x_2 + 11x_3 + 3x_4$
$+ 37x_5 + 19.5x_6 + 14x_7$
$+ 30.5x_8 - 7.5x_9 = -17$
$-3x_1 - 3x_2 - 11x_3 - 3x_4$
$- 41.1x_5 - 3.8x_6 - 5.9x_7$
$- 34.1x_8 + 16.4x_9 = 38.3$
$-2.2x_4 + 5.2x_5 - 4.2x_6$
$- 11.6x_7 - 1.4x_8 + 31.2x_9 = 48.2$
$4.2x_1 + 4.2x_2 + 19.4x_3 - 3.2x_4$
$+ 76.4x_5 - 0.2x_6 + 3.4x_7$
$+ 35.8x_8 - 9.6x_9 = -23.2$

## 1.2 Matrices and Matrix Operations

In the previous section we introduced augmented matrices for systems of linear equations as a convenient way of representing these systems. This is one of many uses of matrices. In this section we will look at matrices from a general point of view.

We should be explicit about exactly what a matrix is, so let us begin with a definition. A **matrix** is a rectangular array of objects called the **entries** of the matrix. (For us, the objects will be numbers, but they do not have to be. For example, we could have matrices whose entries are automobiles or members of a marching band.) We write matrices down by enclosing their entries within brackets (some use parentheses instead) and, if we wish to give a matrix a name, we will do so by using capital letters such as $A$, $B$, or $C$. Here are some examples of matrices:

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

$$C = \begin{bmatrix} 0 \\ -1 \\ 1/2 \\ 4 \end{bmatrix}, \qquad D = \begin{bmatrix} 0 & -2 & \pi & 8 \\ -1 & 12 & 3/8 & \ln 2 \\ 1 & 1 & 1 & -1 \\ \sqrt{2} & -7 & 0.9 & -391/629 \end{bmatrix}$$

Augmented matrices of systems of linear equations have these forms if we delete the dashed line. In fact, the dashed line is included merely as a convenience to help distinguish the left- and right-hand sides of the equations. If a matrix has $m$ rows (which go across) and $n$ columns (which go up and down), we say the **size** (or **dimensions**) of the matrix is (or are) $m \times n$ (read "$m$ by $n$"). Thus, for the matrices just given, $A$ is a $2 \times 3$ matrix, $B$ is a $1 \times 5$ matrix, $C$ is a $4 \times 1$ matrix, and $D$ is a $4 \times 4$ matrix. A matrix such as $B$ that has one row is called a **row matrix** or **row vector**; a matrix such as $C$ that has one column is called a **column matrix** or **column vector.** Matrices that have the same number of rows as columns (that is, $n \times n$ matrices) are called **square matrices.** The matrix $D$ is an example of a square matrix.

As you would expect, we consider two matrices $A$ and $B$ to be equal, written $A = B$, if they have the same size and entries. For example,

$$\begin{bmatrix} -1 & 2 \\ 1 & 12 \end{bmatrix} = \begin{bmatrix} -1 & 8/4 \\ 2-1 & 3 \cdot 4 \end{bmatrix}$$

while

$$\begin{bmatrix} -1 & 2 \\ 1 & 12 \end{bmatrix} \neq \begin{bmatrix} 5 & 2 \\ 1 & 12 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq \begin{bmatrix} 1 & 2 \end{bmatrix}.$$

The general form of an $m \times n$ matrix $A$ is

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}. \tag{1}$$

Notice that in this notation the first subscript $i$ of an entry $a_{ij}$ is the row in which the entry appears and the second subscript $j$ is the column in which it appears. To save writing, we shall often indicate a matrix such as this by simply writing

$$A = [a_{ij}].$$

If we wish to single out the $ij$-entry of a matrix $A$, we will write

$$\text{ent}_{ij}(A).$$

For instance, if $B$ is the matrix

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

then

$$\text{ent}_{23}(B) = -9.$$

If $A = [a_{ij}]$ is an $n \times n$ matrix, the entries $a_{11}, a_{22}, \ldots, a_{nn}$ are called the **diagonal entries** of $A$. The matrix $B$ has diagonal entries $-1, 4, 7$.

We will use the symbol $\mathbb{R}$ to denote the set of real numbers. The set of $m \times n$ matrices with entries from $\mathbb{R}$ will be denoted

$$\boxed{M_{m\times n}(\mathbb{R}).}$$

Thus, for example, in set notation

$$M_{2\times 2}(\mathbb{R}) = \left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\}.$$[5]

You have encountered two-dimensional vectors in two-dimensional space (which we will denote by $\mathbb{R}^2$) and three-dimensional vectors in three-dimensional space (which we will denote by $\mathbb{R}^3$) in previous courses. One standard notation for indicating such vectors is to use ordered pairs $(a, b)$ for two-dimensional vectors and ordered triples $(a, b, c)$ for three-dimensional vectors. Notice that these ordered pairs and triples are in fact row matrices or row vectors. However, we will be notationally better off if we use column matrices for two- and three-dimensional vectors. We also will identify the set of two-dimensional vectors with $\mathbb{R}^2$ and the set of three-dimensional vectors with $\mathbb{R}^3$; in other words,

$$\mathbb{R}^2 = M_{2\times 1}(\mathbb{R}) = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \mid a, b \in \mathbb{R} \right\}$$

and

$$\mathbb{R}^3 = M_{3\times 1}(\mathbb{R}) = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \mid a, b, c \in \mathbb{R} \right\}$$

in this book. More generally, the set of $n \times 1$ column matrices $M_{n\times 1}(\mathbb{R})$ will be denoted $\mathbb{R}^n$ and we will refer to the elements of $\mathbb{R}^n$ as **vectors in $\mathbb{R}^n$** or **$n$-dimensional vectors**.

We next turn our attention to the "arithmetic" of matrices beginning with the operations of **addition** and a multiplication by numbers called **scalar multiplication**.[6] If $A$ and $B$ are matrices of the same size, we add $A$ and $B$ by adding their corresponding

---

[5] In set notation, the vertical bar, |, denotes "such that" (some use a colon, :, instead of a vertical bar) and the symbol $\in$ denotes "element of" (or "member of"). One way of reading

$$\left\{ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \mid a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R} \right\}$$

is as "the set of matrices

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

such that $a_{11}, a_{12}, a_{21}, a_{22}$ are elements of the set of real numbers."

[6] These two operations are extensions of the ones you already know for vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ to matrices in general.

entries; that is, if

$$A = [a_{ij}] \quad \text{and} \quad B = [b_{ij}]$$

are matrices in $M_{m \times n}(\mathbb{R})$, the sum of $A$ and $B$ is the $m \times n$ matrix

$$\boxed{A + B = [a_{ij} + b_{ij}].}$$

For instance, if

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 8 & 9 \\ 10 & 11 \\ 12 & 13 \end{bmatrix},$$

then

$$A + B = \begin{bmatrix} 1+8 & 2+9 \\ 3+10 & 4+11 \\ 5+12 & 6+13 \end{bmatrix} = \begin{bmatrix} 9 & 11 \\ 13 & 15 \\ 17 & 19 \end{bmatrix}.$$

Note that we have only defined sums of matrices of the same size. The sum of matrices of different sizes is undefined. For example, the sum

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

is undefined. If $c$ is a real number (which we call a **scalar** in this setting) and $A = [a_{ij}]$ is an $m \times n$ matrix, the **scalar product** $cA$ is the $m \times n$ matrix obtained by multiplying $c$ times each entry of $A$:

$$\boxed{cA = c[a_{ij}] = [ca_{ij}].}$$

For example,

$$5\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 & 5 \cdot 2 \\ 5 \cdot 3 & 5 \cdot 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix}.$$

The following theorem lists some elementary properties involving addition and scalar multiplication of matrices.

**THEOREM 1.2** If $A$, $B$, and $C$ are matrices of the same size and if $c$ and $d$ are scalars, then:

1. $A + B = B + A$ (commutative law of addition).
2. $A + (B + C) = (A + B) + C$ (associative law of addition).
3. $c(dA) = (cd)A$.
4. $c(A + B) = cA + cB$.
5. $(c + d)A = cA + dA$.

*Proof* We prove these equalities by showing that the matrices on each side have the same entries. Let us prove parts (1) and (4) here. The proofs of the remaining parts will be left as exercises (Exercise 24). For notational purposes, we set

$$A = [a_{ij}] \quad \text{and} \quad B = [b_{ij}].$$

Part (1) follows since

$$\text{ent}_{ij}(A + B) = a_{ij} + b_{ij} = b_{ij} + a_{ij} = \text{ent}_{ij}(B + A).$$

To obtain part (4),

$$\text{ent}_{ij}(c(A + B)) = c(a_{ij} + b_{ij}) = ca_{ij} + cb_{ij} = \text{ent}_{ij}(cA + cB).$$ ●

One special type of matrix is the set of **zero matrices**. The $m \times n$ zero matrix, denoted $O_{m\times n}$, is the $m \times n$ matrix that has all of its entries zero. For example,

$$O_{2\times 2} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad O_{4\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Notice that if $A$ is an $m \times n$ matrix, then:

1. $A + O_{m\times n} = A$.
2. $0 \cdot A = O_{m\times n}$.

We often will indicate a zero matrix by simply writing $\boldsymbol{O}$. (To avoid confusion with the number zero, we put this in boldface print in this book.) For instance, we might write the first property as $A + \boldsymbol{O} = A$. The second property could be written as $0 \cdot A = \boldsymbol{O}$.

The **negative** of a matrix $A = [a_{ij}]$, denoted $-A$, is the matrix whose entries are the negatives of those of $A$:

$$\boxed{-A = [-a_{ij}].}$$

Notice that

$$-A = (-1)A \quad \text{and} \quad A + (-A) = \boldsymbol{O}.$$

Subtraction of matrices $A$ and $B$ of the same size can be defined in terms of adding the negative of $B$:

$$\boxed{A - B = A + (-B).}$$

Of course, notice that $A - B$ could also be found by subtracting the entries of $B$ from the corresponding entries of $A$.

Up to this point, all of the operations we have introduced on matrices should seem relatively natural. Our final operation will be matrix multiplication, which upon first glance may not seem to be the natural way to multiply matrices. However, the manner of multiplying matrices you are about to see is the one that we will need as we use matrix multiplication in our future work.

Here is how we do matrix multiplication: Suppose that $A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times l$ matrix. The product of $A$ and $B$ is defined to be the $m \times l$ matrix

$$\boxed{AB = [p_{ij}]}$$

where

$$\boxed{p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.}$$

In other words, for each $1 \leq i \leq m$ and $1 \leq j \leq l$ the $ij$-entry of $AB$ is found by multiplying each entry of row $i$ of $A$ times its corresponding entry of column $j$ of $B$ and then summing these products.

Here is an example illustrating our matrix multiplication procedure.

**EXAMPLE 1** Find the product $AB$ for

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

***Solution*** The product $AB$ is

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

$$= \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}.$$

Once you practice this sum of row entries times column entries a few times, you should find yourself getting the hang of it.[7] Let us do another example of matrix multiplication.

**EXAMPLE 2** Find the product $CD$ for

$$C = \begin{bmatrix} -1 & 2 & -3 \\ 0 & -1 & 1 \\ 4 & 2 & -1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & -2 \\ -3 & 4 \\ 1 & 1 \end{bmatrix}.$$

[7] You might find it convenient to note that the $ij$-entry of $AB$ is much like the dot product of the vector formed by row $i$ of $A$ with the vector formed by column $j$ of $B$. We will discuss dot products more fully in Chapter 9.

*Solution*

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

$$= \begin{bmatrix} (-1)\cdot 1 + 2(-3) - 3\cdot 1 & (-1)(-2) + 2\cdot 4 - 3\cdot 1 \\ 0\cdot 1 - 1(-3) + 1\cdot 1 & 0(-2) - 1\cdot 4 + 1\cdot 1 \\ 4\cdot 1 + 2(-3) - 1\cdot 1 & 4(-2) + 2\cdot 4 - 1\cdot 1 \end{bmatrix} = \begin{bmatrix} -10 & 7 \\ 4 & -3 \\ -3 & -1 \end{bmatrix}$$

Notice that for the product $AB$ of two matrices $A$ and $B$ to be defined, it is necessary that the number of columns of $A$ be the same as the number of rows of $B$. If this is not the case, the product is not defined. For instance, the product $DC$ for the matrices in Example 2 is not defined. In particular, $CD$ is not the same as $DC$. This is an illustration of the fact that *matrix multiplication is not commutative;* that is, $AB$ is not in general the same as $BA$ for matrices $A$ and $B$. Sometimes these products are not the same because one is defined while the other is not, as the matrices $C$ and $D$ illustrate. But even if both products are defined, it is often the case that they are not the same. If you compute the product $BA$ for the matrices in Example 1, you will find (try it)

$$BA = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix},$$

which is not the same as $AB$.[8] In the case when $AB = BA$ for two matrices $A$ and $B$, we say $A$ and $B$ **commute.**

While matrix multiplication is not commutative, some properties that you are used to having for multiplication of numbers do carry over to matrices when the products are defined.

**THEOREM 1.3** Provided that the indicated sums and products are defined, the following properties hold where $A$, $B$, and $C$ are matrices and $d$ is a scalar.

1. $A(BC) = (AB)C$ (associative law of multiplication)
2. $A(B + C) = AB + AC$ (left-hand distributive law)
3. $(A + B)C = AC + BC$ (right-hand distributive law)
4. $d(AB) = (dA)B = A(dB)$

***Proof*** We will prove the first two parts here and leave proofs of the remaining parts as exercises (Exercise 25). For notational purposes, suppose

$$A = [a_{ij}], \qquad B = [b_{ij}], \qquad \text{and} \qquad C = [c_{ij}].$$

[8] This is not the first time you have encountered an example of a noncommutative operation. Composition of functions is noncommutative. The cross product of two three-dimensional vectors is another example of a noncommutative operation.

To prove part (1), we also have to introduce some notation for the sizes of $A$, $B$, and $C$. Suppose $A$ is an $m \times n$ matrix, $B$ is an $n \times l$ matrix, and $C$ is an $l \times h$ matrix. Both $A(BC)$ and $(AB)C$ are $m \times h$ matrices. (Why?) To see that these products are the same, we work out the $ij$-entry of each. For $A(BC)$, this is

$$\mathrm{ent}_{ij}(A(BC)) = \sum_{k=1}^{n} a_{ik}\mathrm{ent}_{kj}(BC) = \sum_{k=1}^{n} a_{ik}\left(\sum_{q=1}^{l} b_{kq}c_{qj}\right) = \sum_{k=1}^{n}\left(\sum_{q=1}^{l} a_{ik}b_{kq}c_{qj}\right).$$

Carrying out the same steps for $(AB)C$,

$$\mathrm{ent}_{ij}((AB)C)) = \sum_{q=1}^{l} \mathrm{ent}_{iq}(AB)c_{qj} = \sum_{q=1}^{l}\left(\sum_{k=1}^{n} a_{ik}b_{kq}\right)c_{qj} = \sum_{q=1}^{l}\left(\sum_{k=1}^{n} a_{ik}b_{kq}c_{qj}\right).$$

Since the summations over $k$ and $q$ are interchangeable, we see that the $ij$-entries of $A(BC)$ and $(AB)C$ are the same and hence $A(BC) = (AB)C$.

To prove part (2), we again introduce notation for the sizes of our matrices. Suppose $A$ is an $m \times n$ matrix and $B$ and $C$ are $n \times l$ matrices. Both $A(B + C)$ and $AB + AC$ are $m \times l$ matrices. We have

$$\begin{aligned}\mathrm{ent}_{ij}(A(B+C)) &= \sum_{k=1}^{n} a_{ik}(\mathrm{ent}_{kj}(B+C)) = \sum_{k=1}^{n} a_{ik}(b_{kj}+c_{kj})\\ &= \sum_{k=1}^{n}(a_{ik}b_{kj}+a_{ik}c_{kj})\end{aligned}$$

and

$$\begin{aligned}\mathrm{ent}_{ij}(AB+AC) = \mathrm{ent}_{ij}(AB) + \mathrm{ent}_{ij}(AC) &= \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}\\ &= \sum_{k=1}^{n}(a_{ik}b_{kj}+a_{ik}c_{kj}).\end{aligned}$$

Thus $A(B + C) = AB + AC$ since they have the same entries. ●

If $A$ is a square matrix, we can define positive integer powers of $A$ in the same manner as we do for real numbers; that is,

$$A^1 = A, \qquad A^2 = AA, \qquad \text{and} \qquad A^3 = A^2A = AAA, \ldots.$$

Such powers are not defined, however, if $A$ is not a square matrix. (Why is this the case?)

If $A$ is an $m \times n$ matrix, it is easy to see that

$$O_{l\times m}A = O_{l\times n} \quad \text{and} \quad AO_{n\times l} = O_{m\times l}.$$

Besides zero matrices, another special type of matrices is the set of **identity matrices.** The $n \times n$ identity matrix, denoted $I_n$, has diagonal entries 1 and all other entries 0. For example,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Identity matrices play the role the number 1 plays for the real numbers with respect to multiplication in the sense that

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

for any $m \times n$ matrix $A$. (Convince yourself of these statements.)

One use (among many more to come) of matrix multiplication arises in connection with systems of linear equations. Given a system of linear equations

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

we will let $A$ denote the coefficient matrix of this system,

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

$X$ denote the column of variables,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

and $B$ denote the column,

$$B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Observe that our system can then be conveniently written as the **matrix equation**

$$\boxed{AX = B.}$$

For instance, the system

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

would be written

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & -7 & 1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

as a matrix equation. Notice that a homogeneous linear system takes on the form $AX = O$ when written as a matrix equation.

## EXERCISES 1.2

In Exercises 1–18, either perform the indicated operations or state that the expression is undefined where $A$, $B$, $C$, $D$, $E$, and $F$ are the matrices:

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

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

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

**1.** $A + B$ **2.** $D - C$ **3.** $2B$

**4.** $-\frac{3}{4}F$ **5.** $A - 4B$ **6.** $3D + 2C$

**7.** $CD$ **8.** $DC$ **9.** $EF$

**10.** $FE$ **11.** $AE$ **12.** $EA$

**13.** $(E + F)A$ **14.** $B(C + D)$ **15.** $3AC$

**16.** $F(-2B)$ **17.** $C^2$ **18.** $A^3$

Write the systems of equations in Exercises 19 and 20 in the matrix form $AX = B$.

**19.**
$$\begin{aligned} 2x - y + 4z &= 1 \\ x + y - z &= 4 \\ y + 3z &= 5 \\ x + y &= 2 \end{aligned}$$

**20.**
$$\begin{aligned} x_1 - 3x_2 + x_3 - 5x_4 &= 2 \\ x_1 + x_2 - x_3 + x_4 &= 1 \\ x_1 - x_2 - x_3 + 6x_4 &= 6 \end{aligned}$$

Write the matrix equations as systems of equations in Exercises 21 and 22.

**21.** $\begin{bmatrix} 2 & -2 & 5 & 7 \\ 4 & 5 & -11 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 12 \\ -3 \end{bmatrix}$

**22.** $\begin{bmatrix} 2 & 2 & -11 \\ 0 & -1 & -5 \\ 2 & -3 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 51 \\ -33 \\ 1/2 \end{bmatrix}$

**23.** Suppose that $A$ and $B$ are $n \times n$ matrices.

**a)** Show that $(A + B)^2 = A^2 + AB + BA + B^2$.

**b)** Explain why $(A + B)^2$ is not equal to $A^2 + 2AB + B^2$ in general.

**24.** Prove the following parts of Theorem 1.2.

**a)** Part (2)

**b)** Part (3)

**c)** Part (5)

**25.** Prove the following parts of Theorem 1.3.

**a)** Part (3)

**b)** Part (4)

**26.** Suppose $A$ is an $m \times n$ matrix and $B$ is an $n \times l$ matrix. Further, suppose that $A$ has a row of zeros. Does $AB$ have a row of zeros? Why or why not? Does this also hold if $B$ has a row of zeros? Why or why not?

**27.** Suppose $A$ is an $m \times n$ matrix and $B$ is an $n \times l$ matrix. Further, suppose that $B$ has a column of zeros. Does $AB$ have a column of zeros? Why or why not? Does this also hold if $A$ has a column of zeros? Why or why not?

**28.** Give an example of two matrices $A$ and $B$ for which $AB = \mathbf{O}$ with $A \neq \mathbf{O}$ and $B \neq \mathbf{O}$.

**29. a)** Suppose that $A$ is the row vector

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

and $B$ is an $n \times l$ matrix. View $B$ as the column of row vectors

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_n \end{bmatrix}$$

where $B_1, B_2, \ldots, B_n$ are the rows of $B$. Show that

$$AB = a_1 B_1 + a_2 B_2 + \cdots + a_n B_n.$$

**b)** Use the result of part (a) to find $AB$ for

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

**30. a)** Suppose that $B$ is the column vector

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

and $A$ is an $m \times n$ matrix. View $A$ as the row of column vectors

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$$

where $A_1, A_2, \ldots, A_n$ are the columns of $A$. Show that

$$AB = b_1 A_1 + b_2 A_2 + \cdots + b_n A_n.$$

**b)** Use the result of part (a) to find $AB$ for

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

**31.** The **trace** of a square matrix $A$, denoted $\operatorname{tr}(A)$, is the sum of the diagonal entries of $A$. Find $\operatorname{tr}(A)$ for

$$A = \begin{bmatrix} 5 & 0 & -4 \\ 2 & -11 & 6 \\ 2 & 10 & 3 \end{bmatrix}.$$

**32.** Prove the following where $A$ and $B$ are square matrices of the same size and $c$ is a scalar.

**a)** $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$

**b)** $\operatorname{tr}(cA) = c \operatorname{tr}(A)$

**c)** $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

The *matrix* command introduced in the previous section is one way of entering matrices on a Maple worksheet. Maple uses the *evalm* command along with `+`, `-`, `*`, `&*`, and `∧` to find sums, differences, scalar products, matrix products, and matrix powers, respectively. For instance, to find $A - B + 4C + AB - C^3$ where $A$, $B$, and $C$ are matrices already entered on a Maple worksheet, we would type and enter

```
evalm(A-B+4*C+A&*B-C∧3);
```

at the command prompt. A scalar product $cA$ also may be found with the *scalarmul* command by typing

```
scalarmul(A,c);
```

at the command prompt. Products of two or more matrices can be found by using the *multiply* command. For instance, typing and entering

```
multiply(B,A,C);
```

will give us the product $BAC$. Use these Maple commands or appropriate commands in another suitable software package (keep in mind that corresponding Mathematica and MATLAB commands can be found in the *Technology Resource Manual*) to find the indicated expression (if possible) where

$$A = \begin{bmatrix} 4 & -2 & 16 & 27 & -11 \\ 9 & 43 & 9 & -8 & -1 \\ 34 & 20 & -3 & 0 & 21 \\ -5 & 4 & 4 & 7 & 41 \\ 0 & 12 & -2 & -2 & 3 \end{bmatrix},$$

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

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

in Exercises 33–40.

**33.** $A - 2B$

**34.** $5A + 6C$

**35.** $ABC$

**36.** $CB + C$

**37.** $(A + B)^2$

**38.** $4A + CB$

**39.** $4CA - 5CB - 2C$

**40.** $B^2 - 4AB + 2A^2$

## 1.3 Inverses of Matrices

If $A$ is an $n \times n$ matrix, we say that an $n \times n$ matrix $B$ is an **inverse** of $A$ if

$$\boxed{AB = BA = I}$$

where $I$ is the $n \times n$ identity matrix. To illustrate, the matrix

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

has the matrix

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

as an inverse since

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

and

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

(How we obtain $B$ will be seen later.)

Not all square matrices have inverses. Certainly, square zero matrices do not have inverses. (The products $\boldsymbol{O}B$ and $B\boldsymbol{O}$ are $\boldsymbol{O}$, not $I$.) But even a nonzero square matrix may fail to have an inverse. As a simple example, the matrix

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

cannot have an inverse since for any $2 \times 2$ matrix

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

we have

$$AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Square matrices that have inverses are called **invertible** or **nonsingular** matrices; those that do not have inverses are called **noninvertible** or **singular** matrices.

When a matrix has an inverse, it has only one inverse.

