1.1 Introduction

Linear algebra is an important mathematical subject independent of its connection to differential equations. However, it is the relation between differential equations and linear algebra that is important to us and so we will use this relationship to motivate our study of linear algebra.

In Chapter 3 we studied second order linear differential equations:

\[ y'' + p(x)y' + q(x)y = f(x) \]  \hspace{1cm} (1)

where \( p, q \) and \( f \) are continuous functions on an interval \( I \). The focus of our study was on the homogeneous equation

\[ y'' + p(x)y' + q(x)y = 0, \]  \hspace{1cm} (2)

the so-called reduced equation of equation (1). We saw that the general solution of (2) is given by \( y = C_1 y_1(x) + C_2 y_2(x) \) where \( C_1 \) and \( C_2 \) are arbitrary constants and \( y_1 \) and \( y_2 \) are linearly independent solutions of the equation. We also saw that an initial-value problem

\[ y'' + p(x)y' + q(x)y = 0; \quad y(a) = \alpha, \quad y'(a) = \beta \]

required us to solve the system of equations

\[ y_1(a)C_1 + y_2(a)C_2 = \alpha \]
\[ y_1'(a)C_1 + y_2'(a)C_2 = \beta \]

In the next chapter we will consider higher-order linear differential equations as a lead in to our study of systems of linear differential equations. An \( n^{th} \)-order linear differential equation is an equation of the form

\[ y^{(n)} + p_{n-1}(x)y^{(n-1)} + p_{n-2}(x)y^{(n-2)} + \cdots + p_1(x)y' + p_0(x)y = f(x) \]  \hspace{1cm} (3)

where \( p_0, p_1, \ldots, p_{n-1} \) and \( f \) are continuous functions on some interval \( I \). As in the case of second order equations, the focus is on homogeneous equations:

\[ y^{(n)} + p_{n-1}(x)y^{(n-1)} + p_{n-2}(x)y^{(n-2)} + \cdots + p_1(x)y' + p_0(x)y = 0. \]  \hspace{1cm} (4)

All of the theorems in Chapter 3 carry over verbatim to \( n^{th} \)-order equations. In particular, to obtain the general solution of (4) we need to find \( n \) linearly independent solutions \( y_1, y_2, \ldots, y_n \) of (4). The general solution is then given by

\[ y = C_1 y_1(x) + C_2 y_2(x) + \cdots + C_n y_n(x). \]
This is the first major question. We know what it means for two functions to be linearly independent (one is not a multiple of the other), but what does it mean for \( n \) functions to be linearly independent? And, given a set of \( n \) functions, how can we determine whether or not they are linearly independent?

To solve the \( n^{th} \)-order initial-value problem

\[
y^{(n)} + p_{n-1}(x)y^{(n-1)} + p_{n-2}(x)y^{(n-2)} + \cdots + p_1(x)y' + p_0(x)y = 0;
\]

\[
y(a) = \alpha_0, \quad y'(a) = \alpha_1, \quad \ldots, \quad y^{(n-1)}(a) = \alpha_{n-1}
\]

we would need to solve the system of \( n \) equations in \( n \) unknowns

\[
y_1(a)C_1 + y_2(a)C_2 + \cdots + y_n(a)C_n = \alpha_0
\]

\[
y_1'(a)C_1 + y_2'(a)C_2 + \cdots + y_n'(a)C_n = \alpha_1
\]

\[
\vdots
\]

\[
y_1^{(n-1)}(a)C_1 + y_2^{(n-1)}(a)C_2 + \cdots + y_n^{(n-1)}(a)C_n = \alpha_{n-1}.
\]

This is the second major question. While it is easy to solve two equations in two unknowns, how do we solve “large” systems consisting of \( m \) equations in \( n \) unknowns?

The primary purpose of this chapter is to address these two questions along with a number of other topics that are closely related to these questions.
1.2 Systems of Linear Equations, Some Geometry

A linear (algebraic) equation in \( n \) unknowns, \( x_1, x_2, \ldots, x_n \), is an equation of the form
\[
a_1x_1 + a_2x_2 + \cdots + a_nx_n = b
\]
where \( a_1, a_2, \ldots, a_n \) and \( b \) are real numbers. In particular
\[
ax = b, \quad a, b \text{ given real numbers}
\]
is a linear equation in one unknown;
\[
ax + by = c, \quad a, b, c \text{ given real numbers}
\]
is a linear equation in two unknowns (if \( a \) and \( b \) are not both 0, the graph of the equation is a straight line); and
\[
ax + by + cz = d, \quad a, b, c, d \text{ given real numbers}
\]
is a linear equation in three unknowns (if \( a, b \) and \( c \) are not all 0, then the graph is a plane in 3-space).

Our interest in this section is in solving systems of linear equations.

**Linear equations in one unknown**  We begin with simplest case: one equation in one unknown. If you were asked to find a real number \( x \) such that
\[
ax = b
\]
most people would say “that’s easy,” \( x = \frac{b}{a} \). But the fact is, this “solution” is not necessarily correct. For example, consider the three equations
\[
(1) \ 2x = 6, \quad (2) \ 0x = 6, \quad (3) \ 0x = 0.
\]
For equation (1), the solution \( x = 6/2 = 3 \) is correct. However, consider equation (2); there is no real number that satisfies this equation! Now look at equation (3); every real number satisfies (3).

In general, it is easy to see that for the equation \( ax = b \), exactly one of three things happens: either there is precisely one solution \( (x = b/a, \text{ when } a \neq 0) \), or there are no solutions \( (a = 0, b \neq 0) \), or there are infinitely many solutions \( (a = b = 0) \). As we will see, this simple case illustrates what happens in general. **For any system of \( m \) linear equations in \( n \) unknowns, exactly one of three possibilities occurs: a unique solution, no solution, or infinitely many solutions.**

**Linear equations in two unknowns**  We begin with one equation:
\[
ax + by = c.
\]
Here we are looking for ordered pairs of real numbers \((x, y)\) which satisfy the equation. If \(a = b = 0\) and \(c \neq 0\), then there are no solutions. If \(a = b = c = 0\), then every ordered pair \((x, y)\) satisfies the equation. If at least one of \(a\) and \(b\) is different from 0, then the equation \(ax + by = c\) represents a straight line in the \(xy\)-plane and the equation has infinitely many solutions, the set of all points on the line. Note that in this case it is not possible to have a unique solution; we either have no solution or infinitely many solutions.

Two linear equations in two unknowns is a more interesting case. The pair of equations

\[
ax + by = \alpha \\
cx + dy = \beta
\]

represents a pair of lines in the \(xy\)-plane. We are looking for ordered pairs \((x, y)\) of real numbers that satisfy both equations simultaneously. Since two lines in the plane either

(a) have a unique point of intersection (this occurs when the lines have different slopes), or

(b) are parallel (the lines have the same slope but, for example, different \(y\)-intercepts), or

(c) coincide (same slope, same \(y\)-intercept).

If (a) occurs, the system of equations has a unique solution; if (b) occurs, the system has no solution; if (c) occurs, the system has infinitely many solutions.

**Example 1.**

\[
\begin{align*}
\text{(a)} & \quad x + 2y &= 2 \\
& -2x + y &= 6 \\
\text{(b)} & \quad x + 2y &= 2 \\
& -2x - 4y &= -8 \\
\text{(c)} & \quad x + 2y &= 2 \\
& 2x + 4y &= 4
\end{align*}
\]

Three linear equations in two unknowns represent three lines in the \(xy\)-plane. It’s unlikely that three lines chosen at random will all go through the same point. Therefore, we should not expect a system of three equations in two unknowns to have a solution; it’s possible, but not likely. The most likely occurrence is that there will be no solution. Here is a typical example.
**Example 2.**

\[
\begin{align*}
  x + y &= 2 \\
  -2x + y &= 2 \\
  4x + y &= 11
\end{align*}
\]

**Linear equations in three unknowns**  A linear equation in three unknowns has the form

\[ ax + by + cz = d. \]

Here we are looking for ordered triples \((x, y, z)\) that satisfy the equation. The cases \(a = b = c = 0, \ d \neq 0\) and \(a = b = c = d = 0\) should be obvious to you. In the first case: no solutions; in the second case: infinitely many solutions, namely all of 3-space. If \(a, b\) and \(c\) are not all zero, then the equation represents a plane in three space. The solutions of the equation are the points of the plane; the equation has infinitely many solutions, a two-dimensional set. The figure shows the plane \(2x - 3y + z = 2\).

A system of two linear equations in three unknowns

\[
\begin{align*}
  a_{11}x + a_{12}y + a_{13}z &= b_1 \\
  a_{21}x + a_{22}y + a_{23}z &= b_2
\end{align*}
\]

(we’ve switched to subscripts because we’re running out of distinct letters) represents two planes in 3-space. Either the two planes are parallel (the system has no solutions), or they coincide (infinitely many solutions, a whole plane of solutions), or they intersect in a straight line (again, infinitely many solutions, but this time only a one-dimensional set).
The figure shows planes $2x - 3y + z = 2$ and $2x - 3y - z = -2$ and their line of intersection.

The most interesting case is a system of three linear equations in three unknowns.

$$
\begin{align*}
    a_{11}x + a_{12}y + a_{13}z &= b_1 \\
    a_{21}x + a_{22}y + a_{23}z &= b_2 \\
    a_{31}x + a_{32}y + a_{33}z &= b_3
\end{align*}
$$

Geometrically, the system represents three planes in 3-space. We still have the three mutually exclusive cases:

(a) The system has a unique solution; the three planes have a unique point of intersection;

(b) The system has infinitely many solutions; the three planes intersect in a line, or the three planes intersect in a plane.

(c) The system has no solution; there is no point the lies on all three planes.

Try to picture the possibilities here. While we still have the three basic cases, the geometry is considerably more complicated. This is where linear algebra will help us understand the geometry.

We could go on to a system of four (or more) equations in three unknowns but, like the case of three equations in two unknowns, it is unlikely that such a system will have a solution.

Figures and graphs in the plane are standard. Figures and graphs in 3-space are possible, but are often difficult to draw. Figures and graphs are not possible in dimensions higher than three.

**Exercises 1.2**

Solve the system of equations. Then graph the equations to illustrate your solution.
1. \(x - 2y = 2\)
   \(x + y = 5\)

2. \(x + 2y = -4\)
   \(2x + 4y = 8\)

3. \(2x + 4y = 8\)
   \(x + 2y = 4\)

4. \(2x - 2y = -4\)
   \(6x - 3y = -18\)

5. \(-x + 2y = 5\)
   \(2x + 3y = -3\)
   \(x - 2y = -6\)

6. \(2x + 3y = 1\)
   \(3x - y = 7\)
   \(x + y = 1\)

7. \(2x + y = 8\)
   \(x + 2y = -2\)
   \(3x - 6y = -9\)

8. \(x - 2y = -8\)
   \(3x + y = -3\)
   \(4x - 3y = -24\)

9. \(-2x + 4y = 6\)
   \(-\frac{1}{2}x + y = \frac{3}{2}\)

10. \(3x + 8y - 13z = 7\)

Describe the solution set of system of equations. That is, is the solution set a point in 3-space, a line in 3-space, a plane in 3-space, or is there no solution? The graphs of the equations are planes in 3-space. Use “technology” to graph the equations to illustrate your solutions.

11. \(x - 2y + z = 3\)
    \(3x + y - 2z = 2\)

12. \(2x - 4y + 2z = -6\)
    \(-3x + 6y - z = 9\)

13. \(x + 3y - 4z = -2\)
    \(-3x - 9y + 12z = 4\)

14. \(x - 2y + z = 3\)
    \(3x + y - 2z = 2\)
    \(x + 2y - z = 3\)

15. \(2x + 5y - 4z = 5\)
    \(3x + 4y + 2z = 12\)
    \(x + 2y - 3z = 1\)

16. \(2x + 5y - 8z = 4\)
    \(3x + 8y - 13z = 7\)
1.3 Solving Systems of Linear Equations

In this section we will develop a systematic method for solving systems of linear equations. We’ll begin with a simple case, two equations in two unknowns:

\[
\begin{align*}
ax + by &= \alpha \\
(cx + dy) &= \beta
\end{align*}
\]

Our objective is to solve this system of equations. This means that we want to find all pairs of numbers \(x, y\) that satisfy both equations simultaneously. As you probably know, there is a variety of ways to do this. We’ll illustrate an approach which we’ll extend to systems of \(m\) equations in \(n\) unknowns.

Example 1. Solve the system

\[
\begin{align*}
3x + 12y &= 6 \\
2x - 3y &= -7
\end{align*}
\]

SOLUTION We multiply the first equation by \(1/3\) (divide by 3). This results in the system

\[
\begin{align*}
x + 4y &= 2 \\
2x - 3y &= -7
\end{align*}
\]

This system has the same solution set as the original system (multiplying an equation by a nonzero number produces an equivalent equation).

Next, we multiply the first equation by \(-2\), add it to the second equation, and use the result as the second equation. This produces the system

\[
\begin{align*}
x + 4y &= 2 \\
-11y &= -11
\end{align*}
\]

As we will show below, this system also has the same solution set as the original system.

Finally, we multiply the second equation by \(-1/11\) (divide by \(-11\)) to get

\[
\begin{align*}
x + 4y &= 2 \\
y &= 1
\end{align*}
\]

and this system has the same solution set as the original. Notice the “triangular” form of our final system. The advantage of this system is that it is easy to solve. From the second equation, \(y = 1\). Substituting \(y = 1\) in the first equation gives

\[
x + 4(1) = 2 \quad \text{which implies} \quad x = -2.
\]

The system has the unique solution \(x = -2, \ y = 1\). ■
The basic idea is this: given a system of linear equations, perform a sequence of operations to produce a new system which has the same solution set as the given system, and which is easy to solve.

Two systems of equations are *equivalent* if they have the same solution set.

**The Elementary Operations**

In Example 1, we performed a sequence of operations on a system to produce an equivalent system which was easy to solve.

The operations that produce equivalent systems are called *elementary operations*. The elementary operations are

1. Multiply an equation by a nonzero number.
2. Interchange two equations.
3. Multiply an equation by a number and add it to another equation.

It is obvious that the first two operations preserve the solution set; that is, produce equivalent systems.

Using two equations in two unknowns, we’ll justify that the third operation also preserves the solution set. Exactly the same proof will work in the general case of $m$ equations in $n$ unknowns.

Given the system

\[
\begin{align*}
ax + by &= \alpha \\
ax + cy + kby + dy &= \beta.
\end{align*}
\]  

(a)

Multiply the first equation by $k$ and add the result to the second equation. Replace the second equation in the given system with this new equation. We then have

\[
\begin{align*}
ax + by &= \alpha \\
kax + cx + kby + dy &= k\alpha + \beta
\end{align*}
\]

which is the same as

\[
\begin{align*}
ax + by &= \alpha \\
(ax + cy + kby + dy) &= k\alpha + \beta
\end{align*}
\]

(b)

Suppose that $(x_0, y_0)$ is a solution of system (a). To show that (a) and (b) are
equivalent, we need only show that \((x_0, y_0)\) satisfies the second equation in system (b):

\[
(ka + c)x_0 + (kb + d)y_0 = kax_0 + cxy_0 + kby_0 + dy_0 \\
= kax_0 + kby_0 + cxy_0 + dy_0 \\
= k(ax_0 + by_0) + (cxy_0 + dy_0) \\
= k\alpha + \beta.
\]

Thus, \((x_0, y_0)\) is a solution of (b).

Now suppose that \((x_0, y_0)\) is a solution of system (b). In this case, we only need to show that \((x_0, y_0)\) satisfies the second equation in (a). We have

\[
(ka + c)x_0 + (kb + d)y_0 = k\alpha + \beta \\
ka x_0 + kby_0 + cxy_0 + dy_0 = k\alpha + \beta \\
k(ax_0 + by_0) + cxy_0 + dy_0 = k\alpha + \beta \\
ka + cxy_0 + dy_0 = k\alpha + \beta \\
cxy_0 + dy_0 = \beta.
\]

Thus, \((x_0, y_0)\) is a solution of (a). ■

The following examples illustrate the use of the elementary operations to transform a given system of linear equations into an equivalent system which is in a triangular form from which it is easy to determine the set of solutions. To keep track of the steps involved, we’ll use the notations:

- \(kE_i \to E_i\) to mean “multiply equation \((i)\) by the nonzero number \(k\).”
- \(E_i \leftrightarrow E_j\) to mean “interchange equations \(i\) and \(j\).”
- \(kE_i + E_j \to E_j\) to mean “multiply equation \((i)\) by \(k\) and add the result to equation \((j)\).”

The “interchange equations” operation may seem silly to you, but you’ll soon see its value.

**Example 2.** Solve the system

\[
\begin{align*}
x + 2y - 5z &= -1 \\
-3x - 9y + 21z &= 0 \\
x + 6y - 11z &= 1
\end{align*}
\]

**SOLUTION** We’ll use the elementary operations to produce an equivalent system in a “triangular” form.
\[
\begin{align*}
   x + 2y - 5z &= -1 & \quad x + 2y - 5z &= -1 \\
-3x - 9y + 21z &= 0 & \quad 3E_1 + E_2 \rightarrow E_2, (\text{1})E_3 \rightarrow E_3 \\
   x + 6y - 11z &= 1 & \quad -3y + 6z &= -3 \\
\end{align*}
\]

\[
\begin{align*}
   x + 2y - 5z &= -1 & \quad x + 2y - 5z &= -1 \\
\rightarrow & \quad y - 2z = 1 & \quad y - 2z &= 1 \\
(1/3)E_2 \rightarrow E_2 & \quad 4y - 6z = 2 & \quad -4E_2 + E_3 \rightarrow E_3 \\
\end{align*}
\]

\[
\begin{align*}
   x + 2y - 5z &= -1 & \quad x + 2y - 5z &= -1 \\
\rightarrow & \quad y - 2z = 1 & \quad y - 2z &= 1 \\
(1/2)E_3 \rightarrow E_3 & \quad z &= -1
\end{align*}
\]

From the last equation, \( z = -1 \). Substituting this value into the second equation gives \( y = -1 \), and substituting these two values into the first equation gives \( x = -4 \). The system has the unique solution \( x = -4, y = -1, z = -1 \).

Note the “strategy:” use the first equation to eliminate \( x \) from the second and third equations. Then use the (new) second equation to eliminate \( y \) from the third equation. This results in an equivalent system in which the third equation is easily solved for \( z \). Putting that value in the second equation gives \( y \), substituting the values for \( y \) and \( z \) in the first equation gives \( x \). \( \blacksquare \)

We continue with examples that illustrate the solution method as well as the other two possibilities for solution sets: no solution, infinitely many solutions.

