## 1.2 Gaussian Elimination

In this section we will develop a systematic procedure for solving systems of linear equations. The procedure is based on the idea of performing certain operations on the rows of the augmented matrix that simplify it to a form from which the solution of the system can be ascertained by inspection.

*Considerations in Solving Linear Systems*

When considering methods for solving systems of linear equations, it is important to distinguish between large systems that must be solved by computer and small systems that can be solved by hand. For example, there are many applications that lead to linear systems in thousands or even millions of unknowns. Large systems require special techniques to deal with issues of memory size, roundoff errors, solution time, and so forth. Such techniques are studied in the field of ***numerical analysis*** and will only be touched on in this text. However, almost all of the methods that are used for large systems are based on the ideas that we will develop in this section.

*Echelon Forms*

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

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

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

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

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

▶ EXAMPLE 1 **Row Echelon and Reduced Row Echelon Form**

The following matrices are in reduced row echelon form.

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

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

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



▶ EXAMPLE 2 **More on Row Echelon and Reduced Row Echelon Form**

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

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

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

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

If, by a sequence of elementary row operations, the augmented matrix for a system of linear equations is put in *reduced* row echelon form, then the solution set can be obtained either by inspection or by converting certain linear equations to parametric form. Here are some examples.

▶ EXAMPLE 3 **Unique Solution**

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

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

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

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

In Example 3 we could, if desired, express the solution more succinctly as the 4-tuple $(3, -1, 0, 5)$.

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

▶ EXAMPLE 4 **Linear Systems in Three Unknowns**

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

and setting $t = 1$ yields the solution

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

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

$$x - 5y + z = 4 \qquad (1)$$

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

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

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

$$x = 4 + 5s - t, \quad y = s, \quad z = t \; ◄ \qquad (2)$$

We will usually denote parameters in a general solution by the letters $r, s, t, \ldots$, but any letters that do not conflict with the names of the unknowns can be used. For systems with more than three unknowns, subscripted letters such as $t_1, t_2, t_3, \ldots$ are convenient.

Formulas, such as (2), that express the solution set of a linear system parametrically have some associated terminology.

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



*Elimination Methods*

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

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

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

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

Leftmost nonzero column

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

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

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

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

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

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

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

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

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

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

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

Leftmost nonzero column in the submatrix

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

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

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

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

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

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

**Leftmost nonzero column in the new submatrix**

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

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

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

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

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

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

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

⟵ −6 times the third row was added to the first row

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

⟵ 5 times the second row was added to the first row

The last matrix is in reduced row echelon form.

The procedure (or algorithm) we have just described for reducing a matrix to reduced row echelon form is called ***Gauss–Jordan elimination***. This algorithm consists of two parts, a ***forward phase*** in which zeros are introduced below the leading 1's and a ***backward phase*** in which zeros are introduced above the leading 1's. If only theforward phase is

Carl Friedrich Gauss (1777–1855)

Wilhelm Jordan (1842–1899)

Historical Note Although versions of Gaussian elimination were known much earlier, its importance in scientific computation became clear when the great German mathematician Carl Friedrich Gauss used it to help compute the orbit of the asteroid Ceres from limited data. What happened was this: On January 1, 1801 the Sicilian astronomer and Catholic priest Giuseppe Piazzi (1746–1826) noticed a dim celestial object that he believed might be a "missing planet." He named the object Ceres and made a limited number of positional observations but then lost the object as it neared the Sun. Gauss, then only 24 years old, undertook the problem of computing the orbit of Ceres from the limited data using a technique called "least squares," the equations of which he solved by the method that we now call "Gaussian elimination." The work of Gauss created a sensation when Ceres reappeared a year later in the constellation Virgo at almost the precise position that he predicted! The basic idea of the method was further popularized by the German engineer Wilhelm Jordan in his book on geodesy (the science of measuring Earth shapes) entitled *Handbuch der Vermessungskunde* and published in 1888.

[*Images: Photo Inc/Photo Researchers/Getty Images* (Gauss); *Leemage/Universal Images Group/Getty Images* (Jordan)]



used, then the procedure produces a row echelon form and is called ***Gaussian elimination***. For example, in the preceding computations a row echelon form was obtained at the end of Step 5.

▶ **EXAMPLE 5 Gauss–Jordan Elimination**

Solve by Gauss–Jordan elimination.

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

**Solution** The augmented matrix for the system is

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

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

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

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

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

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

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

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

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

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

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

Note that in constructing the linear system in (3) we ignored the row of zeros in the corresponding augmented matrix. Why is this justified?

The corresponding system of equations is

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

Solving for the leading variables, we obtain

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

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

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

*Homogeneous Linear Systems*

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



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

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

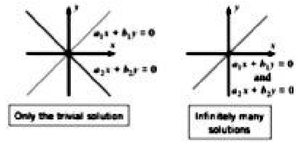
Because a homogeneous linear system always has the trivial solution, there are only two possibilities for its solutions:

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

In the special case of a homogeneous linear system of two equations in two unknowns, say

$$a_1x + b_1y = 0 \quad (a_1, b_1 \text{ not both zero})$$
$$a_2x + b_2y = 0 \quad (a_2, b_2 \text{ not both zero})$$

the graphs of the equations are lines through the origin, and the trivial solution corresponds to the point of intersection at the origin (Figure 1.2.1).




▶ Figure 1.2.1

There is one case in which a homogeneous system is assured of having nontrivial solutions—namely, whenever the system involves more unknowns than equations. To see why, consider the following example of four equations in six unknowns.



▶ EXAMPLE 6 **A Homogeneous System**

Use Gauss–Jordan elimination to solve the homogeneous linear system

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

*Solution* Observe first that the coefficients of the unknowns in this system are the same as those in Example 5; that is, the two systems differ only in the constants on the right side. The augmented matrix for the given homogeneous system is

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

which is the same as the augmented matrix for the system in Example 5, except for zeros in the last column. Thus, the reduced row echelon form of this matrix will be the same as that of the augmented matrix in Example 5, except for the last column. However, a moment's reflection will make it evident that a column of zeros is not changed by an elementary row operation, so the reduced row echelon form of (5) is

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

The corresponding system of equations is

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

Solving for the leading variables, we obtain

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

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

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

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

*Free Variables in Homogeneous Linear Systems*

Example 6 illustrates two important points about solving homogeneous linear systems:

1. Elementary row operations do not alter columns of zeros in a matrix, so the reduced row echelon form of the augmented matrix for a homogeneous linear system has a final column of zeros. This implies that the linear system corresponding to the reduced row echelon form is homogeneous, just like the original system.

2. When we constructed the homogeneous linear system corresponding to augmented matrix (6), we ignored the row of zeros because the corresponding equation

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 = 0$$

does not impose any conditions on the unknowns. Thus, depending on whether or not the reduced row echelon form of the augmented matrix for a homogeneous linear system has any rows of zero, the linear system corresponding to that reduced row echelon form will either have the same number of equations as the original system or it will have fewer.

Now consider a general homogeneous linear system with $n$ unknowns, and suppose that the reduced row echelon form of the augmented matrix has $r$ nonzero rows. Since each nonzero row has a leading 1, and since each leading 1 corresponds to a leading variable, the homogeneous system corresponding to the reduced row echelon form of the augmented matrix must have $r$ leading variables and $n - r$ free variables. Thus, this system is of the form

$$\begin{aligned} x_{k_1} \qquad\qquad & + \textstyle\sum(\ ) = 0 \\ x_{k_2} \qquad & + \textstyle\sum(\ ) = 0 \\ \ddots \qquad & \qquad \vdots \\ x_{k_r} & + \textstyle\sum(\ ) = 0 \end{aligned} \tag{8}$$

where in each equation the expression $\sum(\ )$ denotes a sum that involves the free variables, if any [see (7), for example]. In summary, we have the following result.

**THEOREM 1.2.1 Free Variable Theorem for Homogeneous Systems**

*If a homogeneous linear system has $n$ unknowns, and if the reduced row echelon form of its augmented matrix has $r$ nonzero rows, then the system has $n - r$ free variables.*

Theorem 1.2.1 has an important implication for homogeneous linear systems with more unknowns than equations. Specifically, if a homogeneous linear system has $m$ equations in $n$ unknowns, and if $m < n$, then it must also be true that $r < n$ (why?). This being the case, the theorem implies that there is at least one free variable, and this implies that the system has infinitely many solutions. Thus, we have the following result.

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

Note that Theorem 1.2.2 applies only to homogeneous systems—a *nonhomogeneous* system with more unknowns than equations need not be consistent. However, we will prove later that if a nonhomogeneous system with more unknowns then equations is consistent, then it has infinitely many solutions.

In retrospect, we could have anticipated that the homogeneous system in Example 6 would have infinitely many solutions since it has four equations in six unknowns.

*Gaussian Elimination and Back-Substitution*

For small linear systems that are solved by hand (such as most of those in this text), Gauss–Jordan elimination (reduction to reduced row echelon form) is a good procedure to use. However, for large linear systems that require a computer solution, it is generally more efficient to use Gaussian elimination (reduction to row echelon form) followed by a technique known as ***back-substitution*** to complete the process of solving the system. The next example illustrates this technique.



▶ **EXAMPLE 7 Example 5 Solved by Back-Substitution**

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

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

To solve the corresponding system of equations

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

we proceed as follows:

*Step 1.* Solve the equations for the leading variables.

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

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

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

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

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

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

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

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

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

This agrees with the solution obtained in Example 5.

▶ **EXAMPLE 8**

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

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

*Solution (a)* The last row corresponds to the equation

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

from which it is evident that the system is inconsistent.

*Solution (b)* The last row corresponds to the equation

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

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

*Solution (c)* The last row corresponds to the equation

$$x_4 = 0$$

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

$$x_3 + 6x_4 = 9$$

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

*Some Facts About Echelon Forms*

There are three facts about row echelon forms and reduced row echelon forms that are important to know but we will not prove:

1. Every matrix has a unique reduced row echelon form; that is, regardless of whether you use Gauss–Jordan elimination or some other sequence of elementary row operations, the same reduced row echelon form will result in the end.*
2. Row echelon forms are not unique; that is, different sequences of elementary row operations can result in different row echelon forms.
3. Although row echelon forms are not unique, the reduced row echelon form and all row echelon forms of a matrix $A$ have the same number of zero rows, and the leading 1's always occur in the same positions. Those are called the ***pivot positions*** of $A$. A column that contains a pivot position is called a ***pivot column*** of $A$.