**THEOREM 1.4** If $A$ is an invertible matrix, then the inverse of $A$ is unique.

***Proof*** Suppose that $A$ did have two inverses $B$ and $C$. Consider the product $BAC$. If we group $B$ and $A$ together,

$$BAC = (BA)C = IC = C$$

since $BA = I$. If we group $A$ and $C$ together,

$$BAC = B(AC) = BI = B$$

since $AC = I$. Thus,

$$C = B$$

●

The uniqueness of the inverse of an invertible matrix $A$ allows us to speak of *the* inverse of $A$ rather than *an* inverse of $A$. It also allows us to introduce a symbol for the inverse of $A$. Henceforth we shall denote the inverse of $A$ by

$$\boxed{A^{-1}}$$

in much the same manner as we use the exponent $-1$ for denoting inverses of functions.[9]

Let us now turn our attention to a method for finding inverses of square matrices. Consider again the matrix

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

Let us think of the inverse of $A$ as an unknown matrix

$$A^{-1} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}.$$

We want to find the entries so that

$$AA^{-1} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} x_{11} + 2x_{21} & x_{12} + 2x_{22} \\ 3x_{11} + 5x_{21} & 3x_{12} + 5x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This gives us a system of equations in $x_{11}$ and $x_{21}$,

$$\begin{aligned} x_{11} + 2x_{21} &= 1 \\ 3x_{11} + 5x_{21} &= 0, \end{aligned}$$

and a system in $x_{12}$ and $x_{22}$,

$$\begin{aligned} x_{12} + 2x_{22} &= 0 \\ 3x_{12} + 5x_{22} &= 1. \end{aligned}$$

---

[9] Do note that $A^{-1}$ does not stand for $1/A$ any more than $\sin^{-1} x$ stands for $1/\sin x$; indeed writing $1/A$ for a matrix $A$ amounts to writing nonsense.

We then will have a unique matrix $A^{-1}$ so that $AA^{-1} = I$ if and only if each of these systems of equations has a unique solution (which occurs if and only if the reduced row-echelon form of $A$ is $I$). Let us solve these systems to see if this is the case. To save writing, notice that since both of these systems have the same coefficient matrices, any set of row operations that leads to the solution of one system leads to the solution of the other system too. Thus we can simultaneously solve these two systems by forming the augmented matrix

$$[A|I] = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 5 & 0 & 1 \end{array}\right]$$

and then using row operations to reduce its left-hand portion $A$ to reduced row-echelon form (which will be $I$):

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

(We have not indicated the row operations here. Can you determine the ones we used?) The right-hand portion of our final augmented matrix tells us $x_{11} = -5$, $x_{21} = 3$, $x_{12} = 2$, $x_{22} = -1$, and hence

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

We must be honest, however. There is a gap in our development here. The procedure we have just illustrated produces a matrix $B$ so that $AB = I$. (We describe this by saying $B$ is a right-hand inverse.) But the inverse of $A$ must also have the property that $BA = I$. (When $BA = I$, we say $B$ is a left-hand inverse.) You can check that the right-hand inverse we have just found for the given $2 \times 2$ matrix $A$ is also a left-hand inverse and hence is $A^{-1}$. Shortly we will fill in this left-hand inverse gap. Once we do so, we then will have that the inverse of a square matrix $A$ (if any) can be found by the following procedure:

1. Form the augmented matrix $[A|I]$ where $I$ is the identity matrix with the same size as $A$.
2. Use row operations to reduce the left-hand portion into reduced row-echelon form.
3. If the augmented matrix after step 2 has the form $[I|B]$, then $B = A^{-1}$; if it does not have this form (or equivalently, if the reduced row-echelon form of $A$ contains a zero row), $A$ does not have an inverse.

Examples 1 and 2 illustrate our procedure for finding inverses.

**EXAMPLE 1** If possible, find the inverse of the following matrix.

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

*Solution* We first form the augmented matrix and then apply row operations:

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

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

$$\to \left[\begin{array}{ccc|ccc} 6 & 0 & 0 & -2 & 7 & -1 \\ 0 & 6 & 0 & 1 & -5 & 2 \\ 0 & 0 & -2 & -1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/3 & 7/6 & -1/6 \\ 0 & 1 & 0 & 1/6 & -5/6 & 1/3 \\ 0 & 0 & 1 & 1/2 & -1/2 & 0 \end{array}\right].$$

Thus the inverse of this matrix is

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

●

**EXAMPLE 2** If possible, find the inverse of the following matrix.

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

*Solution* We again form the augmented matrix and apply row operations:

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

At this point it is apparent that the left-hand portion cannot be reduced to $I$ and hence the matrix in this example does not have an inverse. ●

When the inverse of a square matrix $A$ is known, we can easily find the solutions to a system of linear equations

$$AX = B.$$

If we multiply this matrix equation by $A^{-1}$ on the left, we have

$$A^{-1}AX = A^{-1}B$$

and hence the solution is given by

$$X = A^{-1}B.$$

We use this approach to solve the system in the next example.

**EXAMPLE 3** Solve the system

$$\begin{aligned} 2x + y + 3z &= 6 \\ 2x + y + z &= -12 \\ 4x + 5y + z &= 3. \end{aligned}$$

*Solution* From Example 1, we have that the inverse of the coefficient matrix

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

of this system is

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

The solution is then given by

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B = \begin{bmatrix} -1/3 & 7/6 & -1/6 \\ 1/6 & -5/6 & 1/3 \\ 1/2 & -1/2 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ -12 \\ 3 \end{bmatrix} = \begin{bmatrix} -33/2 \\ 12 \\ 9 \end{bmatrix};$$

that is, $x = -33/2$, $y = 12$, $z = 9$. ●

The following theorem gives us a characterization of when a system of $n$ linear equations in $n$ unknowns has a unique solution.

**THEOREM 1.5** A system $AX = B$ of $n$ linear equations in $n$ unknowns has a unique solution if and only if $A$ is invertible.

*Proof* If $A$ is invertible, the solutions to the system are given by $X = A^{-1}B$ and hence are unique. Conversely, suppose $AX = B$ has a unique solution. Considering the result of Gauss-Jordan elimination on the system, it follows that the reduced row-echelon form of $A$ is $I$. Hence $A$ is invertible. ●

We now develop some mathematics that will justify why $B = A^{-1}$ when we are able to reduce $[A|I]$ to $[I|B]$.

Matrices obtained from an identity matrix $I$ by applying an elementary row operation to $I$ are called **elementary matrices.** We classify elementary matrices into the following three types.

Type 1: An elementary matrix obtained by interchanging two rows of $I$

Type 2: An elementary matrix obtained by multiplying a row of $I$ by a nonzero number

Type 3: An elementary matrix obtained by replacing a row of $I$ by itself plus a multiple of another row of $I$

Some examples of elementary matrices of each of these respective types obtained from the $2 \times 2$ identity matrix are

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{(interchange rows 1 and 2)},$$

$$E_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{(multiply row 1 by 2)},$$

$$E_3 = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \quad \text{(add 3 times row 1 to row 2)}.$$

An interesting fact is that multiplication by elementary matrices on the left of another matrix performs the corresponding row operation on the other matrix. Notice how this works when we multiply $E_1$, $E_2$, and $E_3$ times a $2 \times 2$ matrix:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2a & 2b \\ c & d \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 3a + c & 3b + d \end{bmatrix}.$$

These illustrate the following theorem whose proof we leave as an exercise (Exercise 13).

**THEOREM 1.6** Suppose that $A$ is an $m \times n$ matrix and $E$ is an $m \times m$ elementary matrix.

1. If $E$ is obtained by interchanging rows $i$ and $j$ of $I$, then $EA$ is the matrix obtained from $A$ by interchanging rows $i$ and $j$ of $A$.
2. If $E$ is obtained by multiplying row $i$ of $I$ by $c$, then $EA$ is the matrix obtained from $A$ by multiplying row $i$ of $A$ by $c$.
3. If $E$ is obtained by replacing row $i$ of $I$ by itself plus $c$ times row $j$ of $I$, then $EA$ is the matrix obtained from $A$ by replacing row $i$ of $A$ by itself plus $c$ times row $j$ of $A$.

Of course, we would not use multiplication by elementary matrices to perform row operations—certainly we would just do the row operations! Nevertheless, they do serve as a useful theoretical tool from time to time. Our first instance of this involves seeing why our procedure for finding inverses does in fact produce the inverse. Look at our procedure in the following way. We begin with the augmented matrix $[A|I]$ and use elementary row operations to reduce it to $[I|B]$. Suppose this takes $k$ elementary row operations and $E_1, E_2, \ldots, E_k$ are the elementary matrices that perform the successive row operations. Since performing these elementary operations on $[A|I]$ is the same as performing them on $A$ and $I$ individually, it follows that

$$E_k \cdots E_2 E_1 [A|I] = [E_k \cdots E_2 E_1 A | E_k \cdots E_2 E_1 I] = [I|B].$$

From the right-hand portion of this augmented matrix, we see

$$B = E_k \cdots E_2 E_1.$$

From the left-hand portion, we see

$$E_k \cdots E_2 E_1 A = BA = I.$$

Thus $B$ is not only the right-hand inverse of $A$ as we saw from conception of our method for finding inverses, but is the necessary left-hand inverse too.

Let us proceed to further develop the theory of invertible matrices. We begin with the following theorem.

**THEOREM 1.7** If $A$ and $B$ are invertible matrices of the same size, then $AB$ is invertible and

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

***Proof*** It suffices to show that $B^{-1}A^{-1}$ is the inverse of $AB$. This we do by showing the necessary products are $I$:

$$ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

and

$$B^{-1}A^{-1}AB = B^{-1}IB = B^{-1}B = I.$$ ●

Notice that $(AB)^{-1}$ is not $A^{-1}B^{-1}$. The result of Theorem 1.7 generalizes to products of invertible matrices with more factors as follows: If $A_1, A_2, \ldots, A_n$ are invertible matrices of the same size, then

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

since

$$\begin{aligned}(A_1A_2 \cdots A_n)^{-1} &= (A_2A_3 \cdots A_n)^{-1}A_1^{-1} \\ &= (A_3 \cdots A_n)^{-1}A_2^{-1}A_1^{-1} = \ldots = A_n^{-1} \cdots A_2^{-1}A_1^{-1}.\end{aligned}$$

Next, we consider the invertibility of elementary matrices.

**THEOREM 1.8** If $E$ is an elementary matrix, then $E$ is invertible and:

1. If $E$ is obtained by interchanging two rows of $I$, then $E^{-1} = E$;

2. If $E$ is obtained by multiplying row $i$ of $I$ by a nonzero scalar $c$, then $E^{-1}$ is the matrix obtained by multiplying row $i$ of $I$ by $1/c$;
3. If $E$ is obtained by replacing row $i$ of $I$ by itself plus $c$ times row $j$ of $I$, then $E^{-1}$ is the matrix obtained by replacing row $i$ of $I$ by itself plus $-c$ times row $j$ of $I$.

*Proof* In each part, let $B$ denote the described matrix. We can then prove each part by showing that $EB = I$ and $BE = I$. This can be done by either directly calculating these products or by using Theorem 1.6. We leave these details as exercises (Exercise 14). ●

Up to this point we have been careful to show that $B$ is both a right-hand inverse (that is, $AB = I$) and a left-hand inverse (that is, $BA = I$) when verifying that a square matrix $B$ is the inverse of a square matrix $A$. There are places where a right-hand inverse need not be a left-hand inverse or vice versa. The next theorem tells us that this is not the case for square matrices, however.

**THEOREM 1.9** Suppose that $A$ and $B$ are $n \times n$ matrices such that either $AB = I$ or $BA = I$. Then $A$ is an invertible matrix and $A^{-1} = B$.

*Proof* Let us prove this in the case when $AB = I$; the case $BA = I$ will be left as an exercise (Exercise 17). Suppose $A$ is not invertible. Since the reduced row-echelon form of $A$ is not $I$, there are then elementary matrices $E_1, E_2, \ldots, E_m$ so that $E_1 E_2 \cdots E_m A$ contains a zero row. Consequently

$$E_1 E_2 \cdots E_m AB = E_1 E_2 \cdots E_m \tag{1}$$

contains a zero row. But a matrix with a zero row is not invertible (Exercise 16) and hence $E_1 E_2 \cdots E_m$ is not invertible. This gives us a contradiction since each $E_i$ is invertible (Theorem 1.8) and products of invertible matrices are invertible (Theorem 1.7). Now that we know $A$ is invertible, we can choose $E_1, E_2, \ldots, E_m$ so that

$$E_1 E_2 \cdots E_m A = I.$$

This along with Equation (1) gives us that $B = E_1 E_2 \cdots E_m = A^{-1}$. ●

Because of Theorem 1.9, from now on we will only have to verify one of $AB = I$ or $BA = I$ to see if a square matrix $B$ is the inverse of a square matrix $A$.

Our final result gives a characterization of invertible matrices in terms of elementary matrices.

**THEOREM 1.10** A square matrix $A$ is invertible if and only if $A$ is a product of elementary matrices.

*Proof* If $A$ is a product of elementary matrices, then $A$ is invertible by Theorems 1.7 and 1.8. Conversely, if $A$ is invertible there are elementary matrices $E_1, E_2, \ldots E_m$ so that

$$E_1 E_2 \cdots E_m A = I.$$

Thus,

$$A = E_m^{-1} \cdots E_2^{-1} E_1^{-1} E_1 E_2 \cdots E_m A = E_m^{-1} \cdots E_2^{-1} E_1^{-1}.$$

Since each $E_i^{-1}$ is an elementary matrix by Theorem 1.8, $A$ is a product of elementary matrices. ●

## EXERCISES 1.3

For each of the following matrices, either find the inverse of the matrix or determine that the matrix is not invertible.

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

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

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

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

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

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

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

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

**9.** Use the inverse of the matrix in Exercise 5 to solve the system.

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

**10.** Use the inverse of the matrix in Exercise 7 to solve the system.

$$\begin{aligned} x_1 - x_2 + x_3 + 2x_4 &= 3 \\ x_1 + 2x_2 - x_3 - x_4 &= 5 \\ x_1 - 4x_2 + x_3 + 5x_4 &= 1 \\ 3x_1 + x_2 + x_3 + 6x_4 &= 2 \end{aligned}$$

**11.** For $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, find an elementary matrix $E$ so that:

**a)** $EA = \begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix}$.

**b)** $EA = \begin{bmatrix} 7 & 10 \\ 3 & 4 \end{bmatrix}$.

**c)** $EA = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$

**12.** Express the matrix in Exercise 5 as a product of elementary matrices.

**13.** Prove the following parts of Theorem 1.6.

**a)** Part (1)

**b)** Part (2)

**c)** Part (3)

**14.** Prove the following parts of Theorem 1.8.

**a)** Part (1)

**b)** Part (2)

**c)** Part (3)

**15.** Show that if $A$ is an invertible matrix, then so is $A^{-1}$ and $(A^{-1})^{-1} = A$.

**16.** Show that a square matrix containing a zero row or a zero column is not invertible.

**17.** Complete the proof of Theorem 1.9 by showing if $BA = I$, then $A$ is invertible and $B = A^{-1}$.

**18.** Suppose that $A$ is a noninvertible square matrix. Show that the homogeneous system $AX = 0$ has nontrivial solutions.

**19.** Suppose that $A$ is an invertible matrix and $m$ is a positive integer. Show that $(A^m)^{-1} = (A^{-1})^m$.

**20.** **a)** Suppose that $A$ is an invertible $n \times n$ matrix and $B$ and $C$ are $n \times l$ matrices such that $AB = AC$. Show that $B = C$.

**b)** Give an example to show that the result of part (a) need not hold if $A$ is not invertible.

**21.** Suppose that $A$ and $B$ are $n \times n$ matrices such that $AB$ is invertible. Show that $A$ and $B$ are invertible.

Use the *inverse* command in Maple or the appropriate command in another suitable software package to find the inverses of the matrices in Exercises 22–25 (if possible).

**22.** $\begin{bmatrix} 14 & -25 & 39 & 29 & 6 \\ 58 & -41 & 88 & 24 & 18 \\ 15 & -6 & 31 & -23 & 12 \\ -3 & -22 & -25 & 73 & -24 \\ 3 & 6 & 12 & -24 & 9 \end{bmatrix}$

**23.** $\begin{bmatrix} 8 & -21 & 14 & 26 & -3 \\ 7 & -28 & -6 & 66 & -18 \\ 23 & -12 & 45 & -13 & 15 \\ -9 & 14 & -10 & 26 & 3 \\ 2 & 4 & 8 & -16 & 6 \end{bmatrix}$

**24.** $\begin{bmatrix} 13.2 & -11 & 13 & 27 \\ -4.4 & -10.2 & 5 & -9 \\ 15 & -17.6 & 21.2 & 32.4 \\ -10 & 12 & -14 & -18 \end{bmatrix}$

**25.**[10] $\begin{bmatrix} 18 & -24 & 25 & 27 \\ -12-2\pi & 14-\pi & -17 & -27 \\ 15+3\sqrt{3} & -18-4\sqrt{3} & 21+4\sqrt{3} & 27 \\ -10 & 12 & -14 & -18 \end{bmatrix}$

**26.** Use Maple or another appropriate software package to find $P^{-1}AP$ where

$$A = \begin{bmatrix} -46 & 192 & 36 & -23 & -84 \\ -122 & 437 & 73 & -45 & -194 \\ 45 & -191 & -37 & 22 & 84 \\ -120 & 438 & 74 & -48 & -193 \\ -200 & 686 & 110 & -67 & -306 \end{bmatrix},$$

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

## 1.4 Special Matrices and Additional Properties of Matrices

You already have seen some special types of matrices: zero matrices, identity matrices, and elementary matrices. There are some other special forms of matrices that will come up in our future work. One such type are **diagonal matrices,** which are square matrices whose off diagonal entries are zero. The matrix

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

is an example of a diagonal matrix. We will write

$$\boxed{\operatorname{diag}(d_1, d_2, \dots, d_n)}$$

[10] In Maple a square root such as $\sqrt{3}$ is indicated by typing `sqrt(3)`; $\pi$ is indicated by typing `pi`. Products require a multiplication star so that $-12 - 2\pi$ is typed as `-12-2*pi`; likewise, $15 + 3\sqrt{3}$ is typed as `15+3*sqrt(3)`.

to indicate an $n \times n$ diagonal matrix

$$\begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix}.$$

For instance, the diagonal matrix $A$ would be written as

$$A = \operatorname{diag}(2, -4, 7, 3)$$

in this notation. Some easily verified properties of diagonal matrices whose proofs we leave as exercises (Exercise 21) are listed in Theorem 1.11.

**THEOREM 1.11** Suppose that $A$ and $B$ are $n \times n$ diagonal matrices

$$A = \operatorname{diag}(a_1, a_2, \dots, a_n) \quad \text{and} \quad B = \operatorname{diag}(b_1, b_2, \dots, b_n).$$

1. $A + B = \operatorname{diag}(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$.
2. $AB = \operatorname{diag}(a_1 b_1, a_2 b_2, \dots, a_n b_n)$.
3. $A$ is invertible if and only if each $a_i \neq 0$. Further, if each $a_i \neq 0$, $A^{-1} = \operatorname{diag}(1/a_1, 1/a_2, \dots, 1/a_n)$.

Two other special types of square matrices are **triangular matrices,** which come in two forms: **upper triangular matrices** in which all entries below the diagonal are zero and **lower triangular matrices** in which all entries above the diagonal are zero. The matrix

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

is an example of an upper triangular matrix, and the matrix

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

is an example of a lower triangular matrix. Some easily seen properties of triangular matrices whose proofs are left as exercises (Exercise 22) are listed in Theorem 1.12.

**THEOREM 1.12** Suppose that $A$ and $B$ are $n \times n$ triangular matrices.

1. If $A$ and $B$ are both upper triangular, then so is $A + B$; if $A$ and $B$ are both lower triangular, then so is $A + B$.

2. If $A$ and $B$ are both upper triangular, then so is $AB$; if $A$ and $B$ are both lower triangular, then so is $AB$.
3. $A$ is invertible if and only if each of the diagonal entries of $A$ is nonzero.

The **transpose** of a matrix $A$, denoted $A^T$, is the matrix obtained by interchanging the rows and columns of $A$; to put it another way, if $A = [a_{ij}]$ is an $m \times n$ matrix, then $A^T$ is the $n \times m$ matrix with entries

$$\boxed{\text{ent}_{ij}(A^T) = a_{ji}.}$$

For instance, the transpose of the matrix

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

is

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

In the next theorem we list some basic properties of transposes of matrices.

**THEOREM 1.13** If $A$ and $B$ are matrices so that the indicated sum or product is defined and $c$ is a scalar, then:

1. $(A^T)^T = A$.
2. $(A + B)^T = A^T + B^T$.
3. $(cA)^T = cA^T$.
4. $(AB)^T = B^T A^T$.
5. $(A^T)^{-1} = (A^{-1})^T$.

*Proof* Part (4) is the most difficult to prove. We will prove it here and leave proofs of the remaining parts as exercises (Exercise 23). Suppose $A$ is an $m \times n$ matrix and $B$ is an $n \times l$ matrix. The matrix $(AB)^T$ is an $l \times m$ matrix whose $ij$-entry is

$$\text{ent}_{ij}((AB)^T) = \text{ent}_{ji}(AB) = \sum_{k=1}^{n} a_{jk} b_{ki}. \tag{1}$$

$B^T A^T$ is also an $l \times m$ matrix whose $ij$-entry is

$$\text{ent}_{ij}(B^T A^T) = \sum_{k=1}^{n} \text{ent}_{ik}(B^T)\text{ent}_{kj}(A^T) = \sum_{k=1}^{n} b_{ki} a_{jk}. \tag{2}$$

As the results in Equations (1) and (2) are the same, we have $(AB)^T = B^T A^T$. ●

A matrix $A$ is called a **symmetric matrix** if

$$A^T = A.$$

The matrix

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

is a symmetric matrix; the matrix

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

is not symmetric. Notice that a symmetric matrix must necessarily be a square matrix. We leave the proofs of the properties of symmetric matrices listed in the following theorem as exercises (Exercise 24).

**THEOREM 1.14** Suppose $A$ and $B$ are matrices of the same size.

1. If $A$ and $B$ are symmetric matrices, then so is $A + B$.
2. If $A$ is a symmetric matrix, then so is $cA$ for any scalar $c$.
3. $A^T A$ and $AA^T$ are symmetric matrices.
4. If $A$ is an invertible symmetric matrix, then $A^{-1}$ is a symmetric matrix.

We often have applied a finite number of elementary row operations to a matrix $A$ obtaining a matrix $B$. In this setting, we say that the matrix $A$ is **row equivalent** to the matrix $B$. We frequently will use this terminology making statements such as "a square matrix $A$ is invertible if and only if $A$ is row equivalent to $I$" or "a system of linear equations $AX = B$ has no solution if and only if the augmented matrix $[A|B]$ is row equivalent to an augmented matrix containing a row that consists of zero entries in the left-hand portion and a nonzero entry in the right-hand portion." To give a couple more illustrations, notice that the terminology of reduced row-echelon form that we introduced for coefficient matrices of linear systems can be applied to any matrix; that is, a matrix is in reduced row-echelon form if:

1. Any zero rows appear at the bottom.
2. The first nonzero entry of a nonzero row is 1.
3. The leading 1 of a nonzero row appears to the right of the leading 1 of any preceding row.
4. All other entries of a column containing a leading 1 are zero.

Using ideas we shall develop in the next chapter, it can be shown that the reduced row-echelon form of a matrix is unique, so we may speak of the reduced row-echelon form of a matrix and say "a matrix is row equivalent to its reduced row-echelon form." When we do not require property 4, the matrix is in row-echelon form. Row-echelon form is not unique, so we would have to say "a matrix is row equivalent to any of its row-echelon forms."

The notion of row equivalence gives us a relationship between matrices of the same size that possesses the properties listed in Theorem 1.15.

**THEOREM 1.15**

1. Every matrix $A$ is row equivalent to itself.
2. If a matrix $A$ is row equivalent to a matrix $B$, then $B$ is row equivalent to $A$.
3. If a matrix $A$ is row equivalent to a matrix $B$ and $B$ is row equivalent to a matrix $C$, then $A$ is row equivalent to $C$.

We will leave the proof of Theorem 1.15 as another exercise (Exercise 25). In Theorem 1.15, the first property is called the **reflexive property,** the second is called the **symmetric property,** and the third is called the **transitive property** of row equivalence. A relation that has all three of these properties is called an **equivalence relation.** Equivalence relations are important types of relations occurring frequently throughout mathematics. A couple of other important equivalence relations you have encountered before are congruence and similarity of triangles. Not all relations are equivalence relations. The inequality $<$ on the set of real numbers $\mathbb{R}$ is not an equivalence relation since it is neither reflexive nor symmetric (although it is transitive).

We conclude this section by pointing out that just as we perform elementary row operations, it is also possible to perform **elementary column operations** on a matrix. As you might expect, these are the following:

1. Interchange two columns.
2. Multiply a column by a nonzero number.
3. Replace a column by itself plus a multiple of another column.

When we apply a finite number of elementary column operations to a matrix $A$ obtaining a matrix $B$, we say $A$ is **column equivalent** to $B$.