**Example 3.** Solve the system

\[
\begin{align*}
   3x - 4y - z &= 3 \\
   2x - 3y + z &= 1 \\
   x - 2y + 3z &= 2
\end{align*}
\]

**SOLUTION**

\[
\begin{align*}
   3x - 4y - z &= 3 & \quad x - 2y + 3z &= 2 \\
   2x - 3y + z &= 1 & \quad 2x - 3y + z &= 1 \\
   x - 2y + 3z &= 2 & \quad 3x - 4y - z &= 3
\end{align*}
\]
Having $x$ with coefficient 1 in the first equation makes it much easier to eliminate $x$ from the remaining equations; we want to avoid fractions as long as we can in order to keep the calculations as simple as possible. This is the value of the “interchange” operation.

\[ \begin{align*}
  x - 2y + 3z &= 2 \\
  2x - 3y + z &= 1 \\
  3x - 4y - z &= 3
\end{align*} \]

\[ \rightarrow \begin{align*}
  (-2)E_1 + E_2 \rightarrow E_2, (-3)E_1 + E_3 \rightarrow E_3 \\
  y - 5z &= -3 \\
  2y - 10z &= -3
\end{align*} \]

\[ \begin{align*}
  x - 2y + 3z &= 2 \\
  y - 5z &= -3 \\
  0z &= 3
\end{align*} \]

Clearly, the third equation in this system has no solution. Therefore the system has no solution. Since this system is equivalent to the original system, the original system has no solution. ■

**Example 4.** Solve the system

\[ \begin{align*}
  x + y - 3z &= 1 \\
  2x + y - 4z &= 0 \\
  -3x + 2y - z &= 7
\end{align*} \]

**SOLUTION**

\[ \begin{align*}
  x + y - 3z &= 1 \\
  2x + y - 4z &= 0 \rightarrow (-2)E_1 + E_2 \rightarrow E_2, 3E_1 + E_3 \rightarrow E_3 \\
  -3x + 2y - z &= 7 \rightarrow (-5)E_2 + E_3 \rightarrow E_3
\end{align*} \]

\[ \begin{align*}
  x + y - 3z &= 1 \\
  y - 2z &= 2 \\
  5y - 10z &= 10 \rightarrow (-1)E_2 \rightarrow E_2
\end{align*} \]

Since every real number satisfies the third equation, the system has infinitely many solutions. Set $z = a$, $a$ any real number. Then, from the second equation, we get $y = 2 + 2a$ and, from the first equation, $x = 1 - y + 3a = 1 - (2 + 2a) + 3a = -1 + a$. Thus, the solution set is:

\[ x = -1 + a, \quad y = 2 + 2a, \quad z = a, \quad a \text{ any real number.} \]

If we regard $a$ as a parameter, then we can say that the system has a one-parameter family of solutions. ■

In our examples thus far our systems have been “square” – the number of unknowns equals the number of equations. As we’ll see, this is the most interesting case. However having the number of equations equal the number of unknowns is certainly not a requirement; the same method can be used on a system of $m$ linear equations in $n$ unknowns.
Example 5. Solve the system

\[
\begin{align*}
  x_1 - 2x_2 + x_3 - x_4 &= -2 \\
  -2x_1 + 5x_2 - x_3 + 4x_4 &= 1 \\
  3x_1 - 7x_2 + 4x_3 - 4x_4 &= -4 \\
\end{align*}
\]

Note: We typically use subscript notation when the number of unknowns is greater than 3.

**SOLUTION**

\[
\begin{align*}
  x_1 - 2x_2 + x_3 - x_4 &= -2 \\
  -2x_1 + 5x_2 - x_3 + 4x_4 &= 1 & \rightarrow E_1 + 2E_2 - E_2, (-3)E_1 + E_3 & \rightarrow x_2 + x_3 + 2x_4 = 1 \\
  3x_1 - 7x_2 + 4x_3 - 4x_4 &= -4 & \rightarrow E_2 + E_3 - E_4 & \rightarrow x_2 + x_3 + 2x_4 = -3 \\
  & \rightarrow E_2 + E_3 - E_4 & \rightarrow x_2 + x_3 + 2x_4 = -3 \\
  & \rightarrow x_3 + x_4 = -1 & \rightarrow (1/2)E_3 - E_3 & \rightarrow x_3 + \frac{1}{2}x_4 = -\frac{1}{2} \\
\end{align*}
\]

This system has infinitely many solutions: set \( x_4 = a \), \( a \) any real number. Then, from the third equation, \( x_3 = \frac{-1}{2} - \frac{1}{2}a \). Substituting into the second equation, we’ll get \( x_2 \), and then substituting into the first equation we’ll get \( x_1 \). The resulting solution set is:

\[
\begin{align*}
  x_1 &= -\frac{13}{2} - \frac{3}{2}a, \quad x_2 = -\frac{5}{2} - \frac{3}{2}a, \quad x_3 = \frac{-1}{2} - \frac{1}{2}a, \quad x_4 = a, \quad a \text{ any real number.}
\end{align*}
\]

This system has a one-parameter family of solutions. ■

**Some terminology** A system of linear equations is said to be *consistent* if it has at least one solution; that is, a system is consistent if it has either a unique solution of infinitely many solutions. A system that has no solutions is *inconsistent*.

This method of using the elementary operations to “reduce” a given system to an equivalent system in triangular form, and then solving for the unknowns by working backwards from the last equation up to the first equation is called *Gaussian elimination with back substitution*. ■

**Matrix Representation of Systems of Equations**

Look carefully at the examples we’ve done. Note that the operations we performed on the systems of equations had the effect of changing the coefficients of the unknowns and the numbers on the right-hand side. The unknowns themselves played no role in the calculations, they were merely “place-holders.” With this in mind, we can save ourselves some time and effort if we simply write down the numbers in the order in which they appear, and then manipulate the numbers using the elementary operations.
Consider Example 2. The system is
\[
\begin{align*}
  x + 2y - 5z &= -1 \\
  -3x - 9y + 21z &= 0 \\
  x + 6y - 12z &= 1
\end{align*}
\]
Writing down the numbers in the order in which they appear, we get the rectangular array
\[
\begin{pmatrix}
  1 & 2 & -5 & | & -1 \\
  -3 & -9 & 21 & | & 0 \\
  1 & 6 & -12 & | & 1
\end{pmatrix}
\]
The vertical bar locates the “=” sign. The rows represent the equations. Each column to the left of the bar gives the coefficients of the corresponding unknown (the first column gives the coefficients of \( x \), etc.); the numbers to the right of the bar are the numbers on the right sides of the equations. This rectangular array of numbers is called the \textit{augmented matrix} for the system of equations; it is a short-hand representation of the system.

In general, a \textit{matrix} is a rectangular array of objects arranged in rows and columns. The objects are called the \textit{entries} of the matrix. A matrix with \( m \) rows and \( n \) columns is an \( m \times n \) matrix.

The matrices that we will be considering in this chapter will have numbers as entries. In the next chapter we will see matrices with functions as entries. Right now we are concerned with the augmented matrix of a system of linear equations.

Reproducing Example 2 in terms of augmented matrices, we have the sequence
\[
\begin{pmatrix}
  1 & 2 & -5 & | & -1 \\
  -3 & -9 & 21 & | & 0 \\
  1 & 6 & -12 & | & 1
\end{pmatrix} \rightarrow \begin{pmatrix}
  1 & 2 & -5 & | & -1 \\
  0 & -3 & 6 & | & -3 \\
  0 & 4 & -6 & | & 2
\end{pmatrix} \rightarrow \begin{pmatrix}
  1 & 2 & -5 & | & -1 \\
  0 & 1 & -2 & | & 1 \\
  0 & 0 & 2 & | & 1
\end{pmatrix}
\]
The final augmented matrix corresponds to the system
\[
\begin{align*}
  x + 2y - 5z &= -1 \\
  y - 2z &= 1 \\
  z &= -1
\end{align*}
\]
from which we can obtain the solutions \( x = -4, \ y = -1, \ z = -1 \) as we did before.
It’s time to look at systems of linear equations in general.

A system of \( m \) linear equations in \( n \) unknowns has the form

\[
\begin{align*}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
 a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
 a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots ..
\end{align*}
\]

\( a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m \)

The augmented matrix for the system is the \( m \times (n + 1) \) matrix

\[
\begin{pmatrix}
 a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\
 a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & b_2 \\
 a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & b_3 \\
 \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & b_m
\end{pmatrix}.
\]

The \( m \times n \) matrix whose entries are the coefficients of the unknowns

\[
\begin{pmatrix}
 a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
 a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
 a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}
\end{pmatrix}
\]

is called the matrix of coefficients.

**Example 6.** Given the system of equations

\[
\begin{align*}
 x_1 + 2x_2 - 3x_3 - 4x_4 &= 2 \\
 2x_1 + 4x_2 - 5x_3 - 7x_4 &= 7 \\
 -3x_1 - 6x_2 + 11x_3 + 14x_4 &= 0
\end{align*}
\]

The augmented matrix for the system is the \( 3 \times 5 \) matrix

\[
\begin{pmatrix}
 1 & 2 & -3 & -4 & 2 \\
 2 & 4 & -5 & -7 & 7 \\
 -3 & -6 & 11 & 14 & 0
\end{pmatrix}.
\]
and the matrix of coefficients is the $3 \times 4$ matrix

$$
\begin{pmatrix}
1 & 2 & -3 & -4 \\
2 & 4 & -5 & -7 \\
-3 & -6 & 11 & 14
\end{pmatrix}.
$$

**Example 7.** The matrix

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

is the augmented matrix of a system of linear equations. What is the system?

**SOLUTION** The system of equations is

\[
\begin{align*}
2x - 3y + z &= 6 \\
5y - 2z &= -1 \\
-3x + 4z &= 10 \\
7x + 2y - 2z &= 3
\end{align*}
\]

In the method of Gaussian elimination we use the elementary operations to reduce the system to an equivalent system in triangular form. The elementary operations on the equations can be viewed as operations on the rows of the augmented matrix. Rather than using the elementary operations on the equations, we’ll use corresponding operations on the rows of the augmented matrix. The *elementary row operations* on a matrix are:

1. Multiply a row by a nonzero number.
2. Interchange two rows.
3. Multiply a row by a number and add it to another row.

Obviously these correspond exactly to the elementary operations on equations. We’ll use the following notation to denote the row operations:

- $R_i \rightarrow R_i$ means “multiply row $(i)$ by the nonzero number $k$.
- $R_i \leftarrow R_j$ means “interchange rows $i$ and $j$.
- $kR_i + R_j \rightarrow R_j$ means “multiply row $(i)$ by $k$ and add the result to row $(j)$.

Now we’ll re-do Examples 3 and 4 using elementary row operations on the augmented matrix.
Example 8. (Example 3) Solve the system

\[
\begin{align*}
3x - 4y - z &= 3 \\
2x - 3y + z &= 1 \\
x - 2y + 3z &= 2
\end{align*}
\]

\textit{SOLUTION} The augmented matrix for the system of equations is

\[
\begin{pmatrix}
3 & -4 & -1 & 3 \\
2 & -3 & 1 & 1 \\
1 & -2 & 3 & 2
\end{pmatrix}
\]

Mimicking Example 3, we have

\[
\begin{pmatrix}
3 & -4 & -1 & 3 \\
2 & -3 & 1 & 1 \\
1 & -2 & 3 & 2
\end{pmatrix} \rightarrow R_1 \leftrightarrow R_3
\begin{pmatrix}
1 & -2 & 3 & 2 \\
2 & -3 & 1 & 1 \\
3 & -4 & -1 & 3
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 & -2 & 3 & 2 \\
0 & 1 & -5 & -3 \\
0 & 2 & -10 & -3
\end{pmatrix} \rightarrow -2R_2 + R_3 \rightarrow E_3
\begin{pmatrix}
1 & -2 & 3 & 2 \\
0 & 1 & -5 & -3 \\
0 & 0 & 0 & 3
\end{pmatrix}
\]

The last augmented matrix represents the system of equations

\[
\begin{align*}
x - 2y + 3z &= 2 \\
0x + y - 5z &= -3 \\
0x + 0y + 0z &= 3
\end{align*}
\]

As we saw in Example 3, the third equation in this system has no solution which implies that the original system has no solution. 

Example 9. (Example 4) Solve the system

\[
\begin{align*}
x + y - 3z &= 1 \\
2x + y - 4z &= 0 \\
-3x + 2y - z &= 7
\end{align*}
\]

\textit{SOLUTION} The augmented matrix for the system is

\[
\begin{pmatrix}
1 & 1 & -3 & 1 \\
2 & 1 & -4 & 0 \\
-3 & 2 & -1 & 7
\end{pmatrix}
\]

171
Performing the row operations to reduce the augmented matrix to triangular form, we have

\[
\begin{pmatrix}
1 & 1 & -3 & 1 \\
2 & 1 & -4 & 0 \\
-3 & 2 & -1 & 7
\end{pmatrix} \rightarrow
\begin{pmatrix}
2 & 1 & -4 & 0 \\
-3 & 2 & -1 & 7 \\
0 & 1 & -2 & 2
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 & 1 & -3 & 1 \\
0 & 1 & -2 & 2 \\
0 & 5 & -10 & 10
\end{pmatrix} \rightarrow
\begin{pmatrix}
1 & 0 & -1 & 2 \\
0 & 1 & -2 & 2 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

The last augmented matrix represents the system of equations

\[
\begin{align*}
x + y - 3z &= 1 \\
0x + y - 2z &= 2 \\
0x + 0y + 0z &= 0
\end{align*}
\]

As we saw in Example 4, this system has infinitely many solutions given by:

\[
x = -1 + a, \quad y = 2 + 2a, \quad z = a, \quad a \text{ any real number.}
\]

**Example 10.** Solve the system of equations

\[
\begin{align*}
x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\
3x_1 - 9x_2 + 7x_3 - x_4 + 3x_5 &= 7 \\
2x_1 - 6x_2 + 7x_3 + 4x_4 - 5x_5 &= 7
\end{align*}
\]

**SOLUTION** The augmented matrix is

\[
\begin{pmatrix}
1 & -3 & 2 & -1 & 2 & 2 \\
3 & -9 & 7 & -1 & 3 & 7 \\
2 & -6 & 7 & 4 & -5 & 7
\end{pmatrix}
\]

Perform elementary row operations to reduce this matrix to triangular form:

\[
\begin{pmatrix}
1 & -3 & 2 & -1 & 2 & 2 \\
3 & -9 & 7 & -1 & 3 & 7 \\
2 & -6 & 7 & 4 & -5 & 7
\end{pmatrix} \rightarrow
\begin{pmatrix}
1 & -3 & 2 & -1 & 2 & 2 \\
0 & 0 & 1 & 2 & -3 & 1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

172
The system of equations corresponding to this augmented matrix is:

\[
\begin{align*}
    x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\
    0x_1 + 0x_2 + x_3 + 2x_4 - 3x_5 &= 1 \\
    0x_1 + 0x_2 + x_3 + 0x_4 + 0x_5 &= 0.
\end{align*}
\]

Since all values of the unknowns satisfy the third equation, the system can be written equivalently as

\[
\begin{align*}
    x_1 - 3x_2 + 2x_3 - x_4 + 2x_5 &= 2 \\
    x_3 + 2x_4 - 3x_5 &= 1
\end{align*}
\]

From the second equation, \(x_3 = 1 - 2x_4 + 3x_5\). Substituting into the first equation, we get

\[
x_1 = 2 + 3x_2 - 2x_3 + x_4 - 2x_5 = 3x_2 + 5x_4 - 8x_5.
\]

If we set \(x_2 = a, x_4 = b, x_5 = c\). Then the solution set can be written as

\[
x_1 = 3a + 5b - 8c, \quad x_2 = a, \quad x_3 = 1 - 2b + 3c, \quad x_4 = b, \quad x_5 = c, \quad a, b, c \text{ arbitrary real numbers}
\]

The system has a three-parameter family of solutions. ■

**Row Echelon Form and Rank** The Gaussian elimination procedure applied to the augmented matrix of a system of linear equations consists of row operations on the matrix which convert it to a “triangular” form. Look at the examples above. This “triangular” form is called the **row-echelon form** of the matrix. A matrix is in row-echelon form if

1. Rows consisting entirely of zeros are at the bottom of the matrix.
2. The first nonzero entry in a nonzero row is a 1. This is called the **leading** 1.
3. If row \(i\) and row \(i + 1\) are nonzero rows, then the leading 1 in row \(i + 1\) is to the right of the leading 1 in row \(i\). (This implies that all the entries below a leading 1 are zero.)

**NOTE:** Because the leading 1’s in the row echelon form of a matrix move to the right as you move down the matrix, the number of leading 1’s cannot exceed the number of columns in the matrix. To put this another way, the number of nonzero rows in the row echelon form of a matrix is always less than or equal the number of columns in the matrix.

Stated in simple terms, the procedure for solving a system of \(m\) linear equations in \(n\) unknowns is:

1. Write down the augmented matrix for the system.
2. Use the elementary row operations to “reduce” the matrix to row-echelon form.

3. Write down the system of equations corresponding to the row-echelon form matrix.

4. Write down the solutions beginning with the last equation and working upwards.

It should be clear from the examples we’ve worked that it is the nonzero rows in the row-echelon form of the augmented matrix that determine the solution set of the system of equations. The zero rows, if any, represent redundant equations in the original system; the nonzero rows represent the “independent” equations in the system.

If an \( m \times n \) matrix \( A \) is reduced to row echelon form, then the number of non-zero rows in its row echelon form is called the \textit{rank} of \( A \). It is obvious that the rank of a matrix is less than or equal to the number of rows in the matrix. Also, as we noted above, the number of nonzero rows in the row echelon form is always less than or equal to the number of columns. Therefore the rank of a matrix is also less than or equal to the number of columns in the matrix.

The last nonzero row of the augmented matrix usually determines the nature of the solution set of the system.

\textbf{Case 1:} If the last nonzero row has the form
\[
(0 \ 0 \ 0 \ \cdots \ 0 \ | \ b), \quad b \neq 0,
\]
then the row corresponds to the equation
\[
0x_1 + 0x_2 + 0x_3 + \cdots + 0x_n = b, \quad b \neq 0
\]
which has no solutions. Therefore, the system has no solutions. See Example 8.

\textbf{Case 2:} If the last nonzero row has the form
\[
(0 \ 0 \ 0 \ \cdots \ 1 \ * \ * \ * \ | \ b),
\]
where the “1” is in the \( k^{th} \), \( k < n \) column and the *’s represent numbers in the \( k + 1\)-st through the \( n^{th} \) columns, then the row corresponds to the equation
\[
0x_1 + 0x_2 + \cdots + 0x_{k-1} + x_k + (\ast)x_{k+1} + \cdots + (\ast)x_n = b
\]
which has infinitely many solutions. Therefore, the system has infinitely many solutions. See Examples 9 And 10.

\textbf{Case 3:} If the last nonzero row has the form
\[
(0 \ 0 \ 0 \ \cdots \ 0 \ 1 \ | \ b),
\]
then the row corresponds to the equation
\[0x_1 + 0x_2 + 0x_3 + \cdots + 0x_{n-1} + x_n = b\]

which has the unique solution \(x_n = b\). The system itself either will have a unique solution or infinitely many solutions depending upon the solutions determined by the rows above.

Note that in Case 1, the rank of the augmented matrix is greater (by 1) than the rank of the matrix of coefficients. In Cases 2 and 3, the rank of the augmented matrix equals the rank of the matrix of coefficients. Thus, we have:

1. If “rank of the augmented matrix ≠ rank of matrix of coefficients,” the system has no solutions; the system is inconsistent.
2. If “rank of the augmented matrix = rank of matrix of coefficients,” the system either has a unique solution or infinitely many solutions; the system is consistent.

Exercises 1.3

Each of the following matrices is the row echelon form of the augmented matrix of a system of linear equations. Give the ranks of the matrix of coefficients and the augmented matrix, and find all solutions of the system.

1. \[\begin{pmatrix} 1 & -2 & 0 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & -1 \end{pmatrix}\]

2. \[\begin{pmatrix} 1 & -2 & 1 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}\]

3. \[\begin{pmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}\]

4. \[\begin{pmatrix} 1 & -2 & 0 & -1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}\]

5. \[\begin{pmatrix} 1 & 0 & 1 & -1 & 2 \\ 0 & 1 & 0 & 2 & -1 \\ 0 & 0 & 1 & -1 & 3 \end{pmatrix}\]
6. \[
\begin{pmatrix}
1 & -2 & 1 & -1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
2 \\
-1 \\
5
\end{pmatrix}.
\]

7. \[
\begin{pmatrix}
1 & -2 & 0 & 3 & 2 \\
0 & 0 & 1 & 2 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
2 \\
-1 \\
-3
\end{pmatrix}.
\]

8. \[
\begin{pmatrix}
1 & 0 & 2 & 0 & 3 \\
0 & 1 & 5 & 0 & 4 \\
0 & 0 & 0 & 1 & 9
\end{pmatrix}
\begin{pmatrix}
6 \\
7 \\
-3
\end{pmatrix}.
\]

Solve the following systems of equations.