*A proof of this result can be found in the article "The Reduced Row Echelon Form of a Matrix Is Unique: A Simple Proof," by Thomas Yuster, *Mathematics Magazine*, Vol. 57, No. 2, 1984, pp. 93–94.



▶ **EXAMPLE 9 Pivot Positions and Columns**

Earlier in this section (immediately after Definition 1) we found a row echelon form of

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

to be

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

The leading 1's occur in positions (row 1, column 1), (row 2, column 3), and (row 3, column 5). These are the pivot positions. The pivot columns are columns 1, 3, and 5. ◀

If $A$ is the augmented matrix for a linear system, then the pivot columns identify the leading variables. As an illustration, in Example 5 the pivot columns are 1, 3, and 6, and the leading variables are $x_1$, $x_3$, and $x_6$.

*Roundoff Error and Instability*

There is often a gap between mathematical theory and its practical implementation—Gauss–Jordan elimination and Gaussian elimination being good examples. The problem is that computers generally approximate numbers, thereby introducing ***roundoff*** errors, so unless precautions are taken, successive calculations may degrade an answer to a degree that makes it useless. Algorithms (procedures) in which this happens are called ***unstable***. There are various techniques for minimizing roundoff error and instability. For example, it can be shown that for large linear systems Gauss–Jordan elimination involves roughly 50% more operations than Gaussian elimination, so most computer algorithms are based on the latter method. Some of these matters will be considered in Chapter 9.

## Exercise Set 1.2

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

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

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

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

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

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

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

▶ In Exercises 3–4, suppose that the augmented matrix for a linear system has been reduced by row operations to the given row echelon form. Solve the system. ◀

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

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

(c) $\begin{bmatrix} 1 & 7 & -2 & 0 & -8 & -3 \\ 0 & 0 & 1 & 1 & 6 & 5 \\ 0 & 0 & 0 & 1 & 3 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & -3 & 7 & 1 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

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

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

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

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

► In Exercises 5–8, solve the linear system by Gaussian elimination. ◄

5. $\begin{aligned} x_1 + x_2 + 2x_3 &= 8 \\ -x_1 - 2x_2 + 3x_3 &= 1 \\ 3x_1 - 7x_2 + 4x_3 &= 10 \end{aligned}$

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

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

8. $\begin{aligned} -2b + 3c &= 1 \\ 3a + 6b - 3c &= -2 \\ 6a + 6b + 3c &= 5 \end{aligned}$

► In Exercises 9–12, solve the linear system by Gauss–Jordan elimination. ◄

9. Exercise 5

10. Exercise 6

11. Exercise 7

12. Exercise 8

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

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

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

► In Exercises 15–22, solve the given linear system by any method. ◄

15. $\begin{aligned} 2x_1 + x_2 + 3x_3 &= 0 \\ x_1 + 2x_2 \qquad &= 0 \\ x_2 + x_3 &= 0 \end{aligned}$

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

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

18. $\begin{aligned} v + 3w - 2x &= 0 \\ 2u + v - 4w + 3x &= 0 \\ 2u + 3v + 2w - x &= 0 \\ -4u - 3v + 5w - 4x &= 0 \end{aligned}$

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

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

21. $\begin{aligned} 2I_1 - I_2 + 3I_3 + 4I_4 &= 9 \\ I_1 \qquad - 2I_3 + 7I_4 &= 11 \\ 3I_1 - 3I_2 + I_3 + 5I_4 &= 8 \\ 2I_1 + I_2 + 4I_3 + 4I_4 &= 10 \end{aligned}$

22. $\begin{aligned} Z_3 + Z_4 + Z_5 &= 0 \\ -Z_1 - Z_2 + 2Z_3 - 3Z_4 + Z_5 &= 0 \\ Z_1 + Z_2 - 2Z_3 \qquad - Z_5 &= 0 \\ 2Z_1 + 2Z_2 - Z_3 \qquad + Z_5 &= 0 \end{aligned}$

► In each part of Exercises 23–24, the augmented matrix for a linear system is given in which the asterisk represents an unspecified real number. Determine whether the system is consistent, and if so whether the solution is unique. Answer "inconclusive" if there is not enough information to make a decision. ◄

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

(c) $\begin{bmatrix} 1 & * & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 1 & * & * & * \\ 0 & 0 & * & 0 \\ 0 & 0 & 1 & * \end{bmatrix}$

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

(c) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & * & * & * \end{bmatrix}$ (d) $\begin{bmatrix} 1 & * & * & * \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

► In Exercises 25–26, determine the values of $a$ for which the system has no solutions, exactly one solution, or infinitely many solutions. ◄

25. $\begin{aligned} x + 2y - \qquad 3z &= 4 \\ 3x - y + \qquad 5z &= 2 \\ 4x + y + (a^2 - 14)z &= a + 2 \end{aligned}$



26. $\begin{aligned} x + 2y + \qquad z &= 2 \\ 2x - 2y + \qquad 3z &= 1 \\ x + 2y - (a^2 - 3)z &= a \end{aligned}$

► In Exercises 27–28, what condition, if any, must $a$, $b$, and $c$ satisfy for the linear system to be consistent? ◄

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

28. $\begin{aligned} x + 3y + z &= a \\ -x - 2y + z &= b \\ 3x + 7y - z &= c \end{aligned}$

► In Exercises 29–30, solve the following systems, where $a$, $b$, and $c$ are constants. ◄

29. $\begin{aligned} 2x + y &= a \\ 3x + 6y &= b \end{aligned}$

30. $\begin{aligned} x_1 + x_2 + x_3 &= a \\ 2x_1 \qquad + 2x_3 &= b \\ 3x_2 + 3x_3 &= c \end{aligned}$

31. Find two different row echelon forms of

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

This exercise shows that a matrix can have multiple row echelon forms.

32. Reduce

$$\begin{bmatrix} 2 & 1 & 3 \\ 0 & -2 & -29 \\ 3 & 4 & 5 \end{bmatrix}$$

to reduced row echelon form without introducing fractions at any intermediate stage.

33. Show that the following nonlinear system has 18 solutions if $0 \le \alpha \le 2\pi$, $0 \le \beta \le 2\pi$, and $0 \le \gamma \le 2\pi$.

$$\begin{aligned} \sin\alpha + 2\cos\beta + 3\tan\gamma &= 0 \\ 2\sin\alpha + 5\cos\beta + 3\tan\gamma &= 0 \\ -\sin\alpha - 5\cos\beta + 5\tan\gamma &= 0 \end{aligned}$$

[*Hint:* Begin by making the substitutions $x = \sin\alpha$, $y = \cos\beta$, and $z = \tan\gamma$.]

34. Solve the following system of nonlinear equations for the unknown angles $\alpha$, $\beta$, and $\gamma$, where $0 \le \alpha \le 2\pi$, $0 \le \beta \le 2\pi$, and $0 \le \gamma < \pi$.

$$\begin{aligned} 2\sin\alpha - \cos\beta + 3\tan\gamma &= 3 \\ 4\sin\alpha + 2\cos\beta - 2\tan\gamma &= 2 \\ 6\sin\alpha - 3\cos\beta + \tan\gamma &= 9 \end{aligned}$$

35. Solve the following system of nonlinear equations for $x$, $y$, and $z$.

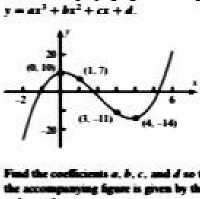
$$\begin{aligned} x^2 + y^2 + z^2 &= 6 \\ x^2 - y^2 + 2z^2 &= 2 \\ 2x^2 + y^2 - z^2 &= 3 \end{aligned}$$

[*Hint:* Begin by making the substitutions $X = x^2$, $Y = y^2$, $Z = z^2$.]

36. Solve the following system for $x$, $y$, and $z$.

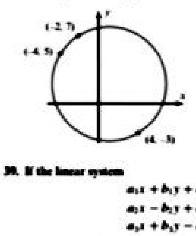
$$\begin{aligned} \frac{1}{x} + \frac{2}{y} - \frac{4}{z} &= 1 \\ \frac{2}{x} + \frac{3}{y} + \frac{8}{z} &= 0 \\ -\frac{1}{x} + \frac{9}{y} + \frac{10}{z} &= 5 \end{aligned}$$

37. Find the coefficients $a$, $b$, $c$, and $d$ so that the curve shown in the accompanying figure is the graph of the equation $y = ax^3 + bx^2 + cx + d$.




◄ Figure Ex-37

38. Find the coefficients $a$, $b$, $c$, and $d$ so that the circle shown in the accompanying figure is given by the equation $ax^2 + ay^2 + bx + cy + d = 0$.




◄ Figure Ex-38

39. If the linear system

$$\begin{aligned} a_1x + b_1y + c_1z &= 0 \\ a_2x - b_2y + c_2z &= 0 \\ a_3x + b_3y - c_3z &= 0 \end{aligned}$$

has only the trivial solution, what can be said about the solutions of the following system?

$$\begin{aligned} a_1x + b_1y + c_1z &= 3 \\ a_2x - b_2y + c_2z &= 7 \\ a_3x + b_3y - c_3z &= 11 \end{aligned}$$

40. (a) If $A$ is a matrix with three rows and five columns, then what is the maximum possible number of leading 1's in its reduced row echelon form?

(b) If $B$ is a matrix with three rows and six columns, then what is the maximum possible number of parameters in the general solution of the linear system with augmented matrix $B$?

(c) If $C$ is a matrix with five rows and three columns, then what is the minimum possible number of rows of zeros in any row echelon form of $C$?

**41.** Describe all possible reduced row echelon forms of

(a) $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ (b) $\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & p & q \end{bmatrix}$

**42.** Consider the system of equations

$$\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \\ ex + fy &= 0 \end{aligned}$$

Discuss the relative positions of the lines $ax + by = 0$, $cx + dy = 0$, and $ex + fy = 0$ when the system has only the trivial solution and when it has nontrivial solutions.