Many (we authors included) are so used to performing row operations that they find it awkward to perform column operations. For the most part, we will avoid using column operations in this book. But we will see one place in the next chapter where column operations arise. If you too feel uncomfortable doing them, notice that column operations may be performed on a matrix $A$ by first performing the corresponding row operations on $A^T$ and then transposing again.

## EXERCISES 1.4

Let $A$ be the matrix

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

Find:

**1.** $A^2$

**2.** $A^{-1}$

**3.** $A^5$

**4.** $(A^{-1})^4$

Let $A$ and $B$ be the matrices

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

Find:

**5.** $AB$

**6.** $A^{-1}$.

Let $A$ and $B$ be the matrices

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

If possible, find:

**7.** $A^T$.

**8.** $B^T$.

**9.** $A^T + 4B$.

**10.** $2A - 5B^T$.

**11.** $(AB)^T$.

**12.** $B^T A^T$.

**13.** $A^T B^T$.

**14.** $A^T A$.

Let $A$ and $B$ be the matrices

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

Determine which of the following are symmetric matrices in Exercises 15–20.

**15.** $A$

**16.** $B$

**17.** $A + B$

**18.** $A^{-1}$

**19.** $BB^T$

**20.** $B^T B$

**21.** Prove the following parts of Theorem 1.11.

**a)** Part (1)
**b)** Part (2)
**c)** Part (3)

**22.** Prove the following parts of Theorem 1.12.

**a)** Part (1)
**b)** Part (2)
**c)** Part (3)

**23.** Prove the following parts of Theorem 1.13.

**a)** Part (1)
**b)** Part (2)
**c)** Part (3)
**d)** Part (5)

**24.** Prove the following parts of Theorem 1.14.

**a)** Part (1)
**b)** Part (2)
**c)** Part (3)
**d)** Part (4)

**25.** Prove the following parts of Theorem 1.15.

**a)** Part (1)
**b)** Part (2)
**c)** Part (3)

**26.** Show that two matrices $A$ and $B$ are row equivalent if and only if there is an invertible matrix $C$ so that $CA = B$.

**27.** Show that two matrices $A$ and $B$ are row equivalent if and only if they have the same reduced row-echelon form.

**28.** Use the result of Exercise 27 to show that the matrices

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

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

are row equivalent.

**29.** Show that any two invertible matrices of the same size are row equivalent.

**30.** Just as we speak of the reduced row-echelon form of a matrix, we may speak of the reduced column-echelon form of a matrix. Write a statement that describes the form of a reduced column-echelon matrix.

**31.** Find the reduced column-echelon form of the matrix

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

**32.** Find $A^3$ for

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

**33.** A square matrix $A$ is called a **nilpotent** matrix if there is a positive integer $m$ so that $A^m = 0$. Prove that if $A$ is a triangular $n \times n$ matrix whose diagonal entries are all zero, then $A$ is a nilpotent matrix by showing $A^n = 0$.

**34.** The Maple command for finding the transpose of a matrix is *transpose*. Use Maple or another appropriate software package to find $AA^T - 3A^T$ for

$$A = \begin{bmatrix} 2 & 1 & -3 & 4 & 1 \\ -2 & 1 & 1 & 0 & 7 \\ 6 & 2 & -8 & 4 & 9 \\ -2 & 0 & 2 & -6 & -2 \\ 10 & 3 & -13 & 14 & 12 \end{bmatrix}.$$

**35. a)** Either of the Maple commands *rref* or *gaussjord* can be used to find the reduced row-echelon form of a matrix. Use them or corresponding commands in another appropriate software package to find the reduced row-echelon form of matrix $A$ in Exercise 34.

**b)** Apply the *gausselim* command of Maple to matrix $A$. Describe the form of the answer Maple is giving us.

**c)** Is $A$ an invertible matrix?

**36.** How could Maple commands or commands of another appropriate software package be used to find the reduced column-echelon form of a matrix? Use them to find the reduced column-echelon form of matrix $A$ in Exercise 34.

## 1.5 DETERMINANTS

You already may have had some exposure to determinants. For instance, you might have encountered them for finding cross products of three-dimensional vectors. Or perhaps you have learned a method for finding solutions to some systems of linear equations involving determinants called Cramer's rule. (We shall discuss this rule in the next section.) The Jacobian of a transformation is yet another example of a determinant you might have encountered. Even if you have had some prior experience with determinants, however, it is likely that your knowledge of them is not thorough. The purpose of this and the remaining sections of this chapter is to a give a thorough treatment of determinants.

There are a number of equivalent ways of defining determinants. We are going to begin with a process called a cofactor expansion approach. Suppose that $A$ is a square matrix:

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

The **minor** of the entry $a_{ij}$ of $A$, denoted $M_{ij}$, is the matrix obtained from $A$ by deleting row $i$ and column $j$ from $A$.[11] (Of course, this only makes sense if $n \geq 2$.) For instance, if $A$ is a $3 \times 3$ matrix,

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

[11] There are two different ways in which the concept of a minor is commonly defined. Quite a few books as well as Maple take our approach. Many other books, however, prefer to define the minor of $a_{ij}$ as the determinant of the matrix obtained from $A$ by deleting row $i$ and column $j$. In other words, minors in these other books are the determinants of our minors.

some minors of $A$ are:

$$M_{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{32} = \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}.$$

We are going to define the determinant of an $n \times n$ matrix $A$ (also called an $n \times n$ determinant), denoted $\det(A)$, with what is called an inductive definition as follows: If $A$ is a $1 \times 1$ matrix,

$$\boxed{A = [a_{11}],}$$

we define the determinant of $A$ to be its entry,

$$\boxed{\det(A) = a_{11}.}$$

If $n \geq 2$, the determinant of $A = [a_{ij}]$ is defined to be

$$\boxed{\begin{aligned} \det(A) &= a_{11}\det(M_{11}) - a_{12}\det(M_{12}) + a_{13}\det(M_{13}) - \cdots + (-1)^{1+n}a_{1n}\det(M_{1n}) \\ &= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(M_{1j}). \end{aligned}}$$

In effect, we have reduced $\det(A)$ to the determinants of the smaller matrices $M_{1j}$, which (by repeating this reduction procedure if necessary) we already know how to find.

To illustrate, if $A$ is a $2 \times 2$ matrix,

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

our definition tells us that

$$\det(A) = a_{11}\det([a_{22}]) - a_{12}\det([a_{21}]) = a_{11}a_{22} - a_{12}a_{21}.$$

The products $a_{11}a_{22}$ and $a_{12}a_{21}$ are often referred to as cross products; the determinant of a $2 \times 2$ matrix can then be easily remembered as being the cross product of the diagonal entries minus the cross product of the off diagonal entries. Determinants of $n \times n$ matrices when $n \geq 2$ are also indicated by putting vertical bars around the entries of the matrix.[12] If we do this for a $2 \times 2$ matrix, our determinant formula becomes

$$\boxed{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.} \qquad \textbf{(1)}$$

[12] We do not do this for a $1 \times 1$ matrix $A = [a_{11}]$ to avoid confusion with the absolute value of $a_{11}$.

For instance,

$$\begin{vmatrix} 2 & -3 \\ 5 & 6 \end{vmatrix} = 2 \cdot 6 - (-3) \cdot 5 = 27.$$

For a $3 \times 3$ matrix, our definition tells us

$$\boxed{\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}}$$

where the remaining $2 \times 2$ determinants can be found by the formula in Equation (1). For instance,

$$\begin{aligned} \begin{vmatrix} 2 & 3 & -2 \\ -1 & 6 & 3 \\ 4 & -2 & 1 \end{vmatrix} &= 2 \begin{vmatrix} 6 & 3 \\ -2 & 1 \end{vmatrix} - 3 \begin{vmatrix} -1 & 3 \\ 4 & 1 \end{vmatrix} + (-2) \begin{vmatrix} -1 & 6 \\ 4 & -2 \end{vmatrix} \\ &= 2(6+6) - 3(-1-12) - 2(2-24) = 107. \end{aligned}$$

Continuing, a $4 \times 4$ determinant can be found by forming the alternating sum and difference of the entries of the first row of the matrix times the determinants of their respective minors, which are $3 \times 3$ determinants and so on.

The **cofactor** of an entry $a_{ij}$ of an $n \times n$ matrix $A$ where $n \geq 2$, denoted $C_{ij}$, is

$$\boxed{C_{ij} = (-1)^{i+j} \det(M_{ij})}$$

where $M_{ij}$ is the minor of $a_{ij}$. Some cofactors of a $3 \times 3$ matrix,

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

are:

$$\begin{aligned} C_{11} &= (-1)^{1+1} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \\ C_{12} &= (-1)^{1+2} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} = - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}. \\ C_{32} &= (-1)^{3+2} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \end{aligned}$$

The signs of the cofactors can be easily remembered by noting that they form the checkerboard pattern:

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

Our formula for finding the determinant of an $n \times n$ matrix $A = [a_{ij}]$ where $n \geq 2$ can be written in terms of cofactors as

$$\det(A) = \sum_{j=1}^{n} a_{1j} C_{1j},$$

which we will refer to as the *cofactor expansion about the first row* or simply the *expansion about the first row*. What is remarkable is that the same procedure may be followed for any row or column.

**THEOREM 1.16** If $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$, then for any $1 \leq i \leq n$

$$\boxed{\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \text{(cofactor expansion about the } i\text{th row)}}$$

or any $1 \leq j \leq n$,

$$\boxed{\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{(cofactor expansion about the } j\text{th column).}}$$

To illustrate this theorem, earlier we had found

$$\begin{vmatrix} 2 & 3 & -2 \\ -1 & 6 & 3 \\ 4 & -2 & 1 \end{vmatrix} = 107$$

by using the cofactor expansion about the first row. Notice that we get the same result if, for instance, we expand about the third row,

$$\begin{vmatrix} 2 & 3 & -2 \\ -1 & 6 & 3 \\ 4 & -2 & 1 \end{vmatrix} = 4 \begin{vmatrix} 3 & -2 \\ 6 & 3 \end{vmatrix} - (-2) \begin{vmatrix} 2 & -2 \\ -1 & 3 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ -1 & 6 \end{vmatrix}$$
$$= 4(21) + 2(4) + 15 = 107,$$

or the second column,

$$\begin{vmatrix} 2 & 3 & -2 \\ -1 & 6 & 3 \\ 4 & -2 & 1 \end{vmatrix} = -3 \begin{vmatrix} -1 & 3 \\ 4 & 1 \end{vmatrix} + 6 \begin{vmatrix} 2 & -2 \\ 4 & 1 \end{vmatrix} - (-2) \begin{vmatrix} 2 & -2 \\ -1 & 3 \end{vmatrix}$$
$$= -3(-13) + 6(10) + 2(4) = 107.$$

So that you first gain an overview of the theory of determinants, we are going to postpone many of the proofs of our results about determinants to the end of this chapter in Section 1.7. The proof of Theorem 1.16 is the first of these we postpone.

You may raise the question: Why is it important to be able to expand a determinant about any row or column? One reason is that sometimes we can make the work easier by choosing a particular row or column. Consider the following example.

**EXAMPLE 1** Evaluate the determinant

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

*Solution* Since the third column contains three zeros, let us begin by expanding about it obtaining

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

The remaining $3 \times 3$ determinant is now quickly found by expanding about its first column. Doing so, we get our answer:

$$(-2) \cdot 7 \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = 84.$$

●

The fact that Theorem 1.16 allows us to expand about any row or column gives us some cases in which determinants can be quickly found. One of these is described in the following corollary.

**COROLLARY 1.17** If an $n \times n$ matrix $A$ has a zero row or zero column, then $\det(A) = 0$.

*Proof* This result is immediate if $A$ is a $1 \times 1$ matrix, so assume $n \geq 2$. Expanding the determinant about the row or column whose entries are all zero, the result follows. ●

Corollary 1.18 describes another case in which we can quickly see the value of a determinant.

**COROLLARY 1.18** The determinant of a triangular matrix is the product of its diagonal entries.

*Proof* We will do the upper triangular case here. The lower triangular case will be left as an exercise (Exercise 15). Suppose $A$ is an upper triangular matrix:

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

Again the result is immediate if $n = 1$, so assume $n \geq 2$. Expanding $\det(A)$ about the first column, we have

$$\det(A) = a_{11} \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix}.$$

If we continue to expand each remaining determinant about the first column, we obtain

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

as desired.[13] ●

In Section 1.7 we shall use the fact that the determinant of a square matrix can be found by performing a cofactor expansion about either its first row or first column to obtain the result stated in Theorem 1.19.

**THEOREM 1.19** If $A$ is an $n \times n$ matrix, then

$$\det(A^T) = \det(A).$$

As the square matrices grow in size, the calculations of determinants using cofactor expansions become lengthy. We next develop a more efficient method for calculating determinants of large square matrices involving row operations. To use this approach, we will have to know the effects of elementary row operations on a determinant. These effects are listed in the following theorem.

**THEOREM 1.20** Suppose that $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$.

1. If $B$ is a matrix obtained from $A$ by interchanging two rows of $A$, then $\det(B) = -\det(A)$.
2. If $B$ is a matrix obtained from $A$ by multiplying a row of $A$ by a scalar $c$, then $\det(B) = c\det(A)$.
3. If $B$ is a matrix obtained from $A$ by replacing a row of $A$ by itself plus a multiple of another row of $A$, then $\det(B) = \det(A)$.

Because $\det(A) = \det(A^T)$, we can replace the elementary row operation by the corresponding elementary column operation in each part of Theorem 1.20 and obtain

[13] If you are familiar with mathematical induction, you will notice that this proof could be more effectively written by using induction on $n$.

the same result. The proof of Theorem 1.20 is another one that we postpone until Section 1.7.

We use row operations to calculate the determinant of a square matrix $A$ in a manner similar to the way we use them to solve a system of linear equations with Gaussian elimination: Use row operations to reduce to row-echelon form with the exception of making the leading entries one. This reduced matrix is an upper triangular matrix whose determinant is easily found by Corollary 1.18.

Of course, when we apply the row operations, we must be careful to compensate for their effects. The following example illustrates how we may do this.

**EXAMPLE 2** Find the determinant of the matrix

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

***Solution*** We begin with a first set of row operations toward the goal of getting $A$ into an upper triangular form. To help you follow our work, we have indicated the row operations that will be performed.

$$\det(A) = \begin{vmatrix} 1 & -1 & 2 & 3 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & 1 & 4 \end{vmatrix} \begin{matrix} \\ R_2 - 2R_1 \\ R_3 - R_1 \\ R_4 - R_1 \end{matrix}$$

By part (3) of Theorem 1.20, performing these elementary row operations does not affect the determinant and hence

$$\det(A) = \begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -2 & -5 \\ 0 & 2 & -3 & -5 \\ 0 & 0 & -1 & 1 \end{vmatrix} \begin{matrix} \\ \\ 3R_3 - 2R_2 \\ \\ \end{matrix}$$

where again we have indicated the row operation we will perform in the next step. This row operation is a combination of two elementary row operations: (1) multiplying the third row by 3 and (2) adding $-2$ times the second row. The second of these has no effect, but the first changes the determinant by a factor of 3 by part (2) of Theorem 1.20. Notice how we multiply by 1/3 to compensate, obtaining

$$\det(A) = \frac{1}{3} \begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -2 & -5 \\ 0 & 0 & -5 & -5 \\ 0 & 0 & -1 & 1 \end{vmatrix} \begin{matrix} \\ \\ R_3 \leftrightarrow R_4 \\ \\ \end{matrix}.$$

The next indicated row operation we will perform changes the sign by part (1) of Theorem 1.20. Observe how we compensate:

$$\det(A) = -\frac{1}{3}\begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -2 & -5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -5 & -5 \end{vmatrix} \begin{matrix} \\ \\ \\ R_4 - 5R_3 \end{matrix}.$$

Performing our last indicated elementary row operation, which has no effect on the determinant, we have a desired upper triangular form from which we easily compute the determinant:

$$\det(A) = -\frac{1}{3}\begin{vmatrix} 1 & -1 & 2 & 3 \\ 0 & 3 & -2 & -5 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -10 \end{vmatrix} = -\frac{1}{3} \cdot 1 \cdot 3(-1)(-10) = -10.$$

## EXERCISES 1.5

Find $\det(A)$ for

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

by expanding about the indicated row or column in Exercises 1–6.

**1.** Row 1 **2.** Row 2 **3.** Row 3

**4.** Column 1 **5.** Column 2 **6.** Column 3

Find the determinants in Exercises 7–10 by expanding about appropriate rows or columns.

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

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

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

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

Use row operations to find the determinants in Exercises 11–14.

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

**12.** $\begin{vmatrix} 2 & -1 & 3 & 1 \\ -1 & 2 & -1 & 4 \\ 1 & -1 & 3 & 1 \\ 3 & 2 & -1 & 5 \end{vmatrix}$

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

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

**15.** Complete the proof of Corollary 1.18 by showing that the determinant of a lower triangular matrix is the product of its diagonal entries.

**16.** Suppose that $A$ is a square matrix in which one of the rows is a scalar multiple of another. Show that $\det(A) = 0$. Does the same result hold if $A$ has one of its columns being a scalar multiple of another? Why or why not?

The Maple command for finding the determinant of a square matrix is *det*. Use this Maple command or the corresponding command in another appropriate software package to find the determinants in Exercises 17 and 18.

**17.** $\begin{vmatrix} 7 & -11 & 4 & 6 & 3 \\ -5 & 5 & 4 & -2 & 3 \\ 6 & 12 & -14 & 3 & 5 \\ 13 & -1 & 3 & 2 & -4 \\ 3 & -2 & 8 & -7 & 6 \end{vmatrix}$

**18.** $\begin{vmatrix} \pi - 3 & 7 & -4 & 6 \\ 12 & \pi - 4 & 2 & -5 \\ 3 & -8 & \pi - 8 & 7 \\ 6 & 4 & -5 & \pi + 2 \end{vmatrix}$

## 1.6 Further Properties of Determinants

In this section we shall see some additional properties of determinants. Our first such property gives us a relationship between the value of the determinant of a square matrix and its invertibility.

**THEOREM 1.21** A square matrix $A$ is invertible if and only if $\det(A) \neq 0$.

*Proof* Suppose that $A$ is invertible. Then $A$ is row equivalent to $I$. From Theorem 1.20, we can see that the determinants of two row equivalent matrices are nonzero multiples of one another. Thus $\det(A)$ is a nonzero multiple of $\det(I) = 1$ and hence $\det(A) \neq 0$. Conversely, suppose $\det(A) \neq 0$. Were $A$ not invertible, the reduced row-echelon form of $A$, let us call this matrix $B$, would have a zero row. But then $\det(B) = 0$ by Corollary 1.17. Since $\det(A)$ is a multiple of $\det(B)$, we obtain $\det(A) = 0$, which is a contradiction. ●

Theorem 1.21 is useful for determining if a square matrix $A$ is invertible when $\det(A)$ can be quickly calculated. For instance, since

$$\begin{vmatrix} -2 & 3 & 0 \\ 4 & 10 & 2 \\ -5 & 7 & 0 \end{vmatrix} = -2 \begin{vmatrix} -2 & 3 \\ -5 & 7 \end{vmatrix} = -2 \neq 0,$$

the matrix

$$\begin{bmatrix} -2 & 3 & 0 \\ 4 & 10 & 2 \\ -5 & 7 & 0 \end{bmatrix}$$

is invertible.

Our next major objective will be to obtain a result about the determinant of a product of matrices. To reach this objective, we first prove a couple of lemmas about elementary matrices.

**LEMMA 1.22** Suppose that $E$ is an elementary matrix.

1. If $E$ is obtained from $I$ by interchanging two rows of $I$, then $\det(E) = -1$.
2. If $E$ is obtained from $I$ by multiplying a row of $I$ by a nonzero scalar $c$, then $\det(E) = c$.
3. If $E$ is obtained from $I$ by replacing a row of $I$ by itself plus a multiple of another row of $I$, then $\det(E) = 1$.

***Proof*** These are all consequences of Theorem 1.20. For example, part (1) follows because $\det(E) = -\det(I) = -1$ by part (1) of Theorem 1.20. We leave the proofs of the remaining two parts as exercises (Exercise 13). ●

**LEMMA 1.23** If $A$ is an $n \times n$ matrix and $E$ is an $n \times n$ elementary matrix, then

$$\det(EA) = \det(E)\det(A).$$

More generally, if $E_1, E_2, \dots, E_m$ are $n \times n$ elementary matrices, then

$$\det(E_1 E_2 \cdots E_m A) = \det(E_1 E_2 \cdots E_m)\det(A).$$

***Proof*** The first part is an immediate consequence of Theorem 1.6, which tells us that left multiplication by an elementary matrix performs an elementary row operation, Theorem 1.20, which tells us the effect of an elementary row operation on a determinant, and Lemma 1.22. The second part follows by repeated use of the first part:

$$\begin{aligned}
\det(E_1 E_2 \cdots E_m A) &= \det(E_1)\det(E_2 \cdots E_m A) \\
&\ \ \vdots \\
&= \det(E_1)\det(E_2)\cdots\det(E_m)\det(A) \\
&= \det(E_1 E_2)\det(E_3)\cdots\det(E_m)\det(A) \\
&\ \ \vdots \\
&= \det(E_1 E_2 \cdots E_m)\det(A).
\end{aligned}$$ ●

Now we are ready to prove Theorem 1.24.

**THEOREM 1.24** If $A$ and $B$ are $n \times n$ matrices,

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

***Proof*** We will break our proof into two cases: one in which $A$ is invertible and the other in which it is not. If $A$ is invertible, then $A$ is a product of elementary matrices by Theorem 1.10 and the result now follows from Lemma 1.23.

Suppose $A$ is not invertible. Then $\det(A) = 0$ by Theorem 1.21 and consequently

$$\det(A)\det(B) = 0.$$

Since $A$ is not invertible, $A$ is row equivalent to a matrix with a zero row. Thus there are elementary matrices $E_1, E_2, \ldots, E_m$ so that $E_1E_2\cdots E_mA$ has a zero row. Then $E_1E_2\cdots E_mAB$ has a zero row and consequently

$$\det(E_1E_2\cdots E_mAB) = 0.$$

Hence

$$\det(E_1E_2\cdots E_m)\det(AB) = 0,$$

which implies

$$\det(AB) = 0$$

since $\det(E_1E_2\cdots E_m) \neq 0$, and we again have the desired result that $\det(AB) = \det(A)\det(B)$. ●

As a consequence of Theorem 1.24, we have the following corollary.

**COROLLARY 1.25** If $A$ is an invertible matrix, $\det(A^{-1}) = 1/\det(A)$.

*Proof* Since

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

it follows that $\det(A^{-1}) = 1/\det(A)$. ●

If $A = [a_{ij}]$ is an $n \times n$ matrix, the $n \times n$ matrix with entries the cofactors of $A$,

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

is called the **cofactor matrix** of $A$. The transpose of this cofactor matrix is called the **adjoint** of $A$ and is denoted $\operatorname{adj}(A)$; that is,

$$\boxed{\operatorname{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.}$$

For example, the cofactor matrix of

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

is

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

and

$$\text{adj}(A) = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}.$$

A curious feature of the adjoint of a square matrix is Theorem 1.26.

**THEOREM 1.26** If $A$ is a square matrix,

$$A\text{adj}(A) = \text{adj}(A)A = \det(A)I.$$

The proof of Theorem 1.26 is another one we postpone until the next section. Notice this theorem is telling us that $\text{adj}(A)$ is almost the inverse of $A$. Indeed, we have Corollary 1.27.

**COROLLARY 1.27** If $A$ is an invertible matrix, then

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

For instance, for the matrix

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

just prior to Theorem 1.26, Corollary 1.27 tells us that

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A) = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$$

As a rule, however, Corollary 1.27 is not an efficient way of finding inverses of matrices because of all the determinants that must be calculated. The approach used in Section 1.3 is usually the better way to go. This adjoint method for finding inverses is used primarily as a theoretical tool.