9. \[
\begin{align*}
x - 2y &= -8 \\
2x - 3y &= -11
\end{align*}
\]

10. \[
\begin{align*}
x + z &= 3 \\
2y - 2z &= -4 \\ y - 2z &= 5
\end{align*}
\]

11. \[
\begin{align*}
x + z &= 3 \\
2x + 5y - 8z &= 4 \\ 3x + 8y - 13z &= 7
\end{align*}
\]

12. \[
\begin{align*}
x + z &= 6 \\
x + 2y + 4z &= 9 \\
2x + y + 6z &= 11
\end{align*}
\]

13. \[
\begin{align*}
x + y + z &= 6 \\
3x - y + 2z &= 7 \\ 5x + 3y - 4z &= 2
\end{align*}
\]

14. \[
\begin{align*}
x + y + z &= 10 \\
4x + 2y + 3z &= 14
\end{align*}
\]

15. \[
\begin{align*}
x_1 + 2x_2 - 3x_3 - 4x_4 &= 2 \\
2x_1 + 4x_2 - 5x_3 - 7x_4 &= 7 \\ -3x_1 - 6x_2 + 11x_3 + 14x_4 &= 0
\end{align*}
\]

16. \[
\begin{align*}
x_1 + 3x_2 - x_3 - 4x_4 &= -12 \\
x_1 - x_2 + x_3 + 2x_4 &= 8 \\
2x_1 + 3x_2 &= 8
\end{align*}
\]
17. \[
x_1 + 2x_2 + 2x_3 + 5x_4 = 11
2x_1 + 4x_2 + 2x_3 + 8x_4 = 14
x_1 + 3x_2 + 4x_3 + 8x_4 = 19
x_1 - x_2 + x_3 = 2
\]
\[
x_1 + 2x_2 - 3x_3 + 4x_4 = 2
\]
18. \[
2x_1 + 5x_2 - 2x_3 + x_4 = 1
5x_1 + 12x_2 - 7x_3 + 6x_4 = 7
x_1 + 2x_2 - x_3 - x_4 = 0
\]
19. \[
x_1 + 2x_2 + x_4 = 4
-x_1 - 2x_2 + 2x_3 + 4x_4 = 5
\]
\[
2x_1 - 4x_2 + 16x_3 - 14x_4 = 10
-x_1 + 5x_2 - 17x_3 + 19x_4 = -2
x_1 - 3x_2 + 11x_3 - 11x_4 = 4
3x_1 - 4x_2 + 18x_3 - 13x_4 = 17
\]
20. \[
x_1 - x_2 + 2x_3 = 7
2x_1 - 2x_2 + 2x_3 - 4x_4 = 12
-x_1 + x_2 - x_3 + 2x_4 = -4
-3x_1 + x_2 - 8x_3 - 10x_4 = -29
\]
\[
2x_1 - 5x_2 + 3x_3 - 4x_4 + 2x_5 = 4
3x_1 - 7x_2 + 2x_3 - 5x_4 + 4x_5 = 9
5x_1 - 10x_2 - 5x_3 - 4x_4 + 7x_5 = 22
\]
23. Determine the values of $k$ so that the system of equations has: (i) a unique solution, (ii) no solutions, (iii) infinitely many solutions:
\[
x + y - z = 1
2x + 3y + kz = 3
x + ky + 3z = 2
\]
24. What condition must be placed on $a$, $b$, $c$ so that the system
\[
x + 2y - 3z = a
2x + 6y - 11z = b
x - 2y + 7z = c
\]
has at least one solution.

25. Consider two systems of linear equations having augmented matrices $(A | b_1)$ and $(A | b_2)$ where the matrix of coefficients of both systems is the same $3 \times 3$ matrix $A$. 

177
(a) Is it possible for \((A \mid b_1)\) to have a unique solution and \((A \mid b_2)\) to have infinitely many solutions? Explain.

(b) Is it possible for \((A \mid b_1)\) to have a unique solution and \((A \mid b_2)\) to have no solution? Explain.

(c) Is it possible for \((A \mid b_1)\) to have infinitely many solutions and \((A \mid b_2)\) to have no solutions? Explain.

26. Suppose that you wanted to solve the systems of equations

\[
\begin{align*}
x + 2y + 4z &= a \\
x + 3y + 2z &= b \\
2x + 3y + 11z &= c
\end{align*}
\]

for

\[
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix} = \begin{pmatrix}
-1 \\
2 \\
3
\end{pmatrix}, \quad \begin{pmatrix}
3 \\
2 \\
4
\end{pmatrix}, \quad \begin{pmatrix}
0 \\
-2 \\
1
\end{pmatrix}
\]

respectively. Show that you can solve all three systems simultaneously by row reducing the matrix

\[
\begin{pmatrix}
1 & 2 & 4 & -1 & 3 & 0 \\
1 & 3 & 2 & 2 & 2 & -2 \\
2 & 3 & 11 & 3 & 4 & 1
\end{pmatrix}
\]
1.4 Solving Systems of Equations, Part 2

Reduced Row-Echelon Form

There is an alternative to Gaussian elimination with back substitution which we'll illustrate here.

Let's look again at Example 2 of the previous section:

\[
\begin{align*}
  x + 2y - 5z &= -1 \\
  -3x - 9y + 21z &= 0 \\
  x + 6y - 12z &= 1
\end{align*}
\]

The augmented matrix is
\[
\begin{bmatrix}
  1 & 2 & -5 & | & -1 \\
  -3 & -9 & 21 & | & 0 \\
  1 & 6 & -12 & | & 1
\end{bmatrix}
\]

which row reduces to
\[
\begin{bmatrix}
  1 & 2 & -5 & | & -1 \\
  0 & 1 & -2 & | & 1 \\
  0 & 0 & 1 & | & -1
\end{bmatrix}
\]

Now, instead of writing down the corresponding system of equations and back substituting to find the solutions, suppose we continue with row operations, eliminating the nonzero entries in the row-echelon matrix, starting with the 1 in the 3,3 position:

\[
\begin{align*}
  \begin{bmatrix}
    1 & 2 & -5 & | & -1 \\
    0 & 1 & -2 & | & 1 \\
    0 & 0 & 1 & | & -1
  \end{bmatrix}
  & \xrightarrow{2R_3+R_2\rightarrow R_2} \xrightarrow{5R_3+R_1\rightarrow R_1} \\
  \begin{bmatrix}
    1 & 0 & 0 & | & -4 \\
    0 & 1 & 0 & | & -1 \\
    0 & 0 & 1 & | & -1
  \end{bmatrix}
\end{align*}
\]

The corresponding system, which is equivalent to the original system since we’ve used only row operations, is

\[
\begin{align*}
  x &= -4 \\
  y &= -1 \\
  z &= -1
\end{align*}
\]

and the solutions are obvious: \( x = -4, \ y = -1, \ z = -1. \)

We’ll re-work Example 5 of the preceding Section using this approach.
Example 1.

\[
\begin{align*}
x_1 - 2x_2 + x_3 - x_4 &= -2 \\
-2x_1 + 5x_2 - x_3 + 4x_4 &= 1 \\
3x_1 - 7x_2 + 4x_3 - 4x_4 &= -4
\end{align*}
\]

The augmented matrix

\[
\begin{pmatrix}
1 & -2 & 1 & -1 & -2 \\
-2 & 5 & -1 & 4 & 1 \\
3 & -7 & 4 & -4 & -4
\end{pmatrix}
\]

reduces to

\[
\begin{pmatrix}
1 & -2 & 1 & -1 & -2 \\
0 & 1 & 1 & 2 & -3 \\
0 & 0 & 1 & 1/2 & -1/2
\end{pmatrix}
\]

Again, instead of back substituting, we’ll continue with row operations to eliminate nonzero entries, beginning with the leading 1 in the third row.

\[
\begin{pmatrix}
1 & -2 & 1 & -1 & -2 \\
0 & 1 & 1 & 2 & -3 \\
0 & 0 & 1 & 1/2 & -1/2
\end{pmatrix} \rightarrow \begin{pmatrix}
1 & 0 & 0 & 3/2 & -13/2 \\
0 & 1 & 0 & 3/2 & -5/2 \\
0 & 0 & 1 & 1/2 & -1/2
\end{pmatrix}
\]

The corresponding system of equations is

\[
\begin{align*}
x_1 + \frac{3}{2} x_4 &= -\frac{13}{2} \\
x_2 + \frac{3}{2} x_4 &= -1 \\
x_3 + \frac{1}{2} x_4 &= -\frac{1}{2}
\end{align*}
\]

If we let \( x_4 = a \), \( a \) any real number, then the solution set is

\[
x_1 = -\frac{13}{2} - \frac{3}{2} a, \ x_2 = -\frac{5}{2} - \frac{3}{2} a, \ x_3 = -\frac{1}{2} - \frac{1}{2} a, \ x_4 = a.
\]

as we saw before. ■

The final matrices in the two examples just given are said to be in reduced row-echelon form. In general, a matrix is in **reduced row-echelon form** if

1. Rows consisting entirely of zeros are at the bottom of the matrix.
2. The first nonzero entry in a nonzero row is a 1. This is called the **leading 1**.
3. If row \( i \) and row \( i + 1 \) are nonzero rows, then the leading 1 in row \( i + 1 \) is to the right of the leading 1 in row \( i \).

4. A leading 1 is the only nonzero entry in its column.

Since the number of nonzero rows in the row-echelon and reduced row-echelon form is the same, we can say that the rank of a matrix \( A \) is the number of nonzero rows in its reduced row-echelon form.

**Example 2.** Let

\[
A = \begin{pmatrix} 1 & -1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -3 \\ 0 & 0 & 0 & -1 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 5 & 0 & 2 \\ 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 1 & 7 \end{pmatrix},
\]

\[
C = \begin{pmatrix} 1 & -1 & 0 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 1 & 7 & -5 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\]

Which of these matrices are in reduced row-echelon form?

**SOLUTION** \( A \) is not in reduced row echelon form, the first nonzero entry in row 3 is not a 1; \( B \) is in reduced row-echelon form; \( C \) is not in reduced row-echelon form, the leading 1 in row 3 is not the only nonzero entry in its column; \( D \) is in reduced row-echelon form.

The steps involved in solving a system of linear equations using this alternative method are:

1. Write down the augmented matrix for the system.
2. Use the elementary row operations to “reduce” the matrix to reduced row-echelon form.
3. Write down the system of equations corresponding to the reduced row-echelon form matrix.
4. Write down the solutions of the system.

**Example 3.** Solve the system of equations

\[
x_1 + 2x_2 + 4x_3 + x_4 - x_5 = 1
\]
\[
2x_1 + 4x_2 + 8x_3 + 3x_4 - 4x_5 = 2
\]
\[
x_1 + 3x_2 + 7x_3 + 3x_5 = -2
\]
SOLUTION  We’ll reduce the augmented matrix to reduced row-echelon form:

\[
\begin{pmatrix}
1 & 2 & 4 & 1 & -1 & | & 1 \\
2 & 4 & 8 & 3 & -4 & | & 2 \\
1 & 3 & 7 & 0 & 3 & | & -2 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 4 & 1 & -1 & | & 1 \\
0 & 0 & 0 & 1 & -2 & | & 0 \\
0 & 1 & 3 & -1 & 4 & | & -3 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 2 & 4 & | & 1 \\
0 & 1 & 3 & | & 2 \\
0 & 0 & 1 & | & -2 \\
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & 0 & -2 & 0 & -3 & | & 7 \\
0 & 1 & 3 & 0 & 2 & | & -3 \\
0 & 0 & 0 & 1 & -2 & | & 0 \\
\end{pmatrix}
\]

The system of equations corresponding to this matrix is

\[
\begin{align*}
x_1 - 2x_3 - 3x_5 &= 7 \\
x_2 + 3x_3 &= -1 \\
x_4 - 2x_5 &= 0 \\
\end{align*}
\]

Let \(x_3 = a, \ x_5 = b, \ a, \ b \) any real numbers. Then the solution set is

\[
x_1 = 2a + 3b + 7, \ x_2 = -3a - 2b - 3, \ x_3 = a, \ x_4 = 2b, \ x_5 = b. \quad \blacksquare
\]

Homogeneous Systems

As we have seen, a system of \(m\) linear equations in \(n\) unknowns has the form

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]

The system is \textit{nonhomogeneous} if at least one of \(b_1, b_2, b_3, \ldots, b_m\) is different from zero. The system is \textit{homogeneous} if \(b_1 = b_2 = b_3 = \cdots = b_m = 0\). Thus, a homogeneous system has the form

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= 0 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= 0 \\
\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= 0
\end{align*}
\]
Compare with homogeneous and nonhomogeneous linear differential equations — same concept, same terminology.

The significant fact about homogeneous systems is that they are always consistent; a homogeneous system always has at least one solution, namely \( x_1 = x_2 = x_3 = \cdots = x_n = 0 \). This solution is called the \textit{trivial solution}. So, the basic question for a homogeneous system is: Are there any nontrivial (i.e., nonzero) solutions?

Since homogeneous systems are simply a special case of general linear systems, our methods of solution still apply.

\textbf{Example 4.} Find the solution set of the homogeneous system

\[
\begin{align*}
x - 2y + 2z &= 0 \\
4x - 7y + 3z &= 0 \\
2x - y + 2z &= 0
\end{align*}
\]

\textit{SOLUTION} The augmented matrix for the system is

\[
\begin{pmatrix}
1 & -2 & 2 & 0 \\
4 & -7 & 3 & 0 \\
2 & -1 & 2 & 0
\end{pmatrix}
\]

We use row operations to reduce this matrix to row-echelon form.

\[
\begin{pmatrix}
1 & -2 & 2 & 0 \\
4 & -7 & 3 & 0 \\
2 & -1 & 2 & 0
\end{pmatrix} \rightarrow
\begin{pmatrix}
1 & -2 & 2 & 0 \\
0 & 1 & -5 & 0 \\
0 & 0 & 13 & 0
\end{pmatrix} \rightarrow
\begin{pmatrix}
1 & -2 & 2 & 0 \\
0 & 1 & -5 & 0 \\
0 & 0 & 1 & 0
\end{pmatrix}
\]

This is the augmented matrix for the system of equations

\[
\begin{align*}
x - 2y + 2z &= 0 \\
y - 5z &= 0 \\
z &= 0.
\end{align*}
\]

This system has the unique solution \( x = y = z = 0 \); the trivial solution is the only solution.

\textbf{Example 5.} Find the solution set of the homogeneous system

\[
\begin{align*}
3x - 2y + z &= 0 \\
x + 4y + 2z &= 0 \\
7x + 4z &= 0
\end{align*}
\]
**SOLUTION**  The augmented matrix for the system is
\[
\begin{pmatrix}
3 & -2 & 1 & 0 \\
1 & 4 & 2 & 0 \\
7 & 0 & 4 & 0 \\
\end{pmatrix}
\]
We use row operations to reduce this matrix to row-echelon form. Note that, since every entry in the last column of the augmented matrix is 0, we only need to reduce the matrix of coefficients. Confirm this by re-doing Example 1 using only the matrix of coefficients.

\[
\begin{pmatrix}
3 & -2 & 1 \\
1 & 4 & 2 \\
7 & 0 & 4 \\
\end{pmatrix}
\]

\[ R_1 \leftrightarrow R_2 \]

\[
\begin{pmatrix}
1 & 4 & 2 \\
0 & -14 & -5 \\
0 & -28 & -10 \\
\end{pmatrix}
\]

\[ ( -2 ) R_2 + R_3 \rightarrow R_3 \]

\[
\begin{pmatrix}
1 & 4 & 2 \\
0 & -14 & -5 \\
0 & 0 & 0 \\
\end{pmatrix}
\]

\[ ( -1 / 14 ) R_3 \rightarrow R_3 \]

This is the augmented matrix for the system of equations
\[
\begin{align*}
x + 4y + 2z &= 0 \\
y + \frac{5}{14}z &= 0 \\
z &= 0.
\end{align*}
\]
This system has infinitely many solutions:
\[
x = -\frac{2}{7}a, \quad y = -\frac{5}{14}a, \quad z = a, \quad \text{any real number}.
\]

Let’s look at homogeneous systems in general:
\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= 0 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= 0
\end{align*}
\]
We know that the system either has a unique solution — the trivial solution — or it has infinitely many nontrivial solutions, and this can be determined by reducing the augmented matrix
\[
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & 0 \\
a_{21} & a_{22} & a_{23} & \cdots & a_{2n} & 0 \\
a_{31} & a_{32} & a_{33} & \cdots & a_{3n} & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} & 0
\end{pmatrix}
\]
to row echelon form. We know, also, that the number \( k \) of nonzero rows (the rank of the matrix) cannot exceed the number of columns. Since the last column is all zeros, it follows that \( k \leq n \). Thus, there are two cases to consider.

Case 1: \( k = n \)  
In this case, the augmented matrix row reduces to

\[
\begin{pmatrix}
1 & * & * & \cdots & * & 0 \\
0 & 1 & * & \cdots & * & 0 \\
0 & 0 & 1 & \cdots & * & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0
\end{pmatrix}
\]

(we are disregarding rows of zeros at the bottom, which will occur if \( m > n \).) The only solution to the corresponding system of equations is the trivial solution.

Case 2: \( k < n \)  
In this case, the augmented matrix row reduces to

\[
\begin{pmatrix}
1 & * & * & \cdots & * & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & \cdots & 0
\end{pmatrix}
\]

Here there are \( k \) rows and the leading 1 is in the \( j^{th} \) column. Again, we have disregarded the zero rows at the bottom. In this case there are infinitely many solutions.

Therefore, if the number of nonzero rows in the row echelon form of the augmented matrix is \( n \), then the trivial solution is the only solution. If the number of nonzero rows is less than \( n \), then there are infinitely many nontrivial solutions.

There is a very important consequence of this result. A homogeneous system with more unknowns than equations always has infinitely many nontrivial solutions.

**Nonhomogeneous Systems and Associated Homogeneous Systems**  
In Example 5 of the previous section, we solved the (nonhomogeneous) system

\[
\begin{align*}
x_1 - 2x_2 + x_3 - x_4 &= -2 \\
-2x_1 + 5x_2 - x_3 + 4x_4 &= 1 \\
3x_1 - 7x_2 + 4x_3 - 4x_4 &= -4
\end{align*}
\]

We found that the system has infinitely many solutions:

\[
x_1 = -\frac{13}{2} - \frac{3}{2}a, \ x_2 = -\frac{5}{2} - \frac{3}{2}a, \ x_3 = -\frac{1}{2} - \frac{1}{2}a, \ x_4 = a, \ \text{a any real number.} \quad (*)
\]
Let’s solve the corresponding homogeneous system
\[
\begin{align*}
    x_1 - 2x_2 + x_3 - x_4 &= 0 \\
    -2x_1 + 5x_2 - x_3 + 4x_4 &= 0 \\
    3x_1 - 7x_2 + 4x_3 - 4x_4 &= 0
\end{align*}
\]
Using row operations to reduce the coefficient matrix to row-echelon form, we have
\[
\begin{pmatrix}
    1 & -2 & 1 & 1 \\
    -2 & 5 & -1 & 4 \\
    3 & -7 & 4 & -4
\end{pmatrix} \rightarrow
\begin{pmatrix}
    1 & -2 & 1 & 1 \\
    0 & 1 & 1 & 2 \\
    0 & 0 & 1 & -1/2
\end{pmatrix}
\]
As you can verify, the solution set here is:
\[
x_1 = -\frac{3}{2}a, \ x_2 = -\frac{3}{2}a, \ x_3 = -\frac{1}{2}a, \ x_4 = a, \ \text{a any real number.}
\]
Now we want to compare the solutions of the nonhomogeneous system with the solutions of the corresponding homogeneous system. Writing the solutions of the nonhomogeneous system as an ordered quadruple, we have
\[
(x_1, x_2, x_3, x_4) = \left( \begin{array}{c}
    \frac{-13}{2} - \frac{3}{2}a, \\
    -\frac{5}{2} - \frac{3}{2}a, \\
    -\frac{1}{2} - \frac{1}{2}a, \\
    a
\end{array} \right)
\]
Note that \( \left( -\frac{13}{2}, -\frac{5}{2}, -\frac{1}{2}, 0 \right) \) is a solution of the nonhomogeneous system [set \( a = 0 \) in (*)] and \( \left( -\frac{3}{2}a, -\frac{3}{2}a, -\frac{1}{2}a, a \right) \) is the set of all solutions of the corresponding homogeneous system.
This example illustrates a general fact:

The set of solutions of the nonhomogeneous system (N) can be represented as the sum of one solution of (N) and the set of all solutions of the corresponding homogeneous system (H).

This is exactly the same result that we obtained for second order linear differential equations.
Exercises 1.4

Determine whether the matrix is in reduced row echelon form. If it is not, give reasons why not.