### Working with Proofs

**43.** (a) Prove that if $ad - bc \neq 0$, then the reduced row echelon form of

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

(b) Use the result in part (a) to prove that if $ad - bc \neq 0$, then the linear system

$$\begin{aligned} ax + by &= k \\ cx + dy &= l \end{aligned}$$

has exactly one solution.

### True-False Exercises

**TF.** In parts (a)–(i) determine whether the statement is true or false, and justify your answer.

(a) If a matrix is in reduced row echelon form, then it is also in row echelon form.

(b) If an elementary row operation is applied to a matrix that is in row echelon form, the resulting matrix will still be in row echelon form.

(c) Every matrix has a unique row echelon form.

(d) A homogeneous linear system in $n$ unknowns whose corresponding augmented matrix has a reduced row echelon form with $r$ leading 1's has $n - r$ free variables.

(e) All leading 1's in a matrix in row echelon form must occur in different columns.

(f) If every column of a matrix in row echelon form has a leading 1, then all entries that are not leading 1's are zero.

(g) If a homogeneous linear system of $n$ equations in $n$ unknowns has a corresponding augmented matrix with a reduced row echelon form containing $n$ leading 1's, then the linear system has only the trivial solution.

(h) If the reduced row echelon form of the augmented matrix for a linear system has a row of zeros, then the system must have infinitely many solutions.

(i) If a linear system has more unknowns than equations, then it must have infinitely many solutions.

### Working with Technology

**T1.** Find the reduced row echelon form of the augmented matrix for the linear system:

$$\begin{aligned} 6x_1 + x_2 \phantom{{}+3x_3} + 4x_4 &= -3 \\ -9x_1 + 2x_2 + 3x_3 - 8x_4 &= 1 \\ 7x_1 \phantom{{}+2x_2} - 4x_3 + 5x_4 &= 2 \end{aligned}$$

Use your result to determine whether the system is consistent and, if so, find its solution.

**T2.** Find values of the constants $A$, $B$, $C$, and $D$ that make the following equation an identity (i.e., true for all values of $x$).

$$\frac{3x^3 + 4x^2 - 6x}{(x^2 + 2x + 2)(x^2 - 1)} = \frac{Ax + B}{x^2 + 2x + 2} + \frac{C}{x - 1} + \frac{D}{x + 1}$$

[*Hint:* Obtain a common denominator on the right, and then equate corresponding coefficients of the various powers of $x$ in the two numerators. Students of calculus will recognize this as a problem in partial fractions.]

## 1.3 Matrices and Matrix Operations

Rectangular arrays of real numbers arise in contexts other than as augmented matrices for linear systems. In this section we will begin to study matrices as objects in their own right by defining operations of addition, subtraction, and multiplication on them.

*Matrix Notation and Terminology*

In Section 1.2 we used rectangular arrays of numbers, called *augmented matrices*, to abbreviate systems of linear equations. However, rectangular arrays of numbers occur in other contexts as well. For example, the following rectangular array with three rows and seven columns might describe the number of hours that a student spent studying three subjects during a certain week:



| | Mon. | Tues. | Wed. | Thurs. | Fri. | Sat. | Sun. |
|---|---|---|---|---|---|---|---|
| Math | 2 | 3 | 2 | 4 | 1 | 4 | 2 |
| History | 0 | 3 | 1 | 4 | 3 | 2 | 2 |
| Language | 4 | 1 | 3 | 1 | 0 | 0 | 2 |

If we suppress the headings, then we are left with the following rectangular array of numbers with three rows and seven columns, called a "matrix":

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

More generally, we make the following definition.

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

▶ **EXAMPLE 1 Examples of Matrices**

Some examples of matrices are

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

Matrix brackets are often omitted from $1 \times 1$ matrices, making it impossible to tell, for example, whether the symbol 4 denotes the number "four" or the matrix [4]. This rarely causes problems because it is usually possible to tell which is meant from the context.

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

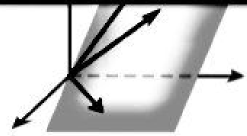
A matrix with only one row, such as the second in Example 1, is called a *row vector* (or a *row matrix*), and a matrix with only one column, such as the fourth in that example, is called a *column vector* (or a *column matrix*). The fifth matrix in that example is both a row vector and a column vector.

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

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

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

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



# General Vector Spaces

CHAPTER CONTENTS



INTRODUCTION Recall that we began our study of vectors by viewing them as directed line segments (arrows). We then extended this idea by introducing rectangular coordinate systems, which enabled us to view vectors as ordered pairs and ordered triples of real numbers. As we developed properties of these vectors we noticed patterns in various formulas that enabled us to extend the notion of a vector to an $n$-tuple of real numbers. Although $n$-tuples took us outside the realm of our "visual experience," it gave us a valuable tool for understanding and studying systems of linear equations. In this chapter we will extend the concept of a vector yet again by using the most important algebraic properties of vectors in $R^n$ as axioms. These axioms, if satisfied by a set of objects, will enable us to think of those objects as vectors.

## 4.1 Real Vector Spaces

In this section we will extend the concept of a vector by using the basic properties of vectors in $R^n$ as axioms, which if satisfied by a set of objects, guarantee that those objects behave like familiar vectors.

### *Vector Space Axioms*

The following definition consists of ten axioms, eight of which are properties of vectors in $R^n$ that were stated in Theorem 3.1.1. It is important to keep in mind that one does not *prove* axioms; rather, they are assumptions that serve as the starting point for proving theorems.





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

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

In this text scalars will be either real numbers or complex numbers. Vector spaces with real scalars will be called ***real vector spaces*** and those with complex scalars will be called ***complex vector spaces***. There is a more general notion of a vector space in which scalars can come from a mathematical structure known as a "field," but we will not be concerned with that level of generality. For now, we will focus exclusively on real vector spaces, which we will refer to simply as "vector spaces." We will consider complex vector spaces later.

Observe that the definition of a vector space does not specify the nature of the vectors or the operations. Any kind of object can be a vector, and the operations of addition and scalar multiplication need not have any relationship to those on $R^n$. The only requirement is that the ten vector space axioms be satisfied. In the examples that follow we will use four basic steps to show that a set with two operations is a vector space.

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

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

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

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

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

Hermann Günther

**Historical Note** The notion of an "abstract vector space" evolved over many years and had many contributors. The idea crystallized with the work of the German mathematician H. G. Grassmann, who published a paper in 1862 in which he considered abstract systems of unspecified elements on which he defined formal operations of addition and scalar multiplication. Grassmann's work was controversial, and others, including Augustin Cauchy (p. 121), laid reasonable claim to the idea.

*[Image: © Sueddeutsche Zeitung Photo/The Image Works]*

Our first example is the simplest of all vector spaces in that it contains only one object. Since Axiom 4 requires that every vector space contain a zero vector, the object will have to be that vector.

▶ EXAMPLE 1 **The Zero Vector Space**

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

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

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

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

▶ EXAMPLE 2 **$R^n$ Is a Vector Space**

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

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

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

Our next example is a generalization of $R^n$ in which we allow vectors to have infinitely many components.



▶ EXAMPLE 3 **The Vector Space of Infinite Sequences of Real Numbers**

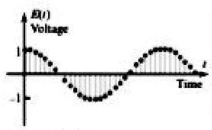
Let $V$ consist of objects of the form

$$\mathbf{u}=(u_1, u_2, \ldots, u_n, \ldots)$$

in which $u_1, u_2, \ldots, u_n, \ldots$ is an infinite sequence of real numbers. We define two infinite sequences to be ***equal*** if their corresponding components are equal, and we define addition and scalar multiplication componentwise by

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

In the exercises we ask you to confirm that $V$ with these operations is a vector space. We will denote this vector space by the symbol $R^\infty$. ◀




▲ Figure 4.1.1

Vector spaces of the type in Example 3 arise when a transmitted signal of indefinite duration is digitized by sampling its values at discrete time intervals (Figure 4.1.1).

In the next example our vectors will be matrices. This may be a little confusing at first because matrices are composed of rows and columns, which are themselves vectors (row vectors and column vectors). However, from the vector space viewpoint we are not



concerned with the individual rows and columns but rather with the properties of the matrix operations as they relate to the matrix as a whole.

▶ EXAMPLE 4 **The Vector Space of 2 × 2 Matrices**

Note that Equation (1) involves *three* different addition operations: the addition operation on vectors, the addition operation on matrices, and the addition operation on real numbers.

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

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

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

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

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

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

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

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

With this definition,

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

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

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

With this definition,

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

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

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

**▶ EXAMPLE 6 The Vector Space of Real-Valued Functions**

Let $V$ be the set of real-valued functions that are defined at each $x$ in the interval $(-\infty, \infty)$. If $\mathbf{f} = f(x)$ and $\mathbf{g} = g(x)$ are two functions in $V$ and if $k$ is any scalar, then define the operations of addition and scalar multiplication by

$$(\mathbf{f} + \mathbf{g})(x) = f(x) + g(x) \tag{2}$$

$$(k\mathbf{f})(x) = kf(x) \tag{3}$$

One way to think about these operations is to view the numbers $f(x)$ and $g(x)$ as "components" of $\mathbf{f}$ and $\mathbf{g}$ at the point $x$, in which case Equations (2) and (3) state that two functions are added by adding corresponding components, and a function is multiplied by a scalar by multiplying each component by that scalar—exactly as in $R^n$ and $R^\infty$. This idea is illustrated in parts ($a$) and ($b$) of Figure 4.1.2. The set $V$ with these operations is denoted by the symbol $F(-\infty, \infty)$. We can prove that this is a vector space as follows:

**Axioms 1 and 6:** These closure axioms require that if we add two functions that are defined at each $x$ in the interval $(-\infty, \infty)$, then sums and scalar multiples of those functions must also be defined at each $x$ in the interval $(-\infty, \infty)$. This follows from Formulas (2) and (3).

**Axiom 4:** This axiom requires that there exists a function $\mathbf{0}$ in $F(-\infty, \infty)$, which when added to any other function $\mathbf{f}$ in $F(-\infty, \infty)$ produces $\mathbf{f}$ back again as the result. The function whose value at every point $x$ in the interval $(-\infty, \infty)$ is zero has this property. Geometrically, the graph of the function $\mathbf{0}$ is the line that coincides with the $x$-axis.