To develop our final property of this section, consider a linear system with two equations and two unknowns:

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2. \end{aligned}$$

Let us start to solve this system. We could use our matrix method, but we will not bother for such a small system. Instead, let us eliminate $y$ by multiplying the first equation by $a_{22}$ and subtracting $a_{12}$ times the second equation. This gives us

$$(a_{11}a_{22} - a_{21}a_{12})x = a_{22}b_1 - a_{12}b_2.$$

If $a_{11}a_{22} - a_{12}a_{21} \neq 0$, we then have

$$x = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}.$$

Carrying out similar steps to find $y$ (try it), we find

$$y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}.$$

Our formulas for $x$ and $y$ can be conveniently expressed in terms of determinants. If we let $A$ denote the coefficient matrix of the system,

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

and $A_1$ and $A_2$ be the matrices obtained from $A$ by replacing the first and second columns, respectively, of $A$ by the column

$$B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

so that

$$A_1 = \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix},$$

then

$$x = \frac{\det(A_1)}{\det(A)}, \qquad y = \frac{\det(A_2)}{\det(A)}.$$

We have just discovered what is known as **Cramer's rule,** named after Gabriel Cramer (1704–1752). Variations of this rule were apparently known prior to Cramer, but his name became attached to it when it appeared in his 1750 work *Introduction à l'analyse des lignes courbes algébriques.* The rule extends to any system of $n$ linear equations in $n$ unknowns provided the determinant of the coefficient matrix is nonzero (or equivalently, provided the coefficient matrix is invertible).

**THEOREM 1.28** **(Cramer's Rule)** Suppose that $AX = B$ is a system of $n$ linear equations in $n$ unknowns such that $\det(A) \neq 0$. Let $A_1$ be the matrix obtained from $A$ by replacing the first column of $A$ by $B$, $A_2$ be the matrix obtained from $A$ by replacing the second column of $A$ by $B, \ldots, A_n$ be the matrix obtained from $A$ by replacing the $n$th column of $A$ by $B$. Then

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

The proof of Cramer's rule for a general positive integer $n$ will be given in the next section. Let us look at an example using it.

**EXAMPLE 1** Use Cramer's rule to solve the system

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

***Solution*** We first calculate the necessary determinants (the details of which we leave out).

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

$$\det(A_1) = \begin{vmatrix} 2 & 1 & -1 \\ 3 & -1 & 1 \\ 1 & -2 & 1 \end{vmatrix} = 5$$

$$\det(A_2) = \begin{vmatrix} 1 & 2 & -1 \\ 2 & 3 & 1 \\ 1 & 1 & 1 \end{vmatrix} = 1$$

$$\det(A_3) = \begin{vmatrix} 1 & 1 & 2 \\ 2 & -1 & 3 \\ 1 & -2 & 1 \end{vmatrix} = 0$$

Our solution is then:

$$x = \frac{5}{3}, \qquad y = \frac{1}{3}, \qquad z = \frac{0}{3} = 0.$$

●

As is the case with the adjoint method for finding inverses, Cramer's rule is usually not a very efficient way to solve a linear system. The Gauss-Jordan or Gaussian elimination methods are normally much better. Note too that Cramer's rule may not be used should the system not have the same number of equations as unknowns or, if it does have as many equations as unknowns, should the coefficient matrix not be invertible. One place where it is often convenient to use Cramer's rule, however, is if the coefficients involve functions such as in the following example.

**EXAMPLE 2** Solve the following system for $x$ and $y$:

$$\begin{aligned} xe^{2t}\sin t - ye^{2t}\cos t &= 1 \\ 2xe^{2t}\cos t + 2ye^{2t}\sin t &= t. \end{aligned}$$

***Solution*** In this example, we have:

$$\det(A) = \begin{vmatrix} e^{2t}\sin t & -e^{2t}\cos t \\ 2e^{2t}\cos t & 2e^{2t}\sin t \end{vmatrix} = 2e^{4t}\sin^2 t + 2e^{4t}\cos^2 t = 2e^{4t}$$

$$\det(A_1) = \begin{vmatrix} 1 & -e^{2t}\cos t \\ t & 2e^{2t}\sin t \end{vmatrix} = 2e^{2t}\sin t + te^{2t}\cos t$$

$$\det(A_2) = \begin{vmatrix} e^{2t}\sin t & 1 \\ 2e^{2t}\cos t & t \end{vmatrix} = te^{2t}\sin t - 2e^{2t}\cos t$$

$$x = e^{-2t}\sin t + \frac{1}{2}te^{-2t}\cos t, \qquad y = \frac{1}{2}te^{-2t}\sin t - e^{-2t}\cos t$$

## EXERCISES 1.6

Use Theorem 1.21 to determine whether the following matrices are invertible.

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

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

Use the adjoint method to find the inverse of the following matrices.

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

Use Cramer's rule to solve the following systems.

**7.** $\begin{aligned} 3x - 4y &= 1 \\ 2x + 3y &= 2 \end{aligned}$ **8.** $\begin{aligned} 7x + y &= 4 \\ 2x - 5y &= 8 \end{aligned}$

**9.** $\begin{aligned} 3x - y + z &= 1 \\ 2x + y - 3z &= 3 \\ x - 2y + z &= 7 \end{aligned}$ **10.** $\begin{aligned} 5x - 4y + z &= 2 \\ 2x - 3y - 2z &= 4 \\ 3x + y + 3z &= 2 \end{aligned}$

**11.** Use Cramer's rule to solve the following system for $x$ and $y$.

$$\begin{aligned} xe^t\sin 2t + ye^t\cos 2t &= t \\ 2xe^t\cos 2t - 2ye^t\sin 2t &= t^2 \end{aligned}$$

**12.** Use Cramer's rule to solve the following system for $x$, $y$, and $z$.

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

**13.** Prove the following parts of Lemma 1.22.

**a)** Part (2)

**b)** Part (3)

**14. a)** An invertible matrix $A$ with integer entries is said to be **unimodular** if $A^{-1}$ also has integer entries. Show that if $A$ is a square matrix with integer entries such that $\det(A) = \pm 1$, then $A$ is a unimodular matrix.

**b)** Prove the converse of the result in part (a); i.e., prove that if $A$ is a unimodular matrix, then $\det(A) = \pm 1$.

**15. a)** Find the determinants of the following matrices.

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

**b)** Find $\det(AB)$, $\det(A^{-1})$, and $\det(B^T A^{-1})$ without finding $AB$, $A^{-1}$, or $B^T A^{-1}$.

**c)** Show that $\det(A + B)$ is not the same as $\det(A) + \det(B)$.

**16.** Show that if $A$ and $B$ are square matrices of the same size, then $\det(AB) = \det(BA)$.

**17. a)** Either of the commands *adj* or *adjoint* can be used in Maple to find the adjoint of a square matrix. Use either one of these Maple commands or corresponding commands in another appropriate software package to find the adjoint of the matrix

$$A = \begin{bmatrix} 1.2 & 2 & -3.1 & -2 \\ 1 & 1.2 & -2 & 2.6 \\ -2.1 & 3.7 & 1 & -4 \\ 2.3 & 4 & -3 & 6.5 \end{bmatrix}.$$

**b)** Use your software package and the result of part (a) to find $adj(A)A$.

**c)** By part (b), what is the value of $\det(A)$?

## 1.7 Proofs of Theorems on Determinants

In this section we will prove those results about determinants whose proofs were omitted in the previous two sections. Many of these proofs will use the technique of mathematical induction, a technique of proof with which we will assume you are familiar.

Recall that we defined the determinant of an $n \times n$ matrix $A = [a_{ij}]$ as the cofactor expansion about the first row:

$$\det(A) = \sum_{j=1}^{n} a_{1j}C_{1j} = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(M_{1j}).$$

As we prove some of our results, we will sometimes have minors of minors; that is, we will have matrices obtained by deleting two rows and two columns from $A$. For notational purposes, let us use

$$M(ij, kl)$$

to denote the matrix obtained from $A$ by deleting rows $i$ and $k$ ($i \neq k$) and columns $j$ and $l$ ($j \neq l$).

Our first theorem about determinants (Theorem 1.16) was that we could expand about any row or column. As a first step toward obtaining this result, we show that we can expand about any row.

**LEMMA 1.29** If $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$, then for any $i$, $1 \leq i \leq n$,

$$\det(A) = \sum_{j=1}^{n} a_{ij}C_{ij}.$$

*Proof* The verification is easy for $n = 2$ and is left as an exercise (Exercise 1). Assume the result is valid for all $k \times k$ matrices and suppose that $A = [a_{ij}]$ is a $(k+1) \times (k+1)$ matrix. There is nothing to show if $i = 1$, so assume $i > 1$. By definition,

$$\det(A) = \sum_{j=1}^{k+1}(-1)^{1+j}a_{1j}\det(M_{1j}).$$

Using the induction hypothesis, we may expand each $\det(M_{1j})$ about its row and obtain

$$\begin{aligned}\det(A) = \sum_{j=1}^{k+1}(-1)^{1+j}a_{1j}\left\{\sum_{l=1}^{j-1}(-1)^{i-1+l}a_{il}\det(M(1j,il))\right.\\ \left. + \sum_{l=j+1}^{k+1}(-1)^{i-1+l-1}a_{il}\det(M(1j,il))\right\}\\ = \sum_{j=1}^{k+1}\sum_{l=1}^{j-1}(-1)^{i+j+l}a_{1j}a_{il}\det(M(1j,il))\\ + \sum_{j=1}^{k+1}\sum_{l=j+1}^{k+1}(-1)^{i+j+l-1}a_{1j}a_{il}\det(M(1j,il))\end{aligned} \tag{1}$$

(On the second summation, $l-1$ occurs in the exponent of $-1$ instead of $l$ since the $l$th column of $A$ becomes the $(l-1)$st column of $M_{1j}$ when $l > j$.) Now consider the cofactor expansion about the $i$th row, which we write as

$$\sum_{l=1}^{k+1}(-1)^{i+l}a_{il}\det(M_{il}).$$

Let us expand each $\det(M_{il})$ about the first row:

$$\begin{aligned}\sum_{l=1}^{k+1}(-1)^{i+l}a_{il}\left\{\sum_{j=1}^{l-1}(-1)^{1+j}a_{1j}\det(M(il,1j))\right.\\ \left. + \sum_{j=l+1}^{k+1}(-1)^{1+j-1}a_{1j}\det(M(il,1j))\right\}\\ = \sum_{l=1}^{k+1}\sum_{j=1}^{l-1}(-1)^{i+j+l+1}a_{il}a_{1j}\det(M(il,1j))\\ + \sum_{l=1}^{k+1}\sum_{j=l+1}^{k+1}(-1)^{i+j+l}a_{il}a_{1j}\det(M(il,1j)).\end{aligned} \tag{2}$$

While they may look different, the results in Equations (1) and (2) are the same. To see why, consider a term with a $j$ and an $l$. If $j > l$, this term in Equation (1) is

$$(-1)^{i+j+l}a_{1j}a_{il}\det(M(1j,il)),$$

which is exactly the same term as we have in Equation (2) for $j > l$. We leave it as an exercise (Exercise 2) to show that the terms with $j < l$ in Equations (1) and (2) are the same. ●

We next show that we can expand about the first column.

**LEMMA 1.30** If $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$, then

$$\det(A) = \sum_{i=1}^{n} a_{i1} C_{i1}.$$

*Proof* We again use induction on $n$. We will let you verify this for $n = 2$ (Exercise 3). Assume this result holds for all $k \times k$ matrices and let $A = [a_{ij}]$ be a $(k+1) \times (k+1)$ matrix. By definition

$$\begin{aligned} \det(A) &= \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} \det(M_{1j}) \\ &= a_{11} \det(M_{11}) + \sum_{j=2}^{k+1} (-1)^{1+j} a_{1j} \det(M_{1j}). \end{aligned}$$

Using the induction hypothesis, we expand each $\det(M_{1j})$ about its first column for $j \geq 2$ and obtain

$$\begin{aligned} \det(A) &= a_{11} \det(M_{11}) + \sum_{j=2}^{k+1} (-1)^{1+j} a_{1j} \left\{ \sum_{i=2}^{k+1} (-1)^{1+i-1} a_{i1} \det(M(1j, i1)) \right\} \\ &= a_{11} \det(M_{11}) + \sum_{j=2}^{k+1} \sum_{i=2}^{k+1} (-1)^{1+j+i} a_{1j} a_{i1} \det(M(1j, i1)). \end{aligned} \tag{3}$$

Writing the cofactor expansion about the first column as

$$\sum_{i=1}^{k+1} (-1)^{i+1} a_{i1} \det(M_{i1}) = a_{11} \det(M_{11}) + \sum_{i=2}^{k+1} (-1)^{i+1} a_{i1} \det(M_{i1})$$

and then expanding each $\det(M_{i1})$ for $i \geq 2$ about its first row, we obtain

$$\begin{aligned} & a_{11} \det(M_{11}) + \sum_{i=2}^{k+1} (-1)^{i+1} a_{i1} \left\{ \sum_{j=2}^{k+1} (-1)^{1+j-1} a_{1j} \det(M(i1, 1j)) \right\} \\ &= a_{11} \det(M_{11}) + \sum_{i=2}^{k+1} \sum_{j=2}^{k+1} (-1)^{i+1+j} a_{i1} a_{1j} \det(M(i1, 1j)) \end{aligned} \tag{4}$$

Since the results of Equations (3) and (4) are the same, our proof is complete. ●

Before completing the proof of Theorem 1.16, we use Lemma 1.30 to prove Theorem 1.19. For convenience, let us restate Theorem 1.19 as Theorem 1.31.

**THEOREM 1.31** If $A$ is an $n \times n$ matrix,

$$\det(A^T) = \det(A).$$

*Proof* Here we use induction too, only we may start with $n = 1$ where the result is trivial for a $1 \times 1$ matrix $A = [a_{11}]$. Assume the result holds for any $k \times k$ matrix and let $A = [a_{ij}]$

be a $(k+1) \times (k+1)$ matrix. Note that the $j1$-entry of $A^T$ is $a_{1j}$ and its minor is $M_{1j}^T$ where $M_{1j}$ is the minor of the entry $a_{1j}$ of $A$. Thus if we expand $\det(A^T)$ about the first column,

$$\det(A^T) = \sum_{j=1}^{k+1} (-1)^{j+1} a_{1j} \det(M_{1j}^T).$$

By the induction hypothesis, $\det(M_{1j}^T) = \det(M_{1j})$ and hence

$$\det(A^T) = \sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} \det(M_{1j}) = \det(A).$$

To complete the proof of Theorem 1.16, we must show that for any $1 \leq j \leq n$,

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij})$$

where $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$. To obtain this, we first expand $\det(A^T)$ about row $j$ (which we may do by Lemma 1.29). This gives us

$$\det(A^T) = \sum_{i=1}^{n} (-1)^{j+i} a_{ij} \det(M_{ij}^T).$$

Now applying the result of Theorem 1.31 to $\det(A^T)$ and each $\det(M_{ij}^T)$, we obtain the desired result.

Another result we have not proved that we now prove is Theorem 1.20, which we restate as Theorem 1.32.

**THEOREM 1.32** Suppose that $A = [a_{ij}]$ is an $n \times n$ matrix with $n \geq 2$.

1. If $B$ is a matrix obtained from $A$ by interchanging two rows of $A$, then $\det(B) = -\det(A)$.
2. If $B$ is a matrix obtained from $A$ by multiplying a row of $A$ by a scalar $c$, then $\det(B) = c\det(A)$.
3. If $B$ is a matrix obtained from $A$ by replacing a row of $A$ by itself plus a multiple of another row of $A$, then $\det(B) = \det(A)$.

***Proof***

1. We proceed by induction on $n$ leaving the first case with $n = 2$ as an exercise (Exercise 4.) Assume part (1) holds for all $k \times k$ matrices and let $A = [a_{ij}]$ be a $(k+1) \times (k+1)$ matrix. Suppose that $B$ is obtained from $A$ by interchanging rows $i$ and $l$. We are going to expand $\det(B)$ about a row other than row $i$ or row $l$. Pick such a row. Let us call this the $m$th row. We have $\mathrm{ent}_{mj}(B) = a_{mj}$. If we interchange the rows of the minor $M_{mj}$ of the entry $a_{mj}$ of $A$ that come from rows $i$ and $l$ of $A$, we obtain the minor of $\mathrm{ent}_{mj}(B)$ in $B$. By the

induction hypothesis, the determinant of this minor of $B$ is $-\det(M_{mj})$. Thus,

$$\det(B) = \sum_{j=1}^{k+1} (-1)^{m+j} a_{mj} (-\det(M_{mj})) = -\det(A).$$

2. Suppose that $B$ is obtained from $A$ by multiplying row $i$ of $A$ by $c$. The minor of $\mathrm{ent}_{ij}(B)$ of $B$ is the same as the minor $M_{ij}$ of the entry $a_{ij}$ of $A$. Hence if we expand about the $i$th row,

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j} c a_{ij} \det(M_{ij}) = c \det(A).$$

Before proving part (3), we prove the following lemma.

**LEMMA 1.33** If $A$ is an $n \times n$ matrix where $n \geq 2$ with two rows that have the same entries, then $\det(A) = 0$.

***Proof*** Suppose that row $i$ and row $j$ of $A$ have the same entries. Let $B$ be the matrix obtained from $A$ by interchanging rows $i$ and $j$. On the one hand, by part (1) of Theorem 1.32, we have

$$\det(B) = -\det(A).$$

On the other hand, $B = A$ and hence

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

Thus $\det(A) = -\det(A)$, which implies $\det(A) = 0$. ●

***Proof of Theorem 1.32, Part (3)*** Suppose that $B$ is obtained from $A$ by replacing row $i$ of $A$ by itself plus $c$ times row $l$ of $A$. Then $\mathrm{ent}_{ij}(B) = a_{ij} + c a_{lj}$ and the minor of $\mathrm{ent}_{ij}(B)$ of $B$ is the same as the minor $M_{ij}$ of the entry $a_{ij}$ of $A$. If we expand $\det(B)$ about row $i$,

$$\begin{aligned}\det(B) &= \sum_{j=1}^{n} (-1)^{i+j} (a_{ij} + c a_{mj}) \det(M_{ij}) \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}) + c \sum_{j=1}^{n} (-1)^{i+j} a_{mj} \det(M_{ij}).\end{aligned}$$

The second sum is the same as the determinant of the matrix obtained from $A$ by replacing the $i$th row of $A$ by row $m$ and hence is zero by Lemma 1.33. Thus

$$\det(B) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(M_{ij}) = \det(A).$$

The proof of our result about the product of a square matrix and its adjoint (Theorem 1.26) is another place where Lemma 1.33 is used. We restate this result as Theorem 1.34.

**THEOREM 1.34** If $A$ is a square matrix, then

$$A\ \mathrm{adj}(A) = \mathrm{adj}(A)A = \det(A)I.$$

*Proof* We will prove that $A\ \mathrm{adj}(A) = \det(A)I$ and leave the proof for $\mathrm{adj}(A)A$ as an exercise (Exercise 5). Suppose $A$ is an $n \times n$ matrix. Notice that

$$\mathrm{ent}_{ij}(A\ \mathrm{adj}(A)) = \sum_{k=1}^{n} a_{ik}C_{jk}.$$

If $i = j$,

$$\mathrm{ent}_{ii}(A\ \mathrm{adj}(A)) = \sum_{k=1}^{n} a_{ik}C_{ik} = \det(A).$$

If $i \neq j$,

$$\mathrm{ent}_{ij}(A\ \mathrm{adj}(A)) = \sum_{k=1}^{n} a_{ik}C_{jk}$$

is the determinant of the matrix obtained from $A$ by replacing the $j$th row of $A$ by the $i$th row of $A$. Since this determinant contains two rows with the same entries, we have

$$\mathrm{ent}_{ij}(A\ \mathrm{adj}(A)) = \sum_{k=1}^{n} a_{ik}C_{jk} = 0$$

when $i \neq j$. This gives us $A\ \mathrm{adj}(A) = \det(A)I$. ●

The final result about determinants we have yet to prove is Cramer's rule, restated as Theorem 1.35.

**THEOREM 1.35** Suppose that $AX = B$ is a system of $n$ linear equations in $n$ unknowns such that $\det(A) \neq 0$. Let $A_1$ be the matrix obtained from $A$ by replacing the first column of $A$ by $B$, $A_2$ be the matrix obtained from $A$ by replacing the second column of $A$ by $B, \ldots, A_n$ be the matrix obtained from $A$ by replacing the $n$th column of $A$ by $B$. Then

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

*Proof* Since

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

we have

$$X = A^{-1}B = \frac{1}{\det(A)}\mathrm{adj}(A)B.$$

Thus for each $1 \le i \le n$,

$$x_i = \frac{1}{\det(A)} \sum_{k=1}^{n} C_{ki} b_k.$$

The summation is exactly the determinant of $A_i$ expanded about the $i$th column and hence

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

●

## EXERCISES 1.7

1. Prove Lemma 1.29 for $n = 2$.
2. Show that the terms with $j < l$ in Equations (1) and (2) are the same.
3. Prove Lemma 1.30 for $n = 2$.
4. Prove part (1) of Theorem 1.32 for $n = 2$.
5. Prove that $\operatorname{adj}(A)A = \det(A)I$.
6. If $A$ is an $n \times n$ matrix and $c$ is a scalar, show that $\det(cA) = c^n \det(A)$.
7. A matrix of the form

$$V = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}$$

is called a **Vandermonde** matrix.[14]

**a)** Show that if $n = 2$,

$$\det(V) = x_2 - x_1.$$

**b)** Use row operations to show that if $n = 3$,

$$\det(V) = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

**c)** Use row operations to show that if $n = 4$,

$$\det(V) = (x_2 - x_1)(x_3 - x_1)(x_4 - x_1)(x_3 - x_2)$$
$$(x_4 - x_2)(x_4 - x_3).$$

**d)** In general,

$$\det(V) = \prod_{j=2}^{n} \left\{ \prod_{i=1}^{j-1} (x_j - x_i) \right\}.$$

Prove this result.

[14]Named for Alexandre Théophile Vandermonde (1735–1796) who studied the theory of equations and determinants.

# 2 Vector Spaces

Your first encounter with vectors in two or three dimensions likely was for modeling physical situations. For example, winds blowing with speeds of 5 and 10 miles per hour 45° east of north may be illustrated by the velocity vectors **u** and **v** in Figure 2.1 drawn as directed line segments of lengths 5 and 10 units pointing 45° east of north. A force pushing a block up an inclined plane might be illustrated by drawing a force vector **F** as in Figure 2.2. In your calculus courses you should have encountered many uses of vectors in two and three dimensions in the study of equations of lines and planes, tangent and normal vectors to curves, and gradients, just to name a few.


**Figure 2.1**

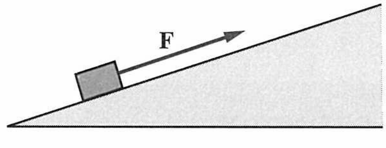


**Figure 2.2**

Were someone to ask you to briefly tell them about vectors you might well respond by saying simply that a vector $\mathbf{v}$ is a directed line segment in two or three dimensions. If we place a vector so that its initial point is at the origin (choice of the initial point does not matter since we are only trying to indicate magnitude and direction with a vector) and its terminal point is $(a, b)$ in two dimensions or $(a, b, c)$ in three dimensions, we can denote the vector by its terminal point as

$$\mathbf{v} = (a, b) \quad \text{or} \quad \mathbf{v} = (a, b, c)$$

in two or three dimensions, respectively. (Other standard notations you might have used instead are $\mathbf{v} = \langle a, b \rangle$ or $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$ in two dimensions and $\mathbf{v} = \langle a, b, c \rangle$ or $\mathbf{v} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ in three dimensions. In Chapter 1 we mentioned that in this text we will write our vectors in two, three, and even $n$ dimensions as column matrices or column vectors.) Vectors are added by the rules

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

or

$$(a_1, b_1, c_1) + (a_2, b_2, c_2) = (a_1 + a_2, b_1 + b_2, c_1 + c_2)$$

and we have a scalar multiplication defined as

$$k(a, b) = (ka, kb) \quad \text{or} \quad k(a, b, c) = (ka, kb, kc)$$

(which, of course, are special cases of matrix addition and scalar multiplication).

We could continue discussing things such as the geometric impact of vector addition (the parallelogram rule), the geometric impact of scalar multiplication (stretching, shrinking, and reflecting), dot products, cross products, and so on, but that is not our purpose here. Our purpose is to study sets of vectors forming a type of structure called a vector space from an algebraic point of view rather than a geometric one. To us a vector space will be a set on which we have defined an addition and a scalar multiplication satisfying certain properties. Two old friends, vectors in two dimensions and vectors in three dimensions, are two examples of vector spaces, but they are not the only ones as you are about to see.

## 2.1 Vector Spaces

As just mentioned, a vector space will be a set of objects on which we have an addition and a scalar multiplication satisfying properties. The formal definition of a vector space is as follows.