1. \[
\begin{pmatrix}
1 & 3 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{pmatrix}
\]

2. \[
\begin{pmatrix}
1 & 3 & 0 & -1 \\
0 & 0 & 2 & 4
\end{pmatrix}
\]

3. \[
\begin{pmatrix}
1 & 0 & 3 & -2 \\
0 & 0 & 1 & 4 \\
0 & 1 & 2 & 6
\end{pmatrix}
\]

4. \[
\begin{pmatrix}
1 & 3 & 0 & 0 & 5 \\
0 & 0 & 1 & 0 & -8 \\
0 & 0 & 0 & 1 & -5
\end{pmatrix}
\]

5. \[
\begin{pmatrix}
1 & 3 & 0 & 0 & -2 & 5 \\
0 & 0 & 1 & 0 & 4 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

6. \[
\begin{pmatrix}
1 & 0 & 3 & 2 & 0 \\
0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

7. \[
\begin{pmatrix}
1 & 0 & 6 & 0 & 0 \\
0 & 1 & -3 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix}
\]

8. \[
\begin{pmatrix}
1 & 0 & -2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 1 & 2 & 0 & -1 \\
0 & 0 & 0 & 1 & 5
\end{pmatrix}
\]

Solve the system by reducing the augmented matrix to its reduced row echelon form.

\[x + z = 3\]

9. \[2y - 2z = -4\]

\[y - 2z = 5\]
10. \[ x + y + z = 6 \]
    \[ x + 2y + 4z = 9 \]
    \[ 2x + y + 6z = 11 \]
    \[ x + 2y - 3z = 1 \]

11. \[ 2x + 5y - 8z = 4 \]
    \[ 3x + 8y - 13z = 7 \]
    \[ x + 2y - 2z = -1 \]

12. \[ 3x - y + 2z = 7 \]
    \[ 5x + 3y - 4z = 2 \]
    \[ x_1 + 2x_2 - 3x_3 - 4x_4 = 2 \]

13. \[ 2x_1 + 4x_2 - 5x_3 - 7x_4 = 7 \]
    \[ -3x_1 - 6x_2 + 11x_3 + 14x_4 = 0 \]
    \[ x_1 + 2x_2 + 2x_3 + 5x_4 = 11 \]

14. \[ 2x_1 + 4x_2 + 2x_3 + 8x_4 = 14 \]
    \[ x_1 + 3x_2 + 4x_3 + 8x_4 = 19 \]
    \[ x_1 - x_2 + x_3 = 2 \]
    \[ 2x_1 - 4x_2 + 16x_3 - 14x_4 = 10 \]

15. \[ -x_1 + 5x_2 - 17x_3 + 19x_4 = -2 \]
    \[ x_1 - 3x_2 + 11x_3 - 11x_4 = 4 \]
    \[ 3x_1 - 4x_2 + 18x_3 - 13x_4 = 17 \]
    \[ x_1 - x_2 + 2x_3 = 7 \]

16. \[ 2x_1 - 2x_2 + 2x_3 - 4x_4 = 12 \]
    \[ -x_1 + x_2 - x_3 + 2x_4 = -4 \]
    \[ -3x_1 + x_2 - 8x_3 - 10x_4 = -29 \]

Solve the homogeneous system.

17. \[ x - 3y = 0 \]
    \[ -4x + 6y = 0 \]
    \[ 6x - 9y = 0 \]

18. \[ 3x - y + z = 0 \]
    \[ x - y - z = 0 \]
    \[ x + y + z = 0 \]
    \[ x - y - 3z = 0 \]

19. \[ x + y + z = 0 \]
    \[ 2x + 2y + z = 0 \]
    \[ 3x - 3y - z = 0 \]

20. \[ 2x - y - z = 0 \]

188
21. \[3x_1 + x_2 - 5x_3 - x_4 = 0\]
\[2x_1 + x_2 - 3x_3 - 2x_4 = 0\]
\[x_1 + x_2 - x_3 - 3x_4 = 0\]
\[2x_1 - 2x_2 - x_3 + x_4 = 0\]

22. \[-x_1 + x_2 + x_3 - 2x_4 = 0\]
\[3x_1 - 3x_2 + x_3 - 6x_4 = 0\]
\[2x_1 - 2x_2 - 2x_4 = 0\]
\[3x_1 + 6x_2 - 3x_4 = 0\]

23. \[x_1 + 3x_2 - x_3 - 4x_4 = 0\]
\[x_1 - x_2 + x_3 + 2x_4 = 0\]
\[2x_1 + 3x_2 = 0\]
\[x_1 - 2x_2 + x_3 - x_4 + 2x_5 = 0\]

24. \[2x_1 - 4x_2 + 2x_3 - x_4 + x_5 = 0\]
\[x_1 - 2x_2 + x_3 + 2x_4 - 7x_5 = 0\]

25. If a homogeneous system has more equations than unknowns, is it possible for the system to have nontrivial solutions? Justify your answer.

26. For what values of \(a\) does the system
\[x + ay = 0\]
\[-3x + 2y = 0\]

have nontrivial solutions?

27. For what values of \(a\) and \(b\) does the system
\[x - 2y = a\]
\[-3x + 6y = b\]

have a solution?

28. For what values of \(a\) and \(b\) does the system
\[-x - 2z = a\]
\[2x + y + x = 0\]
\[x + y - z = b\]

have a solution?

29. For what values of \(a\) does the system
\[x + ay - 2z = 0\]
\[2x - y - z = 0\]
\[-x - y + z = 0\]
have nontrivial solutions?

30. Given the system of equations

\[
\begin{align*}
x_1 - 3x_2 - 2x_3 + 4x_4 &= 5 \\
3x_1 - 8x_2 - 3x_3 + 8x_4 &= 18 \\
2x_1 - 3x_2 + 5x_3 - 4x_4 &= 19
\end{align*}
\]

(a) Find the set of all solutions of the system.

(b) Show that the result in (a) can be written in the form “one solution of the given system + the set of all solutions of the corresponding homogeneous system.”
1.5 Matrices and Vectors

In the preceding section we introduced the concept of a matrix and we used matrices (augmented matrices) as a shorthand notation for systems of linear equations. In this and the following sections, we will develop this concept further.

Recall that a matrix is a rectangular array of objects arranged in rows and columns. The objects are called the entries of the matrix. A matrix with $m$ rows and $n$ columns is called an $m \times n$ matrix. The expression $m \times n$ is called the size of the matrix, and the numbers $m$ and $n$ are its dimensions. A matrix in which the number of rows equals the number of columns, $m = n$, is called a square matrix of order $n$.

Until further notice, we will be considering matrices whose entries are real numbers. We will use capital letters to denote matrices and lower case letters to denote its entries. If $A$ is an $m \times n$ matrix, then $a_{ij}$ denotes the element in the $i$th row and $j$th column of $A$:

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

We will also use the notation $A = (a_{ij})$ to represent this display.

There are two important special cases that we need to define. A $1 \times n$ matrix

$$(a_1 \ a_2 \ \ldots \ a_n) \quad \text{also written as} \quad (a_1, \ a_2, \ \ldots, \ a_n)$$

(since there is only one row, we use only one subscript) is called an $n$-dimensional row vector. An $m \times 1$ matrix

$$\begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_m
\end{pmatrix}$$

is called an $m$-dimensional column vector. The entries of a row or column vector are often called the components of the vector. On occasion we'll regard an $m \times n$ matrix as a set of $m$ row vectors, or as a set of $n$ column vectors. To distinguish vectors from matrices in general, we'll use lower case boldface letters to denote vectors. Thus we'll write

$$\mathbf{u} = (a_1 \ a_2 \ \ldots \ a_n) \quad \text{and} \quad \mathbf{v} = \begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_m
\end{pmatrix}$$
Arithmetic of Matrices

Before we can define arithmetic operations for matrices (addition, subtraction, etc.), we have to define what it means for two matrices to be equal.

**DEFINITION** (Equality for Matrices) Let \( A = (a_{ij}) \) be an \( m \times n \) matrix and let \( B = (b_{ij}) \) be a \( p \times q \) matrix. Then \( A = B \) if and only if

1. \( m = p \) and \( n = q \); that is, \( A \) and \( B \) must have exactly the same size; and
2. \( a_{ij} = b_{ij} \) for all \( i \) and \( j \).

In short, \( A = B \) if and only if \( A \) and \( B \) are identical.

**Example 1.**

\[
\begin{pmatrix}
a & b & 3 \\
2 & c & 0
\end{pmatrix}
= \begin{pmatrix}
7 & -1 & x \\
2 & 4 & 0
\end{pmatrix}
\text{ if and only if } a = 7, b = -1, c = 4, x = 3.
\]

**Matrix Addition**

Let \( A = (a_{ij}) \) and \( B = (b_{ij}) \) be matrices of the same size, say \( m \times n \). Then \( A + B \) is the \( m \times n \) matrix \( C = (c_{ij}) \) where \( c_{ij} = a_{ij} + b_{ij} \) for all \( i \) and \( j \). That is, you add two matrices of the same size simply by adding their corresponding entries; \( A + B = (a_{ij} + b_{ij}) \).

*Addition of matrices is not defined for matrices of different sizes.*

**Example 2.**

(a) \[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix} + \begin{pmatrix}
x & y \\
z & w
\end{pmatrix} = \begin{pmatrix}
a + x & b + y \\
c + z & d + w
\end{pmatrix}
\]

(b) \[
\begin{pmatrix}
2 & 4 & -3 \\
2 & 5 & 0
\end{pmatrix} + \begin{pmatrix}
-4 & 0 & 6 \\
-1 & 2 & 0
\end{pmatrix} = \begin{pmatrix}
-2 & 4 & 3 \\
1 & 7 & 0
\end{pmatrix}
\]

(c) \[
\begin{pmatrix}
2 & 4 & -3 \\
2 & 5 & 0
\end{pmatrix} + \begin{pmatrix}
1 & 3 \\
5 & -3 \\
0 & 6
\end{pmatrix}
\text{ is not defined.}
\]

Since we add two matrices simply by adding their corresponding entries, it follows from the properties of the real numbers that matrix addition is commutative and associative. That is, if \( A, B, \) and \( C \) are matrices of the same size, then

\[
A + B = B + A \quad \text{Commutative}
\]

\[
(A + B) + C = A + (B + C) \quad \text{Associative}
\]
A matrix with all entries equal to 0 is called a zero matrix. For example,
\[
\begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
0 & 0 \\
0 & 0 \\
0 & 0
\end{pmatrix}
\]
are zero matrices. Often the symbol 0 will be used to denote the zero matrix of arbitrary size. The zero matrix acts like the number zero in the sense that if \( A \) is an \( m \times n \) matrix and 0 is the \( m \times n \) zero matrix, then
\[A + 0 = 0 + A = A.\]
A zero matrix is an additive identity.

The negative of a matrix \( A \), denoted by \(-A\) is the matrix whose entries are the negatives of the entries of \( A \). For example, if
\[
A = \begin{pmatrix}
a & b & c \\
d & e & f
\end{pmatrix}
\]
then \(-A = \begin{pmatrix}
-a & -b & -c \\
-d & -e & -f
\end{pmatrix}\)

Subtraction Let \( A = (a_{ij}) \) and \( B = (b_{ij}) \) be matrices of the same size, say \( m \times n \). Then
\[
A - B = A + (-B).
\]
To put this simply, \( A - B \) is the \( m \times n \) matrix \( C = (c_{ij}) \) where \( c_{ij} = a_{ij} - b_{ij} \) for all \( i \) and \( j \); you subtract two matrices of the same size by subtracting their corresponding entries; \( A - B = (a_{ij} - b_{ij}) \). Note that \( A - A = 0 \).

Subtraction of matrices is not defined for matrices of different sizes.

Example 3.
\[
\begin{pmatrix}
2 & -4 & -3 \\
2 & 5 & 0
\end{pmatrix}
- \begin{pmatrix}
-4 & 0 & 6 \\
-1 & 2 & 0
\end{pmatrix}
= \begin{pmatrix}
6 & 4 & -9 \\
3 & 3 & 0
\end{pmatrix}.
\]

Multiplication of a Matrix by a Number

The product of a number \( k \) and a matrix \( A \), denoted by \( kA \), is the matrix formed by multiplying each element of \( A \) by the number \( k \). That is, \( kA = (ka_{ij}) \). This product is also called multiplication by a scalar.

Example 4.
\[
-3 \begin{pmatrix}
2 & -1 & 4 \\
1 & 5 & -2 \\
4 & 0 & 3
\end{pmatrix}
= \begin{pmatrix}
-6 & 3 & -12 \\
-3 & -15 & 6 \\
-12 & 0 & -9
\end{pmatrix}.
\]
Matrix Multiplication

We now want to define matrix multiplication. While addition and subtraction of matrices is a natural extension of addition and subtraction of real numbers, matrix multiplication is a much different sort of product.

We’ll start by defining the product of an $n$-dimensional row vector and an $n$-dimensional column vector, in that specific order — the row on the left and the column on the right.

**The Product of a Row Vector and a Column Vector** The product of a $1 \times n$ row vector and an $n \times 1$ column vector is the number given by

\[
\left( a_1, a_2, a_3, \ldots, a_n \right) \left( \begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array} \right) = a_1b_1 + a_2b_2 + a_3b_3 + \cdots + a_nb_n.
\]

This product has several names, including *scalar product* (because the result is a number (scalar)), *dot product*, and *inner product*. It is important to understand and remember that the product of a row vector and a column vector (of the same dimension and in that order!) is a *number*.

*The product of a row vector and a column vector of different dimensions is not defined.*

**Example 5.**

\[
(3, -2, 5) \left( \begin{array}{c} -1 \\ -4 \\ 1 \end{array} \right) = 3(-1) + (-2)(-4) + 5(1) = 10
\]

\[
(-2, 3, -1, 4) \left( \begin{array}{cccc} 2 & \cdots & b_{1j} & \cdots & b_{1n} \\ 4 & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{array} \right) = \left( \begin{array}{c} c_{ij} \end{array} \right)
\]

**Matrix Multiplication** If $A = (a_{ij})$ is an $m \times p$ matrix and $B = (b_{ij})$ is a $p \times n$ matrix, then the matrix product of $A$ and $B$ (in that order), denoted $AB$, is the $m \times n$ matrix $C = (c_{ij})$ where $c_{ij}$ is the product of the $i^{th}$ row of $A$ and the $j^{th}$ column of $B$. 

\[
\left( \begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{array} \right) \left( \begin{array}{cccc} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{array} \right) = \left( \begin{array}{c} c_{ij} \end{array} \right)
\]
where

\[ c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj}. \]

**NOTE:** Let \( A \) and \( B \) be given matrices. The product \( AB \), in that order, is defined if and only if the number of columns of \( A \) equals the number of rows of \( B \). If the product \( AB \) is defined, then the size of the product is: \((\text{no. of rows of } A) \times (\text{no. of columns of } B)\). That is

\[ A_{m \times p} B_{p \times n} = C_{m \times n} \]

**Example 6.** Let

\[ A = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & -2 \end{pmatrix} \]

Since \( A \) is \( 2 \times 3 \) and \( B \) is \( 3 \times 2 \), we can calculate the product \( AB \), which will be a \( 2 \times 2 \) matrix.

\[
AB = \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 \cdot 3 + 4 \cdot (-1) + 2 \cdot 1 & 1 \cdot 0 + 4 \cdot 2 + 2 \cdot (-2) \\ 3 \cdot 3 + 1 \cdot (-1) + 5 \cdot 1 & 3 \cdot 0 + 1 \cdot 2 + 5 \cdot (-2) \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 13 & -8 \end{pmatrix}
\]

We can also calculate the product \( BA \) since \( B \) is \( 3 \times 2 \) and \( A \) is \( 2 \times 3 \). This product will be \( 3 \times 3 \).

\[
BA = \begin{pmatrix} 3 & 0 \\ -1 & 2 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 4 & 2 \\ 3 & 1 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 12 & 6 \\ 5 & -2 & 6 \\ -5 & 2 & -6 \end{pmatrix}.
\]

**Example 6** illustrates a significant fact about matrix multiplication. **Matrix multiplication is not commutative;** \( AB \neq BA \) in general. While matrix addition and subtraction mimic addition and subtraction of real numbers, matrix multiplication is distinctly different.

**Example 7.** (a) In Example 6 we calculated both \( AB \) and \( BA \), but they were not equal because the products were of different size. Consider the matrices

\[ C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} -1 & 0 & 3 \\ 5 & 7 & 2 \end{pmatrix} \]

Here the product \( CD \) exists and equals \( \begin{pmatrix} 9 & 14 & 7 \\ 17 & 28 & 17 \end{pmatrix} \). You should verify this.

On the other hand, the product \( DC \) does not exist; you cannot multiply \( \begin{pmatrix} 2 & 3 \end{pmatrix} \) \( C \).
(b) Consider the matrices

\[ E = \begin{pmatrix} 4 & -1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 9 & -5 \\ 3 & 0 \end{pmatrix}. \]

In this case \( EF \) and \( FE \) both exist and have the same size, \( 2 \times 2 \). But

\[ EF = \begin{pmatrix} 33 & -20 \\ 6 & 0 \end{pmatrix} \quad \text{and} \quad FE = \begin{pmatrix} 36 & -19 \\ 12 & -3 \end{pmatrix} \quad \text{(verify)} \]

so \( EF \neq FE \).

While matrix multiplication is not commutative, it is associative. Let \( A \) be an \( m \times p \) matrix, \( B \) a \( p \times q \) matrix, and \( C \) a \( q \times n \) matrix. Then

\[ A(BC) = (AB)C. \]

By indicating the dimensions of the products we can see that the left and right sides do have the same size, namely \( m \times n \):

\[ A_{m \times p} (BC)_{p \times n} = (AB)_{m \times q} C_{q \times n}. \]

A straightforward, but tedious, manipulation of double subscripts and sums shows that the left and right sides are, in fact, equal.

There is another significant departure from the real number system. If \( a \) and \( b \) are real numbers and \( ab = 0 \), then either \( a = 0 \), \( b = 0 \), or they are both \( 0 \).

**Example 8.** Consider the matrices

\[ A = \begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 5 \\ 0 & 0 \end{pmatrix}. \]

Neither \( A \) nor \( B \) is the zero matrix yet, as you can verify,

\[ AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \]

Also,

\[ BA = \begin{pmatrix} 0 & 7 \\ 0 & 0 \end{pmatrix}. \quad \blacksquare \]

As we saw above, zero matrices act like the number zero with respect to matrix addition,

\[ A + 0 = 0 + A = A. \]

Are there matrices which act like the number 1 with respect to multiplication \((a \cdot 1 = 1 \cdot a = a \) for any real number \( a \))? Since matrix multiplication is not commutative, we cannot expect to find a matrix \( I \) such that \( AI = IA = A \) for an arbitrary \( m \times n \) matrix \( A \). However there are matrices which do act like the number 1 with respect to multiplication
Example 9. Consider the $2 \times 3$ matrix
\[
A = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}
\]
Let $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Then
\[
I_2 A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}
\]
and
\[
AI_3 = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.
\]

Identity Matrices Let $A$ be a square matrix of order $n$. The entries from the upper left corner of $A$ to the lower right corner; that is, the entries $a_{11}$, $a_{22}$, $a_{33}$, . . . , $a_{nn}$ form the main diagonal of $A$.

For each positive integer $n > 1$, let $I_n$ denote the square matrix of order $n$ whose entries on the main diagonal are all 1, and all other entries are 0. The matrix $I_n$ is called the $n \times n$ identity matrix. In particular,
\[
I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
\]
and so on.

If $A$ is an $m \times n$ matrix, then
\[
I_m A = A \quad \text{and} \quad AI_n = A.
\]
In particular, if $A$ is a square matrix of order $n$, then
\[
AI_n = I_n A = A,
\]
so $I_n$ does act like the number 1 for the set of square matrices of order $n$.

Matrix addition, multiplication of a matrix by a number, and matrix multiplication are connected by the following properties. In each case, assume that the sums and products are defined.

1. $A(B + C) = AB + AC$ \text{ This is called the left distributive law.}
2. \((A + B)C = AC + BC\) This is called the right distributive law.

3. \(k(AB) = (kA)B = A(kB)\)

Another way to look at systems of linear equations.