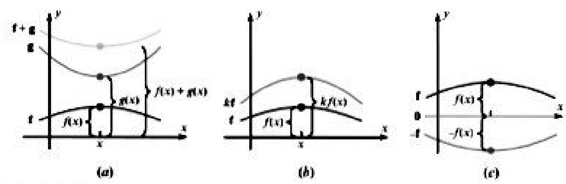
In Example 6 the functions were defined on the entire interval $(-\infty, \infty)$. However, the arguments used in that example apply as well on all subintervals of $(-\infty, \infty)$, such as a closed interval $[a, b]$ or an open interval $(a, b)$. We will denote the vector spaces of functions on these intervals by $F[a, b]$ and $F(a, b)$, respectively.

**Axiom 5:** This axiom requires that for each function $\mathbf{f}$ in $F(-\infty, \infty)$ there exists a function $-\mathbf{f}$ in $F(-\infty, \infty)$, which when added to $\mathbf{f}$ produces the function $\mathbf{0}$. The function defined by $-\mathbf{f}(x) = -f(x)$ has this property. The graph of $-\mathbf{f}$ can be obtained by reflecting the graph of $\mathbf{f}$ about the $x$-axis (Figure 4.1.2$c$).

**Axioms 2, 3, 7, 8, 9, 10:** The validity of each of these axioms follows from properties of real numbers. For example, if $\mathbf{f}$ and $\mathbf{g}$ are functions in $F(-\infty, \infty)$, then Axiom 2 requires that $\mathbf{f} + \mathbf{g} = \mathbf{g} + \mathbf{f}$. This follows from the computation

$$(\mathbf{f} + \mathbf{g})(x) = \mathbf{f}(x) + \mathbf{g}(x) = \mathbf{g}(x) + \mathbf{f}(x) = (\mathbf{g} + \mathbf{f})(x)$$

in which the first and last equalities follow from (2), and the middle equality is a property of real numbers. We will leave the proofs of the remaining parts as exercises. ◀




▲ Figure 4.1.2

It is important to recognize that you cannot impose any two operations on any set $V$ and expect the vector space axioms to hold. For example, if $V$ is the set of $n$-tuples with *positive* components, and if the standard operations from $R^n$ are used, then $V$ is not closed under scalar multiplication, because if $\mathbf{u}$ is a nonzero $n$-tuple in $V$, then $(-1)\mathbf{u}$ has



at least one negative component and hence is not in $V$. The following is a less obvious example in which only one of the ten vector space axioms fails to hold.

**▶ EXAMPLE 7 A Set That Is Not a Vector Space**

Let $V = R^2$ and define addition and scalar multiplication operations as follows: If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$, then define

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2)$$

and if $k$ is any real number, then define

$$k\mathbf{u} = (ku_1, 0)$$

For example, if $\mathbf{u} = (2, 4)$, $\mathbf{v} = (-3, 5)$, and $k = 7$, then

$$\mathbf{u} + \mathbf{v} = (2 + (-3), 4 + 5) = (-1, 9)$$
$$k\mathbf{u} = 7\mathbf{u} = (7 \cdot 2, 0) = (14, 0)$$

The addition operation is the standard one from $R^2$, but the scalar multiplication is not. In the exercises we will ask you to show that the first nine vector space axioms are satisfied. However, Axiom 10 fails to hold for certain vectors. For example, if $\mathbf{u} = (u_1, u_2)$ is such that $u_2 \neq 0$, then

$$1\mathbf{u} = 1(u_1, u_2) = (1 \cdot u_1, 0) = (u_1, 0) \neq \mathbf{u}$$

Thus, $V$ is not a vector space with the stated operations. ◀

Our final example will be an unusual vector space that we have included to illustrate how varied vector spaces can be. Since the vectors in this space will be real numbers, it will be important for you to keep track of which operations are intended as vector operations and which ones as ordinary operations on real numbers.

**▶ EXAMPLE 8 An Unusual Vector Space**

Let $V$ be the set of positive real numbers, let $\mathbf{u} = u$ and $\mathbf{v} = v$ be any vectors (i.e., positive real numbers) in $V$, and let $k$ be any scalar. Define the operations on $V$ to be

$$u + v = uv \quad \text{[Vector addition is numerical multiplication.]}$$
$$ku = u^k \quad \text{[Scalar multiplication is numerical exponentiation.]}$$

Thus, for example, $1 + 1 = 1$ and $(2)(1) = 1^2 = 1$—strange indeed, but nevertheless the set $V$ with these operations satisfies the ten vector space axioms and hence is a vector space. We will confirm Axioms 4, 5, and 7, and leave the others as exercises.

- Axiom 4—The zero vector in this space is the number 1 (i.e., $\mathbf{0} = 1$) since

$$u + 1 = u \cdot 1 = u$$

- Axiom 5—The negative of a vector $u$ is its reciprocal (i.e., $-u = 1/u$) since

$$u + \frac{1}{u} = u\left(\frac{1}{u}\right) = 1\ (= \mathbf{0})$$

- Axiom 7—$k(u + v) = (uv)^k = u^k v^k = (ku) + (kv)$. ◀

*Some Properties of Vectors*

The following is our first theorem about vector spaces. The proof is very formal with each step being justified by a vector space axiom or a known property of real numbers. There will not be many rigidly formal proofs of this type in the text, but we have included this one to reinforce the idea that the familiar properties of vectors can all be derived from the vector space axioms.

**THEOREM 4.1.1** *Let $V$ be a vector space, $\mathbf{u}$ a vector in $V$, and $k$ a scalar; then:*

(*a*) $0\mathbf{u} = \mathbf{0}$

(*b*) $k\mathbf{0} = \mathbf{0}$

(*c*) $(-1)\mathbf{u} = -\mathbf{u}$

(*d*) *If* $k\mathbf{u} = \mathbf{0}$, *then* $k = 0$ *or* $\mathbf{u} = \mathbf{0}$.

We will prove parts (*a*) and (*c*) and leave proofs of the remaining parts as exercises.

***Proof (a)*** We can write

$$\begin{aligned} 0\mathbf{u} + 0\mathbf{u} &= (0 + 0)\mathbf{u} && \text{[Axiom 8]} \\ &= 0\mathbf{u} && \text{[Property of the number 0]} \end{aligned}$$

By Axiom 5 the vector $0\mathbf{u}$ has a negative, $-0\mathbf{u}$. Adding this negative to both sides above yields

$$[0\mathbf{u} + 0\mathbf{u}] + (-0\mathbf{u}) = 0\mathbf{u} + (-0\mathbf{u})$$

or

$$\begin{aligned} 0\mathbf{u} + [0\mathbf{u} + (-0\mathbf{u})] &= 0\mathbf{u} + (-0\mathbf{u}) && \text{[Axiom 3]} \\ 0\mathbf{u} + \mathbf{0} &= \mathbf{0} && \text{[Axiom 5]} \\ 0\mathbf{u} &= \mathbf{0} && \text{[Axiom 4]} \end{aligned}$$

***Proof (c)*** To prove that $(-1)\mathbf{u} = -\mathbf{u}$, we must show that $\mathbf{u} + (-1)\mathbf{u} = \mathbf{0}$. The proof is as follows:

$$\begin{aligned} \mathbf{u} + (-1)\mathbf{u} &= 1\mathbf{u} + (-1)\mathbf{u} && \text{[Axiom 10]} \\ &= (1 + (-1))\mathbf{u} && \text{[Axiom 8]} \\ &= 0\mathbf{u} && \text{[Property of numbers]} \\ &= \mathbf{0} && \text{[Part (a) of this theorem]} \end{aligned}$$



*A Closing Observation*

This section of the text is important to the overall plan of linear algebra in that it establishes a common thread among such diverse mathematical objects as geometric vectors, vectors in $R^n$, infinite sequences, matrices, and real-valued functions, to name a few. As a result, whenever we discover a new theorem about general vector spaces, we will at the same time be discovering a theorem about geometric vectors, vectors in $R^n$, sequences, matrices, real-valued functions, and about any new kinds of vectors that we might discover.

To illustrate this idea, consider what the rather innocent-looking result in part (*a*) of Theorem 4.1.1 says about the vector space in Example 8. Keeping in mind that the vectors in that space are positive real numbers, that scalar multiplication means numerical exponentiation, and that the zero vector is the number 1, the equation

$$0\mathbf{u} = \mathbf{0}$$

is really a statement of the familiar fact that if $u$ is a positive real number, then

$$u^0 = 1$$



## Exercise Set 4.1

**1.** Let $V$ be the set of all ordered pairs of real numbers, and consider the following addition and scalar multiplication operations on $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2), \quad k\mathbf{u} = (0, ku_2)$$

(a) Compute $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for $\mathbf{u} = (-1, 2)$, $\mathbf{v} = (3, 4)$, and $k = 3$.

(b) In words, explain why $V$ is closed under addition and scalar multiplication.

(c) Since addition on $V$ is the standard addition operation on $R^2$, certain vector space axioms hold for $V$ because they are known to hold for $R^2$. Which axioms are they?

(d) Show that Axioms 7, 8, and 9 hold.

(e) Show that Axiom 10 fails and hence that $V$ is not a vector space under the given operations.

**2.** Let $V$ be the set of all ordered pairs of real numbers, and consider the following addition and scalar multiplication operations on $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$:

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1 + 1, u_2 + v_2 + 1), \quad k\mathbf{u} = (ku_1, ku_2)$$

(a) Compute $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for $\mathbf{u} = (0, 4)$, $\mathbf{v} = (1, -3)$, and $k = 2$.

(b) Show that $(0, 0) \neq \mathbf{0}$.

(c) Show that $(-1, -1) = \mathbf{0}$.

(d) Show that Axiom 5 holds by producing an ordered pair $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ for $\mathbf{u} = (u_1, u_2)$.

(e) Find two vector space axioms that fail to hold.

► In Exercises 3–12, determine whether each set equipped with the given operations is a vector space. For those that are not vector spaces identify the vector space axioms that fail. ◄

**3.** The set of all real numbers with the standard operations of addition and multiplication.

**4.** The set of all pairs of real numbers of the form $(x, 0)$ with the standard operations on $R^2$.

**5.** The set of all pairs of real numbers of the form $(x, y)$, where $x \geq 0$, with the standard operations on $R^2$.