DEFINITION A nonempty set $V$ is called a **vector space** if there are operations of addition and scalar multiplication on $V$ such that the following eight properties are satisfied:

**1.** $u + v = v + u$ for all $u$ and $v$ in $V$.
**2.** $u + (v + w) = (u + v) + w$ for all $u$, $v$, and $w$ in $V$.

3. There is an element 0 in $V$ so that $v + 0 = v$ for all $v$ in $V$.
4. For each $v$ in $V$ there is an element $-v$ in $V$ so that $v + (-v) = 0$.
5. $c(u + v) = cu + cv$ for all real numbers $c$ and for all $u$ and $v$ in $V$.
6. $(c + d)v = cv + dv$ for all real numbers $c$ and $d$ and for all $v$ in $V$.
7. $c(dv) = (cd)v$ for all real numbers $c$ and $d$ and for all $v$ in $V$.
8. $1 \cdot v = v$ for all $v$ in $V$.

The eight properties of a vector space are also called the *laws*, *axioms*, or *postulates* of a vector space. The elements of the set $V$ when $V$ is a vector space are called the **vectors** of $V$ and, as we have done already with matrices, real numbers are called **scalars** in connection with the scalar multiplication on $V$.[1] Actually, not all vector spaces are formed using real numbers for the scalars. Later we shall work with some vector spaces where the complex numbers are used as scalars. But for now, all scalars will be real numbers.

Some terminology is associated with the vector space properties. Property 1 is called the **commutative law of addition,** and property 2 is called the **associative law of addition.** The element 0 of $V$ in property 3 is called an **additive identity** or a **zero vector,**[2] and the element $-v$ of $V$ in property 4 is called an **additive inverse** or a **negative** of the vector $v$. Because of commutativity of addition, we could have equally well put our zero and negative vectors on the left in the equations in properties 3 and 4, writing them as

$$0 + v = v \quad \text{and} \quad -v + v = 0.$$

Properties 5 and 6 are distributive properties: Property 5 is a left-hand distributive property saying that scalar multiplication distributes over vector addition, and property 6 is a right-hand distributive property saying that scalar multiplication distributes over scalar addition. Property 7 is an associative property for scalar multiplication.

Let us now look at some examples of vector spaces.

**EXAMPLE 1** From our matrix addition and scalar multiplication properties in Chapter 1, we immediately see the set of $n \times 1$ column vectors or $n$-dimensional vectors $\mathbb{R}^n$ satisfies the eight properties of a vector space under our addition and scalar multiplication of column vectors,

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} \quad \text{and} \quad c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}.$$

[1] In print, vectors are often set in boldface type and, in handwritten work, marked with an arrow over the top to distinguish them from scalars. Such confusion will not arise in this text, however, since we will reserve the lowercase letters $u$, $v$, and $w$ for vectors. Scalars usually will be denoted by letters such as $a$, $b$, $c$, and $d$.

[2] In cases where zero vectors could be confused with the scalar zero, we will put zero vectors in boldface print as **0** as we did for zero matrices in Chapter 1.

Hence $\mathbb{R}^n$ (of which vectors in two and three dimensions are special cases when we write these vectors in column form) is a vector space for each positive integer $n$. ●

Row vectors of a fixed length $n$ would also form a vector space under our addition and scalar multiplication of row vectors. More generally, so would matrices of a fixed size under matrix addition and scalar multiplication. Let us make this our second example.

**EXAMPLE 2** The set of $m \times n$ matrices $M_{m\times n}(\mathbb{R})$ satisfies the eight properties of a vector space under matrix addition and scalar multiplication and hence is a vector space. ●

Making our definition of a vector space as general as we have done will prove valuable to us in the future. For instance, various sets of real-valued functions[3] form vector spaces as is illustrated in the next example. Because of this, sets of real-valued functions will have many properties similar to those enjoyed by our matrix vector spaces, which we shall exploit later in our study of differential equations.

**EXAMPLE 3** Let $F(a, b)$ denote the set of all real-valued functions defined on the open interval $(a, b)$.[4] We can define an addition on $F(a, b)$ as follows: If $f$ and $g$ are two functions in $F(a, b)$, we let $f + g$ be the function defined on $(a, b)$ by

$$(f + g)(x) = f(x) + g(x).$$

We can also define a scalar multiplication on $F(a, b)$: If $c$ is a real number and $f$ is a function in $F(a, b)$, we let $cf$ be the function defined on $(a, b)$ by

$$(cf)(x) = cf(x).$$

Show that $F(a, b)$ is a vector space under this addition and this scalar multiplication.

***Solution*** We verify that the eight properties of a vector space are satisfied.

1. Do we have equality of the function $f + g$ and $g + f$ for all $f$ and $g$ in $F(a, b)$? To see, we have to check if $(f + g)(x)$ is the same as $(g + f)(x)$ for any $x$ in $(a, b)$. If $x$ is in $(a, b)$,

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

(the second equality holds since addition of the real numbers $f(x)$ and $g(x)$ is commutative) and we do have $f + g = g + f$.

2. Do we have equality of the functions $f + (g + h)$ and $(f + g) + h$ for any $f$, $g$, and $h$ in $F(a, b)$? We proceed in much the same manner as in the previous

---

[3] A real-valued function is a function whose range is contained in the set of real numbers. For instance, the function $f$ from $\mathbb{R}$ to $\mathbb{R}$ defined by $f(x) = x^2$ is real-valued; so is the function $f$ from $\mathbb{R}^2$ to $\mathbb{R}$ defined by $f(x, y) = x^2 + y^2$. For a fixed positive integer $n$ the determinant is a real-valued function from $M_{n\times n}(\mathbb{R})$ to $\mathbb{R}$.

[4] For instance, the functions given by $f(x) = x^2$, $f(x) = |x|$, $f(x) = \sin x$, $f(x) = e^x$, and the greatest integer function would be elements of $F(-\infty, \infty)$; these five functions along with the functions defined by $f(x) = 1/x$, $f(x) = \ln x$, and $f(x) = \cot x$ would be elements of $F(0, \pi)$.

part. For any $x$ in $(a, b)$,

$$\begin{aligned}(f + (g + h))(x) &= f(x) + (g + h)(x) = f(x) + (g(x) + h(x)) \\ &= (f(x) + g(x)) + h(x) = (f + g)(x) + h(x) \\ &= ((f + g) + h)(x)\end{aligned}$$

(the third equality holds since addition of the real numbers $f(x)$, $g(x)$, and $h(x)$ is associative) and we have $f + (g + h) = (f + g) + h$.

**3.** Do we have a zero vector? How about if we use the constant function that is 0 for each $x$, which we shall also denote by 0 and call the *zero function*? That is, the zero function is

$$0(x) = 0.$$

The zero function is then an element of $F(a, b)$, and for any $f$ in $F(a, b)$ and $x$ in $(a, b)$,

$$(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x).$$

Hence $f + 0 = f$ and the zero function serves as a zero vector for $F(a, b)$.

**4.** What could we use for the negative of a function $f$ in $F(a, b)$? How about the function $-f$ defined by

$$(-f)(x) = -f(x)?$$

The function $-f$ is in $F(a, b)$, and for any $x$ in $(a, b)$ we have

$$(f + (-f))(x) = f(x) + (-f)(x) = f(x) + (-f(x)) = 0 = 0(x).$$

Hence $-f$ serves as a negative of $f$.

**5.** This and the remaining properties are verified in much the same manner as the first two, so we will go a little faster now. For any real number $c$ and any functions $f$ and $g$ in $F(a, b)$,

$$\begin{aligned}(c(f + g))(x) &= c(f + g)(x) = c(f(x) + g(x)) = cf(x) + cg(x) \\ &= (cf)(x) + (cg)(x)\end{aligned}$$

for any $x$ in $(a, b)$, and hence $c(f + g) = cf + cg$.

**6.** For any real numbers $c$ and $d$ and any function $f$ in $F(a, b)$,

$$((c + d)f)(x) = (c + d)f(x) = cf(x) + df(x) = (cf)(x) + (df)(x)$$

for any $x$ in $(a, b)$, and hence $(c + d)f = cf + df$.

Lest we be accused of doing everything for you, we will let you verify the last two properties as an exercise (Exercise 1). ●

There is nothing special about using an open interval in Example 3. The set of real-valued functions defined on a closed interval $[a, b]$, which we will denote by $F[a, b]$, also forms a vector space under the addition and scalar multiplication of functions we used in Example 3. More generally, the set of real-valued functions defined on any set

$S$, which we will denote by $F(S)$, is a vector space under the types of addition and scalar multiplication of functions we used in Example 3.

The next example illustrates that not every set on which is defined an addition and a scalar multiplication is a vector space.

**EXAMPLE 4** On the set of pairs of real numbers $(x, y)$, define an addition by

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2 + 1, y_1 + y_2)$$

and a scalar multiplication by

$$c(x, y) = (cx, cy).$$

Determine if this is a vector space.

*Solution* Let us start checking the eight properties.

1. We have

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2 + 1, y_1 + y_2)$$

and

$$(x_2, y_2) + (x_1, y_1) = (x_2 + x_1 + 1, y_2 + y_1).$$

Since these ordered pairs are the same, addition is commutative.

2. For three pairs $(x_1, y_1)$, $(x_2, y_2)$, and $(x_3, y_3)$ of real numbers, we have

$$\begin{aligned}(x_1, y_1) + ((x_2, y_2) + (x_3, y_3)) &= (x_1, y_1) + (x_2 + x_3 + 1, y_2 + y_3)\\ &= (x_1 + x_2 + x_3 + 2, y_1 + y_2 + y_3)\end{aligned}$$

while

$$\begin{aligned}((x_1, y_1) + (x_2, y_2)) + (x_3, y_3) &= (x_1 + x_2 + 1, y_1 + y_2) + (x_3, y_3)\\ &= (x_1 + x_2 + x_3 + 2, y_1 + y_2 + y_3).\end{aligned}$$

Since these ordered pairs are again the same, addition is associative.

3. The pair $(-1, 0)$ serves as an additive identity here since

$$(x, y) + (-1, 0) = (x + (-1) + 1, y + 0) = (x, y)$$

so we have property 3. (Notice that an additive identity does not have to be $(0, 0)$!)

4. An additive inverse of $(x, y)$ is $(-x - 2, -y)$ (and is not $(-x, -y)$!) since

$$(x, y) + (-x - 2, -y) = (x + (-x - 2) + 1, y + (-y)) = (-1, 0)$$

and hence we have property 4.

5. Since

$$c((x_1, y_1) + (x_2, y_2)) = c(x_1 + x_2 + 1, y_1 + y_2) = (cx_1 + cx_2 + c, cy_1 + cy_2)$$

while

$$c(x_1, y_1) + c(x_2, y_2) = (cx_1, cy_1) + (cx_2, cy_2) = (cx_1 + cx_2 + 1, cy_1 + cy_2),$$

we see that property 5 does not hold for every real number $c$. (In fact, it holds only when $c = 1$.) Thus, this is not a vector space. Once we see one property that does not hold we are done with determining whether we have a vector space in this example. Were we to continue going through the properties, we would also find that property 6 does not hold. (Try it.) Properties 7 and 8 will hold. (Verify this.) ●

Do not let Example 4 mislead you into thinking that unusual operations for addition or scalar multiplication will not produce a vector space. Consider the next example.

**EXAMPLE 5** Let $\mathbb{R}^+$ denote the set of positive real numbers. Define addition on $\mathbb{R}^+$ by

$$x \oplus y = xy$$

and scalar multiplication by

$$c \odot x = x^c$$

where $x$ and $y$ are in $\mathbb{R}^+$ and $c$ is a real number. (We use the symbols $\oplus$ and $\odot$ to avoid confusion with usual addition and multiplication.) Determine if $\mathbb{R}^+$ is a vector space under this addition and scalar multiplication.

*Solution*

1. Since

$$x \oplus y = xy = yx = y \oplus x$$

this addition is commutative.

2. Since

$$x \oplus (y \oplus z) = x(yz) = (xy)z = (x \oplus y) \oplus z$$

this addition is associative.

3. The positive real number 1 is an additive identity since

$$x \oplus 1 = x \cdot 1 = x.$$

4. An additive inverse of $x$ is $1/x$. (You verify this one.)
5. Since

$$c \odot (x \oplus y) = c \odot (xy) = (xy)^c = x^c y^c = x^c \oplus y^c = c \odot x \oplus c \odot y$$

we have property 5.

We will let you verify that

6. $(c + d) \odot x = c \odot x \oplus d \odot x$,

7. $c \odot (d \odot x) = (cd) \odot x$, and
8. $1 \odot x = x$

as an exercise (Exercise 2) and hence we have a vector space. ●

We conclude this section by discussing some properties of vector spaces. The first theorem deals with the uniqueness of zero and negative vectors.

**THEOREM 2.1** Suppose that $V$ is a vector space.

1. A zero vector of $V$ is unique.
2. A negative of a vector $v$ in $V$ is unique.

***Proof***

1. Suppose that $0$ and $0'$ are zero vectors of $V$. On the one hand, since $0$ is a zero vector, we have that
$$0' + 0 = 0'.$$
On the other hand,
$$0' + 0 = 0$$
since $0'$ is a zero vector. Thus we see $0' = 0$.
2. Suppose that $-v$ and $-v'$ are negatives of $v$. Notice that
$$-v' + (v + (-v)) = -v' + 0 = -v' = (-v' + v) + (-v) = 0 + (-v) = -v,$$
and hence we see $-v = -v'$. ●

Because of Theorem 2.1, we may now say *the* zero vector instead of *a* zero vector and *the* negative of a vector instead of *a* negative of a vector.

Theorem 2.2 contains some more properties that we shall use often.

**THEOREM 2.2** Let $V$ be a vector space.

1. For any vector $v$ in $V$, $0 \cdot v = \mathbf{0}$.[5]
2. For any real number $c$, $c\mathbf{0} = \mathbf{0}$.
3. For any vector $v$ in $V$, $(-1)v = -v$.

***Proof*** We will prove parts (1) and (3) and leave the proof of part (2) as an exercise (Exercise 10).

1. One way to prove this is to first notice that since
$$0 \cdot v = (0 + 0)v = 0 \cdot v + 0 \cdot v$$

[5] Here is a place where we have put the zero vector in boldface print to distinguish it from the scalar zero.

we have

$$0 \cdot v = 0 \cdot v + 0 \cdot v.$$

Adding $-(0 \cdot v)$ to each side of the preceding equation, we obtain

$$0 \cdot v + (-(0 \cdot v)) = 0 \cdot v + 0 \cdot v + (-(0 \cdot v))$$

from which the desired equation

$$\mathbf{0} = 0 \cdot v$$

now follows.

2. Noting on the one hand that

$$(1 + (-1))v = 1 \cdot v + (-1)v = v + (-1)v$$

and on the other hand that

$$(1 + (-1))v = 0 \cdot v = \mathbf{0}$$

by part (1), we have

$$v + (-1)v = \mathbf{0}.$$

Adding $-v$ to each side of the preceding equation, we obtain

$$-v + v + (-1)v = -v + \mathbf{0}$$

and hence

$$(-1)v = -v.$$ ●

Finally, we point out that we can define **subtraction** on a vector space $V$ by setting the difference of two vectors $u$ and $v$ in $V$ to be

$$u - v = u + (-v).$$

You might notice that we could have equally well subtracted the vector whenever we added its negative in the proofs of parts (1) and (3) in Theorem 2.2.

## EXERCISES 2.1

1. Complete Example 3 by showing that properties 7 and 8 of a vector space hold.
2. Complete Example 5 by showing that properties 6, 7, and 8 of a vector space hold.
3. In each of the following, determine whether the indicated addition and scalar multiplication on ordered pairs of real numbers yields a vector space. For those that are not vector spaces, determine which properties of a vector space fail to hold.
   **a)** $(x_1, y_1) + (x_2, y_2) = (x_1, y_2)$, $c(x, y) = (cx, cy)$
   **b)** $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$, $c(x, y) = (c + x, c + y)$
   **c)** $(x_1, y_1) + (x_2, y_2) = (x_1 + y_2, x_2 + y_1)$, $c(x, y) = (cx, cy)$

4. In each of the following, determine whether the indicated addition and scalar multiplication of ordered triples of real numbers yields a vector space. For those that are not vector spaces, determine which properties of a vector space fail to hold.

   **a)** $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$, $c(x, y, z) = (cx, y, cz)$

   **b)** $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (z_1 + z_2, y_1 + y_2, x_1 + x_2)$, $c(x, y, z) = (cx, cy, cz)$

   **c)** $(x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2 - 2, z_1 + z_2)$, $c(x, y, z) = (cx, y, z)$

5. Show that the set of ordered pairs of positive real numbers is a vector space under the addition and scalar multiplication

$$(x_1, y_1)+(x_2, y_2) = (x_1x_2, y_1y_2), c(x, y) = (x^c, y^c).$$

6. Does the set of complex numbers under the addition and scalar multiplication

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$c(a + bi) = ca + cbi$$

   where $a$, $b$, $c$, and $d$ are real numbers form a vector space? If not, why not?

7. Let $C$ denote the set of all convergent sequences of real numbers $\{a_n\}$. Is $C$ a vector space under the addition and scalar multiplication

$$\{a_n\} + \{b_n\} = \{a_n + b_n\}, c\{a_n\} = \{ca_n\}?$$

   If not, why not?

8. Let $S$ denote the set of all convergent series of real numbers $\sum_{n=1}^{\infty} a_n$. Is $S$ a vector space under the addition and scalar multiplication

$$\sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} (a_n + b_n),$$
$$c\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} ca_n?$$

   If not, why not?

9. Let $V$ be a set consisting of a single element $z$. Define addition and scalar multiplication on $V$ by

$$z + z = z, \qquad cz = z.$$

   Show that $V$ is a vector space. Such a vector space is called a **zero vector space.**

10. Prove part (2) of Theorem 2.2.

11. Prove that if $c$ is a real number and $v$ is a vector in a vector space $V$ such that $cv = \mathbf{0}$, then either $c = 0$ or $v = \mathbf{0}$.

12. Show that subtraction is not an associative operation on a vector space.

## 2.2 SUBSPACES AND SPANNING SETS

We begin this section with subspaces. Roughly speaking, by a subspace we mean a vector space sitting within a larger vector space. The following definition states this precisely.

**DEFINITION** 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 of $V$ restricted to $W$.

**EXAMPLE 1** Let $W$ be the set of all column vectors of the form

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

The set $W$ is a subset of $\mathbb{R}^3$. In fact, $W$ is a subspace of $\mathbb{R}^3$. To see this, first notice that the addition of $\mathbb{R}^3$ on elements of $W$ gives us an addition on $W$: For two elements

$$\begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} x_2 \\ y_2 \\ 0 \end{bmatrix}$$

of $W$, we have the sum

$$\begin{bmatrix} x_1 \\ y_1 \\ 0 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ 0 \end{bmatrix},$$

which is also an element of $W$. The fact that the sum of two elements of $W$ is again an element of $W$ is usually described by saying that $W$ is **closed under addition.** Next notice that the scalar multiplication of $\mathbb{R}^3$ gives us a multiplication on $W$: If $c$ is a scalar and

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

is an element of $W$, then

$$c \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = \begin{bmatrix} cx \\ cy \\ 0 \end{bmatrix}$$

is an element of $W$. Here we say $W$ is **closed under scalar multiplication.** So we have two of the ingredients we need (an addition and a scalar multiplication) for $W$ to be a vector space.

Let us move on to the eight properties. Since addition on $\mathbb{R}^3$ is commutative and associative, it certainly is on $W$ too since the elements of $W$ are elements of $\mathbb{R}^3$. Hence properties 1 and 2 of a vector space hold for $W$. Property 3 holds since

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

is an element of $W$. Because columns of the form

$$\begin{bmatrix} -x \\ -y \\ 0 \end{bmatrix}$$

are in $W$, property 4 holds. As was the case with commutativity and associativity of addition, the scalar multiplication properties 5–8 will carry over from $\mathbb{R}^3$ to the subset $W$. Hence $W$ is a vector space. ●

Looking back at Example 1, notice that properties 1, 2, and 5–7 are immediately inherited by any subset of a vector space, so we really do not need to check for them. In fact, the next theorem tells us that the closure properties are really the crucial ones in determining whether a nonempty subset of a vector space is a subspace.

**THEOREM 2.3** Let $W$ be a nonempty subset of a vector space $V$. Then $W$ is a subspace of $V$ if and only if for all $u$ and $w$ in $W$ and for all scalars $c$, $u + w$ is in $W$ and $cu$ is in $W$.

*Proof* If $W$ is a subspace, $W$ is a vector space and hence we immediately have $W$ is closed under addition and scalar multiplication. (Otherwise, $W$ would not have an addition or scalar multiplication.) The main part of this proof is then to show the converse: If $W$ is closed under addition and scalar multiplication, then $W$ is a subspace. As already noted, properties 1, 2, and 5–8 carry over to $W$ from $V$, so we only have to do some work to get properties 3 and 4. To get 3, pick an element $v$ in $W$. (We can do this because $W$ is nonempty.) Since $W$ is closed under scalar multiplication, $(-1)v = -v$ is in $W$. Now since $W$ is closed under addition,

$$v + (-v) = 0$$

lies in $W$ and we have property 3. The multiplying by $-1$ trick also gives us negatives: For any $u$ in $W$, $(-1)u = -u$ is in $W$ by the closure under scalar multiplication and hence property 4 holds. ●

Let us do some more examples determining whether subsets of vector spaces are subspaces, but now applying Theorem 2.3 by only checking the closure properties.

**EXAMPLE 2** Do the vectors of the form

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

form a subspace of $\mathbb{R}^2$?

*Solution* Since the sum of two such vectors,

$$\begin{bmatrix} x_1 \\ 1 \end{bmatrix} + \begin{bmatrix} x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ 2 \end{bmatrix},$$

has 2 not 1 for its second entry, the set of such vectors is not closed under addition and hence is not a subspace. It is also easily seen that this set of vectors is not closed under scalar multiplication. ●

**EXAMPLE 3** Do the vectors of the form

$$\begin{bmatrix} x \\ y \\ x - 2y \end{bmatrix}$$

form a subspace of $\mathbb{R}^3$?

*Solution* Adding two such vectors, we obtain a vector of the same form:

$$\begin{bmatrix} x_1 \\ y_1 \\ x_1 - 2y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ x_2 - 2y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ x_1 - 2y_1 + x_2 - 2y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ x_1 + x_2 - 2(y_1 + y_2) \end{bmatrix}.$$

Hence we have closure under addition. We also have closure under scalar multiplication since

$$c\begin{bmatrix} x \\ y \\ x - 2y \end{bmatrix} = \begin{bmatrix} cx \\ cy \\ cx - 2cy \end{bmatrix}.$$

Thus these vectors do form a subspace.

Solutions to systems of homogeneous linear equations form subspaces. Indeed this will be such an important fact for us that we record it as a theorem.

**THEOREM 2.4** If $A$ is an $m \times n$ matrix, then the solutions to the system of homogeneous linear equations $AX = 0$ is a subspace of $\mathbb{R}^n$.

*Proof* First notice that the set of solutions contains the trivial solution $X = 0$ and hence is a nonempty subset of $\mathbb{R}^n$. If $X_1$ and $X_2$ are two solutions of $AX = 0$, then $AX_1 = 0$ and $AX_2 = 0$ so that

$$A(X_1 + X_2) = AX_1 + AX_2 = 0 + 0 = 0$$

and hence the set of solutions is closed under addition. If $X$ is a solution and $c$ is a scalar, then

$$A(cX) = cAX = c0 = 0$$

and hence the set of solutions is closed under scalar multiplication. Thus the set of solutions to $AX = 0$ is a subspace of $\mathbb{R}^n$.

In Section 2.1, we noted that sets of real-valued functions on intervals form vector spaces. There are numerous examples of subspaces of such function spaces that will come up in our future work listed in Examples 4–10.

**EXAMPLE 4** Let $C(a, b)$ denote the set of continuous real-valued functions on the open interval $(a, b)$, which is a nonempty subset of $F(a, b)$. From calculus, we know that sums of continuous functions and constant multiples of continuous functions are continuous. Hence $C(a, b)$ is closed under addition and scalar multiplication of functions and is a subspace of $F(a, b)$.