Now that we have defined matrix multiplication, we have another way to represent the system of \(m\) linear equations in \(n\) unknowns

\[
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n = b_3 \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n = b_m
\]

We can write this system as

\[
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
\vdots \\
x_n
\end{pmatrix}
=
\begin{pmatrix}
b_1 \\
b_2 \\
b_3 \\
\vdots \\
b_m
\end{pmatrix}
\]

or, more formally, in the vector-matrix form

\[
Ax = b
\]  

(1)

where \(A\) is the matrix of coefficients, \(x\) is the column vector of unknowns, and \(b\) is the column vector whose components are the numbers on the right-hand sides of the equations. This “looks like” the linear equation in one unknown \(ax = b\) where \(a\) and \(b\) are real numbers. By writing the system in this form we are tempted to “divide” both sides of equation (1) by \(A\), provided \(A\) “is not zero.” We’ll take these ideas up in the next section.

The set of vectors with \(n\) components, either row vectors or column vectors, is said to be a vector space of dimension \(n\), and it is denoted by \(\mathbb{R}^n\). Thus, \(\mathbb{R}^2\) is the vector space of vectors with two components, \(\mathbb{R}^3\) is the vector space of vectors with three components, and so on. The \(xy\)-plane can be used to give a geometric representation of \(\mathbb{R}^2\), 3-dimensional space gives a geometric representation of \(\mathbb{R}^3\).

Suppose that \(A\) is an \(m \times n\) matrix. If \(u\) is an \(n\)-component vector (an element in \(\mathbb{R}^n\)), then

\[
Au = v
\]
is an \( m \)-component vector (an element of \( \mathbb{R}^m \)). Thus, the matrix \( A \) can be viewed as a transformation (a function, a mapping) that takes an element in \( \mathbb{R}^n \) to an element in \( \mathbb{R}^m \); \( A \) “maps” \( \mathbb{R}^n \) into \( \mathbb{R}^m \).

**Example 10.** The \( 2 \times 3 \) matrix
\[
A = \begin{pmatrix}
1 & -2 & 4 \\
-1 & 0 & 3
\end{pmatrix}
\]
maps \( \mathbb{R}^3 \) into \( \mathbb{R}^2 \). For example
\[
\begin{pmatrix}
1 & -2 & 4 \\
-1 & 0 & 3
\end{pmatrix}
\begin{pmatrix}
2 \\
-4 \\
-1
\end{pmatrix} = \begin{pmatrix}
6 \\
-5
\end{pmatrix},
\begin{pmatrix}
1 & -2 & 4 \\
-1 & 0 & 3
\end{pmatrix}
\begin{pmatrix}
-3 \\
7 \\
2
\end{pmatrix} = \begin{pmatrix}
-9 \\
9
\end{pmatrix}
\]
and, in general,
\[
\begin{pmatrix}
1 & -2 & 4 \\
-1 & 0 & 3
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix} = \begin{pmatrix}
a - 2b + 4c \\
-a + 3c
\end{pmatrix}.
\]

If we view the \( m \times n \) matrix \( A \) as a transformation from \( \mathbb{R}^n \) to \( \mathbb{R}^m \), then the question of solving the equation
\[
Ax = b
\]
can be stated as: Find the set of vectors \( x \) in \( \mathbb{R}^n \) which are transformed to the given vector \( b \) in \( \mathbb{R}^m \).

As a final remark, an \( m \times n \) matrix \( A \) regarded as a transformation, is a **linear transformation** since
\[
A(x + y) = Ax + Ay \quad \text{and} \quad A(kx) = kAx.
\]
These two properties follow from the connections between matrix addition, multiplication by a number, and matrix multiplication given above.

**Exercises 1.5**

1. Let
\[
A = \begin{pmatrix}
2 & 4 \\
-1 & 3 \\
4 & -2
\end{pmatrix}, \quad B = \begin{pmatrix}
-2 & 0 \\
4 & 2 \\
-3 & 1
\end{pmatrix}, \quad C = \begin{pmatrix}
1 & -2 \\
-2 & 4
\end{pmatrix}, \quad D = \begin{pmatrix}
7 & -2 \\
3 & 0
\end{pmatrix}.
\]

Compute the following (if possible).

(a) \( A + B \) \quad (b) \( -2B \) \quad (c) \( C + D \)

(d) \( A + D \) \quad (e) \( 2A + B \) \quad (f) \( A - B \)

2. Let
\[
A = \begin{pmatrix}
-1 \\
2 \\
4
\end{pmatrix}, \quad B = \begin{pmatrix}
0 & -2 & 3 \\
3 & -2 & 0 \\
2 & 3 & 1
\end{pmatrix}, \quad C = \begin{pmatrix}
-2, & 4, & -3
\end{pmatrix}, \quad D = \begin{pmatrix}
-6 & 5 \\
3 & 0 \\
-4 & -4
\end{pmatrix}.
\]

Compute the following (if possible).
3. Let \( A = \begin{pmatrix} 0 & -1 \\ 0 & 2 \\ 4 & -2 \end{pmatrix} \), \( B = \begin{pmatrix} -1 & 0 \\ 3 & -2 \\ 2 & 5 \end{pmatrix} \), \( C = \begin{pmatrix} -2 & 0 \\ 4 & -3 \end{pmatrix} \), \( D = \begin{pmatrix} 5 & 0 \\ 3 & -3 \end{pmatrix} \).

Compute the following (if possible).

(a) \( 3A - 2BC \)  
(b) \( AB \)  
(c) \( CA \)  
(d) \( CA \)  
(e) \( DA \)  
(f) \( DB \)  
(g) \( AC \)  
(h) \( B^2(= BB) \)
8. Let \( A = \begin{pmatrix} 1 & -3 & 1 \\ -2 & 4 & 0 \\ 3 & -1 & -4 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 1 & -3 \\ -1 & 3 & 2 \end{pmatrix}, \) and \( C = \begin{pmatrix} 2 & 0 & -2 \\ 4 & 5 & -1 \\ 1 & 0 & -2 \end{pmatrix}. \)

Put \( D = 2B - 3C. \)

Determine the following elements of \( D \) without calculating the entire matrices.

(a) \( d_{11} \)  
(b) \( d_{23} \)  
(c) \( d_{32} \)

9. Let \( A \) be a matrix whose second row is all zeros. Let \( B \) be a matrix such that the product exists \( AB \) exists. Prove that the second row of \( AB \) is all zeros.

10. Let \( B \) be a matrix whose third column is all zeros. Let \( A \) be a matrix such that the product exists \( AB \) exists. Prove that the third column of \( AB \) is all zeros.

11. Let \( A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 & 5 & 4 \\ 2 & 1 & 3 \end{pmatrix}, \ C = \begin{pmatrix} 2 & 3 \\ 6 & 1 \end{pmatrix}, \) and \( D = \begin{pmatrix} 2 & -2 \\ 1 & 3 \end{pmatrix}. \)

Calculate (if possible).

(a) \( AB \) and \( BA \)  
(b) \( AC \) and \( CA \)  
(c) \( AD \) and \( DA. \)

This illustrates all the possibilities when order is reversed in matrix multiplication.

12. Let \( A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \ B = \begin{pmatrix} 2 & 4 & 0 \\ -2 & 1 & 3 \end{pmatrix}, \ C = \begin{pmatrix} 2 & 1 \\ 3 & 4 \\ -1 & 2 \end{pmatrix}. \)

Calculate \( A(BC) \) and \( (AB)C. \) This illustrates the associative property of matrix multiplication.

13. Let \( A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ B = \begin{pmatrix} 0 & -2 \\ 3 & 5 \end{pmatrix}, \ C = \begin{pmatrix} -2 & 3 & -2 \\ 3 & 0 & -4 \end{pmatrix}. \)

Calculate \( A(BC) \) and \( (AB)C. \) This illustrates the associative property of matrix multiplication.

14. Let \( A \) be a \( 4 \times 2 \) matrix, \( B \) a \( 2 \times 6 \) matrix, \( C \) a \( 3 \times 4 \) matrix, \( D \) a \( 6 \times 3 \) matrix.

Determine which of the following matrix expressions exist, and give the sizes of the resulting matrix when they do exist.

(a) \( ABC \)  
(b) \( ABD \)  
(c) \( CAB \)

(d) \( DCAB \)  
(e) \( A^2BDC \)

15. Let \( A \) be a \( 3 \times 2 \) matrix, \( B \) a \( 2 \times 1 \) matrix, \( C \) a \( 1 \times 3 \) matrix, \( D \) a \( 3 \times 1 \) matrix, and \( E \) a \( 3 \times 3 \) matrix. Determine which of the following matrix expressions exist, and give the sizes of the resulting matrix when they do exist.

(a) \( ABC \)  
(b) \( AB + EAB \)  
(c) \( DC - EAC \)

(d) \( DAB + AB \)  
(e) \( BCD + BC \)
1.6 Square Matrices; Inverse of a Matrix and Determinants

Recall that a matrix $A$ is said to be square if the number of rows equals the number of columns; that is, $A$ is an $n \times n$ matrix for some integer $n \geq 2$. The material in this section is restricted to square matrices.

From the previous section we know that we can represent a system of $n$ linear equations in $n$ unknowns

\[
\begin{align*}
    a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
    a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
    a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
    \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
    a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n &= b_n
\end{align*}
\]

as

\[
A\mathbf{x} = \mathbf{b}
\]

where $A$ is the $n \times n$ matrix of coefficients

\[
A = \begin{pmatrix}
    a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\
    a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\
    a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}
\end{pmatrix}
\]

As you know, we would solve one equation in one unknown, $ax = b$, by dividing the equation by $a$, provided $a \neq 0$. Dividing by $a$, $a \neq 0$, is equivalent to multiplying by $a^{-1} = 1/a$. For $a \neq 0$, the number $a^{-1}$ is called the multiplicative inverse of $a$. So, to solve $A\mathbf{x} = \mathbf{b}$ we would like to multiply the equation by “multiplicative inverse” of $A$. This idea has to be defined.

**DEFINITION** (Multiplicative Inverse of a Square Matrix) Let $A$ be an $n \times n$ matrix. An $n \times n$ matrix $B$ with the property that

\[
AB = BA = I_n
\]

is called the multiplicative inverse of $A$ or, more simply, the inverse of $A$. ■

Not every $n \times n$ matrix has an inverse. For example, a matrix with a row or a column of zeros cannot have an inverse. It’s sufficient to look at a $2 \times 2$ example. If the matrix

\[
\begin{pmatrix}
    a & b \\
    0 & 0
\end{pmatrix}
\]
had an inverse
\[
\begin{pmatrix}
x & y \\
z & w
\end{pmatrix},
\]
then
\[
\begin{pmatrix}
a & b \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
x & y \\
z & w
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]
But
\[
\begin{pmatrix}
a & b \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
x & y \\
z & w
\end{pmatrix} = \begin{pmatrix}
ax + bz & ay + bw \\
0 & 0
\end{pmatrix} \neq \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]
for all \(x, y, z, w\); look at the entries in the (2,2)-position. A similar contradiction is obtained with a matrix that has a column of zeros.

Here’s another example.

**Example 1.** Determine whether the matrix \(A = \begin{pmatrix} 2 & -1 \\ -4 & 2 \end{pmatrix}\) has an inverse.

**SOLUTION** To find the inverse of \(A\), we need to find a matrix \(B = \begin{pmatrix} x & y \\ z & w \end{pmatrix}\) such that
\[
\begin{pmatrix}
2 & -1 \\
-4 & 2
\end{pmatrix}
\begin{pmatrix}
x & y \\
z & w
\end{pmatrix} = \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}.
\]
This matrix equation yields the two systems
\[
\begin{align*}
2x - z &= 1 \\
-4x + 2z &= 0
\end{align*}
\]
and
\[
\begin{align*}
2y - w &= 0 \\
-4y + 2w &= 1
\end{align*}
\]
It is easy to see that neither of these systems has a solution. Therefore, \(A\) does not have an inverse. ■

If \(A\) has an inverse matrix \(B\), then it is unique. That is, there cannot exist two different matrices \(B\) and \(C\) such that
\[
AB = BA = I_n \quad \text{and} \quad AC = CA = I_n.
\]
For if \(AB = I_n\), then \(C(AB) = CI_n = C\);
but
\[
C(AB) = (CA)B = I_nB = B.
\]
Therefore, \(B = C\).

The inverse of a matrix \(A\), if it exists, is denoted by \(A^{-1}\).

**Finding the Inverse of a Matrix**
We'll illustrate a general method for finding the inverse of a matrix by finding the inverse of a $2 \times 2$ matrix.

**Example 2.** Find the inverse of $A = \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix}$.

**SOLUTION** We are looking for a matrix $A^{-1} = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$ such that

$$\begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This matrix equation yields the two systems of equations

\begin{align*}
x - 2z &= 1 \\
3x - 4z &= 0 \\
y - 2w &= 0 \\
3y - 4w &= 1
\end{align*}

Writing down the augmented matrices for these two systems, we have

\begin{align*}
\begin{pmatrix} 1 & -2 & 1 \\ 3 & -4 & 0 \end{pmatrix} & \quad \text{and} \quad \begin{pmatrix} 1 & -2 & 0 \\ 3 & -4 & 1 \end{pmatrix}
\end{align*}

We'll solve the systems by reducing each of these matrices to reduced row-echelon form.

\begin{align*}
\begin{pmatrix} 1 & -2 & 1 \\ 3 & -4 & 0 \end{pmatrix} & \rightarrow \begin{pmatrix} 1 & -2 & 1 \\ 0 & 2 & -3 \end{pmatrix} \\
& \rightarrow \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -3/2 \end{pmatrix} \\
& \rightarrow \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3/2 \end{pmatrix}
\end{align*}

Therefore, $x = -2$, $z = -\frac{3}{2}$.

Now we'll row reduce the second augmented matrix

\begin{align*}
\begin{pmatrix} 1 & -2 & 0 \\ 3 & -4 & 1 \end{pmatrix} & \rightarrow \begin{pmatrix} 1 & -2 & 0 \\ 0 & 2 & 1 \end{pmatrix} \\
& \rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1/2 \end{pmatrix}
\end{align*}

Therefore, $y = 1$, $w = \frac{1}{2}$.

Putting these two results together, we get $A^{-1} = \begin{pmatrix} -2 & 1 \\ -\frac{3}{2} & 1/2 \end{pmatrix}$. You should verify that

$$\begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ -\frac{3}{2} & 1/2 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -\frac{3}{2} & 1/2 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
Note that the coefficient matrices in each of the systems was exactly the same, and we used exactly the same row operations to row reduce the augmented matrices. We could have saved ourselves some writing by solving the two systems simultaneously (see Exercises 1.3, Problem 26). That is, by reducing the augmented matrix
\[
\begin{pmatrix}
1 & -2 & 1 & 0 \\
3 & -4 & 0 & 1
\end{pmatrix}
\]
to reduced row-echelon form.

\[
\begin{pmatrix}
1 & -2 & 1 & 0 \\
3 & -4 & 0 & 1
\end{pmatrix} \rightarrow
\begin{pmatrix}
-3 \\ 0
\end{pmatrix}
\begin{pmatrix}
R_1 + R_2 \\ R_2
\end{pmatrix}
\rightarrow
\begin{pmatrix}
1 & -2 & 1 & 0 \\
0 & 2 & -3 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
0 & 1 & -3/2 & 1/2
\end{pmatrix}
\begin{pmatrix}
2R_2 + R_1 \\ -3/2 & 1/2 & -3/2 & 1/2
\end{pmatrix}.
\]

The matrix to the right of the bar is \( A^{-1} \).

We use exactly the same method to find the inverse of an \( n \times n \) matrix. That is, to find the inverse of an \( n \times n \) matrix \( A \), form the augmented matrix

\[
(A | I_n)
\]

and use the row operations to reduce this matrix to reduced row-echelon form. If, in the process, you get a row of zeros to the left of the bar, then the matrix does not have an inverse because one of the equations in one of the corresponding systems would have the form

\[0x_1 + 0x_2 + -x_3 + \ldots + 0x_n = b, \quad b \neq 0.
\]

If all the rows are nonzero, then the reduced row-echelon form will be \((I_n | A^{-1})\); the matrix to the right of the bar is the inverse of \( A \).

There is a connection between the existence of the inverse and rank. Recall that the rank of a matrix is the number of nonzero rows in its row-echelon or reduced row-echelon form. Thus, an \( n \times n \) matrix \( A \) has an inverse if and only if its rank is \( n \).

We’ll illustrate the method with some \( 3 \times 3 \) examples.

**Example 3.** Find the inverse of

\[
A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}
\]

**SOLUTION** We form the augmented matrix \((A | I_3)\) and use row operations to obtain the reduced row-echelon form.

\[
\begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 2 & -1 & 3 & | & 0 & 1 & 0 \\ 4 & 1 & 8 & | & 0 & 0 & 1 \end{pmatrix}
\]

\[
(2)R_1 + R_2 \rightarrow R_2, \quad (-4)R_1 + R_3 \rightarrow R_3
\]

\[
\begin{pmatrix} 1 & 0 & 2 & | & 1 & 0 & 0 \\ 0 & -1 & 1 & | & -2 & 1 & 0 \\ 0 & 1 & 0 & | & -4 & 0 & 1 \end{pmatrix}
\]

\[
R_2 \rightarrow R_3
\]
Thus, \( A^{-1} = \begin{pmatrix} -11 & 2 & 2 \\ -4 & 0 & 1 \\ 6 & -1 & -1 \end{pmatrix} \).

You should verify that \( AA^{-1} + A^{-1}A = I_3 \). \( \square \)

**Example 4.** Find the inverse of

\[
A = \begin{pmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{pmatrix}
\]

**SOLUTION** We form the augmented matrix \((A | I_3)\) and use row operations to obtain the reduced row-echelon form.

\[
\begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 1 & 5 & -1 & | & 0 & 1 & 0 \\ 3 & 13 & -6 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow{(-1)R_1 + R_2 \rightarrow R_2} \begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 0 & 2 & -3 & | & -1 & 1 & 0 \\ 3 & 13 & -6 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow{(1/2)R_2 \rightarrow R_2} \begin{pmatrix} 1 & 3/2 & -1/2 & | & 1/2 & 0 & 0 \\ 0 & 3/2 & -1/2 & | & -1/2 & 1/2 & 0 \\ 3 & 13 & -6 & | & 0 & 0 & 1 \end{pmatrix} \xrightarrow{(-4)R_2 + R_3 \rightarrow R_3} \begin{pmatrix} 1 & 3 & -4 & | & 1 & 0 & 0 \\ 0 & 3/2 & -1/2 & | & -1/2 & 1/2 & 0 \\ 0 & 0 & 0 & | & -1 & -2 & 1 \end{pmatrix}
\]

The row of zeros to the left of the bar indicates that \( A \) does not have an inverse. \( \square \)

**Warning** The examples have been chosen fairly carefully to keep the calculations simple; for example, avoiding fractions as much as possible. In general, even though you start with a matrix with integer entries, you will not end up with the inverse having integer entries.

\( \square \)

**Example 5.** Let \( A = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 2 & -2 \\ 3 & 1 & 0 \end{pmatrix} \). Then \( A^{-1} = \begin{pmatrix} 2/5 & 1/5 & 0 \\ -6/5 & -3/5 & 1 \\ -1 & -1 & 1 \end{pmatrix} \).

Let \( B = \begin{pmatrix} 1 & 0 & 2 \\ 2 & -3 & 4 \\ 0 & 2 & 1 \end{pmatrix} \). Then \( B^{-1} = \begin{pmatrix} 11/3 & -4/3 & -2 \\ 2/3 & -1/3 & 0 \\ -4/3 & 2/3 & 1 \end{pmatrix} \). \( \square \)
Application to Systems of Equations  System (1) can be written in the vector-matrix
form (2). If the coefficient matrix  \( A \) has an inverse, \( A^{-1} \), then

\[
A \mathbf{x} = \mathbf{b} \quad \text{implies} \quad A^{-1}(A \mathbf{x}) = A^{-1} \mathbf{b} = (A^{-1}A) \mathbf{x} = A^{-1} \mathbf{b} = I_n \mathbf{x} = A^{-1} \mathbf{b} = \mathbf{x} = A^{-1} \mathbf{b}
\]

Thus, the system has the unique solution (vector) \( \mathbf{x} = A^{-1} \mathbf{b} \). If \( A \) does not have an
inverse, then the system of equations either has no solutions or infinitely many solutions.