**6.** The set of all $n$-tuples of real numbers that have the form $(x, x, \dots, x)$ with the standard operations on $R^n$.

**7.** The set of all triples of real numbers with the standard vector addition but with scalar multiplication defined by

**9.** The set of all $2 \times 2$ matrices of the form

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

with the standard matrix addition and scalar multiplication.

**10.** The set of all real-valued functions $f$ defined everywhere on the real line and such that $f(1) = 0$ with the operations used in Example 6.

**11.** The set of all pairs of real numbers of the form $(1, x)$ with the operations

$$(1, y) + (1, y') = (1, y + y') \quad \text{and} \quad k(1, y) = (1, ky)$$

**12.** The set of polynomials of the form $a_0 + a_1x$ with the operations

$$(a_0 + a_1x) + (b_0 + b_1x) = (a_0 + b_0) + (a_1 + b_1)x$$

and

$$k(a_0 + a_1x) = (ka_0) + (ka_1)x$$

**13.** Verify Axioms 3, 7, 8, and 9 for the vector space given in Example 4.

**14.** Verify Axioms 1, 2, 3, 7, 8, 9, and 10 for the vector space given in Example 6.

**15.** With the addition and scalar multiplication operations defined in Example 7, show that $V = R^2$ satisfies Axioms 1–9.

**16.** Verify Axioms 1, 2, 3, 6, 8, 9, and 10 for the vector space given in Example 8.

**17.** Show that the set of all points in $R^2$ lying on a line is a vector space with respect to the standard operations of vector addition and scalar multiplication if and only if the line passes through the origin.

**18.** Show that the set of all points in $R^3$ lying in a plane is a vector space with respect to the standard operations of vector addition and scalar multiplication if and only if the plane passes through the origin.

► In Exercises 19–20, let $V$ be the vector space of positive real numbers with the vector space operations given in Example 8. Let $\mathbf{u} = u$ be any vector in $V$, and rewrite the vector statement as a statement about real numbers. ◄

**19.** $-\mathbf{u} = (-1)\mathbf{u}$

**20.** $k\mathbf{u} = \mathbf{0}$ if and only if $k = 0$ or $\mathbf{u} = \mathbf{0}$.

### Working with Proofs

## 4.2 Subspaces

It is often the case that some vector space of interest is contained within a larger vector space whose properties are known. In this section we will show how to recognize when this is the case, we will explain how the properties of the larger vector space can be used to obtain properties of the smaller vector space, and we will give a variety of important examples.

We begin with some terminology.

> **DEFINITION 1** A subset $W$ of a vector space $V$ is called a ***subspace*** of $V$ if $W$ is itself a vector space under the addition and scalar multiplication defined on $V$.

In general, to show that a nonempty set $W$ with two operations is a vector space one must verify the ten vector space axioms. However, if $W$ is a subspace of a known vector space $V$, then certain axioms need not be verified because they are "inherited" from $V$. For example, it is *not* necessary to verify that $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$ holds in $W$ because it holds for all vectors in $V$ including those in $W$. On the other hand, it *is* necessary to verify



that $W$ is closed under addition and scalar multiplication since it is possible that adding two vectors in $W$ or multiplying a vector in $W$ by a scalar produces a vector in $V$ that is outside of $W$ (Figure 4.2.1). Those axioms that are *not* inherited by $W$ are

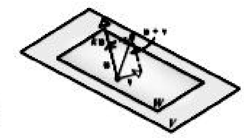
Axiom 1—Closure of $W$ under addition

Axiom 4—Existence of a zero vector in $W$

Axiom 5—Existence of a negative in $W$ for every vector in $W$

Axiom 6—Closure of $W$ under scalar multiplication

so these must be verified to prove that it is a subspace of $V$. However, the next theorem shows that if Axiom 1 and Axiom 6 hold in $W$, then Axioms 4 and 5 hold in $W$ as a consequence and hence need not be verified.




▶ Figure 4.2.1 The vectors $\mathbf{u}$ and $\mathbf{v}$ are in $W$, but the vectors $\mathbf{u}+\mathbf{v}$ and $k\mathbf{u}$ are not.

> **THEOREM 4.2.1** *If $W$ is a set of one or more vectors in a vector space $V$, then $W$ is a subspace of $V$ if and only if the following conditions are satisfied.*
> 
> (*a*) *If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $W$, then $\mathbf{u}+\mathbf{v}$ is in $W$.*
> 
> (*b*) *If $k$ is a scalar and $\mathbf{u}$ is a vector in $W$, then $k\mathbf{u}$ is in $W$.*

**Proof** If $W$ is a subspace of $V$, then all the vector space axioms hold in $W$, including Axioms 1 and 6, which are precisely conditions (*a*) and (*b*).

Conversely, assume that conditions (*a*) and (*b*) hold. Since these are Axioms 1 and 6, and since Axioms 2, 3, 7, 8, 9, and 10 are inherited from $V$, we only need to show that Axioms 4 and 5 hold in $W$. For this purpose, let $\mathbf{u}$ be any vector in $W$. It follows from condition (*b*) that $k\mathbf{u}$ is a vector in $W$ for every scalar $k$. In particular, $0\mathbf{u}=\mathbf{0}$ and $(-1)\mathbf{u}=-\mathbf{u}$ are in $W$, which shows that Axioms 4 and 5 hold in $W$. ◀

**Theorem 4.2.1 states that $W$ is a subspace of $V$ if and only if it is closed under addition and scalar multiplication.**

### ▶ EXAMPLE 1 The Zero Subspace

If $V$ is any vector space, and if $W=\{\mathbf{0}\}$ is the subset of $V$ that consists of the zero vector only, then $W$ is closed under addition and scalar multiplication since

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

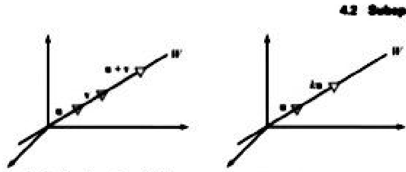
for any scalar $k$. We call $W$ the ***zero subspace*** of $V$.

**Note that every vector space has at least two subspaces, itself and its zero subspace.**

### ▶ EXAMPLE 2 Lines Through the Origin Are Subspaces of $R^2$ and of $R^3$

If $W$ is a line through the origin of either $R^2$ or $R^3$, then adding two vectors on the line or multiplying a vector on the line by a scalar produces another vector on the line, so $W$ is closed under addition and scalar multiplication (see Figure 4.2.2 for an illustration in $R^3$).



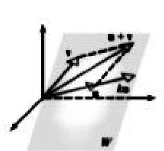



(*a*) $W$ is closed under addition. (*b*) $W$ is closed under scalar multiplication.

▶ Figure 4.2.2

### ▶ EXAMPLE 3 Planes Through the Origin Are Subspaces of $R^3$

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in a plane $W$ through the origin of $R^3$, then it is evident geometrically that $\mathbf{u}+\mathbf{v}$ and $k\mathbf{u}$ also lie in the same plane $W$ for any scalar $k$ (Figure 4.2.3). Thus $W$ is closed under addition and scalar multiplication. ◀




▲ Figure 4.2.3 The vectors $\mathbf{u}+\mathbf{v}$ and $k\mathbf{u}$ both lie in the same plane as $\mathbf{u}$ and $\mathbf{v}$.

Table 1 below gives a list of subspaces of $R^2$ and of $R^3$ that we have encountered thus far. We will see later that these are the only subspaces of $R^2$ and of $R^3$.

Table 1

| Subspaces of $R^2$ | Subspaces of $R^3$ |
|---|---|

CALCULUS REQUIRED

**▶ EXAMPLE 7 The Subspace $C(-\infty, \infty)$**

There is a theorem in calculus which states that a sum of continuous functions is continuous and that a constant times a continuous function is continuous. Rephrased in vector language, the set of continuous functions on $(-\infty, \infty)$ is a subspace of $F(-\infty, \infty)$. We will denote this subspace by $C(-\infty, \infty)$.

CALCULUS REQUIRED

**▶ EXAMPLE 8 Functions with Continuous Derivatives**

A function with a continuous derivative is said to be ***continuously differentiable***. There is a theorem in calculus which states that the sum of two continuously differentiable functions is continuously differentiable and that a constant times a continuously differentiable function is continuously differentiable. Thus, the functions that are continuously differentiable on $(-\infty, \infty)$ form a subspace of $F(-\infty, \infty)$. We will denote this subspace by $C^1(-\infty, \infty)$, where the superscript emphasizes that the *first* derivatives are continuous. To take this a step further, the set of functions with $m$ continuous derivatives on $(-\infty, \infty)$ is a subspace of $F(-\infty, \infty)$ as is the set of functions with derivatives of all orders on $(-\infty, \infty)$. We will denote these subspaces by $C^m(-\infty, \infty)$ and $C^\infty(-\infty, \infty)$, respectively.

**▶ EXAMPLE 9 The Subspace of All Polynomials**

Recall that a ***polynomial*** is a function that can be expressed in the form

$$p(x) = a_0 + a_1x + \cdots + a_nx^n \tag{1}$$

where $a_0, a_1, \ldots, a_n$ are constants. It is evident that the sum of two polynomials is a polynomial and that a constant times a polynomial is a polynomial. Thus, the set $W$ of all polynomials is closed under addition and scalar multiplication and hence is a subspace of $F(-\infty, \infty)$. We will denote this space by $P_\infty$.

In this text we regard all constants to be polynomials of degree zero. Be aware, however, that some authors do not assign a degree to the constant 0.

**▶ EXAMPLE 10 The Subspace of Polynomials of Degree ≤ $n$**

Recall that the ***degree*** of a polynomial is the highest power of the variable that occurs with a nonzero coefficient. Thus, for example, if $a_n \neq 0$ in Formula (1), then that polynomial has degree $n$. It is *not* true that the set $W$ of polynomials with p[illegible] subspace of $F(-\infty, \infty)$ because that set is not closed under additi[illegible] polynomials



$$1 + 2x + 3x^2 \quad \text{and} \quad 5 + 7x - 3x^2$$

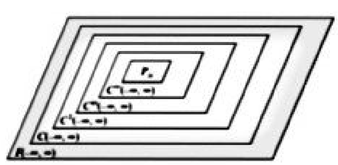
both have degree 2, but their sum has degree 1. What *is* true, however, is that for each nonnegative integer $n$ the polynomials of degree $n$ *or less* form a subspace of $F(-\infty, \infty)$. We will denote this space by $P_n$. ◀

*The Hierarchy of Function Spaces*

It is proved in calculus that polynomials are continuous functions and have continuous derivatives of all orders on $(-\infty, \infty)$. Thus, it follows that $P_\infty$ is not only a subspace of $F(-\infty, \infty)$, as previously observed, but is also a subspace of $C^\infty(-\infty, \infty)$. We leave it for you to convince yourself that the vector spaces discussed in Examples 7 to 10 are "nested" one inside the other as illustrated in Figure 4.2.5.