**EXAMPLE 5** Let $D(a, b)$ denote the set of differentiable functions on $(a, b)$. From calculus we know that $D(a, b)$ is a nonempty subset of $C(a, b)$ that is closed under addition and scalar multiplication of functions. Hence $D(a, b)$ is a subspace of $C(a, b)$. ●

**EXAMPLE 6** Generalizing Example 5, for each positive integer $n$, let $D^n(a, b)$ denote the set of functions that have an $n$th derivative on $(a, b)$. We have that $D^1(a, b) = D(a, b)$ and each $D^{n+1}(a, b)$ is a subspace of $D^n(a, b)$. ●

**EXAMPLE 7** For each nonnegative integer $n$, we will use $C^n(a, b)$ to denote the set of all functions that have a continuous $n$th derivative on $(a, b)$. Notice that $C^0(a, b) = C(a, b)$, each $C^{n+1}(a, b)$ is a subspace of $C^n(a, b)$, and $C^n(a, b)$ is a subspace of $D^n(a, b)$ for each $n \geq 1$. ●

**EXAMPLE 8** We will let $C^\infty(a, b)$ denote the set of functions that have a continuous $n$th derivative for every nonnegative integer $n$. The set $C^\infty(a, b)$ is a subspace of $C^n(a, b)$ for every nonnegative integer $n$. ●

**EXAMPLE 9** We will let $P$ denote the set of all polynomials; that is, $P$ consists of all expressions $p(x)$ of the form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

where $n$ is a nonnegative integer and each $a_i$ is a real number. Each such polynomial $p(x)$ gives us a function $p$ that is an element of $C^\infty(-\infty, \infty)$. Identifying the polynomial $p(x)$ with the function $p$, we may view $P$ as being a subset of $C^\infty(-\infty, \infty)$. Since polynomials are closed under addition and scalar multiplication, $P$ is a subspace of $C^\infty(-\infty, \infty)$. ●

**EXAMPLE 10** For each nonnegative integer $k$, we will let $P_k$ denote the set of all polynomials of degree less than or equal to $k$ along with the polynomial 0. In particular, $P_0$ is the set of all constant functions $p(x) = a$, $P_1$ is the set of all linear functions $p(x) = mx + b$, and $P_2$ is the set of all functions of the form $p(x) = ax^2 + bx + c$. Each $P_k$ is a subspace of $P$. Also, $P_0$ is a subspace of $P_1$, $P_1$ is a subspace of $P_2$, and so on. ●

We could equally well use other types of intervals in Examples 4–8. When doing so, we will adjust the notation accordingly. For example, $C[a, b]$ will denote the set of continuous functions on the closed interval $[a, b]$, which is a subspace of $F[a, b]$.

We next turn our attention to spanning sets. If $V$ is a vector space and $v_1, v_2, \ldots, v_n$ are vectors in $V$, an expression of the form

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

where $c_1, c_2, \ldots, c_n$ are scalars is called a **linear combination** of $v_1, v_2, \ldots, v_n$. Given a fixed collection of vectors, Theorem 2.5 tells us the set of their linear combinations forms a subspace.

**THEOREM 2.5** If $V$ is a vector space and $v_1, v_2, \ldots, v_n$ are vectors in $V$, then the set of all linear combinations of $v_1, v_2, \ldots, v_n$ is a subspace of $V$.

*Proof* Consider two linear combinations

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \quad \text{and} \quad d_1 v_1 + d_2 v_2 + \cdots + d_n v_n$$

of $v_1, v_2, \ldots, v_n$. As

$$\begin{aligned} c_1 v_1 + c_2 v_2 + \cdots + c_n v_n + d_1 v_1 + d_2 v_2 + \cdots + d_n v_n \\ = (c_1 + d_1) v_1 + (c_2 + d_2) v_2 + \cdots + (c_n + d_n) v_n \end{aligned}$$

is a linear combination of $v_1, v_2, \ldots, v_n$, we have closure under addition. Also, for any scalar $c$, the fact that

$$c(c_1 v_1 + c_2 v_2 + \cdots + c_n v_n) = cc_1 v_1 + cc_2 v_2 + \cdots + cc_n v_n$$

shows we have closure under scalar multiplication and completes our proof. ●

The subspace of a vector space $V$ consisting of all linear combinations of vectors $v_1, v_2, \ldots, v_n$ of $V$ will henceforth be called the **subspace of $V$ spanned by** $v_1, v_2, \ldots, v_n$ and will be denoted

$$\boxed{\operatorname{Span}\{v_1, v_2, \ldots, v_n\}.}$$

In Examples 11 and 12 we determine if a vector lies in the subspace spanned by some vectors.

**EXAMPLE 11** Is the vector

$$\begin{bmatrix} 2 \\ -5 \\ 1 \\ 10 \end{bmatrix} \quad \text{in Span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 3 \end{bmatrix} \right\}?$$

*Solution* We need to see if we can find scalars $c_1, c_2, c_3$ so that

$$c_1 \begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -5 \\ 1 \\ 10 \end{bmatrix}.$$

Comparing entries in these columns, we arrive at the system

$$\begin{aligned} c_1 + c_2 - c_3 &= 2 \\ -c_1 - 2c_2 &= -5 \\ 2c_1 - c_2 + c_3 &= 1 \\ 3c_1 + 2c_2 + 3c_3 &= 10 \end{aligned}$$

and our answer will be yes or no depending on whether or not this system has solutions. Reducing the augmented matrix for this system until it becomes clear whether or not we have a solution,

$$\left[\begin{array}{rrr|r} 1 & 1 & -1 & 2 \\ -1 & -2 & 0 & -5 \\ 2 & -1 & 1 & 1 \\ 3 & 2 & 3 & 10 \end{array}\right] \to \left[\begin{array}{rrr|r} 1 & 1 & -1 & 2 \\ 0 & -1 & -1 & -3 \\ 0 & -3 & 3 & -3 \\ 0 & -1 & 6 & 4 \end{array}\right]$$

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

we see the system does have a solution and hence our answer to this problem is yes. ●

**EXAMPLE 12** Is $2x^2 + x + 1$ in $\mathrm{Span}\{x^2 + x, x^2 - 1, x + 1\}$?

*Solution* By comparing coefficients of $x^2$, $x$, and the constant terms, we see that there are scalars $c_1, c_2, c_3$ so that

$$c_1(x^2 + x) + c_2(x^2 - 1) + c_3(x + 1) = 2x^2 + x + 1$$

if and only if the system of equations

$$c_1 + c_2 = 2$$

$$c_1 + c_3 = 1$$

$$-c_2 + c_3 = 1$$

has solutions. Starting to reduce the augmented matrix for this system,

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

we can see that we have arrived at a system with no solution and hence our answer to this problem is no. ●

We say that the vectors $v_1, v_2, \ldots, v_n$ of a vector space $V$ **span** $V$ if $\mathrm{Span}\{v_1, v_2, \ldots, v_n\} = V$. To put it another way, $v_1, v_2, \ldots, v_n$ span $V$ if every vector in $V$ is a linear combination of $v_1, v_2, \ldots, v_n$. In our final two examples of this section we determine if some given vectors span the given vector space.

**EXAMPLE 13** Do

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

span $\mathbb{R}^2$?

*Solution* For an arbitrary vector

$$\begin{bmatrix} a \\ b \end{bmatrix}$$

of $\mathbb{R}^2$, we must determine whether there are scalars $c_1, c_2$ so that

$$c_1 \begin{bmatrix} 1 \\ -2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ -4 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Reducing the augmented matrix for the resulting system of equations,

$$\left[\begin{array}{cc|c} 1 & 2 & a \\ -2 & -4 & b \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & 2 & a \\ 0 & 0 & b+2a \end{array}\right],$$

we see that the system does not have a solution for all $a$ and $b$ and hence the answer to this problem is no. ●

**EXAMPLE 14** Do $x^2 + x - 3, x - 5, 3$ span $P_2$?

*Solution* Here we must determine whether for an arbitrary element $ax^2 + bx + c$ there are scalars $c_1, c_2, c_3$ so that

$$c_1(x^2 + x - 3) + c_2(x - 5) + c_3 \cdot 3 = ax^2 + bx + c.$$

Comparing coefficients, we are led to the system

$$\begin{aligned} c_1 &= a \\ c_1 + c_2 &= b \\ -3c_1 - 5c_2 + 3c_3 &= c, \end{aligned}$$

which obviously has a solution. Thus the answer to this problem is yes. ●

## EXERCISES 2.2

1. Determine which of the following sets of vectors are subspaces of $\mathbb{R}^2$.

   **a)** All vectors of the form $\begin{bmatrix} 0 \\ y \end{bmatrix}$

   **b)** All vectors of the form $\begin{bmatrix} x \\ 3x \end{bmatrix}$

   **c)** All vectors of the form $\begin{bmatrix} x \\ 2 - 5x \end{bmatrix}$

   **d)** All vectors $\begin{bmatrix} x \\ y \end{bmatrix}$ where $x + y = 0$

2. Determine which of the following sets of vectors are subspaces of $\mathbb{R}^3$.

   **a)** All vectors of the form $\begin{bmatrix} x \\ y \\ y - 4x \end{bmatrix}$

   **b)** All vectors of the form $\begin{bmatrix} y + z + 1 \\ y \\ z \end{bmatrix}$

**c)** All vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ where $z = x + y$

**d)** All vectors $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ where $z = x^2 + y^2$

**3.** Determine which of the following sets of functions are subspaces of $F[a, b]$.

**a)** All functions $f$ in $F[a, b]$ for which $f(a) = 0$

**b)** All functions $f$ in $F[a, b]$ for which $f(a) = 1$

**c)** All functions $f$ in $C[a, b]$ for which $\int_a^b f(x)\,dx = 0$

**d)** All functions $f$ in $D[a, b]$ for which $f'(x) = f(x)$

**e)** All functions $f$ in $D[a, b]$ for which $f'(x) = e^x$

**4.** Determine which of the following sets of $n \times n$ matrices are subspaces of $M_{n\times n}(\mathbb{R})$.

**a)** The $n \times n$ diagonal matrices

**b)** The $n \times n$ upper triangular matrices

**c)** The $n \times n$ symmetric matrices

**d)** The $n \times n$ matrices of determinant zero

**e)** The $n \times n$ invertible matrices

**5.** If $A$ is an $m \times n$ matrix and $B$ is a nonzero element of $\mathbb{R}^m$, do the solutions to the system $AX = B$ form a subspace of $\mathbb{R}^n$? Why or why not?

**6.** Complex numbers $a + bi$ where $a$ and $b$ are integers are called *Gaussian integers*. Do the Gaussian integers form a subspace of the vector space of complex numbers? Why or why not?

**7.** Do the sequences that converge to zero form a subspace of the vector space of convergent sequences? How about the sequences that converge to a rational number?

**8.** Do the series that converge to a positive number form a subspace of the vector space of convergent series? How about the series that converge absolutely?

**9.** Is $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ in $\mathrm{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$?

**10.** Is $\begin{bmatrix} -4 \\ 3 \end{bmatrix}$ in $\mathrm{Span}\left\{\begin{bmatrix} 3 \\ -3 \end{bmatrix}, \begin{bmatrix} -4 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix}\right\}$?

**11.** Is $\begin{bmatrix} 1 \\ -5 \\ -3 \end{bmatrix}$ in $\mathrm{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right\}$?

**12.** Is $\begin{bmatrix} 3 & 5 \\ 3 & 4 \end{bmatrix}$ in $\mathrm{Span}\left\{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right\}$?

**13.** Is $3x^2$ in $\mathrm{Span}\{x^2 - x, x^2 + x + 1, x^2 - 1\}$?

**14.** Is $\sin(x + \pi/4)$ in $\mathrm{Span}\{\sin x, \cos x\}$?

**15.** Determine if $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ span $\mathbb{R}^2$.

**16.** Determine if $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$ span $\mathbb{R}^3$.

**17.** Determine if $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$ span $\mathbb{R}^4$.

**18.** Determine if $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}$ span $M_{2\times 2}(\mathbb{R})$.

**19.** Determine if $x^2 - 1, x^2 + 1, x^2 + x$ span $P_2$.

**20.** Determine if $x^3 + x^2, x^2 + x, x + 1$ span $P_3$.

Use the system of linear equation solving capabilities of Maple or another appropriate software package in Exercises 21–24.

**21.** Determine if $\begin{bmatrix} 1 \\ 2 \\ -2 \\ 0 \\ -1 \\ 4 \end{bmatrix}$ is in

$$\mathrm{Span}\left\{\begin{bmatrix} 4 \\ -2 \\ 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \\ -5 \\ 7 \\ 1 \end{bmatrix}, \begin{bmatrix} -17 \\ 22 \\ -1 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -31 \\ -3 \\ -8 \\ 1 \\ 1 \\ 0 \end{bmatrix},\right.$$

$$\left.\begin{bmatrix} 17 \\ 44 \\ 1 \\ -22 \\ 11 \\ 1.9 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 15 \\ -3 \\ 3 \end{bmatrix}\right\}.$$

**22.** Determine if $x^4 + x^2 + 1$ is in

$$\begin{aligned} \text{Span}\{&x^4 - x^3 + 3x - 4, x^4 - x^3 - x^2 + x - 4, \\ &x^3 + x^2 - 3x + 3, \\ &x^4 + x^3 - 2x^2 + 4x - 8, \\ &5x^4 - 7x^3 - 2x^2 - x + 9, 2x^4 - 7x^3 + 1\}. \end{aligned}$$

**23.** Determine if

$$\begin{gathered} x^5 - x^4 + x^3 + x, \\ x^4 - x^3 + 2x - 4, \\ x^5 - 5x^3 + 6x^2 - 8x + 2, \\ x^5 + \tfrac{2}{3}x^4 - x^3 + 2x^2 + 3x - 1, \\ -2x^3 - 4x^2 + 3x - 9, \\ x^4 - 3x^3 + \pi x^2 - 2x + 1 \end{gathered}$$

span $P_5$.

**24.** Determine if

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

$$\begin{bmatrix} 5 & -2 & -1 \\ 2 & 6 & -3 \end{bmatrix}, \begin{bmatrix} \sqrt{3} & 7 & -1 \\ 2 & 1 & \pi \end{bmatrix},$$

$$\begin{bmatrix} -7 & 9 & 21 \\ 14 & -22 & 8 \end{bmatrix}, \begin{bmatrix} 0 & 3 & -1 \\ 1 & 1 & 2 \end{bmatrix}$$

span $M_{2\times 3}(\mathbb{R})$.

**25.** Suppose that $v_1, v_2, \ldots, v_k$ are vectors in $\mathbb{R}^n$. How can we tell from a row-echelon form of the matrix

$$A = [v_1 \quad v_2 \quad \ldots \quad v_k]$$

if $v_1, v_2, \ldots, v_k$ span $\mathbb{R}^n$?

**26.** Use the answer to Exercise 25 and one of the *gausselim*, *gaussjord*, or *rref* commands of Maple or corresponding commands in another appropriate software package to determine if the vectors

$$\begin{bmatrix} 1 \\ -1 \\ 2 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -3 \\ \sqrt{2} \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 5 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 8 \\ 7 \\ -9 \\ 11 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ -4 \\ -4 \end{bmatrix}$$

span $\mathbb{R}^5$.

## 2.3 Linear Independence and Bases

The concept of a spanning set that we encountered in the previous section is a very fundamental one in theory of vector spaces involving linear combinations. Another is the concept of linear independence and its opposite, linear dependence. These concepts are defined as follows.

**DEFINITION** Suppose that $v_1, v_2, \ldots, v_n$ are vectors in a vector space $V$. We say that $v_1, v_2, \ldots, v_n$ are **linearly dependent** if there are scalars $c_1, c_2, \ldots, c_n$ *not all zero* so that

$$\boxed{c_1v_1 + c_2v_2 + \cdots + c_nv_n = 0.}$$

If $v_1, v_2, \ldots, v_n$ are not linearly dependent, we say $v_1, v_2, \ldots, v_n$ are **linearly independent.**

We can always get a linear combination of vectors $v_1, v_2, \ldots, v_n$ equal to the zero vector by using zero for each scalar:

$$0 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_n = 0.$$

Carrying over the terminology we used for solutions of homogenous systems of linear equations, let us call this the **trivial linear combination** of $v_1, v_2, \ldots, v_n$. We then could equally well say $v_1, v_2, \ldots, v_n$ are linearly dependent if there is a nontrivial linear combination of them equal to the zero vector; saying $v_1, v_2, \ldots, v_n$ are linearly independent would mean that the trivial linear combination is the only linear combination of $v_1, v_2, \ldots, v_n$ equal to the zero vector.

**EXAMPLE 1** Are the vectors

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

linearly dependent or linearly independent?

*Solution* Consider a linear combination of these vectors equal to the zero vector of $\mathbb{R}^3$:

$$c_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

This leads us to the system of equations:

$$\begin{aligned} c_1 + 3c_2 - c_3 &= 0 \\ 2c_1 + 2c_2 + 2c_3 &= 0 \\ 3c_1 + c_2 + 5c_3 &= 0. \end{aligned}$$

Beginning to apply row operations to the augmented matrix for this system,

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

we can see that our system has nontrivial solutions. Thus there are nontrivial linear combinations of our three vectors equal to the zero vector and hence they are linearly dependent. ●

**EXAMPLE 2** Are $x^2 + 1, x^2 - x + 1, x + 2$ linearly dependent or linearly independent?

*Solution* Suppose

$$c_1(x^2 + 1) + c_2(x^2 - x + 1) + c_3(x + 2) = 0.$$

Comparing coefficients of $x^2$, $x$ and the constant terms on each side of the preceding equation, we obtain the system:

$$\begin{aligned} c_1 + c_2 &= 0 \\ -c_2 + c_3 &= 0 \\ c_1 + c_2 + 2c_3 &= 0. \end{aligned}$$

We do not need to set up an augmented matrix here. Notice that subtracting the first equation from the third will give us $c_3 = 0$. The second equation then tells us $c_2 = 0$ from which we can now see $c_1 = 0$ from the first equation. Since our system has only the trivial solution, it follows that the three given polynomials are linearly independent. ●

The next theorem gives us another characterization of linear dependence.

**THEOREM 2.6** Suppose $v_1, v_2, \ldots, v_n$ are vectors in a vector space $V$. Then $v_1, v_2, \ldots, v_n$ are linearly dependent if and only if one of $v_1, v_2, \ldots, v_n$ is a linear combination of the others.

***Proof*** Suppose $v_1, v_2, \ldots, v_n$ are linearly dependent. Then there are scalars $c_1, c_2, \ldots, c_n$ not all zero so that

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0. \tag{1}$$

Suppose that $c_i \neq 0$. Then we may solve Equation (1) for $v_i$ as follows

$$\begin{aligned} c_i v_i &= -c_1 v_1 - \cdots - c_{i-1} v_{i-1} - c_{i+1} v_{i+1} - \cdots - c_n v_n \\ v_i &= -\frac{c_1}{c_i} v_1 - \cdots - \frac{c_{i-1}}{c_i} v_{i-1} - \frac{c_{i+1}}{c_i} v_{i+1} - \cdots - \frac{c_n}{c_i} v_n \end{aligned}$$

and hence obtain that $v_i$ is a linear combination of $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

To prove the converse, suppose one of $v_1, v_2, \ldots, v_n$, let us call it $v_i$, is a linear combination of the other $v_1, v_2, \ldots, v_n$:

$$v_i = c_1 v_1 + \cdots + c_{i-1} v_{i-1} + c_{i+1} v_{i+1} + \cdots + c_n v_n.$$

Then

$$-c_1 v_1 - \cdots - c_{i-1} v_{i-1} + v_i - c_{i+1} v_{i+1} - \cdots - c_n v_n = 0.$$

This gives us a nontrivial linear combination of $v_1, v_2, \ldots, v_n$ equal to the zero vector since the scalar with $v_i$ is 1 and hence $v_1, v_2, \ldots, v_n$ are linearly dependent. ●

If we view having one vector as a linear combination of others as a dependence of the vector on the others, Theorem 2.6 in some sense gives us a more natural way to think of linear dependence. Unfortunately, it is not as practical in general. Were we to use it to check if a set of vectors $v_1, v_2, \ldots, v_n$ are linearly dependent, we could start by checking if $v_1$ is a linear combination of $v_2, \ldots, v_n$. If so, we would have linear dependence. But if not, we would then have to look at $v_2$ and see if $v_2$ is a linear combination of $v_1, v_3, \ldots, v_n$. If so, we again would have linear dependence. If not, we would move on to $v_3$ and so on. Notice that checking to see if there is a nontrivial linear combination of $v_1, v_2, \ldots, v_n$ equal to the zero vector is much more efficient.

One exception is in the case of two vectors $v_1$ and $v_2$, for having one vector a linear combination of the other is the same as saying one vector is a scalar multiple of the other, which is often easily seen by inspection. For example,

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} \text{ and } \begin{bmatrix} 3 \\ -6 \end{bmatrix}$$

are linearly dependent since the second vector is 3 times the first (or the first is 1/3 times the second). The polynomials $x^2 + x$ and $x^2 - 1$ are linearly independent since neither is a scalar multiple of the other.

We next introduce the concept of a basis.

**DEFINITION** We say that the vectors $v_1, v_2, \dots, v_n$ of a vector space $V$ are a **basis** for $V$ if both of the following two conditions are satisfied:

1. $v_1, v_2, \dots, v_n$ are linearly independent.
2. $v_1, v_2, \dots, v_n$ span $V$.

**EXAMPLE 3** The vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

are easily seen to be linearly independent since if

$$c_1e_1 + c_2e_2 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

then $c_1 = 0$ and $c_2 = 0$. They also span $\mathbb{R}^2$ since

$$ae_1 + be_2 = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Thus $e_1, e_2$ form a basis for $\mathbb{R}^2$. ●

**EXAMPLE 4** As in Example 3, it is easily seen that the three vectors

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

both are linearly independent and span $\mathbb{R}^3$. Hence $e_1, e_2, e_3$ form a basis for $\mathbb{R}^3$. ●

**EXAMPLE 5** Generalizing Examples 3 and 4 to $\mathbb{R}^n$, let us use $e_i$ to denote the vector in $\mathbb{R}^n$ that has 1 in the $i$th position and 0s elsewhere:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Here again the vectors $e_1, e_2, \dots, e_n$ are easily seen to both be linearly independent and span $\mathbb{R}^n$ and consequently form a basis for $\mathbb{R}^n$.[6]

**EXAMPLE 6** Generalizing even further, the $m \times n$ matrices $E_{ij}$ that have 1 in the $ij$-position and 0s elsewhere for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$ are linearly independent and span $M_{m\times n}(\mathbb{R})$. Hence they are a basis for $M_{m\times n}(\mathbb{R})$. For instance,

$$E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad E_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis for $M_{2\times 2}(\mathbb{R})$.

**EXAMPLE 7** For a nonnegative integer $n$, the $n + 1$ polynomials

$$x^n, x^{n-1}, \dots, x, 1$$

are linearly independent since if

$$c_1x^n + c_2x^{n-1} + \cdots + c_nx + c_{n+1} \cdot 1 = 0,$$

then $c_1 = 0, c_2 = 0, \dots, c_n = 0, c_{n+1} = 0$. Further, they span $P_n$; indeed we typically write the polynomials in $P_n$ as linear combinations of $x^n, x^{n-1}, \dots, x, 1$ when we write them as $a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$. Hence $x^n, x^{n-1}, \dots, x, 1$ form a basis for $P_n$.

The bases given in each of Examples 3–7 are natural bases to use and are called the **standard bases** for each of these respective vector spaces. Standard bases are not the only bases, however, for these vector spaces. Consider Examples 8 and 9.

**EXAMPLE 8** Show that

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

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

[6] The vectors $e_1$ and $e_2$ of $\mathbb{R}^2$ in Example 3 are often denoted by **i** and **j**, respectively; the vectors $e_1$, $e_2$, and $e_3$ of $\mathbb{R}^3$ in Example 4 are often denoted by **i, j,** and **k,** respectively.

*Solution* Let us first show that these three vectors are linearly independent. Suppose

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

Setting up and reducing the augmented matrix for the resulting homogeneous system,

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

we see that we have the trivial solution $c_1 = 0$, $c_2 = 0$, $c_3 = 0$. Hence these vectors are linearly independent. Similar work applied to the system of equations resulting from the vector equation

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

shows us that our three vectors span $\mathbb{R}^3$. Thus these vectors do form a basis for $\mathbb{R}^3$. ●

**EXAMPLE 9** Show that $x^2 + x - 3$, $x - 5$, $3$ form a basis for $P_2$.

*Solution* Let us first check to see if these three polynomials are linearly independent. If

$$c_1(x^2 + x - 3) + c_2(x - 5) + c_3 \cdot 3 = 0,$$

we have the homogeneous system:

$$\begin{aligned} c_1 &= 0 \\ c_1 + c_2 &= 0 \\ -3c_1 - 5c_2 + 3c_3 &= 0. \end{aligned}$$

Since this system has only the trivial solution, the three polynomials are linearly independent. Likewise, the system of equations resulting from

$$c_1(x^2 + x - 3) + c_2(x - 5) + 3c_3 = ax^2 + bx + c$$

has a solution and hence our three polynomials span $P_2$. (Indeed, if you have a good memory, you will note that we already did the spanning part in the last example of the previous section.) Thus we have shown $x^2 + x - 3$, $x - 5$, $3$ form a basis for $\mathbb{R}^3$. ●

Keep in mind that we must have both linear independence and spanning to have a basis. If either one (or both) fails to hold, we do not have a basis.