**Example 6.** Solve the system of equations

\[
\begin{align*}
    x - 2y &= 4 \\
    3x - 4y &= -8
\end{align*}
\]

***SOLUTION*** The system can be written in vector-matrix form as

\[
\begin{pmatrix}
    1 & -2 \\
    3 & -4
\end{pmatrix}
\begin{pmatrix}
    x \\
    y
\end{pmatrix} =
\begin{pmatrix}
    4 \\
    -8
\end{pmatrix}.
\]

In Example 2, we saw that \( \begin{pmatrix}
    -2 & 1 \\
    -3/2 & 1/2
\end{pmatrix} \) is the inverse of the coefficient matrix.

Therefore, the unique solution (vector) of the system is

\[
\begin{pmatrix}
    x \\
    y
\end{pmatrix} =
\begin{pmatrix}
    -2 & 1 \\
    -3/2 & 1/2
\end{pmatrix}
\begin{pmatrix}
    4 \\
    -8
\end{pmatrix} =
\begin{pmatrix}
    -16 \\
    -10
\end{pmatrix}.
\]

The solution set is \( x = -16, \ y = -10. \) ■

**Example 7.** Solve the system of equations

\[
\begin{align*}
    x + 2z &= -1 \\
    2x - y + 3z &= 2 \\
    4x + y + 8z &= 0
\end{align*}
\]

***SOLUTION*** The system can be written in vector-matrix form as

\[
\begin{pmatrix}
    1 & 0 & 2 \\
    2 & -1 & 3 \\
    4 & 1 & 8
\end{pmatrix}
\begin{pmatrix}
    x \\
    y \\
    z
\end{pmatrix} =
\begin{pmatrix}
    -1 \\
    2 \\
    0
\end{pmatrix}.
\]
In Example 3, we saw that \[
\begin{pmatrix}
-11 & 2 & 2 \\
-4 & 0 & 1 \\
6 & -1 & -1
\end{pmatrix}
\]
is the inverse of the coefficient matrix. Therefore, the unique solution (vector) of the system is

\[
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
= \begin{pmatrix}
-11 & 2 & 2 \\
-4 & 0 & 1 \\
6 & -1 & -1
\end{pmatrix}
\begin{pmatrix}
-1 \\
2 \\
0
\end{pmatrix}
= \begin{pmatrix}
15 \\
4 \\
-8
\end{pmatrix}.
\]

The solution set is \(x = 15, y = 4, z = -8\). ■

**Determinants**

Associated with each \(n \times n\) matrix \(A\) is a number called its *determinant*. We will give an inductive development of this concept, beginning with the determinant of a \(2 \times 2\) matrix. Then we'll express a \(3 \times 3\) determinant as a sum of \(2 \times 2\) determinants, a \(4 \times 4\) determinant as a sum of \(3 \times 3\) determinants, and so on.

Consider the system of two linear equations in two unknowns

\[
\begin{align*}
ax + by & = \alpha \\
cx + dy & = \beta
\end{align*}
\]

We eliminate the \(y\) unknown by multiplying the first equation by \(d\), the second equation by \(-b\), and adding. This gives

\[
(ad - bc)x = d\alpha - b\beta.
\]

This equation has the solution \(x = \frac{d\alpha - b\beta}{ad - bc}\), provided \(ad - bc \neq 0\).

Similarly, we can solve the system for the \(y\) unknown by multiplying the first equation by \(-c\), the second equation by \(a\), and adding. This gives

\[
(ad - bc)y = a\beta - c\alpha
\]

which has the solution \(y = \frac{a\beta - c\alpha}{ad - bc}\), again provided \(ad - bc \neq 0\).

The matrix of coefficients of the system is \(A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\). The number \(ad - bc\) is called the *determinant of \(A\).* The determinant of \(A\) is denoted by \(\det A\) and by

\[
\begin{vmatrix} a & b \\ c & d \end{vmatrix}.
\]

The determinant has a geometric interpretation in this \(2 \times 2\) case.
The graph of the first equation \( ax + by = \alpha \) is a straight line with slope \(-a/b\), provided \( b \neq 0\). The graph of the second equation is a straight line with slope \(-c/d\), provided \( d \neq 0\). (If \( b = 0 \) or \( d = 0 \), then the corresponding line is vertical.) Assume that \( b, d \neq 0\). If \[
\frac{-a}{b} \neq \frac{-c}{d},
\]then the lines have different slopes and the system of equations has a unique solution. However, \[
\frac{-a}{b} \neq \frac{-c}{d}
\]is equivalent to \( ad - bc \neq 0 \). Thus, \( \det A \neq 0 \) implies that the system has a unique solution.

On the other hand, if \( ad - bc = 0 \), then \( \frac{-a}{b} = \frac{-c}{d} \) (assuming \( b, d \neq 0 \)), and the two lines have the same slope. In this case, the lines are either parallel (the system has no solutions), or the lines coincide (the system has infinitely many solutions).

In general, an \( n \times n \) matrix \( A \) is said to be nonsingular if \( \det A \neq 0 \); \( A \) is singular if \( \det A = 0 \).

Look again at the solutions
\[
x = \frac{d\alpha - b\beta}{ad - bc}, \quad y = \frac{a\beta - c\alpha}{ad - bc}, \quad ad - bc \neq 0.
\]
The two numerators also have the form of a determinant of a \( 2 \times 2 \) matrix. In particular, these solutions can be written as
\[
x = \frac{\begin{vmatrix} \alpha & b \\ \beta & d \end{vmatrix}}{ad - bc}, \quad y = \frac{\begin{vmatrix} a & \alpha \\ c & \beta \end{vmatrix}}{ad - bc}.
\]
This representation of the solutions of a system of two equations in two unknowns is the \( n = 2 \) version of a general result known as Cramer’s rule.

**Example 8.** Given the system of equations
\[
5x - 2y = 8 \\
3x + 4y = 10
\]
Verify that the determinant of the matrix of coefficients is nonzero and solve the system using Cramer’s rule.

**SOLUTION** The matrix of coefficients is \( A = \begin{pmatrix} 5 & -2 \\ 3 & 4 \end{pmatrix} \) and \( \det A = 26 \). According to Cramer’s rule,
\[
x = \frac{\begin{vmatrix} 8 & -2 \\ 10 & 4 \end{vmatrix}}{26} = \frac{52}{26} = 2, \quad y = \frac{\begin{vmatrix} 5 & 8 \\ 3 & 10 \end{vmatrix}}{26} = \frac{26}{26} = 1
\]
The solution set is $x = 2, \ y = 1$.

Now we’ll go to $3 \times 3$ matrices.

The determinant of a $3 \times 3$ matrix

If

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix},$$

then

$$\text{det} \ A = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$  

The problem with this definition is that it is hard to remember. Fortunately the expression on the right can be written conveniently in terms of $2 \times 2$ determinants as follows:

$$\text{det} \ A = a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1 = a_1(b_2c_3 - b_3c_2) - a_2(b_1c_3 - b_3c_1) + a_3(b_1c_2 - b_2c_1)$$

This representation of a $3 \times 3$ determinant is called the expansion of the determinant across the first row. Notice that the coefficients are the entries $a_1, a_2, a_3$ of the first row, that they occur alternately with $+$ and $-$ signs, and that each is multiplied by a $2 \times 2$ determinant. You can remember the determinant that goes with each entry $a_i$ as follows: in the original matrix, mentally cross out the row and column containing $a_i$ and take the determinant of the $2 \times 2$ matrix that remains.

**Example 9.** Let $A = \begin{pmatrix} 3 & -2 & -4 \\ 2 & 5 & -1 \\ 0 & 6 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 7 & 6 & 5 \\ 1 & 2 & 1 \\ 3 & -2 & 1 \end{pmatrix}$. Calculate $\text{det} \ A$ and $\text{det} \ B$.

**SOLUTION**

$$\text{det} \ A = 3 \begin{vmatrix} 5 & -1 \\ 6 & 1 \end{vmatrix} - (-2) \begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} + (-4) \begin{vmatrix} 2 & 5 \\ 0 & 6 \end{vmatrix}$$

$$= 3[(5)(1) - (-1)(6)] + 2[(2)(1) - (-1)(0)] - 4[(2)(6) - (5)(0)]$$

$$= 3(11) + 2(2) - 4(12) = -11$$

210
\[
\det B = 7 \begin{vmatrix} 2 & 1 & -6 \\ -2 & 1 & 5 \\ 3 & 1 & 3 \\ 1 & 2 & -2 \end{vmatrix} \\
= 7[(2)(1) - (1)(-2)] - 6[(1)(1) - (1)(3)] + 5[(1)(-2) - (2)(3)] \\
= 7(4) - 6(-2) + 5(-8) = 0.
\]

There are other ways to group the terms in the definition. For example
\[
\det A = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1 \\
= -a_2(b_1 c_3 - b_3 c_1) + b_2(a_1 c_3 - a_3 c_1) - c_2(a_1 c_3 - a_3 c_1) \\
= -a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & a_3 \\ c_1 & c_3 \end{vmatrix} - c_2 \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}
\]

This is called the expansion of the determinant down the second column.

In general, depending on how you group the terms in the definition, you can expand across any row or down any column. The signs of the coefficients in the expansion across a row or down a column are alternately +, −, starting with a + in the (1,1)-position. The pattern of signs is:

\[
\begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}
\]

**Example 10.** Let \( A = \begin{pmatrix} 3 & -2 & -4 \\ 2 & 5 & -1 \\ 0 & 6 & 1 \end{pmatrix} \) and \( C = \begin{pmatrix} 7 & 0 & 5 \\ 1 & 0 & 1 \\ 3 & -2 & 1 \end{pmatrix} \).

1. Calculate \( \det A \) by expanding down the first column.

\[
\det A = 3 \begin{vmatrix} 5 & -1 \\ 6 & 1 \end{vmatrix} - 2 \begin{vmatrix} -2 & -4 \\ 6 & 1 \end{vmatrix} + 0 \begin{vmatrix} -2 & -4 \\ 5 & 1 \end{vmatrix} \\
= 3[(5)(1) - (-1)(6)] - 2[(-2)(1) - (-4)(2)] + 0 \\
= 3(11) - 2(22) + 0 = -11
\]

2. Calculate \( \det A \) by expanding across the third row.

\[
\det A = 0 \begin{vmatrix} -2 & -4 \\ 5 & -1 \end{vmatrix} - 6 \begin{vmatrix} 3 & 4 \\ 2 & -1 \end{vmatrix} + (1) \begin{vmatrix} 3 & -2 \\ 2 & 5 \end{vmatrix} \\
= 0 - 6[(3)(-1) - (-4)(2)] + [(3)(5) - (-2)(2)] \\
= -6(5) + (19) = -11
\]
3. Calculate $\det C$ by expanding down the second column.

$$
\det C = -0 \left| \begin{array}{cc} 1 & 1 \\ 3 & 1 \end{array} \right| + 0 \left| \begin{array}{cc} 7 & 5 \\ 3 & 1 \end{array} \right| - (-2) \left| \begin{array}{cc} 7 & 5 \\ 1 & 1 \end{array} \right|
= 0 + 0 + 2(2) = 14
$$

Notice the advantage of expanding across a row or down a column that contains one or more zeros. ■

Now consider the system of three equations in three unknowns

$$
\begin{align*}
\begin{array}{ccc}
\phantom{+}a_{11}x + a_{12}y + a_{13}z &=& b_1 \\
\phantom{+}a_{21}x + a_{22}y + a_{23}z &=& b_2 \\
\phantom{+}a_{31}x + a_{32}y + a_{33}z &=& b_2
\end{array}
\end{align*}
$$

Writing this system in vector-matrix form, we have

$$
\begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
=
\begin{pmatrix}
b_1 \\
b_2 \\
b_3
\end{pmatrix}
$$

It can be shown that if $\det A \neq 0$, then the system has a unique solution which is given by

$$
x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad x_3 = \frac{\det A_3}{\det A}
$$

where

$$
\begin{align*}
A_1 &= \begin{pmatrix}
b_1 & a_{12} & a_{13} \\
b_2 & a_{22} & a_{23} \\
b_3 & a_{32} & a_{33}
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
a_{11} & b_1 & a_{13} \\
a_{21} & b_2 & a_{23} \\
a_{31} & b_3 & a_{33}
\end{pmatrix}, \quad \text{and} \quad A_3 = \begin{pmatrix}
a_{11} & a_{12} & b_1 \\
a_{21} & a_{22} & b_2 \\
a_{31} & a_{32} & b_3
\end{pmatrix}
\end{align*}
$$

This is Cramer’s rule in the $3 \times 3$ case.

If $\det A = 0$, then the system either has infinitely many solutions or no solutions.

**Example 11.** Given the system of equations

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

Verify that the determinant of the matrix of coefficients is nonzero and find the value of $y$ using Cramer’s rule.
SOLUTION The matrix of coefficients is

\[
A = \begin{pmatrix}
2 & 1 & -1 \\
1 & 1 & 1 \\
1 & -2 & -3
\end{pmatrix}
\]

and \( \det A = 5 \).

According to Cramer’s rule

\[
y = \frac{\begin{vmatrix}
2 & 3 & -1 \\
1 & 1 & 1 \\
1 & 4 & -3
\end{vmatrix}}{5} = -\frac{5}{5} = -1.
\]

The determinant of a 4 \( \times \) 4 matrix

Following the pattern suggested by the calculation of a 3 \( \times \) 3 determinant, we’ll express a 4 \( \times \) 4 determinant as the sum of four 3 \( \times \) 3 determinants. For example, the expansion of

\[
det A = \begin{vmatrix}
a_1 & a_2 & a_3 & a_4 \\
b_1 & b_2 & b_3 & b_4 \\
c_1 & c_2 & c_3 & c_4 \\
d_1 & d_2 & d_3 & d_4
\end{vmatrix}
\]

across the first row is

\[
det A = a_1 \begin{vmatrix}
b_2 & b_3 & b_4 \\
c_2 & c_3 & c_4 \\
d_2 & d_3 & d_4
\end{vmatrix} - a_2 \begin{vmatrix}
b_1 & b_3 & b_4 \\
c_1 & c_3 & c_4 \\
d_1 & d_3 & d_4
\end{vmatrix} + a_3 \begin{vmatrix}
b_1 & b_2 & b_4 \\
c_1 & c_2 & c_4 \\
d_1 & d_2 & d_4
\end{vmatrix} + a_4 \begin{vmatrix}
b_1 & b_2 & b_3 \\
c_1 & c_2 & c_3 \\
d_1 & d_2 & d_3
\end{vmatrix}
\]

As in the 3 \( \times \) 3 case, you can calculate a 4 \( \times \) 4 determinant by expanding across any row or down any column. The matrix of signs associated with a 4 \( \times \) 4 determinant is

\[
\begin{pmatrix}
+ & - & - & - \\
- & + & + & - \\
+ & - & - & - \\
- & + & - & +
\end{pmatrix}
\]

Cramer’s rule

Here is the general version of Cramer’s rule: Given the system of \( n \) equations in \( n \) unknowns

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
& \vdots \\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n &= b_n
\end{align*}
\] (1)
If \( \det A \neq 0 \), then the system has a unique solution \( x_1, x_2, \ldots, x_n \) given by

\[
x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \ldots, \quad x_n = \frac{\det A_n}{\det A}
\]

where \( \det A_i \) is the determinant obtained by replacing the \( i^{th} \) column of \( \det A \) by the column

\[
\begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_n
\end{pmatrix}
\]

\( i = 1, 2, \ldots, n. \)

If \( \det A = 0 \), then the system either has no solution or infinitely many solutions. In the special case of a homogeneous system, \( \det A = 0 \) implies that the system has infinitely many nontrivial solutions.

Properties of Determinants It should now be clear how to calculate an \( n \times n \) determinant for any \( n \). However, for \( n > 3 \) the calculations, while simple in theory, tend to be long, tedious, and involved. Although the determinant is a complicated mathematical function (its domain is the set of square matrices, its range is the set of real numbers), it does have certain properties that can be used to simplify calculations.

Before listing the properties of determinants, we give the determinants of some special types of matrices.

1. If an \( n \times n \) matrix \( A \) has a row of zeros, or a column of zeros, then \( \det A = 0; \) an \( n \times n \) matrix with a row of zeros or a column of zeros is singular. (Simply expand across the row or column of zeros.)

2. An \( n \times n \) matrix is said to be upper triangular if all of its entries below the main diagonal are zero. (Recall that the entries \( a_{11}, a_{22}, a_{33}, \ldots, a_{nn} \) form the main diagonal of an \( n \times n \) matrix \( A \).) A \( 4 \times 4 \) upper triangular matrix has the form

\[
T = \begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} \\
0 & a_{22} & a_{23} & a_{24} \\
0 & 0 & a_{33} & a_{34} \\
0 & 0 & 0 & a_{44}
\end{pmatrix}
\]

Note that this upper triangular form is closely related to the row-echelon form that was so important in solving systems of linear equations. Calculating \( \det T \) by expanding down the first column, we get

\[
\det T = a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{vmatrix} = a_{11}a_{22} \begin{vmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{vmatrix} = a_{11}a_{22}a_{33}a_{44}.
\]
In general, we can see that the determinant of an upper triangular matrix is simply the product of its entries on the main diagonal.

3. An $n \times n$ matrix $L$ is **lower triangular** if all its entries above the main diagonal are zero. Just like an upper triangular matrix, the determinant of a lower triangular matrix is the product of the entries on the main diagonal.

**Properties:** We’ll list the properties for general $n$ and illustrate with $2 \times 2$ determinants. Let $A$ be an $n \times n$ matrix

1. If the matrix $B$ is obtained from $A$ by interchanging two rows (or two columns), then $\det B = -\det A$.

   \[
   \begin{vmatrix}
   c & d \\
   a & b
   \end{vmatrix}
   = bc - ad
   = -
   \begin{vmatrix}
   a & b \\
   c & d
   \end{vmatrix}.
   
   Note: An immediate consequence of this property is the fact that if $A$ has two identical rows (or columns), then $\det A = 0$ — interchange the two identical rows, then $\det A = -\det A$ which implies $\det A = 0$.

2. If the matrix $B$ is obtained from $A$ by multiplying a row (or column) by a nonzero number $k$, then $\det B = k \det A$.

   \[
   \begin{vmatrix}
   ka & kb \\
   c & d
   \end{vmatrix}
   = kad - kbc
   = k(ad - bc)
   = k\begin{vmatrix}
   a & b \\
   c & d
   \end{vmatrix}.
   
3. If the matrix $B$ is obtained from $A$ by multiplying a row (or column) by a number $k$ and adding the result to another row (or column), then $\det B = \det A$.

   \[
   \begin{vmatrix}
   a & b \\
   ka + c & kb + d
   \end{vmatrix}
   = a(kb + d) - b(ka + c)
   = ad - bc
   = \begin{vmatrix}
   a & b \\
   c & d
   \end{vmatrix}.
   
   Of course these are the operations we used to row reduce a matrix to row-echelon form. We’ll use them here to row reduce a determinant to upper triangular form. The difference between row reducing a matrix and row reducing a determinant is that with a determinant we have to keep track of the sign if we interchange rows and we have to account for the factor $k$ if we multiply a row by $k$.

**Example 12.** Calculate the determinant

\[
\begin{vmatrix}
1 & 2 & 0 & 1 \\
0 & 2 & 1 & 0 \\
-2 & -3 & 3 & -1 \\
1 & 0 & 5 & 2
\end{vmatrix}
\]
SOLUTION

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

\begin{align*}
2R_1+R_3 &= -R_1+r_4 \\
0 & 1 & 3 & 1
\end{align*}

\begin{align*}
R_2 &\leftrightarrow R_3 \\
-\frac{1}{5}R_3 &\rightarrow R_3 \\
5 &\begin{bmatrix}
1 & 2 & 0 & 1 \\
0 & 1 & 3 & 1 \\
0 & 0 & 1 & 2/5 \\
0 & 0 & 11 & 3
\end{bmatrix}
= \begin{bmatrix}
1 & 2 & 0 & 1 \\
0 & 1 & 3 & 1 \\
0 & 0 & 0 & -7/5 \\
0 & 0 & 0 & -7/5
\end{bmatrix}
= -\frac{7}{5} = -7.
\end{align*}

Note, we could have stopped at

\[
\begin{bmatrix}
1 & 2 & 0 & 1 \\
0 & 1 & 3 & 1 \\
0 & 0 & -5 & -2 \\
0 & 0 & 11 & 3
\end{bmatrix}
= -(1)(1)
\]

and calculated

\[
\begin{bmatrix}
1 & 2 & 0 & 1 \\
0 & 1 & 3 & 1 \\
0 & 0 & -5 & -2 \\
0 & 0 & 11 & 3
\end{bmatrix}
= -(15+22) = -7.
\]

Inverse, determinant and rank

We have seen that a system of equations \( Ax = b \) where \( A \) is an \( n \times n \) matrix has a unique solution if and only if \( A \) has an inverse. We have also seen that the system has a unique solution if and only if \( \det A \neq 0 \); that is, if and only if \( A \) is nonsingular. It now follows that \( A \) has an inverse if and only if \( \det A \neq 0 \). Thus, the determinant provides a test as to whether an \( n \times n \) matrix has an inverse.