**Remark** In our previous examples we considered functions that were defined at all points of the interval $(-\infty, \infty)$. Sometimes we will want to consider functions that are only defined on some subinterval of $(-\infty, \infty)$, say the closed interval $[a, b]$ or the open interval $(a, b)$. In such cases we will make an appropriate notation change. For example, $C[a, b]$ is the space of continuous functions on $[a, b]$ and $C(a, b)$ is the space of continuous functions on $(a, b)$.





▶ Figure 4.2.5

*Building Subspaces*

The following theorem provides a useful way of creating a new subspace from known subspaces.

**THEOREM 4.2.2** *If $W_1, W_2, \ldots, W_r$ are subspaces of a vector space $V$, then the intersection of these subspaces is also a subspace of $V$.*

*Proof* Let $W$ be the intersection of the subspaces $W_1, W_2, \ldots, W_r$. This set is not empty because each of these subspaces contains the zero vector of $V$, and hence so does their intersection. Thus, it remains to show that $W$ is closed under addition and scalar multiplication.

To prove closure under addition, let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $W$. Since $W$ is the intersection of $W_1, W_2, \ldots, W_r$, it follows that $\mathbf{u}$ and $\mathbf{v}$ also lie in each of these subspaces. Moreover, since these subspaces are closed under addition and scalar multiplication, they also all contain the vectors $\mathbf{u} + \mathbf{v}$ and $k\mathbf{u}$ for every scalar $k$, and hence so does their intersection $W$. This proves that $W$ is closed under addition and scalar multiplication. ◀

Note that the first step in proving Theorem 4.2.2 was to establish that $W$ contained at least one vector. This is important, for otherwise the subsequent argument might be logically correct but meaningless.

Sometimes we will want to find the "smallest" subspace of a vector space $V$ that contains all of the vectors in some set of interest. The following definition, which generalizes Definition 4 of Section 3.1, will help us to do that.

**DEFINITION 2** If $\mathbf{w}$ is a vector in a vector space $V$, then $\mathbf{w}$ is said to be a ***linear combination*** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in $V$ if $\mathbf{w}$ can be expressed in the form

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r \tag{2}$$

where $k_1, k_2, \ldots, k_r$ are scalars. These scalars are called the ***coefficients*** of the linear combination.

If $k = 1$, then Equation (2) has the form $\mathbf{w} = k_1\mathbf{v}_1$, in which case the linear combination is just a scalar multiple of $\mathbf{v}_1$.

**THEOREM 4.2.3** *If $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is a nonempty set of vectors in a vector space $V$, then:*

*(a) The set $W$ of all possible linear combinations of the vectors in $S$ is a subspace of $V$.*

*(b) The set $W$ in part (a) is the "smallest" subspace of $V$ that contains all of the vectors in $S$ in the sense that any other subspace that contains those vectors contains $W$.*

*Proof (a)* Let $W$ be the set of all possible linear combinations of the vectors in $S$. We must show that $W$ is closed under addition and scalar multiplication. To prove closure under addition, let

$$\mathbf{u} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_r\mathbf{w}_r \quad \text{and} \quad \mathbf{v} = k_1\mathbf{w}_1 + k_2\mathbf{w}_2 + \cdots + k_r\mathbf{w}_r$$

be two vectors in $W$. It follows that their sum can be written as

$$\mathbf{u} + \mathbf{v} = (c_1 + k_1)\mathbf{w}_1 + (c_2 + k_2)\mathbf{w}_2 + \cdots + (c_r + k_r)\mathbf{w}_r$$

under addition and scalar multiplication.

## 4.3 Linear Independence

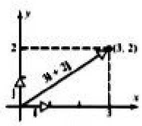
In this section we will consider the question of whether the vectors in a given set are interrelated in the sense that one or more of them can be expressed as a linear combination of the others. This is important to know in applications because the existence of such relationships often signals that some kind of complication is likely to occur.

*Linear Independence and Dependence*

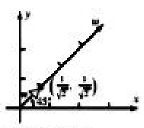
In a rectangular $xy$-coordinate system every vector in the plane can be expressed in exactly one way as a linear combination of the standard unit vectors. For example, the only way to express the vector $(3, 2)$ as a linear combination of $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ is

$$(3, 2) = 3(1, 0) + 2(0, 1) = 3\mathbf{i} + 2\mathbf{j} \tag{1}$$





▲ Figure 4.3.1



▲ Figure 4.3.2

(Figure 4.3.1). Suppose, however, that we were to introduce a third coordinate axis that makes an angle of 45° with the $x$-axis. Call it the $w$-axis. As illustrated in Figure 4.3.2, the unit vector along the $w$-axis is

$$\mathbf{w} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

Whereas Formula (1) shows the only way to express the vector $(3, 2)$ as a linear combination of $\mathbf{i}$ and $\mathbf{j}$, there are infinitely many ways to express this vector as a linear combination of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{w}$. Three possibilities are

$$(3, 2) = 3(1, 0) + 2(0, 1) + 0\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 3\mathbf{i} + 2\mathbf{j} + 0\mathbf{w}$$

$$(3, 2) = 2(1, 0) + (0, 1) + \sqrt{2}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 2\mathbf{i} + \mathbf{j} + \sqrt{2}\mathbf{w}$$

$$(3, 2) = 4(1, 0) + 3(0, 1) - \sqrt{2}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = 4\mathbf{i} + 3\mathbf{j} - \sqrt{2}\mathbf{w}$$

In short, by introducing a superfluous axis we created the complication of having multiple ways of assigning coordinates to points in the plane. What makes the vector $\mathbf{w}$ superfluous is the fact that it can be expressed as a linear combination of the vectors $\mathbf{i}$ and $\mathbf{j}$, namely,

$$\mathbf{w} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}$$

This leads to the following definition.

> **DEFINITION 1** If $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ is a set of two or more vectors in a vector space $V$, then $S$ is said to be a ***linearly independent set*** if no vector in $S$ can be expressed as a linear combination of the others. A set that is not linearly independent is said to be ***linearly dependent***.

**In the case where the set $S$ in Definition 1 has only one vector, we will agree that $S$ is linearly independent if and only if that vector is nonzero.**

In general, the most efficient way to determine whether a set is linearly independent or not is to use the following theorem whose proof is given at the end of this section.

> **THEOREM 4.3.1** *A nonempty set $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r\}$ in a vector space $V$ is linearly independent if and only if the only coefficients satisfying the vector equation*
>
> $$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}$$
>
> *are $k_1 = 0, k_2 = 0, \dots, k_r = 0$.*

▶ **EXAMPLE 1 Linear Independence of the Standard Unit Vectors in $R^n$**

The most basic linearly independent set in $R^n$ is the set of standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0), \dots, \quad \mathbf{e}_n = (0, 0, 0, \dots, 1)$$

To illustrate this in $R^3$, consider the standard unit vectors

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$



To prove linear independence we must show that the only coefficients satisfying the vector equation

$$k_1\mathbf{i} + k_2\mathbf{j} + k_3\mathbf{k} = \mathbf{0}$$

are $k_1 = 0, k_2 = 0, k_3 = 0$. But this becomes evident by writing this equation in its component form

$$(k_1, k_2, k_3) = (0, 0, 0)$$

You should have no trouble adapting this argument to establish the linear independence of the standard unit vectors in $R^n$.

▶ **EXAMPLE 2 Linear Independence in $R^3$**

Determine whether the vectors

$$\mathbf{v}_1 = (1, -2, 3), \quad \mathbf{v}_2 = (5, 6, -1), \quad \mathbf{v}_3 = (3, 2, 1) \tag{2}$$

are linearly independent or linearly dependent in $R^3$.

*Solution* The linear independence or dependence of these vectors is determined by whether the vector equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + k_3\mathbf{v}_3 = \mathbf{0} \tag{3}$$

can be satisfied with coefficients that are not all zero. To see whether this is so, let us rewrite (3) in the component form

$$k_1(1, -2, 3) + k_2(5, 6, -1) + k_3(3, 2, 1) = (0, 0, 0)$$

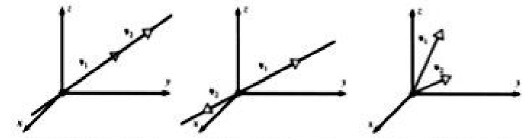
Equating corresponding components on the two sides yields the homogeneous linear system

$$\begin{aligned} k_1 + 5k_2 + 3k_3 &= 0 \\ -2k_1 + 6k_2 + 2k_3 &= 0 \\ 3k_1 - k_2 + k_3 &= 0 \end{aligned} \tag{4}$$

*A Geometric Interpretation of Linear Independence*

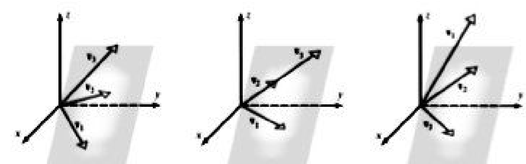
Linear independence has the following useful geometric interpretations in $R^2$ and $R^3$:

- Two vectors in $R^2$ or $R^3$ are linearly independent if and only if they do not lie on the same line when they have their initial points at the origin. Otherwise one would be a scalar multiple of the other (Figure 4.3.3).



► Figure 4.3.3 (*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent

- Three vectors in $R^3$ are linearly independent if and only if they do not lie in the same plane when they have their initial points at the origin. Otherwise at least one would be a linear combination of the other two (Figure 4.3.4).



► Figure 4.3.4 (*a*) Linearly dependent (*b*) Linearly dependent (*c*) Linearly independent

At the beginning of this section we observed that a third coordinate axis in $R^2$ is superfluous by showing that a unit vector along such an axis would have to be expressible as a linear combination of unit vectors along the positive $x$- and $y$-axis. That result is a consequence of the next theorem, which shows that there can be at most $n$ vectors in any linearly independent set $R^n$.