**EXAMPLE 10** Do $x^2 + x - 1$, $x^2 - x + 1$ form a basis for $P_2$?

*Solution* Since neither polynomial is a scalar multiple of the other, these two polynomials are linearly independent. However, they do not span $P_2$. To see this, observe that the system of equations resulting from

$$c_1(x^2 + x - 1) + c_2(x^2 - x + 1) = ax^2 + bx + c$$

is

$$\begin{aligned} c_1 + c_2 &= a \\ c_1 - c_2 &= b \\ -c_1 + c_2 &= c, \end{aligned}$$

which does not have a solution for all $a, b,$ and $c$ since adding the second and third equations gives us

$$0 = b + c.$$

Since $x^2 + x - 1, x^2 - x + 1$ do not span $P_2$, they do not form a basis for $P_2$. ●

**EXAMPLE 11** Do

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

form a basis for $\mathbb{R}^2$?

***Solution*** Let us first see if these vectors are linearly independent. If

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

then

$$\begin{aligned} c_1 - c_2 + 2c_3 &= 0 \\ c_1 + c_2 + 3c_3 &= 0. \end{aligned}$$

Since we know that a homogeneous linear system with more variables than equations always has a nontrivial solution (Theorem 1.1), the three given vectors are linearly dependent and hence do not form a basis for $\mathbb{R}^2$. ●

There are other ways of characterizing bases besides saying they are linearly independent spanning sets. Theorem 2.7 describes one other way.

**THEOREM 2.7** Suppose that $v_1, v_2, \dots, v_n$ are vectors in a vector space $V$. Then $v_1, v_2, \dots, v_n$ form a basis for $V$ if and only if each vector in $V$ is uniquely expressible as a linear combination of $v_1, v_2, \dots, v_n$.

***Proof*** First suppose $v_1, v_2, \dots, v_n$ form a basis for $V$. Let $v$ be a vector in $V$. Since $v_1, v_2, \dots, v_n$ span $V$, there are scalars $c_1, c_2, \dots, c_n$ so that

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

We must show this is unique. To do so, suppose we also have

$$v = d_1 v_1 + d_2 v_2 + \cdots + d_n v_n$$

where $d_1, d_2, \ldots, d_n$ are scalars. Subtracting these two expressions for $v$, we obtain

$$v - v = 0 = (c_1 - d_1)v_1 + (c_2 - d_2)v_2 + \cdots + (c_n - d_n)v_n.$$

Now because $v_1, v_2, \ldots, v_n$ are linearly independent, we have

$$c_1 - d_1 = 0, \qquad c_2 - d_2 = 0, \qquad \ldots, \qquad c_n - d_n = 0$$

or

$$c_1 = d_1, \qquad c_2 = d_2, \qquad \ldots, \qquad c_n = d_n.$$

Hence we have the desired uniqueness of the linear combination.

To prove the converse, first note that if every vector in $V$ is uniquely expressible as a linear combination of $v_1, v_2, \ldots, v_n$, we immediately have that $v_1, v_2, \ldots, v_n$ span $V$. Suppose

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0.$$

Since the trivial linear combination of $v_1, v_2, \ldots, v_n$ is the zero vector, the uniqueness property gives us

$$c_1 = 0, \qquad c_2 = 0, \qquad \ldots, \qquad c_n = 0.$$

Hence $v_1, v_2, \ldots, v_n$ are linearly independent, which completes our proof. ●

As a matter of convenience, we shall sometimes denote bases by lowercase Greek letters in this text. If $\alpha$ is a basis for a vector space $V$ consisting of the vectors $v_1, v_2, \ldots, v_n$ and $v$ is a vector in $V$, when we write $v$ uniquely as

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n,$$

the scalars $c_1, c_2, \ldots, c_n$ are called the **coordinates of $v$ relative to the basis $\alpha$** of $V$. The column vector

$$[v]_\alpha = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

is called the **coordinate vector of $v$ relative to the basis $\alpha$**. To illustrate, suppose $\alpha$ is the standard basis

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for $\mathbb{R}^2$. Since for any vector

$$v = \begin{bmatrix} a \\ b \end{bmatrix},$$

we have

$$v = \begin{bmatrix} a \\ b \end{bmatrix} = ae_1 + be_2,$$

we then get

$$[v]_\alpha = \begin{bmatrix} a \\ b \end{bmatrix}.$$

In other words, the coordinates and coordinate vectors relative to the standard basis of $\mathbb{R}^2$ are just the usual coordinates and column vectors. More generally, the same thing occurs when we use the standard basis for $\mathbb{R}^n$. If we use the standard basis for $M_{m\times n}(\mathbb{R})$, the coordinates of a matrix relative to it are the entries of the matrix. If we use the standard basis for $P_n$, the coordinates of a polynomial in $P_n$ relative to it are the coefficients of the polynomial. If we do not use standard bases, we have to work harder to determine the coordinates relative to the basis. Consider Examples 12 and 13.

**EXAMPLE 12** Find the coordinate vector of

$$v = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}$$

relative to the basis $\beta$ for $\mathbb{R}^3$ in Example 8 consisting of

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

*Solution* We need to find $c_1, c_2, c_3$ so that

$$c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 7 \end{bmatrix}.$$

Reducing the augmented matrix for the resulting system of equations,

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 3 \\ 1 & 1 & 1 & 7 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 2 & 6 \end{array}\right],$$

we see

$$c_3 = 3, \qquad c_2 = 0, \qquad c_1 = 4.$$

Hence the coordinate vector relative to $\beta$ is

$$[v]_\beta = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}.$$

**EXAMPLE 13** Find the coordinate vector of $v = x + 1$ relative to the basis $\gamma$ for $P_2$ in Example 9 consisting of $x^2 + x - 3$, $x - 5$, 3.

*Solution* We need to find $c_1, c_2, c_3$ so that

$$c_1(x^2 + x - 3) + c_2(x - 5) + 3c_3 = x + 1.$$

We then have the system

$$\begin{aligned} c_1 &= 0 \\ c_1 + c_2 &= 1 \\ -3c_1 - 5c_2 + 3c_3 &= 1, \end{aligned}$$

which has the solution

$$c_1 = 0, \qquad c_2 = 1, \qquad c_3 = 2.$$

Hence

$$[v]_\gamma = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}.$$

We can reverse the procedure of translating vectors into coordinate vectors as the following example illustrates.

**EXAMPLE 14** Find $v$ in $P_2$ if

$$[v]_\gamma = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

where $\gamma$ is the basis for $P_2$ in Examples 9 and 13 consisting of $x^2 + x - 3$, $x - 5$, 3.

*Solution* We have

$$v = 1 \cdot (x^2 + x - 3) + 2(x - 5) + (-1) \cdot 3 = x^2 + 3x - 16.$$

Using coordinates relative to a basis gives us a way of translating elements in a vector space into vectors with numbers. One place where this is useful is when working on a computer: Translate vectors into coordinate vectors, then do the desired work on the computer in terms of coordinate vectors, and finally translate the computer's results in coordinate vectors back into the original type of vectors. We will further develop these number translation ideas when we come to the study of linear transformations in Chapter 5.

## EXERCISES 2.3

In each of Exercises 1–10, determine whether the given vectors are linearly dependent or linearly independent.

**1.** $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

**2.** $\begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} -8 \\ 2 \end{bmatrix}$

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

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

**5.** $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

**6.** $\begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -1 \end{bmatrix}$

**7.** $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

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

**9.** $x^2 + x + 2, x^2 + 2x + 1, 2x^2 + 5x + 1$

**10.** $x^3 - 1, x^2 - 1, x - 1, 1$

**11.** Show that $\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ form a basis for $\mathbb{R}^2$.

**12.** Show that $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}$ form a basis for $\mathbb{R}^2$.

**13.** Show that $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ form a basis for $\mathbb{R}^3$.

**14.** Show that $\begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -4 \\ -1 \end{bmatrix}$ form a basis for $\mathbb{R}^3$.

**15.** Show that $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$ form a basis for $M_{2\times2}(\mathbb{R})$.

**16.** Show that $\begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -1 \\ -1 & 2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 3 & 0 \\ 0 & 2 & 3 \end{bmatrix}$ form a basis for $M_{2\times3}(\mathbb{R})$.

**17.** Show that $x^2 + x + 1, x^2 - x + 1, x^2 - 1$ form a basis for $P_2$.

**18.** Show that $x^3 + x, x^2 - x, x + 1, x^3 + 1$ form a basis for $P_3$.

**19.** Show that $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ do not form a basis for $\mathbb{R}^3$.

**20.** Show that $x^2 - 3x, x + 7$ do not form a basis for $P_2$.

**21.** Show that $x + 1, x + 2, x + 3$ do not form a basis for $P_1$.

**22.** Show that $\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ do not form a basis for $\mathbb{R}^3$.

**23.** If $\alpha$ is the basis in Exercise 13, find:

**a)** $[v]_\alpha$ for $v = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$,

**b)** $v$ if $[v]_\alpha = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.

**24.** If $\beta$ is the basis in Exercise 14, find:

**a)** $[v]_\beta$ if $v = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$,

**b)** $v$ if $[v]_\beta = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

**25.** If $\beta$ is the basis in Exercise 17, find:

**a)** $[v]_\beta$ if $v = 2x^2 + 3x$,

**b)** $v$ if $[v]_\beta = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

**26.** If $\gamma$ is the basis of Exercise 18, find:

**a)** $[v]_\gamma$ if $v = x^3 + x^2 + x + 1$,

**b)** $v$ if $[v]_\gamma = \begin{bmatrix} -2 \\ 2 \\ 0 \\ 1 \end{bmatrix}$.

**27.** Show that any set of vectors that contains the zero vector is a linearly dependent set of vectors.

**28.** Show that if $v_1, v_2, \ldots, v_n$ is a linearly independent set of vectors, then any subset of these vectors is also linearly independent.

**29.** Suppose $v_1$ and $v_2$ are nonzero vectors in $\mathbb{R}^3$ and $L$ is a line in $\mathbb{R}^3$ parallel to $v_1$. What are necessary and sufficient conditions in terms of the line $L$ for $v_1$ and $v_2$ to be linearly independent?

**30.** Let $v_1$ and $v_2$ be vectors in $\mathbb{R}^3$ that are linearly independent.

**a)** If all initial points of vectors are placed at the same point in $\mathbb{R}^3$, what geometric object do the linear combinations of $v_1$ and $v_2$ determine?

**b)** If $v_3$ is a third vector in $\mathbb{R}^3$, what are necessary and sufficient conditions for $v_1, v_2, v_3$ to be linearly independent in terms of this geometric object?

Use the system of linear equation solving capabilities of Maple or another appropriate software package in Exercises 31 and 32.

**31. a)** Show that the vectors

$$\begin{bmatrix} 13 \\ 15 \\ 11 \\ -1 \\ 53 \\ 16 \end{bmatrix}, \begin{bmatrix} 1 \\ 22 \\ 12 \\ -9 \\ 18 \\ 77 \end{bmatrix}, \begin{bmatrix} -5 \\ -14 \\ 13 \\ 10 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 81 \\ 14 \\ 101 \\ 11 \\ 15 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 15 \\ 2 \\ 99 \\ -68 \end{bmatrix}, \begin{bmatrix} 7 \\ 3 \\ 16 \\ 88 \\ -49 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^6$.

**b)** Find the coordinates of

$$\begin{bmatrix} 29 \\ 4 \\ -9 \\ 13 \\ 71 \\ -51 \end{bmatrix}$$

relative to the basis in part (a).

**32. a)** Show that

$$x^6 - x^5 + 3x + 1,\ x^5 + x^4 + x^2 + x + 1,$$
$$-x^6 + x^4 + x^3 + 3x^2 + x + 1,$$
$$2x^6 - x^4 - 2x^3 - 5x^2 - 1,$$
$$3x^6 + x^5 + 2x^4 + x^3 - x^2 - x - 1,$$
$$x^6 + 3x^5 + 2x^4 - 3x^3 + 2x^2 - 2x - 1,$$
$$2x^6 - x^5 + 2x^3 + x^2 - 3x + 1$$

form a basis for $P_6$.

**b)** Find the coordinates of $7x^6 + 6x^5 - 5x^4 - 4x^3 - 3x^2 + 2x - 1$ relative to the basis in part (a).

**33.** We saw in the solution to Example 11 that the three vectors given in this example are linearly dependent in $\mathbb{R}^2$. Show that this can be generalized to the following: If $v_1, v_2, \ldots, v_m$ are vectors in $\mathbb{R}^n$ and if $m > n$, then $v_1, v_2, \ldots, v_m$ are linearly dependent.

**34.** Let $v_1, v_2, \ldots, v_n$ be vectors in $\mathbb{R}^n$. Show that $v_1, v_2, \ldots, v_n$ form a basis for $\mathbb{R}^n$ if and only if the matrix $[v_1\ v_2\ \ldots\ v_n]$ is nonsingular.

**35.** Use the result of Exercise 34 and a suitable test for invertibility of a matrix in Maple or another appropriate software package to show that the vectors in Exercise 31(a) form a basis for $\mathbb{R}^6$.

## 2.4 Dimension; Nullspace, Row Space, and Column Space

Looking back through the last section at all the examples of bases for $\mathbb{R}^n$ we did for various values of $n$ or any exercises you did involving bases for these spaces, you might notice every basis had $n$ elements. This is no accident. In fact, our first main objective in this section will be to show that once a vector space has a basis of $n$ vectors, then every other basis also has $n$ vectors. To reach this objective, we first prove the following lemma.

**LEMMA 2.8** If $v_1, v_2, \ldots, v_n$ are a basis for a vector space $V$, then every set of vectors $w_1, w_2, \ldots, w_m$ in $V$ where $m > n$ is linearly dependent.

***Proof*** We must show that there is a nontrivial linear combination

$$c_1 w_1 + c_2 w_2 + \cdots + c_m w_m = 0. \tag{1}$$

To this end, let us first write each $w_i$ as a linear combination of our basis vectors $v_1, v_2, \ldots, v_n$:

$$\begin{aligned} w_1 &= a_{11} v_1 + a_{21} v_2 + \cdots + a_{n1} v_n \\ w_2 &= a_{12} v_1 + a_{22} v_2 + \cdots + a_{n2} v_n \\ &\ \ \vdots \\ w_m &= a_{1m} v_1 + a_{2m} v_2 + \cdots + a_{nm} v_n. \end{aligned} \tag{2}$$

Substituting the results of Equations (2) into Equation (1), we have

$$\begin{aligned} &c_1(a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n) + c_2(a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n) \\ &\quad + \cdots + c_m(a_{1m}v_1 + a_{2m}v_2 + \cdots + a_{nm}v_n) = \\ &(a_{11}c_1 + a_{12}c_2 + \cdots + a_{1m}c_m)v_1 + (a_{21}c_1 + a_{22}c_2 + \cdots + a_{2m}c_m)v_2 \\ &\quad + \cdots + (a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nm}c_m)v_n = 0. \end{aligned}$$

Since $v_1, v_2, \ldots, v_n$ are linearly independent, the last equation tells us that we have the homogeneous system:

$$\begin{aligned}
a_{11}c_1 + a_{12}c_2 + \cdots + a_{1m}c_m &= 0 \\
a_{21}c_1 + a_{22}c_2 + \cdots + a_{2m}c_m &= 0 \\
&\vdots \\
a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nm}c_m &= 0.
\end{aligned}$$

Since $m > n$, we know by Theorem 1.1 that this system has nontrivial solutions. Thus we have nontrivial linear combinations equal to the zero vector in Equation (1) and hence $w_1, w_2, \ldots, w_m$ are linearly dependent. ●

Now we are ready to prove the result referred to at the beginning of this section, which we state as follows.

**THEOREM 2.9** If $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_m$ both form bases for a vector space $V$, then $n = m$.

***Proof*** Applying Lemma 2.8 with $v_1, v_2, \ldots, v_n$ as the basis, we must have $m \leq n$ or else $w_1, w_2, \ldots, w_m$ would be linearly dependent. Interchanging the roles of $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_m$, we obtain $n \leq m$. Hence $n = m$. ●

The number of vectors in a basis (which Theorem 2.9 tells us is always the same) is what we call the dimension of the vector space.

DEFINITION If a vector space $V$ has a basis of $n$ vectors, we say the **dimension** of $V$ is $n$.

We denote the dimension of a vector space $V$ by

$$\boxed{\dim(V).}$$

Thus, for example, since the standard basis $e_1, e_2, \ldots, e_n$ for $\mathbb{R}^n$ has $n$ vectors,

$$\boxed{\dim(\mathbb{R}^n) = n.}$$

Since the standard basis for $M_{m\times n}(\mathbb{R})$ consists of the $mn$ matrices $E_{ij}$ with 1 in the $ij$-position and 0s elsewhere,

$$\boxed{\dim(M_{m\times n}(\mathbb{R})) = mn.}$$

Since the standard basis $x^n, \ldots, x, 1$ for $P_n$ has $n + 1$ elements,

$$\boxed{\dim(P_n) = n + 1.}$$

Not every vector space $V$ has a basis consisting of a finite number of vectors, but we can still introduce dimensions for such vector spaces. One such case occurs when $V$ is the zero vector space consisting of only the zero vector. The zero vector space does not have a basis. In fact, the set consisting of the zero vector is the only spanning set for the zero vector space. But it is not a basis since, as you were asked to show in Exercise 27 in the previous section, no set containing the zero vector is linearly independent. For obvious reasons, if $V$ is the zero vector space, we take the dimension of $V$ to be 0 and write $\dim(V) = 0$. The zero vector space along with vector spaces that have bases with a finite number of vectors are called **finite dimensional** vector spaces. Vector spaces $V$ that are not finite dimensional are called **infinite dimensional** vector spaces and we write $\dim(V) = \infty$. It can be proven that infinite dimensional vector spaces have bases with infinitely many vectors, but we will not attempt to prove this here.[7] The set of all polynomials $P$ is an example of an infinite dimensional vector space. Indeed, the polynomials $1, x, x^2, x^3, \ldots$ form a basis for $P$. Many of the other function spaces we looked at in Sections 2.1 and 2.2, such as $F(a, b)$, $C(a, b)$, $D(a, b)$, and $C^\infty(a, b)$, are infinite dimensional, but we will not attempt to give bases for these vector spaces.

We next develop some facts that will be useful to us from time to time. We begin with the following lemma.

**LEMMA 2.10** Let $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_m$ be vectors in a vector space $V$. Then $\text{Span}\{v_1, v_2, \ldots, v_n\} = \text{Span}\{w_1, w_2, \ldots, w_m\}$ if and only if each $v_i$ is a linear combination of $w_1, w_2, \ldots, w_m$ and each $w_i$ is a linear combination of $v_1, v_2, \ldots, v_n$.

*Proof* Suppose that $\text{Span}\{v_1, v_2, \ldots, v_n\} = \text{Span}\ \{w_1, w_2, \ldots, w_m\}$. Since each $v_i$ is in $\text{Span}\{v_1, v_2, \ldots, v_n\}$ (use 1 for the scalar on $v_i$ and 0 for the scalar on all other $v_j$ to write $v_i$ as a linear combination of $v_1, v_2, \ldots, v_n$), $v_i$ is in $\text{Span}\{w_1, w_2, \ldots, w_m\}$. Hence $v_i$ is a linear combination of $w_1, w_2, \ldots, w_m$. Likewise each $w_i$ is a linear combination of $v_1, v_2, \ldots, v_n$.

To prove the converse, first note that if each $v_i$ is a linear combination of $w_1, w_2, \ldots, w_m$, then each $v_i$ is in $\text{Span}\{w_1, w_2, \ldots, w_m\}$. Since subspaces are closed under addition and scalar multiplication, they are also closed under linear combinations. Hence any linear combination of $v_1, v_2, \ldots, v_n$ lies in $\text{Span}\{w_1, w_2, \ldots, w_m\}$ giving us that $\text{Span}\{v_1, v_2, \ldots, v_n\}$ is contained in $\text{Span}\{w_1, w_2, \ldots, w_m\}$. Likewise we will have $\text{Span}\{w_1, w_2, \ldots, w_m\}$ is contained in $\text{Span}\{v_1, v_2, \ldots, v_n\}$ so that these two subspaces are equal as desired. ●

The next lemma tells us that we can extend linearly independent sets of vectors to bases and reduce spanning sets to bases by eliminating vectors if necessary.

**LEMMA 2.11** Suppose that $V$ is a vector space of positive dimension $n$.

1. If $v_1, v_2, \ldots v_k$ are linearly independent vectors in $V$, then there exist vectors $v_{k+1}, \ldots, v_n$ so that $v_1, \ldots, v_k, v_{k+1}, \ldots, v_n$ form a basis for $V$.
2. If $v_1, v_2, \ldots, v_k$ span $V$, then there exists a subset of $v_1, v_2, \ldots, v_k$ that forms a basis of $V$.

---

[7] Typical proofs involve using a result called Zorn's Lemma, which you may well encounter if you continue your study of mathematics.

*Proof*

1. Notice that $k \leq n$ by Lemma 2.8. If $k < n$, $v_1, v_2, \ldots, v_k$ cannot span $V$; otherwise $\dim(V) = k$ not $n$. Thus there is a vector $v_{k+1}$ not in $\text{Span}\{v_1, v_2, \ldots, v_k\}$. We must have $v_1, v_2, \ldots, v_k, v_{k+1}$ are linearly independent. To see this, suppose
$$c_1v_1 + c_2v_2 + \cdots + c_kv_k + c_{k+1}v_{k+1} = 0$$
where $c_1, c_2, \ldots, c_k, c_{k+1}$ are scalars. Were $c_{k+1} \neq 0$, we could solve this equation for $v_{k+1}$ obtaining $v_{k+1}$ is in $\text{Span}\{v_1, v_2, \ldots, v_k\}$, which we know is not the case. Now that $c_{k+1} = 0$ in this linear combination, the linear independence of $v_1, v_2, \ldots, v_k$ tells us we also have $c_1 = 0, c_2 = 0, \ldots, c_k = 0$. Hence we have the linear independence of $v_1, v_2, \ldots, v_{k+1}$. If $k + 1 < n$, repeat this procedure again. After $n - k$ steps we arrive at a set of $n$ linearly independent vectors $v_1, v_2, \ldots, v_n$. These $n$ vectors must span $V$; otherwise we could repeat our procedure again obtaining $n + 1$ linearly independent vectors $v_1, v_2, \ldots, v_n, v_{n+1}$, which is impossible by Lemma 2.8. Thus we have arrived at the desired basis.
2. If $v_1, v_2, \ldots, v_k$ are linearly independent, they then form the desired basis. If not, one of them must be a linear combination of the others. Relabeling if necessary to make the notation simpler, we may assume $v_k$ is a linear combination of $v_1, v_2, \ldots, v_{k-1}$. Since each $v_1, v_2, \ldots, v_k$ is a linear combination of $v_1, v_2, \ldots, v_{k-1}$ and vice versa, it follows by Lemma 2.10 that $v_1, v_2, \ldots, v_{k-1}$ span $V$. If $v_1, v_2, \ldots, v_{k-1}$ are also linearly independent, we have our desired basis; if not, repeat the procedure we just did again. Since such steps cannot go on forever, we must obtain the desired type of basis after a finite number of steps. ●

At the beginning of this section we noted that the examples and exercises of the previous section suggested that once one basis has $n$ elements, all other bases also have $n$ elements. We then went on to prove this (Theorem 2.9). Another thing you might notice from the examples and exercises in the last section is that every time we had $n$ vectors in a vector space of dimension $n$, having one of the properties of linear independence or spanning seemed to force the other property to hold too. The next theorem tells us this is indeed the case.

**THEOREM 2.12** Suppose that $V$ is a vector space of dimension $n$.

1. If the vectors $v_1, v_2, \ldots, v_n$ are linearly independent, then $v_1, v_2, \ldots, v_n$ form a basis for $V$.
2. If the vectors $v_1, v_2, \ldots, v_n$ span $V$, then $v_1, v_2, \ldots, v_n$ form a basis for $V$.

*Proof*

1. By part (1) of Lemma 2.11, we can extend $v_1, v_2, \ldots, v_n$ to a basis of $V$. But since every basis of $V$ has $n$ elements, $v_1, v_2, \ldots, v_n$ must already form a basis for $V$.