There is also the connection with rank: an \( n \times n \) matrix has an inverse if and only if its rank is \( n \).

Putting all this together we have the following equivalent statements:
1. The system of equations $Ax = b$, $A$ an $n \times n$ matrix, has a unique solution.

2. $A$ has an inverse.

3. $\det A \neq 0$.

4. $A$ has rank $n$.

**Exercises 1.6**

Use the determinant to decide whether the matrix has an inverse. If it exists, find it and verify your answer by calculating $AA^{-1}$.

1. $\begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix}$
2. $\begin{pmatrix} 0 & 1 \\ -2 & 4 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$
5. $\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ 4 & 1 & 8 \end{pmatrix}$
6. $\begin{pmatrix} 1 & 2 & -4 \\ -1 & -1 & 5 \\ 2 & 7 & -3 \end{pmatrix}$
7. $\begin{pmatrix} 1 & 3 & -4 \\ 1 & 5 & -1 \\ 3 & 13 & -6 \end{pmatrix}$
8. $\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 4 & 5 \end{pmatrix}$
9. $\begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 2 & 0 & -1 \\ -1 & 2 & 2 & -2 \\ 0 & -1 & 0 & 1 \end{pmatrix}$
10. \[
\begin{pmatrix}
1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 \\
0 & 0 & 1 & 1
\end{pmatrix}
\]

11. Each of the matrices in Problems 1 - 10 has integer entries. In some cases the inverse matrix also had integer entries, in other cases it didn’t. Suppose \( A \) is an \( n \times n \) matrix with integer entries. Make a conjecture as to when \( A^{-1} \), if it exists, will also have integer entries.

Solve the system of equations by finding the inverse of the matrix of coefficients.

12. \[
\begin{align*}
x + 2y &= 2 \\
3x + 5y &= 4 
\end{align*}
\]

13. \[
\begin{align*}
x + 3y &= 5 \\
2x + y &= 10 
\end{align*}
\]

14. \[
\begin{align*}
2x + 4y &= 2 \\
3x + 8y &= 1 
\end{align*}
\]

15. \[
\begin{align*}
2x + y &= 4 \\
4x + 3y &= 3 \\
x - 3y &= 2 
\end{align*}
\]

16. \[
\begin{align*}
y + z &= -2 \\
2x - y + 4z &= 1 \\
x + 2y - z &= 2 
\end{align*}
\]

17. \[
\begin{align*}
x + y + 2z &= 0 \\
x - y - z &= 1 \\
-x + y &= 5 
\end{align*}
\]

18. \[
\begin{align*}
-x + z &= -2 \\
6x - 2y - 3z &= 1 
\end{align*}
\]

Evaluate the determinant in two ways, using the indicated row and column expansions

19. \[
\begin{vmatrix}
0 & 3 & 2 \\
1 & 5 & 7 \\
-2 & -6 & -1
\end{vmatrix}; \quad \text{across the 2nd row, down the 1st column.}
\]

20. \[
\begin{vmatrix}
1 & 2 & -3 \\
2 & 5 & -8 \\
3 & 8 & -13
\end{vmatrix}; \quad \text{across the 3rd row, down the 2nd column.}
\]
21. \[
\begin{vmatrix}
5 & -1 & 2 \\
3 & 0 & 6 \\
-4 & 3 & 1
\end{vmatrix}; \text{ across the 2nd row, down the 3rd column.}
\]

22. \[
\begin{vmatrix}
5 & -2 & 0 \\
2 & 0 & -1 \\
1 & 0 & 3
\end{vmatrix}; \text{ across the 3rd row, down the 3rd column.}
\]

23. \[
\begin{vmatrix}
2 & -2 & 1 \\
4 & 0 & -3
\end{vmatrix}; \text{ across the 1st row, down the 2nd column.}
\]

Evaluate the determinant using the row or column that minimizes the amount of computation.

24. \[
\begin{vmatrix}
1 & 3 & 0 \\
2 & 5 & -2 \\
3 & 4 & 0
\end{vmatrix}
\]

25. \[
\begin{vmatrix}
2 & -5 & 1 \\
0 & 3 & 0 \\
3 & 4 & -2
\end{vmatrix}
\]

26. \[
\begin{vmatrix}
1 & 3 & 0 \\
2 & 5 & -2 \\
3 & 4 & 0
\end{vmatrix}
\]

27. \[
\begin{vmatrix}
1 & 3 & -4 \\
2 & 0 & -2 \\
0 & 0 & 3
\end{vmatrix}
\]

28. \[
\begin{vmatrix}
1 & -2 & 3 & 0 \\
4 & 0 & 5 & 0 \\
7 & -3 & 2 & 2 \\
-3 & 0 & 4 & 0
\end{vmatrix}
\]

29. \[
\begin{vmatrix}
2 & -1 & 3 & 4 \\
1 & 0 & 5 & 2 \\
-2 & 0 & 0 & 2 \\
-2 & 0 & -1 & 4
\end{vmatrix}
\]

30. Find the values of \( x \) such that \[
\begin{vmatrix}
x + 1 & x \\
3 & x - 2
\end{vmatrix} = 3.
\]

31. Find the values of \( x \) such that \[
\begin{vmatrix}
x & 0 & 2 \\
2x & x - 1 & 4 \\
-x & x - 1 & x + 1
\end{vmatrix} = 0.
\]
Determine whether Cramer’s rule applies. If it does, solve for the indicated unknown

32. \( 2x - y + 3z = 1 \)
\( y + 2z = -3 \); \( x =? \)
\( x + z = 0 \)
\( -4x + y = 3 \)
33. \( 2x + 2y + z = -2 \); \( y =? \)
\( 3x + 4z = 2 \)
\( 3x + z = -2 \)
34. \( x + 2y - z = 0 \); \( z =? \)
\( x - 4y + 3z = 1 \)
\( -2x - y = 3 \)
35. \( x + 3y - z = 0 \); \( z =? \)
\( 5y - 2z = 3 \)
\( 2x + y + 3z = 2 \)
36. \( 3x - 2y + 4z = 2 \); \( y =? \)
\( x + 4y - 2z = 1 \)
\( 2x + 7y + 3z = 7 \)
37. \( x + 2y + z = 2 \); \( x =? \)
\( x + 5y + 2z = 5 \)
\( 3x + 6y - z = 3 \)
38. \( x - 2y + 3z = 2 \); \( z =? \)
\( 4x - 2y + 5z = 5 \)
39. Determine the values of \( \lambda \) for which the system
\[
(1 - \lambda)x + 6y = 0 \\
5x + (2 - \lambda)y = 0
\]
has nontrivial solutions. Find the solutions for each value of \( \lambda \).

40. Determine the values of \( \lambda \) for which the system
\[
(5 - \lambda)x + 4y + 2z = 0 \\
4x + (5 - \lambda)y + 2z = 0 \\
2x + 2y + (2 - \lambda)z = 0
\]
has nontrivial solutions. Find the solutions for each value of \( \lambda \).
1.7 Vectors; Linear Dependence and Linear Independence

In Section 1.4 vectors were defined either as $1 \times n$ matrices, which we called row vectors, or as $m \times 1$ matrices, which we called column vectors. In this section we want to think of vectors as simply being ordered $n$-tuples of real numbers. Most of the time we will write vectors as row vectors, but occasionally it will be convenient to write them as columns. Our main interest is in the concept of linear dependence/linear independence of sets of vectors. At the end of the section we will extend this concept to linear dependence/independence of sets of functions.

We denote the set of ordered $n$-tuples of real numbers by the symbol $\mathbb{R}^n$. That is,

$$\mathbb{R}^n = \{(a_1, a_2, a_3, \ldots, a_n) | a_1, a_2, a_3, \ldots, a_n \text{ are real numbers}\}$$

In particular, $\mathbb{R}^2 = \{(a, b) | a, b \text{ real numbers}\}$, which we can identify with the set of all points in the plane, and $\mathbb{R}^3 = \{(a, b, c) | a, b, c \text{ real numbers}\}$, which we can identify with the set of points in space. We will use lower case boldface letters to denote vectors. The entries of a vector are called the components of the vector.

The operations of addition and multiplication by a number (scalar) that we defined for matrices in Section 1.4 hold automatically for vectors since a vector is a matrix. Addition is defined only for vectors with the same number of components. For any two vectors $u = (a_1, a_2, a_3, \ldots, a_n)$ and $v = (b_1, b_2, b_3, \ldots, b_n)$ in $\mathbb{R}^n$, we have

$$u + v = (a_1, a_2, a_3, \ldots, a_n) + (b_1, b_2, b_3, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots, a_n + b_n)$$

and for any real number $\lambda$,

$$\lambda v = \lambda (a_1, a_2, a_3, \ldots, a_n) = (\lambda a_1, \lambda a_2, \lambda a_3, \ldots, \lambda a_n).$$

Clearly, the sum of two vectors in $\mathbb{R}^n$ is another vector in $\mathbb{R}^n$ and a scalar multiple of a vector in $\mathbb{R}^n$ is a vector in $\mathbb{R}^n$. A sum of the form

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k$$

where $v_1, v_2, \ldots, v_k$ are vectors in $\mathbb{R}^n$ and $c_1, c_2, \ldots, c_k$ are real numbers is called a linear combination of $v_1, v_2, \ldots, v_k$.

The set $\mathbb{R}^n$ together with the operations of addition and multiplication by a number is called a vector space of dimension $n$. The term “dimension $n$” will become clear as we go on.

The zero vector in $\mathbb{R}^n$, which we’ll denote by $\mathbf{0}$ is the vector

$$\mathbf{0} = (0, 0, 0, \ldots, 0).$$

For any vector $v \in \mathbb{R}^n$, we have $v + \mathbf{0} = \mathbf{0} + v = v$; the zero vector is the additive identity in $\mathbb{R}^n$. 

221
Linear Dependence and Linear Independence

In Chapter 3 we said that two functions $f$ and $g$ are linearly dependent if one of the functions is a scalar multiple of the other; $f$ and $g$ are linearly independent if neither is a scalar multiple of the other. Similarly, let $u$ and $v$ be vectors in $\mathbb{R}^n$. Then $u$ and $v$ are linearly dependent if one of the vectors is a scalar multiple of the other (e.g., $u = \lambda v$ for some number $\lambda$); they are linearly independent if neither is a scalar multiple of the other.

Suppose that $u, v \in \mathbb{R}^n$ are linearly dependent with $u = \lambda v$. Then

$$u = \lambda v \quad \text{implies} \quad u - \lambda v = 0.$$ 

This leads to an equivalent definition of linear dependence: $u$ and $v$ are linearly dependent if there exist two numbers $c_1$ and $c_2$, not both zero, such that

$$c_1u + c_2v = 0.$$ 

Note that if $c_1u + c_2v = 0$ and $c_1 \neq 0$, then $u = (c_2/c_1)v = \lambda v$.

This is the idea that we’ll use to define linear dependence/independence in general.

**DEFINITION 1.** The set of vectors $\{v_1, v_2, \ldots, v_k\}$ in $\mathbb{R}^n$ is linearly dependent if there exist $k$ numbers $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0;$$

$0$ is a non-trivial linear combination of $v_1, v_2, \ldots, v_k$. The set of vectors is linearly independent if it is not linearly dependent.

**NOTE:** If there exists one set of $k$ numbers $c_1, c_2, \ldots, c_k$, not all zero, then there exist infinitely many such sets. For example, $2c_1, 2c_2, \ldots, 2c_k$ is another such set; and $\frac{1}{3}c_1, \frac{1}{3}c_2, \ldots, \frac{1}{3}c_k$ is another such set; and so on.

The definition of linear dependence can also be stated as: The vectors $v_1, v_2, \ldots, v_k$ are linearly dependent if one of the vectors can be written as a linear combination of the others. For example if

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$$

and $c_1 \neq 0$, then

$$v_1 = -\frac{c_2}{c_1}v_2 - \frac{c_3}{c_1}v_3 - \cdots - \frac{c_k}{c_1}v_k = \lambda_2 v_2 + \lambda_3 v_3 + \cdots + \lambda_k v_k.$$ 

This form of the definition parallels the definition of the linear dependence of two vectors.

Stated in terms of linear independence, the definition can be stated equivalently as:
The set of vectors \( \{v_1, v_2, \ldots, v_k\} \) is **linearly independent** if

\[
c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0 \quad \text{implies} \quad c_1 = c_2 = \cdots = c_k = 0;
\]

i.e., the only linear combination of \( v_1, v_2, \ldots, v_k \) which equals \( 0 \) is the trivial linear combination. The set of vectors is **linearly dependent** if it is not linearly independent.

**Linearly dependent/independent sets in** \( \mathbb{R}^n \)

**Example 1.** Vectors in \( \mathbb{R}^2 \).

(a) The vectors \( u = (3, 4) \) and \( v = (1, 3) \) are linearly independent because neither is a multiple of the other.

(b) The vectors \( u = (2, -3) \) and \( v = (-6, 9) \) are linearly dependent because \( v = -3u \).

(c) Determine whether the vectors \( v_1 = (1, 2), v_2 = (-1, 3), v_3 = (5, 7) \) are linearly dependent or linearly independent.

**SOLUTION** In this case we need to determine whether or not there are three numbers \( c_1, c_2, c_3 \), not all zero such that

\[
c_1 v_1 + c_2 v_2 + c_3 v_3 = 0.
\]

Equating the components of the vector on the left and the vector on the right, we get the homogeneous system of equations

\[
\begin{align*}
c_1 - c_2 + 5c_3 &= 0 \\
2c_1 + 3c_2 + 7c_3 &= 0
\end{align*}
\]

We know that this system has nontrivial solutions since it is a homogeneous system with more unknowns than equations (see Section 5.4). Thus the vectors must be linearly dependent.

The three parts of Example 1 hold in general: any set of three or more vectors in \( \mathbb{R}^2 \) is linearly dependent. A set of two vectors in \( \mathbb{R}^2 \) is independent if neither vector is a multiple of the other; dependent if one vector is a multiple of the other.

**Example 2.** Vectors in \( \mathbb{R}^3 \).

(a) The vectors \( u = (3, 4, -2) \) and \( v = (2, -6, 7) \) are linearly independent because neither is a multiple of the other.
(b) The vectors \( u = (-4, 6, -2) \) and \( v = (2, -3, 1) \) are linearly dependent because \( v = -2u \).

(c) Determine whether the vectors \( v_1 = (1, -2, 1) \), \( v_2 = (2, 1, -1) \), \( v_3 = (7, -4, 1) \) are linearly dependent or linearly independent.

**SOLUTION**

**Method 1** We need to determine whether or not there are three numbers \( c_1, c_2, c_3 \), not all zero, such that

\[
c_1v_1 + c_2v_2 + c_3v_3 = 0.
\]

Equating the components of the vector on the left and the vector on the right, we get the homogeneous system of equations

\[
\begin{align*}
c_1 + 2c_2 + 7c_3 &= 0 \\ -2c_1 + c_2 - 4c_3 &= 0 \\ c_1 - c_2 + c_3 &= 0
\end{align*}
\]

(A)

Writing the augmented matrix and reducing to row echelon form, we get

\[
\begin{pmatrix}
1 & 2 & 7 & 0 \\
-2 & 1 & -4 & 0 \\
1 & -1 & 1 & 0
\end{pmatrix} \rightarrow \begin{pmatrix}
1 & 2 & 7 & 0 \\
0 & 1 & 2 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

The row echelon form implies that system (A) has infinitely nontrivial solutions. Thus, we can find three numbers \( c_1, c_2, c_3 \), not all zero, such that

\[
c_1v_1 + c_2v_2 + c_3v_3 = 0.
\]

In fact, we can find infinitely many such sets \( c_1, c_2, c_3 \). Note that the vectors \( v_1, v_2, v_3 \) appear as the columns of the coefficient matrix in system (A).

**Method 2** Form the matrix with rows \( v_1, v_2, v_3 \) and reduce to echelon form:

\[
\begin{pmatrix}
1 & -2 & 1 \\
2 & 1 & -1 \\
7 & -4 & 1
\end{pmatrix} \rightarrow \begin{pmatrix}
1 & -2 & 1 \\
0 & 1 & -3/5 \\
0 & 0 & 0
\end{pmatrix}.
\]

The row of zeros indicates that the zero vector is a nontrivial linear combination of \( v_1, v_2, v_3 \). Thus the vectors are linearly dependent.

**Method 3** Calculate the determinant of the matrix of coefficients:

\[
\begin{vmatrix}
1 & 2 & 7 \\
-2 & 1 & -4 \\
1 & -1 & 1
\end{vmatrix} = 0.
\]

224
Therefore, as we saw in the preceding section, system (A) has infinitely many nontrivial solutions.

(d) Determine whether the vectors \( \mathbf{v}_1 = (1, 2, -3) \), \( \mathbf{v}_2 = (1, -3, 2) \), \( \mathbf{v}_3 = (2, -1, 5) \) are linearly dependent or linearly independent.

**SOLUTION** In part (c) we illustrated three methods for determining whether or not a set of vectors is linearly dependent or linearly independent. We could use any one of the three methods here. The determinant method is probably the easiest. Since a determinant can be evaluated by expanding across any row or down any column, it does not make any difference whether we write the vectors as rows or columns. We’ll write them as rows.

\[
\begin{vmatrix}
1 & 2 & -3 \\
1 & -3 & 2 \\
2 & -1 & 5 \\
\end{vmatrix}
= -30.
\]

Since the determinant is nonzero, the only solution to the vector equation

\[ c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3 = \mathbf{0}. \]

is the trivial solution; the vectors are linearly independent.

**Example 3.** Determine whether the vectors \( \mathbf{v}_1 = (1, 2, -4) \), \( \mathbf{v}_2 = (2, 0, 5) \), \( \mathbf{v}_3 = (1, -1, 7) \), \( \mathbf{v}_4 = (2, -2, -6) \) are linearly dependent or linearly independent.

**SOLUTION** In this case we need to determine whether or not there are four numbers \( c_1, c_2, c_3, c_4 \), not all zero such that

\[ c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + c_3 \mathbf{v}_3 + c_4 \mathbf{v}_4 = \mathbf{0}. \]

Equating the components of the vector on the left and the vector on the right, we get the homogeneous system of equations

\[
\begin{align*}
c_1 + 2c_2 + c_3 + 2c_4 &= 0 \\
2c_1 - c_3 - 2c_4 &= 0 \\
-4c_1 + 5c_2 + 7c_3 - 6c_4 &= 0
\end{align*}
\]

We know that this system has nontrivial solutions since it is a homogeneous system with more unknowns than equations. Thus the vectors must be linearly dependent.

**Example 4.** Vectors in \( \mathbb{R}^4 \)

Let \( \mathbf{v}_1 = (2, 0, -1, 4) \), \( \mathbf{v}_2 = (2, -1, 0, 2) \), \( \mathbf{v}_3 = (-2, 4, -3, 4) \), \( \mathbf{v}_4 = (1, -1, 3, 0) \), \( \mathbf{v}_5 = (0, 1, -5, 3) \).

(a) Determine whether the vectors \( \mathbf{v}_1 \), \( \mathbf{v}_2 \), \( \mathbf{v}_3 \), \( \mathbf{v}_4 \), \( \mathbf{v}_5 \) are linearly dependent or linearly independent.
SOLUTION The vectors are linearly dependent because the vector equation

\[ c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 + c_5v_5 = 0. \]

leads to a homogeneous system with more unknowns than equations.