**THEOREM 4.3.3** *Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ be a set of vectors in $R^n$. If $r > n$, then $S$ is linearly dependent.*

*Proof* Suppose that

$$\begin{aligned} \mathbf{v}_1 &= (v_{11}, v_{12}, \ldots, v_{1n}) \\ \mathbf{v}_2 &= (v_{21}, v_{22}, \ldots, v_{2n}) \\ &\ \vdots \\ \mathbf{v}_r &= (v_{r1}, v_{r2}, \ldots, v_{rn}) \end{aligned}$$

and consider the equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0}$$



It follows from Theorem 4.3.3 that a set in $R^2$ with more than two vectors is linearly dependent and a set in $R^3$ with more than three vectors is linearly dependent.

If we express both sides of this equation in terms of components and then equate the corresponding components, we obtain the system

$$\begin{aligned} v_{11}k_1 + v_{21}k_2 + \cdots + v_{r1}k_r &= 0 \\ v_{12}k_1 + v_{22}k_2 + \cdots + v_{r2}k_r &= 0 \\ \vdots \qquad\qquad & \ \ \vdots \\ v_{1n}k_1 + v_{2n}k_2 + \cdots + v_{rn}k_r &= 0 \end{aligned}$$

This is a homogeneous system of $n$ equations in the $r$ unknowns $k_1, \ldots, k_r$. Since $r > n$, it follows from Theorem 1.2.2 that the system has nontrivial solutions. Therefore, $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a linearly dependent set. ◄

**CALCULUS REQUIRED**

*Linear Independence of Functions*

Sometimes linear dependence of functions can be deduced from known identities. For example, the functions

$$\mathbf{f}_1 = \sin^2 x, \quad \mathbf{f}_2 = \cos^2 x, \quad \text{and} \quad \mathbf{f}_3 = 5$$

form a linearly dependent set in $F(-\infty, \infty)$, since the equation

$$\begin{aligned} 5\mathbf{f}_1 + 5\mathbf{f}_2 - \mathbf{f}_3 &= 5\sin^2 x + 5\cos^2 x - 5 \\ &= 5(\sin^2 x + \cos^2 x) - 5 = \mathbf{0} \end{aligned}$$

expresses $\mathbf{0}$ as a linear combination of $\mathbf{f}_1$, $\mathbf{f}_2$, and $\mathbf{f}_3$ with coefficients that are not all zero.

However, it is relatively rare that linear independence or dependence of functions can be ascertained by algebraic or trigonometric methods. To make matters worse, there is no general method for doing that either. That said, there does exist a theorem that can be useful for that purpose in certain cases. The following definition is needed for that theorem.

**DEFINITION 2** If $\mathbf{f}_1 = f_1(x), \mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are functions that are $n-1$ times differentiable on the interval $(-\infty, \infty)$, then the determinant

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

is called the ***Wronskian*** of $f_1, f_2, \ldots, f_n$.

Józef Hoëné de Wroński (1776–1853)

**Historical Note** The Polish-French mathematician Józef Hoëné de Wroński was born Józef Hoëné and adopted the name Wroński after he married. Wroński's life was fraught with controversy and conflict, which some say was due to psychopathic tendencies and his exaggeration of the importance of his own work. Although Wroński's work was dismissed as rubbish for many years, and much of it was indeed erroneous, some of his ideas contained hidden brilliance and have survived. Among other things, Wroński designed a caterpillar vehicle to compete with trains (though it was never manufactured) and did research on the famous problem of determining the longitude of a ship at sea. His final years were spent in poverty.

[Image: © TopFoto/The Image Works]

Suppose for the moment that $\mathbf{f}_1 = f_1(x), \mathbf{f}_2 = f_2(x), \ldots, \mathbf{f}_n = f_n(x)$ are *linearly dependent* vectors in $C^{(n-1)}(-\infty, \infty)$. This implies that the vector equation

$$k_1\mathbf{f}_1 + k_2\mathbf{f}_2 + \cdots + k_n\mathbf{f}_n = \mathbf{0}$$

is satisfied by values of the coefficients $k_1, k_2, \ldots, k_n$ that are not all zero, and for these coefficients the equation

$$k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x) = 0$$

is satisfied for all $x$ in $(-\infty, \infty)$. Using this equation together with those that result by differentiating it $n - 1$ times we obtain the linear system

$$\begin{array}{lllll} k_1 f_1(x) & + k_2 f_2(x) & + \cdots + k_n f_n(x) & = 0 \\ k_1 f_1'(x) & + k_2 f_2'(x) & + \cdots + k_n f_n'(x) & = 0 \\ \vdots & \vdots & \vdots & \vdots \\ k_1 f_1^{(n-1)}(x) & + k_2 f_2^{(n-1)}(x) & + \cdots + k_n f_n^{(n-1)}(x) & = 0 \end{array}$$

Thus, the linear dependence of $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ implies that the linear system

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{10}$$

has a nontrivial solution for every $x$ in the interval $(-\infty, \infty)$, and this in turn implies that the determinant of the coefficient matrix of (10) is zero for every such $x$. Since this determinant is the Wronskian of $f_1, f_2, \ldots, f_n$, we have established the following result.

**WARNING** The converse of Theorem 4.3.4 is false. If the Wronskian of $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ is identically zero on $(-\infty, \infty)$, then no conclusion can be reached about the linear independence of $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$—this set of vectors may be linearly independent or linearly dependent.

**THEOREM 4.3.4** *If the functions $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n$ have $n-1$ continuous derivatives on the interval $(-\infty, \infty)$, and if the Wronskian of these functions is not identically zero on $(-\infty, \infty)$, then these functions form a linearly independent set of vectors in $C^{(n-1)}(-\infty, \infty)$.*

In Example 6 we showed that $x$ and $\sin x$ are linearly independent functions by observing that neither is a scalar multiple of the other. The following example illustrates how to obtain the same result using the Wronskian (though it is a more complicated procedure in this particular case).

▶ **EXAMPLE 7 Linear Independence Using the Wronskian**

Use the Wronskian to show that $\mathbf{f}_1 = x$ and $\mathbf{f}_2 = \sin x$ are linearly independent vectors in $C^\infty(-\infty, \infty)$.

*Solution* The Wronskian is

$$W(x) = \begin{vmatrix} x & \sin x \\ 1 & \cos x \end{vmatrix} = x\cos x - \sin x$$

This function is not identically zero on the interval $(-\infty, \infty)$ since, for example,

$$W\left(\frac{\pi}{2}\right) = \frac{\pi}{2}\cos\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2}\right) = \frac{\pi}{2}$$

Thus, the functions are linearly independent.



▶ **EXAMPLE 8 Linear Independence Using the Wronskian**

Use the Wronskian to show that $\mathbf{f}_1 = 1, \mathbf{f}_2 = e^x$, and $\mathbf{f}_3 = e^{2x}$ are linearly independent vectors in $C^\infty(-\infty, \infty)$.

*Solution* The Wronskian is

$$W(x) = \begin{vmatrix} 1 & e^x & e^{2x} \\ 0 & e^x & 2e^{2x} \\ 0 & e^x & 4e^{2x} \end{vmatrix} = 2e^{3x}$$

This function is obviously not identically zero on $(-\infty, \infty)$, so $\mathbf{f}_1, \mathbf{f}_2$, and $\mathbf{f}_3$ form a linearly independent set. ◀

**OPTIONAL**

We will close this section by proving Theorem 4.3.1.

*Proof of Theorem 4.3.1* We will prove this theorem in the case where the set $S$ has two or more vectors, and leave the case where $S$ has only one vector as an exercise. Assume first that $S$ is linearly independent. We will show that if the equation

$$k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r = \mathbf{0} \tag{11}$$

can be satisfied with coefficients that are not all zero, then at least one of the vectors in $S$ must be expressible as a linear combination of the others, thereby contradicting the assumption of linear independence. To be specific, suppose that $k_1 \neq 0$. Then we can rewrite (11) as

$$\mathbf{v}_1 = \left(-\frac{k_2}{k_1}\right)\mathbf{v}_2 + \cdots + \left(-\frac{k_r}{k_1}\right)\mathbf{v}_r$$

which expresses $\mathbf{v}_1$ as a linear combination of the other vectors in $S$.

Conversely, we must show that if the only coefficients satisfying (11) are

$$k_1 = 0, \quad k_2 = 0, \ldots, \quad k_r = 0$$

then the vectors in $S$ must be linearly independent. But if this were true of the coefficients and the vectors were not linearly independent, then at least one of them would be expressible as a linear combination of the others, say

$$\mathbf{v}_1 = c_2\mathbf{v}_2 + \cdots + c_r\mathbf{v}_r$$

which we can rewrite as

$$\mathbf{v}_1 + (-c_2)\mathbf{v}_2 + \cdots + (-c_r)\mathbf{v}_r = \mathbf{0}$$

But this contradicts our assumption that (11) can only be satisfied by coefficients that are all zero. Thus, the vectors in $S$ must be linearly independent. ◀

## Exercise Set 4.3

1. Explain why the following form linearly dependent sets of vectors. (Solve this problem by inspection.)

   (a) $\mathbf{u}_1 = (-1, 2, 4)$ and $\mathbf{u}_2 = (5, -10, -20)$ in $R^3$

   (b) $\mathbf{u}_1 = (3, -1)$, $\mathbf{u}_2 = (4, 5)$, $\mathbf{u}_3 = (-4, 7)$ in $R^2$

   (c) $\mathbf{p}_1 = 3 - 2x + x^2$ and $\mathbf{p}_2 = 6 - 4x + 2x^2$ in $P_2$

   (d) $A = \begin{bmatrix} -3 & 4 \\ \ldots \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -4 \\ \ldots \end{bmatrix}$ in $M_{22}$