2. By part (2) of Lemma 2.11, we know some subset of the set of vectors $v_1, v_2, \ldots, v_n$ forms a basis of $V$. Again since every basis has $n$ elements, this subset forming the basis must be the entire set of vectors $v_1, v_2, \ldots, v_n$. ●

**EXAMPLE 1** Show that $x^2 - 1, x^2 + 1, x + 1$ form a basis for $P_2$.

*Solution* Since $\dim(P_2) = 3$ and we are given three vectors in $P_2$, Theorem 2.12 tells us we can get by with showing either linear independence or spanning. Let us do linear independence. If

$$c_1(x^2 - 1) + c_2(x^2 + 1) + c_3(x + 1) = 0,$$

we have the system:

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_3 &= 0 \\ -c_1 + c_2 + c_3 &= 0. \end{aligned}$$

It is easily seen that this system has only the trivial solution. Thus $x^2 - 1, x^2 + 1, x + 1$ are linearly independent and hence form a basis for $P_2$. ●

Of course, notice that Theorem 2.12 allows us to get by with checking for linear independence or spanning only when we already know the dimension of a finite dimensional vector space $V$ and are given the same number of vectors as $\dim(V)$. If we do not know $\dim(V)$ for a finite dimensional vector space $V$, we must check both linear independence and spanning. Notice too that if we do know the dimension of a vector space $V$ is $n$, no set of vectors with fewer or more than $n$ elements could form a basis since all bases must have exactly $n$ elements. For example, neither the set of two vectors $(1, 1, 1), (1, -1, 3)$ nor the set of four vectors $(1, 1, 1), (1, -1, 3), (0, 1, 1), (2, 1, -1)$ could form a basis for the three-dimensional space $\mathbb{R}^3$.

We conclude this section with a discussion of techniques for finding bases for three important subspaces associated with a matrix. Recall that the solutions to the homogeneous system $AX = 0$ where $A$ is an $m \times n$ matrix form a subspace of $\mathbb{R}^n$ (Theorem 2.4). This vector space of solutions is called the **nullspace** or **kernel** of the matrix $A$ and we shall denote it by

$$\boxed{NS(A).}$$

The manner in which we wrote our solutions to homogeneous systems in Chapter 1 leads us naturally to a basis for $NS(A)$. Consider the following example.

**EXAMPLE 2** Find a basis for $NS(A)$ if

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

*Solution* Reducing the augmented matrix for the system $AX = 0$ to reduced row-echelon form,

$$\left[\begin{array}{ccccc|c} 1 & 2 & -1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 4 & 1 & 0 \\ 1 & 4 & -3 & 1 & -2 & 0 \end{array}\right] \to \left[\begin{array}{ccccc|c} 1 & 2 & -1 & 3 & 0 & 0 \\ 0 & -1 & 1 & 1 & 1 & 0 \\ 0 & 2 & -2 & -2 & -2 & 0 \end{array}\right]$$
$$\to \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 5 & 2 & 0 \\ 0 & 1 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

we see our solutions are

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -x_3 - 5x_4 - 2x_5 \\ x_3 + x_4 + x_5 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}.$$

If we express this column vector as

$$x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

we immediately have that every solution is a linear combination of

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

so that these three columns span $NS(A)$. They also are easily seen to be linearly independent. (The only such linear combination of them that is the zero column is the one with $x_3 = 0, x_4 = 0, x_5 = 0$.) Hence these three columns form a basis for $NS(A)$. ●

The subspace of $M_{1\times n}(\mathbb{R})$ spanned by the rows of an $m \times n$ matrix $A$ is called the **row space** of $A$ and is denoted

$$\boxed{RS(A).}$$

For instance, if $A$ is the matrix in Example 2, the row space of $A$ is

$$\text{Span}\left\{\begin{bmatrix} 1 & 2 & -1 & 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 & 4 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 4 & -3 & 1 & -2 \end{bmatrix}\right\}.$$

Notice that if we perform an elementary row operation on a matrix $A$, the resulting matrix has its rows being linear combinations of the rows of $A$ and vice versa since elementary row operations are reversible. Consequently, $A$ and the resulting matrix have the same row space by Lemma 2.10. More generally this holds if we repeatedly do elementary row operations and hence we have the following theorem.

**THEOREM 2.13** If $A$ and $B$ are row equivalent matrices, then

$$RS(A) = RS(B).$$

Because of Theorem 2.13, if $B$ is the reduced row-echelon form of a matrix $A$, then $B$ has the same row space as $A$. It is easy to obtain a basis for $RS(B)$ (which equals $RS(A)$) as the following example illustrates.

**EXAMPLE 3** Find a basis for the row space of the matrix

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

of Example 2.

*Solution* From the solution to Example 2, we see the reduced row-echelon form of $A$ is

$$B = \begin{bmatrix} 1 & 0 & 1 & 5 & 2 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The nonzero rows of $B$ span $RS(B) = RS(A)$. They also are linearly independent. (If

$$c_1 \begin{bmatrix} 1 & 0 & 1 & 5 & 2 \end{bmatrix} + c_2 \begin{bmatrix} 0 & 1 & -1 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

we see from the first two entries that $c_1 = 0$ and $c_2 = 0$.) Hence

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

form a basis for $RS(A)$. ●

There is a relationship between the dimensions of $RS(A)$ and $NS(A)$ and the number of columns $n$ of an $m \times n$ matrix $A$. Notice that $\dim(RS(A))$ is the number of nonzero rows of the reduced row-echelon form of $A$, which is the same as the number of nonfree variables in the solutions of the homogeneous system $AX = 0$. Also notice that $\dim(NS(A))$ is the number of free variables in the solutions of the homogeneous system $AX = 0$. Since the number of nonfree variables plus the number of free variables is the total number of variables in the system $AX = 0$, which is $n$, we have the following theorem.

**THEOREM 2.14** If $A$ is an $m \times n$ matrix,

$$\dim(RS(A)) + \dim(NS(A)) = n.$$

The final subspace associated with an $m \times n$ matrix $A$ we shall introduce is the subspace of $\mathbb{R}^m$ spannned by the columns of $A$, which is called the **column space** of $A$ and is denoted

$$\boxed{CS(A).}$$

For the matrix $A$ in Examples 2 and 3, the column space is

$$\text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}\right\}.$$

In the same manner as we use row operations to find bases for the row space of a matrix $A$, we could find a basis for the column space of $A$ by using column operations to get the column reduced echelon form of $A$. The nonzero columns of the column reduced echelon form will form a basis for $CS(A)$. But if you feel more comfortable using row operations as we authors do, notice that a basis for $CS(A)$ can be found by reducing $A^T$ to row-echelon form and then transposing the basis vectors of $RS(A^T)$ back to column vectors. This is the approach we take in the next example.

**EXAMPLE 4** Find a basis for the column space of the matrix

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

of Examples 2 and 3.

***Solution*** Reducing $A^T$ to reduced row-echelon form,

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

from which we see

$$\begin{bmatrix} 1 & 0 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & -2 \end{bmatrix}^T = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$$

form a basis for $CS(A)$. ●

You might note that we actually do not have to go all the way to reduced row-echelon form to see a basis. For instance, it is easy to see that the first two rows in the second matrix of the solution to Example 4 would have to give us a basis so that we could equally well use

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

as a basis for $CS(A)$ in Example 4.

The dimensions of $RS(A)$ and $CS(A)$ are the same for the matrix $A$ in Examples 3 and 4. (Both dimensions are 2.) In Corollary 5.5 of Chapter 5 we shall prove that this is always the case; that is, for any matrix $A$,

$$\boxed{\dim(RS(A)) = \dim(CS(A)).}$$

This common dimension is called the **rank** of the matrix $A$ and is denoted $\operatorname{rank}(A)$. For instance, if $A$ is the matrix in Examples 3 and 4, then $\operatorname{rank}(A) = 2$. Observe that the rank of a matrix is the same as the number of nonzero rows (columns) in its reduced row (column)-echelon form.

Sometimes we have to find a basis for a subspace of $\mathbb{R}^n$ spanned by several vectors of $\mathbb{R}^n$. This is the same as finding a basis for the column space of a matrix as the final example of this section illustrates.

**EXAMPLE 5** Find a basis for the subspace of $\mathbb{R}^3$ spanned by

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

*Solution* The subspace spanned by these three vectors is the same as the column space of the matrix

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

Reducing $A^T$,

$$A^T = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}.$$

It is apparent that

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

form a basis for $CS(A)$ and hence for the subspace spanned by the three given vectors. ●

## EXERCISES 2.4

**1.** In each of parts (a)–(d), determine whether the given vectors form a basis for $\mathbb{R}^2$.

**a)** $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

**b)** $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

**c)** $\begin{bmatrix} -12 \\ 16 \end{bmatrix}, \begin{bmatrix} 9 \\ -12 \end{bmatrix}$

**d)** $\begin{bmatrix} -2 \\ -7 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} -5 \\ 8 \end{bmatrix}$

**2.** In each of parts (a)–(d), determine whether the given vectors form a basis for $\mathbb{R}^3$.

**a)** $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 8 \end{bmatrix}$

**b)** $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$

**c)** $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}$

**d)** $\begin{bmatrix} -7 \\ -9 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

**3.** In each of parts (a)–(d), determine whether the given polynomials form a basis for $P_2$.

**a)** $x^2 + x - 1, 2x^2 - 3, x^2 + x + 2$

**b)** $5 - 4x^2, 3 - 2x$

**c)** $x^2 + x - 1, x^2 + x + 2, x^2 + x + 14$

**d)** $x^2 + x, x + 1, x^2 + 1, 1$

**4.** In each of parts (a)–(d), determine whether the given matrices form a basis for $M_{2\times 2}(\mathbb{R})$.

**a)** $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

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

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

**d)** $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Do the following for the matrices in Exercises 5–12:

**a)** Find a basis for the nullspace of the matrix.

**b)** Find a basis for the row space of the matrix.

**c)** Find a basis for the column space of the matrix.

**d)** Determine the rank of the matrix.

(Parts (a)–(c) do not have unique answers.)

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

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

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

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

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

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

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

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

In Exercises 13–15, find a basis for the subspace spanned by the given vectors. (These do not have unique answers.)

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

14. $\begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ -1 \end{bmatrix}$

15. $\begin{bmatrix} -2 \\ 1 \\ 3 \\ -4 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 9 \\ -11 \end{bmatrix}$

16. Without formally showing that there is a nontrivial linear combination of $x^4-2x$, $x^4+x^3-1$, $x^3+x+3$, $x-x^2-x^4$, $10x-91$, $\pi x^4+\sqrt{3}x^3-7x^2$ equal to the zero polynomial, we can conclude these polynomials are linearly dependent. Why is this?

17. Show that $\det(A) \neq 0$ for an $n \times n$ matrix $A$ if and only if $\operatorname{rank}(A) = n$.

18. Suppose that $A$ is an $m \times n$ matrix. Show that if $m > n$, then the rows of $A$ are linearly dependent.

19. Suppose that $A$ is an $m \times n$ matrix. Show that if $m < n$, then the columns of $A$ are linearly dependent.

20. Consider the linearly independent polynomials $p_1(x) = x^2 + x$, $p_2(x) = x + 1$. Find a polynomial $p_3(x)$ so that $p_1(x)$, $p_2(x)$, $p_3(x)$ form a basis for $P_2$ in the manner of the proof of part (1) of Lemma 2.11.

21. **a)** Show that the polynomials $p_1(x) = x^2 - 1$, $p_2(x) = x^2 + 1$, $p_3(x) = x - 1$, $p_4(x) = x + 1$ span $P_2$.

**b)** Find a subset of $p_1(x)$, $p_2(x)$, $p_3(x)$, $p_4(x)$ that forms a basis for $P_2$ in the manner of the proof of part (2) of Lemma 2.11.

22. If the initial points of vectors in $\mathbb{R}^3$ are placed at the origin, what geometric object is a subspace of $\mathbb{R}^3$ of dimension one? What geometric object is a subspace of $\mathbb{R}^3$ of dimension two?

23. Explain why the reduced row-echelon form of a matrix is unique. Is the reduced column-echelon form unique? Why or why not?

24. Use the system of linear equation solving capabilities of Maple or another appropriate software package to show that the matrices

$$\begin{bmatrix} 1 & 0 \\ -3 & -33 \\ 11 & -2 \end{bmatrix}, \begin{bmatrix} 1 & -3 \\ 1 & 13 \\ 16 & -2 \end{bmatrix}, \begin{bmatrix} 1 & 17 \\ -3 & 46 \\ 32 & -9 \end{bmatrix}, \begin{bmatrix} -1 & 91 \\ 6 & -98 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & -1 \\ 21 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 3 & 2 \\ 5 & -5 \\ -4 & 0 \end{bmatrix}$$

form a basis for $M_{3\times 2}(\mathbb{R})$.

25. Let $A$ be the matrix

$$A = \begin{bmatrix} -2 & 3 & -1 & 4 & -2 \\ 3 & -1 & 2 & 6 & 8 \\ 1 & 7 & -5 & 1 & 4 \\ 0 & 12 & -5 & 15 & 8 \end{bmatrix}.$$

**a)** In Maple, the command *nullspace* (or equivalently, *kernel*) can be used to find a basis for the nullspace of a matrix. (The basis vectors will be given as row vectors instead of column vectors.) Use this command or a corresponding command in an appropriate software package to find a basis for the nullspace of the matrix $A$.

**b)** Bases for the row space of a matrix can be found with Maple by using the *gausselim* or *gaussjord* (or equivalently, *rref*) commands. Such bases may also be found using the *rowspace* and *rowspan* commands. Find bases for row space of the matrix $A$ using each of these four Maple commands or corresponding commands in another appropriate software package. If you are using Maple, compare your results obtained with these four different commands.

c) Bases for the column space of a matrix can be found with Maple by using the *gausselim* or *gaussjord* commands on the transpose of the matrix. Such bases may also be found by using the *colspace* and *colspan* commands. Find bases for the column space of the matrix $A$ using each of these four methods in Maple or corresponding commands in another appropriate software package. If you are using Maple, compare your results obtained with these four different approaches.

**26.** The *basis* command of Maple may be used to find a basis for the subspace spanned by a finite number of row vectors in $M_{1\times n}(\mathbb{R})$. Use this command or the corresponding command in an appropriate software package to find a basis for the subspace of $M_{1\times 5}(\mathbb{R})$ spanned by the vectors

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

$$\begin{bmatrix} 1 & 7 & -5 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 12 & -5 & 15 & 8 \end{bmatrix}.$$

Compare the result you obtain here with your results in Exercise 25(b).

**27.** Suppose that $v_1, v_2, \ldots, v_k$ are linearly independent vectors in $\mathbb{R}^n$. Let $v_{k+1}$ be another vector in $\mathbb{R}^n$. How could the *gausselim* or *gaussjord* commands of Maple or corresponding commands in another appropriate software package be used to determine if $v_{k+1}$ is in $\text{Span}\{v_1, v_2, \ldots, v_k\}$?

**28.** Use the result of Exercise 27 to extend the set consisting of the vectors

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

to a basis of $\mathbb{R}^4$ in the manner of the proof of part (1) of Lemma 2.11.

## 2.5 WRONSKIANS

In our work with linear differential equations in Chapter 4 we will have instances where we will have to determine whether a set of functions forms a linearly independent set. For instance, problems such as the following will arise: Are the functions given by $e^x, \cos x, \sin x$ linearly independent? Consider a linear combination of these functions that equals the zero function:

$$c_1 e^x + c_2 \cos x + c_3 \sin x = 0.$$

The only such linear combination is the trivial one. One way to see this is to choose three values of $x$ such as $x = 0, \pi/2, \pi$. Substituting these values of $x$ into our linear combination, we have the system:

$$\begin{aligned} c_1 + c_2 &= 0 \\ e^{\pi/2} c_1 + c_3 &= 0 \\ e^{\pi} c_1 - c_2 &= 0. \end{aligned}$$

Adding the first and third equations, we obtain $(1 + e^{\pi})c_1 = 0$ from which we get $c_1 = 0$. It then follows that $c_2 = 0$ and $c_3 = 0$. Hence $e^x$, $\cos x$, and $\sin x$ are linearly independent. But there is another way to arrive at a system of equations involving derivatives that is often more convenient to use for showing linear independence.

Suppose that we have $n$ functions $f_1, f_2, \ldots, f_n$ each of which have $(n-1)$st derivatives on an open interval $(a, b)$; that is, assume $f_1, f_2, \ldots, f_n$ all lie in $D^{n-1}(a, b)$.[8] Consider a linear combination of these functions that is equal to the zero function:

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

Considering this equation along with the first $n-1$ derivatives of each side of it, we arrive at the following system:

$$\begin{aligned} c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) &= 0 \\ c_1 f_1'(x) + c_2 f_2'(x) + \cdots + c_n f_n'(x) &= 0 \\ c_1 f_1''(x) + c_2 f_2''(x) + \cdots + c_n f_n''(x) &= 0 \\ &\vdots \\ c_1 f_1^{(n-1)}(x) + c_2 f_2^{(n-1)}(x) + \cdots + c_n f_n^{(n-1)}(x) &= 0. \end{aligned} \tag{1}$$

If there is some $x$ in $(a, b)$ for which this system has only the trivial solution, then $f_1, f_2, \ldots, f_n$ will be linearly independent. Having such an $x$ is the same as having an $x$ in $(a, b)$ for which the the matrix

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{bmatrix}$$

is nonsingular. A convenient way of seeing if there is an $x$ in $(a, b)$ where this matrix is nonsingular is by looking at its determinant, which is called the **Wronskian**[9] of the functions $f_1, f_2, \ldots, f_n$. We will denote the Wronskian of $f_1, f_2, \ldots, f_n$ by $w(f_1(x), f_2(x), \ldots, f_n(x))$:

$$\boxed{w(f_1(x), f_2(x), \ldots, f_n(x)) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\ \vdots & \vdots & \vdots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}.}$$

Since a square matrix is nonsingular if and only if its determinant is nonzero (Theorem 1.21), we have the following theorem.

---

[8] In fact, what we are about to do does not require the interval to be open; it will work on other types of intervals (such as closed ones) as well.

[9] The Wronskian is named in honor of the Polish-French mathematician Jósef Maria Hoëné-Wronski (1776–1853).

**THEOREM 2.15** Suppose that $f_1, f_2, \ldots, f_n$ are functions in $D^{n-1}(a, b)$. If the Wronskian $w(f_1(x), f_2(x), \ldots, f_n(x))$ of $f_1, f_2, \ldots, f_n$ is nonzero for some $x$ in $(a, b)$, then $f_1, f_2, \ldots, f_n$ are linearly independent elements of $D^{n-1}(a, b)$.

Let us use the Wronskian to show $e^x, \cos x, \sin x$ are linearly independent (with $(-\infty, \infty)$ for the interval).

**EXAMPLE 1** Determine if $e^x, \cos x, \sin x$ are linearly independent.

*Solution* We have

$$\begin{aligned} w(e^x, \cos x, \sin x) &= \begin{vmatrix} e^x & \cos x & \sin x \\ e^x & -\sin x & \cos x \\ e^x & -\cos x & -\sin x \end{vmatrix} \\ &= e^x(\sin^2 x + \cos^2 x) - \cos x(-e^x \sin x - e^x \cos x) \\ &\quad + \sin x(-e^x \cos x + e^x \sin x) \\ &= 2e^x. \end{aligned}$$

Since $2e^x$ is not zero for some $x$ (in fact, it is not zero for every $x$), we have that $e^x, \cos x, \sin x$ are linearly independent by Theorem 2.15. ●

Be careful not to read too much into Theorem 2.15. It only tells us that if the Wronskian is nonzero for some $x$, then the functions are linearly independent. It does not tell us that the converse is true; that is, it does not tell us that linearly independent functions have their Wronskian being nonzero for some $x$ (or equivalently, that if the Wronskian is zero for all $x$, then the functions are linearly dependent). In fact, the converse of Theorem 2.15 does not hold in general. Here is an example illustrating this.

**EXAMPLE 2** Show that the Wronskian of the functions $f$ and $g$ where $f(x) = x^2$ and $g(x) = x|x|$ is zero for every $x$ and that $f$ and $g$ are linearly independent on $(-\infty, \infty)$.

*Solution* To calculate the Wronskian of $f$ and $g$, we need their derivatives. This is easy for $f$. To get the derivative of $g$ notice that we can also express $g(x)$ as

$$g(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ -x^2 & \text{if } x < 0 \end{cases}.$$

The graph of $g$ appears in Figure 2.3.

We have

$$g'(x) = \begin{cases} 2x & \text{if } x > 0 \\ -2x & \text{if } x < 0 \end{cases}.$$

At $x = 0$,

$$g'(0) = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} = \lim_{h \to 0} \frac{h|h|}{h} = \lim_{h \to 0} |h| = 0.$$

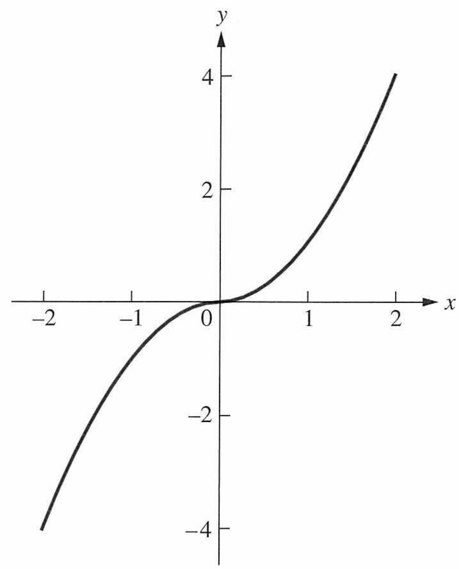


**Figure 2.3**

Thus we can say

$$g'(x) = \begin{cases} 2x & \text{if } x \geq 0 \\ -2x & \text{if } x < 0 \end{cases}.$$

If $x \geq 0$,

$$w(f(x), g(x)) = \begin{vmatrix} x^2 & x^2 \\ 2x & 2x \end{vmatrix} = 0.$$

If $x < 0$,

$$w(f(x), g(x)) = \begin{vmatrix} x^2 & -x^2 \\ 2x & -2x \end{vmatrix} = 0.$$

Hence we have that the Wronskian of $f$ and $g$ is zero for every $x$.

To see that $f$ and $g$ are linearly independent, suppose that

$$c_1 f(x) + c_2 g(x) = 0$$

where $c_1$ and $c_2$ are scalars. Substituting $x = 1$ and $x = -1$ into this equation, we arrive at the system

$$\begin{aligned} c_1 + c_2 &= 0 \\ c_1 - c_2 &= 0, \end{aligned}$$

which has only the trivial solution. Hence $f$ and $g$ are linearly independent as desired. ●

It is possible to obtain a converse of Theorem 2.15 provided we put some additional restrictions on the types of functions being considered. In Chapter 4 we will see a case in which this occurs.

## EXERCISES 2.5

In each of Exercises 1–9, show that the given functions are linearly independent on $(-\infty, \infty)$.

**1.** $e^{3x}, e^{-2x}$

**2.** $\cos 5x$, $\sin 5x$

**3.** $e^{2x}\cos x$, $e^{2x}\sin x$

**4.** $e^{-x}, xe^{-x}$

**5.** $x^2-1, x^2+1, x+1$

**6.** $e^x, e^{2x}, e^{3x}$

**7.** $e^{4x}, xe^{4x}, x^2e^{4x}$

**8.** $e^x, e^x\cos x, e^x\sin x$

**9.** $x^3, |x|^3$

**10.** Show that $1/x$, $x$ are linearly independent on $(0, \infty)$.

**11.** Show that $x+1$, $x-1$, $x$ are linearly dependent on $(-\infty, \infty)$.

**12.** Show that $\sin^2 x$, $\cos^2 x$, $\cos 2x$ are linearly dependent on $(-\infty, \infty)$.

**13.** Show that the two functions in Example 2 are linearly dependent on $[0, \infty)$.

**14.** Suppose that $a < c < d < b$ and that $f_1, f_2, \dots, f_n$ are functions in $F(a, b)$.

**a)** If $f_1, f_2, \dots, f_n$ are linearly independent on $(a, b)$, are $f_1, f_2, \dots, f_n$ necessarily linearly independent on $(c, d)$? Why or why not?

**b)** If $f_1, f_2, \dots, f_n$ are linearly independent on $(c, d)$, are $f_1, f_2, \dots, f_n$ necessarily linearly independent on $(a, b)$? Why or why not?

**15. a)** Find the Wronskian of $1, x, x^2, \dots, x^{n-1}$.

**b)** Show that

$$w(g(x)f_1(x), g(x)f_2(x), \dots, g(x)f_n(x)) = [g(x)]^n w(f_1(x), f_2(x), \dots, f_n(x)).$$

**c)** Show that $e^{rx}, xe^{rx}, x^2e^{rx}, \dots, x^{n-1}e^{rx}$ are linearly independent on $(-\infty, \infty)$.

**16.** Use the *wronskian* and *det* commands in Maple or appropriate commands in another software package to find the Wronskian of the functions $e^{3x}$, $xe^{3x}$, $e^{3x}\cos 2x$, $e^{3x}\sin 2x$, $x\cos x$, $x\sin x$. Are these functions linearly independent?