2. In each part, determine whether the vectors are linearly independent or are linearly dependent in $R^3$.

   (a) $(-3, 0, 4)$, $(5, -1, 2)$, $(1, 1, 3)$

   (b) $(-2, 0, 1)$, $(3, 2, 5)$, $(6, -1, 1)$, $(7, 0, -2)$

3. In each part, determine whether the vectors are linearly independent or are linearly dependent in $R^4$.

   (a) $(3, 8, 7, -3)$, $(1, 5, 3, -1)$, $(2, -1, 2, 6)$, $(4, 2, 6, 4)$

4. In each part, determine whether the vectors are linearly independent or are linearly dependent in $P_2$.

(a) $2 - x + 4x^2$, $3 + 6x + 2x^2$, $2 + 10x - 4x^2$

(b) $1 + 3x + 3x^2$, $x + 4x^2$, $5 + 6x + 3x^2$, $7 + 2x - x^2$

5. In each part, determine whether the matrices are linearly independent or dependent.

(a) $\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ in $M_{22}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ in $M_{23}$

6. Determine all values of $k$ for which the following matrices are linearly independent in $M_{22}$.

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

7. In each part, determine whether the three vectors lie in a plane in $R^3$.

(a) $\mathbf{v}_1 = (2, -2, 0)$, $\mathbf{v}_2 = (6, 1, 4)$, $\mathbf{v}_3 = (2, 0, -4)$

(b) $\mathbf{v}_1 = (-6, 7, 2)$, $\mathbf{v}_2 = (3, 2, 4)$, $\mathbf{v}_3 = (4, -1, 2)$

8. In each part, determine whether the three vectors lie on the same line in $R^3$.

(a) $\mathbf{v}_1 = (-1, 2, 3)$, $\mathbf{v}_2 = (2, -4, -6)$, $\mathbf{v}_3 = (-3, 6, 0)$

(b) $\mathbf{v}_1 = (2, -1, 4)$, $\mathbf{v}_2 = (4, 2, 3)$, $\mathbf{v}_3 = (2, 7, -6)$

(c) $\mathbf{v}_1 = (4, 6, 8)$, $\mathbf{v}_2 = (2, 3, 4)$, $\mathbf{v}_3 = (-2, -3, -4)$

9. (a) Show that the three vectors $\mathbf{v}_1 = (0, 3, 1, -1)$, $\mathbf{v}_2 = (6, 0, 5, 1)$, and $\mathbf{v}_3 = (4, -7, 1, 3)$ form a linearly dependent set in $R^4$.

(b) Express each vector in part (a) as a linear combination of the other two.

10. (a) Show that the vectors $\mathbf{v}_1 = (1, 2, 3, 4)$, $\mathbf{v}_2 = (0, 1, 0, -1)$, and $\mathbf{v}_3 = (1, 3, 3, 3)$ form a linearly dependent set in $R^4$.

(b) Express each vector in part (a) as a linear combination of the other two.

11. For which real values of $\lambda$ do the following vectors form a linearly dependent set in $R^3$?

$$\mathbf{v}_1 = \left(\lambda, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \quad \mathbf{v}_2 = \left(-\tfrac{1}{2}, \lambda, -\tfrac{1}{2}\right), \quad \mathbf{v}_3 = \left(-\tfrac{1}{2}, -\tfrac{1}{2}, \lambda\right)$$

12. Under what conditions is a set with one vector linearly independent?

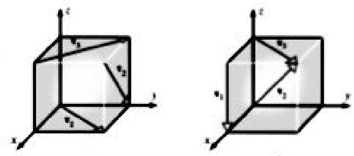
13. In each part, let $T_A: R^2 \to R^2$ be multiplication by $A$, and let $\mathbf{u}_1 = (1, 2)$ and $\mathbf{u}_2 = (-1, 1)$. Determine whether the set $\{T_A(\mathbf{u}_1), T_A(\mathbf{u}_2)\}$ is linearly independent in $R^2$.

(a) $A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$

14. In each part, let $T_A: R^3 \to R^3$ be multiplication by $A$, and let $\mathbf{u}_1 = (1, 0, 0)$, $\mathbf{u}_2 = (2, -1, 1)$, and $\mathbf{u}_3 = (0, 1, 1)$. Determine whether the set $\{T_A(\mathbf{u}_1), T_A(\mathbf{u}_2), T_A(\mathbf{u}_3)\}$ is linearly independent in $R^3$.

(a) $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -3 \\ 2 & 2 & 0 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -3 \\ 2 & 2 & 0 \end{bmatrix}$

15. Are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ in part (a) of the accompanying figure linearly independent? What about those in part (b)? Explain.



(a) (b)

▲ Figure Ex-15

16. By using appropriate identities, where required, determine which of the following sets of vectors in $F(-\infty, \infty)$ are linearly dependent.

(a) $6, 3\sin^2 x, 2\cos^2 x$ (b) $x, \cos x$

(c) $1, \sin x, \sin 2x$ (d) $\cos 2x, \sin^2 x, \cos^2 x$

(e) $(3 - x)^2, x^2 - 6x, 5$ (f) $0, \cos^3 \pi x, \sin^5 3\pi x$

17. (*Calculus required*) The functions

$$f_1(x) = x \quad \text{and} \quad f_2(x) = \cos x$$

are linearly independent in $F(-\infty, \infty)$ because neither function is a scalar multiple of the other. Confirm the linear independence using the Wronskian.

18. (*Calculus required*) The functions

$$f_1(x) = \sin x \quad \text{and} \quad f_2(x) = \cos x$$

are linearly independent in $F(-\infty, \infty)$ because neither function is a scalar multiple of the other. Confirm the linear independence using the Wronskian.

19. (*Calculus required*) Use the Wronskian to show that the following sets of vectors are linearly independent.

(a) $1, x, e^x$ (b) $1, x, x^2$

20. (*Calculus required*) Use the Wronskian to show that the functions $f_1(x) = e^x$, $f_2(x) = xe^x$, and $f_3(x) = x^2e^x$ are linearly independent vectors in $C^\infty(-\infty, \infty)$.

21. (*Calculus required*) Use the Wronskian to show that the functions $f_1(x) = \sin x$, $f_2(x) = \cos x$, and $f_3(x) = x\cos x$ are linearly independent vectors in $C^\infty(-\infty, \infty)$.



22. Show that for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in a vector space $V$, the vectors $\mathbf{u} - \mathbf{v}$, $\mathbf{v} - \mathbf{w}$, and $\mathbf{w} - \mathbf{u}$ form a linearly dependent set.

23. (a) In Example 1 we showed that the mutually orthogonal vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ form a linearly independent set of vectors in $R^3$. Do you think that every set of three nonzero mutually orthogonal vectors in $R^3$ is linearly independent? Justify your conclusion with a geometric argument.

(b) Justify your conclusion with an algebraic argument. [*Hint:* Use dot products.]

### Working with Proofs

24. Prove that if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set of vectors, then so are $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_1, \mathbf{v}_3\}$, $\{\mathbf{v}_2, \mathbf{v}_3\}$, $\{\mathbf{v}_1\}$, $\{\mathbf{v}_2\}$, and $\{\mathbf{v}_3\}$.

25. Prove that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a linearly independent set of vectors, then so is every nonempty subset of $S$.

26. Prove that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly dependent set of vectors in a vector space $V$, and $\mathbf{v}_4$ is any vector in $V$ that is not in $S$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is also linearly dependent.

27. Prove that if $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is a linearly dependent set of vectors in a vector space $V$, and if $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ are any vectors in $V$ that are not in $S$, then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ is also linearly dependent.

28. Prove that in $P_2$ every set with more than three vectors is linearly dependent.

29. Prove that if $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent and $\mathbf{v}_3$ does not lie in span$\{\mathbf{v}_1, \mathbf{v}_2\}$, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

30. Use part (*a*) of Theorem 4.3.1 to prove part (*b*).

31. Prove part (*b*) of Theorem 4.3.2.

32. Prove part (*c*) of Theorem 4.3.2.

### True-False Exercises

TF. In parts (a)–(h) determine whether the statement is true or false, and justify your answer.

(a) A set containing a single vector is linearly independent.

(b) The set of vectors $\{\mathbf{v}, k\mathbf{v}\}$ is linearly dependent for every scalar $k$.

(c) Every linearly dependent set contains the zero vector.

(d) If the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent, then $\{k\mathbf{v}_1, k\mathbf{v}_2, k\mathbf{v}_3\}$ is also linearly independent for every nonzero scalar $k$.

(e) If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent nonzero vectors, then at least one vector $\mathbf{v}_k$ is a unique linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$.

(f) The set of $2 \times 2$ matrices that contain exactly two 1's and two 0's is a linearly independent set in $M_{22}$.

(g) The three polynomials $(x - 1)(x + 2)$, $x(x + 2)$, and $x(x - 1)$ are linearly independent.

(h) The functions $f_1$ and $f_2$ are linearly dependent if there is a real number $x$ such that $k_1 f_1(x) + k_2 f_2(x) = 0$ for some scalars $k_1$ and $k_2$.

### Working with Technology

T1. Devise three different methods for using your technology utility to determine whether a set of vectors in $R^n$ is linearly independent, and then use each of those methods to determine whether the following vectors are linearly independent.

$$\mathbf{v}_1 = (4, -5, 2, 6), \quad \mathbf{v}_2 = (2, -2, 1, 3),$$
$$\mathbf{v}_3 = (6, -3, 3, 9), \quad \mathbf{v}_4 = (4, -1, 5, 6)$$

T2. Show that $S = \{\cos t, \sin t, \cos 2t, \sin 2t\}$ is a linearly independent set in $C(-\infty, \infty)$ by evaluating the left side of the equation

$$c_1 \cos t + c_2 \sin t + c_3 \cos 2t + c_4 \sin 2t = 0$$

at sufficiently many values of $t$ to obtain a linear system whose only solution is $c_1 = c_2 = c_3 = c_4 = 0$.

## 4.4 Coordinates and Basis

We usually think of a line as being one-dimensional, a plane as two-dimensional, and the space around us as three-dimensional. It is the primary goal of this section and the next to make this intuitive notion of dimension precise. In this section we will discuss coordinate systems in general vector spaces and lay the groundwork for a precise definition of dimension in the next section.

*Coordinate Systems in Linear Algebra*

In analytic geometry one uses *rectangular* coordinate systems to create a one-to-one correspondence between points in 2-space and ordered pairs of real numbers and between points in 3-space and ordered triples of real numbers (Figure 4.4.1). Although rectangular coordinate systems are common, they are not essential. For example, Figure 4.4.2