(b) Determine whether the vectors \( v_1, v_2, v_3, v_4 \) are linearly dependent or linearly independent.

SOLUTION To test for dependence/independence in this case, we have three options.

1. Solve the system of equations

\[ c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0. \]

A nontrivial solution implies that the vectors are linearly dependent; if the trivial solution is the only solution, then the vectors are linearly independent.

2. Form the matrix \( A \) having \( v_1, v_2, v_3, v_4 \) as the rows and reduce to row-echelon form. If the row-echelon form has one or more rows of zeros, the vectors are linearly dependent; four nonzero rows means the vectors are linearly independent.

3. Calculate \( \det A \). \( \det A = 0 \) implies that the vectors are linearly dependent; \( \det A \neq 0 \) implies that the vectors are linearly independent.

Options 1 and 2 are essentially equivalent; the difference being that in option 1 the vectors appear as columns. Option 2 requires a little less writing so we’ll use it.

\[ A = \begin{pmatrix} 2 & 0 & -1 & 4 \\ 2 & -1 & 0 & 2 \\ -2 & 4 & -3 & 4 \\ 1 & -1 & 3 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & 3 & 0 \\ 0 & 1 & -6 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \] (verify this).

Therefore the vectors are linearly dependent. You can also check that \( \det A = 0 \).

(c) Determine whether \( v_1, v_2, v_3 \) are linearly dependent or linearly independent.

SOLUTION Calculating a determinant is not an option here; three vectors with four components do not form a square matrix. We’ll row reduce

\[ A = \begin{pmatrix} 2 & 0 & -1 & 4 \\ 2 & -1 & 0 & 2 \\ -2 & 4 & -3 & 4 \end{pmatrix} \]

As you can verify,

\[ \begin{pmatrix} 2 & 0 & -1 & 4 \\ 2 & -1 & 0 & 2 \\ -2 & 4 & -3 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & -1 & 4 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \]
Therefore the vectors are linearly dependent.

(d) Determine whether the vectors \( \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4, \mathbf{v}_5 \) are linearly dependent or linearly independent.

**SOLUTION** You can verify that

\[
\begin{vmatrix}
2 & 0 & -1 & 4 \\
2 & -1 & 0 & 2 \\
1 & -1 & 3 & 0 \\
0 & 1 & -5 & 3 \\
\end{vmatrix} = -5.
\]

Therefore the vectors are linearly independent. ■

In general, suppose that \( \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \) is a set of vectors in \( \mathbb{R}^n \):

1. If \( k > n \), the vectors are linearly dependent.

2. If \( k = n \), write the matrix \( A \) having \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \) as rows. Either reduce \( A \) to row echelon form, or calculate \( \det A \). A row of zeros or \( \det A = 0 \) implies that the vectors are linearly dependent; all rows nonzero or \( \det A \neq 0 \) implies that the vectors are linearly independent.

3. If \( k < n \), write the matrix \( A \) having \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \) as rows and reduce to row echelon form. A row of zeros implies that the vectors are linearly dependent; all rows nonzero implies that the vectors are linearly independent.

**Another look at systems of linear equations**

Consider the system of linear equations

\[
\begin{align*}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m
\end{align*}
\]
Note that we can write this system as the vector equation

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

which is

\[x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_n \mathbf{v}_n = \mathbf{b}\]

where

\[
\mathbf{v}_1 = \begin{bmatrix}
a_{11} \\
a_{21} \\
\vdots \\
a_{m1}
\end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix}
a_{12} \\
a_{22} \\
\vdots \\
a_{m2}
\end{bmatrix}, \quad \ldots, \quad \mathbf{v}_n = \begin{bmatrix}
a_{1n} \\
a_{2n} \\
\vdots \\
a_{mn}
\end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix}
b_1 \\
b_2 \\
\vdots \\
b_m
\end{bmatrix}.
\]

Written in this form, the question of solving the system of equations can be interpreted as asking whether or not the vector \( \mathbf{b} \) can be written as a linear combination of the vectors \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \). As we know, \( \mathbf{b} \) may be written uniquely as a linear combination of \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) (the system has a unique solution); \( \mathbf{b} \) may not be expressible as a linear combination of \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) (the system has no solution); or it may be possible to represent \( \mathbf{b} \) as a linear combination of \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) in infinitely many different ways (the system has infinitely many solutions).

**Linear Dependence and Linear Independence of Functions**

As we saw in Chapter 3, two functions, \( f \) and \( g \), are linearly dependent if one is a multiple of the other; otherwise they are linearly independent.

**DEFINITION 2.** Let \( f_1, f_2, f_3, \ldots, f_n \) be functions defined on an interval \( I \). The functions are *linearly dependent* if there exist \( n \) real numbers \( c_1, c_2, \ldots, c_n \), not all zero, such that

\[c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \cdots + c_n f_n(x) \equiv 0;
\]

that is,

\[c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \cdots + c_n f_n(x) = 0 \quad \text{for all} \quad x \in I.
\]

Otherwise the functions are *linearly independent*.

Equivalently, the functions \( f_1, f_2, f_3, \ldots, f_n \) are linearly independent if

\[c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) + \cdots + c_n f_n(x) \equiv 0
\]

only when \( c_1 = c_2 = \cdots = c_n = 0 \).
Example 5. Let \( f_1(x) = 1, f_2(x) = x, f_3(x) = x^2 \) on \( I = (-\infty, \infty) \). Show that \( f_1, f_2, f_3 \) are linearly independent.

**SOLUTION** Suppose that the functions are linearly dependent. Then there exist three numbers \( c_1, c_2, c_3 \), not all zero, such that

\[
    c_1 1 + c_2 x + c_3 x^2 \equiv 0. \tag{B}
\]

**Method 1** Since (B) holds for all \( x \), we’ll let \( x = 0 \). Then we have

\[
    c_1 + c_2 (0) + c_3 (0) = 0 \quad \text{which implies} \quad c_1 = 0.
\]

Since \( c_1 = 0 \), (B) becomes

\[
    c_2 x + c_3 x^2 \equiv 0 \quad \text{or} \quad x [c_2 + c_3 x] \equiv 0.
\]

Since \( x \) is not identically zero, we must have \( c_2 + c_3 x \equiv 0 \). Letting \( x = 0 \), we have \( c_2 = 0 \). Finally, \( c_1 = c_2 = 0 \) implies \( c_3 = 0 \). This contradicts our assumption that \( f_1, f_2, f_3 \) are linearly dependent. Thus, the functions are linearly independent.

**Method 2** Since (B) holds for all \( x \), we’ll evaluate at \( x = 0, x = 1, x = 2 \). This yields the system of equations

\[
    \begin{align*}
    c_1 &= 0 \\
    c_1 + c_2 + c_3 &= 0 \\
    c_1 + 2c_2 + 4c_3 &= 0.
    \end{align*}
\]

It is easy to verify that the only solution of this system is \( c_1 = c_2 = c_3 = 0 \). Thus, the functions are linearly independent.

**Method 3** Our functions are differentiable, so we’ll differentiate (B) twice to get

\[
    \begin{align*}
    c_1 1 + c_2 x + c_3 x^2 &\equiv 0 \\
    c_2 + 2c_3 x &\equiv 0 \\
    2c_3 &\equiv 0.
    \end{align*}
\]

From the last equation, \( c_3 = 0 \). Substituting \( c_3 = 0 \) in the second equation gives \( c_2 = 0 \). Substituting \( c_2 = 0 \) and \( c_3 = 0 \) in the first equation gives \( c_1 = 0 \). Thus (B) holds only when \( c_1 = c_2 = c_3 = 0 \), which implies that the functions are linearly independent.

Example 6. Let \( f_1(x) = \sin x, f_2(x) = \cos x, f_3(x) = \sin (x - \frac{1}{6} \pi) \), \( x \in (-\infty, \infty) \). Are these functions linearly dependent or linearly independent?

**SOLUTION** By the addition formula for the sine function

\[
    \sin (x - \frac{1}{6} \pi) = \sin x \cos \frac{1}{6} \pi - \cos x \sin \frac{1}{6} \pi = \frac{1}{2} \sqrt{3} \sin x - \frac{1}{2} \cos x,
\]

Since \( f_3 \) is a linear combination of \( f_1 \) and \( f_2 \), we can conclude that \( f_1, f_2, f_3 \) are linearly dependent on \( (-\infty, \infty) \).
A Test for Linear Independence of Functions; Wronskian

Our test for linear independence is an extension of Method 3 in Example 5.

**THEOREM 1.** Suppose that the functions  $f_1, f_2, f_3, \ldots, f_n$ are $n-1$-times differentiable on an interval $I$. If the functions are linearly dependent on $I$, then

\[
\begin{vmatrix}
  f_1(x) & f_2(x) & \cdots & f_n(x) \\
  f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\
  f_1''(x) & f_2''(x) & \cdots & f_n''(x) \\
  \vdots & \vdots & & \vdots \\
  f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x)
\end{vmatrix} \equiv 0 \text{ on } I.
\]

**Proof** Since the functions are linearly dependent on $I$, there exist $n$ numbers $c_1, c_2, \ldots, c_n$, not all zero, such that

\[c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) \equiv 0.\]

Differentiating this equation $n-1$ times, we get the system of equations

\[
\begin{align*}
  c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x) & \equiv 0 \\
  c_1 f_1'(x) + c_2 f_2'(x) + \cdots + c_n f_n'(x) & \equiv 0 \\
  c_1 f_1''(x) + c_2 f_2''(x) + \cdots + c_n f_n''(x) & \equiv 0 \\
  \vdots & \vdots \\
  c_1 f_1^{(n-1)}(x) + c_2 f_2^{(n-1)}(x) + \cdots + c_n f_n^{(n-1)}(x) & \equiv 0
\end{align*}
\]

(A)

Choose any point $a \in I$ and consider the system of $n$ equations in $n$ unknowns $z_1, z_2, \ldots, z_n$:

\[
\begin{align*}
  f_1(a) z_1 + f_2(a) z_2 + \cdots + f_n(a) z_n &= 0 \\
  f_1'(a) z_1 + f_2'(a) z_2 + \cdots + f_n'(a) z_n &= 0 \\
  f_1''(a) z_1 + f_2''(a) z_2 + \cdots + f_n''(a) z_n &= 0 \\
  \vdots & \vdots \\
  f_1^{(n-1)}(a) z_1 + f_2^{(n-1)}(a) z_2 + \cdots + f_n^{(n-1)}(a) z_n &= 0
\end{align*}
\]

This is a homogeneous system which, from (A), has a nontrivial solution $c_1, c_2, \ldots, c_n$. Therefore, as we showed in Section 5.6,

\[
\begin{vmatrix}
  f_1(a) & f_2(a) & \cdots & f_n(a) \\
  f_1'(a) & f_2'(a) & \cdots & f_n'(a) \\
  f_1''(a) & f_2''(a) & \cdots & f_n''(a) \\
  \vdots & \vdots & & \vdots \\
  f_1^{(n-1)}(a) & f_2^{(n-1)}(a) & \cdots & f_n^{(n-1)}(a)
\end{vmatrix} = 0.
\]

230
Since a was any point on I, we conclude that the determinant is zero for all points in I.

Recall from Chapter 3 that the determinant

\[
\begin{vmatrix}
  f_1 & f_2 \\
  f'_1 & f'_2 \\
\end{vmatrix}
\]

is called the Wronskian of \( f_1, f_2 \). The same terminology is used here.

**DEFINITION 3.** Suppose that the functions \( f_1, f_2, f_3, \ldots, f_n \) are \( n-1 \)-times differentiable on an interval \( I \). The determinant

\[
W(x) = \begin{vmatrix}
  f_1(x) & f_2(x) & \cdots & f_n(x) \\
  f'_1(x) & f'_2(x) & \cdots & f'_n(x) \\
  f''_1(x) & f''_2(x) & \cdots & f''_n(x) \\
  \vdots & \vdots & \ddots & \vdots \\
  f^{(n-1)}_1(x) & f^{(n-1)}_2(x) & \cdots & f^{(n-1)}_n(x) \\
\end{vmatrix}
\]

is called the Wronskian of \( f_1, f_2, f_3, \ldots, f_n \).

Theorem 1 can be stated equivalently as:

**COROLLARY** Suppose that the functions \( f_1, f_2, f_3, \ldots, f_n \) are \((n-1)\)-times differentiable on an interval \( I \) and let \( W(x) \) be their Wronskian. If \( W(x) \neq 0 \) for at least one \( x \in I \), then the functions are linearly independent on \( I \).

This is a useful test for determining the linear independence of a set of functions.

**Example 7.** Show that the functions \( f_1(x) \equiv 1, f_2(x) = x, f_3(x) = x^2, f_4(x) = x^3 \) are linearly independent.

**SOLUTION** These functions are three-times differentiable on \(( -\infty, \infty )\). Their Wronskian is

\[
W(x) = \begin{vmatrix}
  1 & x & x^2 & x^3 \\
  0 & 1 & 2x & 3x^2 \\
  0 & 0 & 2 & 6x \\
  0 & 0 & 0 & 6 \\
\end{vmatrix} = 12.
\]

Since \( W \neq 0 \), the functions are linearly independent.

**Note:** You can use the Wronskian to show that any set of distinct powers of \( x \) is a linearly independent set.

**Caution** Theorem 1 says that if a set of (sufficiently differentiable) functions is linearly dependent on an interval \( I \), then their Wronskian is identically zero on \( I \). The theorem...
does not say that if the Wronskian of a set of functions is identically zero on some interval, then the functions are linearly dependent on that interval. Here is an example of a pair of functions which are linearly independent and whose Wronskian is identically zero.

**Example 8.** Let \( f(x) = x^2 \) and let
\[
g(x) = \begin{cases} 
-x^2 & -2 < x < 0 \\
x^2 & 0 \leq x < 2 
\end{cases}
\]
on \((-2, 2)\). The only question is whether \( g \) is differentiable at 0. You can verify that it is. Thus we can form their Wronskian:

For \( x \geq 0 \),
\[
W(x) = \begin{vmatrix} 
x^2 & x^2 \\
2x & 2x 
\end{vmatrix} \equiv 0.
\]

For \( x < 0 \),
\[
W(x) = \begin{vmatrix} 
x^2 & -x^2 \\
2x & -2x 
\end{vmatrix} \equiv 0.
\]

Thus, \( W(x) \equiv 0 \) on \((-2, 2)\).

We can state that \( f \) and \( g \) are linearly independent because neither is a constant multiple of the other (\( f = g \) on \([0, 2)\), \( f = -g \) on \((-2, 0))\). Another way to see this is:

Suppose that \( f \) and \( g \) are linearly dependent. Then there exist two numbers \( c_1, c_2 \), not both zero, such that
\[
c_1 f(x) + c_2 g(x) \equiv 0 \quad \text{on} \quad (-2, 2).
\]

If we evaluate this identity at \( x = 1 \) and \( x = -1 \), we get the pair of equations
\[
c_1 + c_2 = 0 \\
c_1 - c_2 = 0
\]
The only solution of this pair of equations is \( c_1 = c_2 = 0 \). Thus, \( f \) and \( g \) are linearly independent.

**Exercises 1.7**

1. Show that any set of vectors in \( \mathbb{R}^n \) that contains the zero vector is a linearly dependent set. Hint: Show that the set \( \{0, v_2, v_3, \ldots, v_k\} \) is linearly dependent.

2. Show that if the set \( \{v_1, v_2, v_3, \ldots, v_k\} \) is linearly dependent, then one of the vectors can be written as a linear combination of the others.

Determine whether the set of vectors is linearly dependent or linearly independent. If it is linearly dependent, express one of the vectors as a linear combination of the others.
15. Consider the matrix
\[ A = \begin{pmatrix} 1 & 2 & -1 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \]
which is row-echelon form. Show that the row vectors of \( A \) are a linear independent set. Are the nonzero row vectors of any matrix in row-echelon form linearly independent?

16. Let \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) be linearly independent. Prove that \( \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 - \mathbf{v}_2 \) are linearly independent.

17. Let \( S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \) be a linearly independent set of vectors. Prove that every non-empty subset of \( S \) is also linearly independent. Suppose that \( T = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\} \) is a linearly dependent set. Is every non-empty subset of \( T \) linearly dependent?

Calculate the Wronskian of the set of functions. Then determine whether the functions are linearly dependent or linearly independent.

18. \( f_1(x) = e^{ax}, f_2(x) = e^{bx}, \ a \neq b; \ x \in (-\infty, \infty). \)

19. \( f_1(x) = \sin ax, f_2(x) = \cos ax, \ a \neq 0; \ x \in (-\infty, \infty). \)

20. \( f_1(x) = x, f_2(x) = x^2, f_3(x) = x^3; \ x \in (-\infty, \infty). \)

21. \( f_1(x) = 1, f_2(x) = x^{-1}, f_3(x) = x^{-2}; \ x \in (0, \infty). \)

22. \( f_1(x) = x - x^2, f_2(x) = x^2 - 3x, f_3(x) = 2x + x^2; \ x \in (-\infty, \infty). \)
23. \( f_1(x) = x, \ f_2(x) = e^x, \ f_3(x) = xe^x; \ x \in (-\infty, \infty). \)

24. \( f_1(x) = e^x, \ f_2(x) = e^{-x}, \ f_3(x) = e^{2x}, \ f_4(x) = e^{3x}; \ x \in (-\infty, \infty). \)

25. [a] If the Wronskian of a set of functions is identically zero on an interval \( I, \) then the functions are linearly dependent on \( I. \) True or false?

[b] If a set of functions is linearly dependent on an interval \( I, \) then the Wronskian of the functions is identically zero on \( I. \) True or false?

[c] If the Wronskian of a set of functions is nonzero at some points of an interval \( I \) and zero at other points, then the functions are linearly independent on \( I. \) True or false?

26. Show that the functions \( f_0(x) = 1, \ f_1(x) = x, \ f_2(x) = x^2, \ldots, \ f_k(x) = x^k \) are linearly independent.
1.8 Eigenvalues and Eigenvectors

Let $A$ be an $n \times n$ matrix. Then, as we noted in Section 5.5, $A$ can be viewed as a transformation that maps an $n$-component vector to an $n$ component vector; that is, $A$ can be thought of as a function from $\mathbb{R}^n$ into $\mathbb{R}^n$. Here’s an example.

**Example 1.** Let $A$ be the $3 \times 3$ matrix

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

Then $A$ maps the vector $(2, -1, 3)$ to $(11, 3, 9)$:

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
2 \\
-1 \\
3
\end{pmatrix} =
\begin{pmatrix}
11 \\
3 \\
9
\end{pmatrix};
$$

$A$ maps the vector $(-1, 2, 3)$ to the vector $(10, 7, -6)$:

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
-1 \\
2 \\
3
\end{pmatrix} =
\begin{pmatrix}
10 \\
7 \\
-6
\end{pmatrix};
$$

and, in general,

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
c
\end{pmatrix} =
\begin{pmatrix}
2a + 2b + 3c \\
a + 2b + c \\
2a - 2b + c
\end{pmatrix}.
$$

Now consider the vector $v = (8, 5, 2)$. It has a special property relative to $A$:

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
8 \\
5 \\
2
\end{pmatrix} =
\begin{pmatrix}
32 \\
20 \\
8
\end{pmatrix} =
\begin{pmatrix}
8 \\
5 \\
2
\end{pmatrix};
$$

$A$ maps $v = (8, 5, 2)$ to a multiple of itself; $Av = 4v$.

You can also verify that

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
2 \\
3 \\
-2
\end{pmatrix} =
2
\begin{pmatrix}
2 \\
3 \\
-2
\end{pmatrix}
$$

and

$$
\begin{pmatrix}
2 & 2 & 3 \\
1 & 2 & 1 \\
2 & -2 & 1
\end{pmatrix}
\begin{pmatrix}
1 \\
0 \\
-1
\end{pmatrix} =
-1
\begin{pmatrix}
1 \\
0 \\
-1
\end{pmatrix}.
$$

This is the property that we will study in this section.

**DEFINITION 1.** Let $A$ be an $n \times n$ matrix. A number $\lambda$ is an *eigenvalue* of $A$ if there exists a *nonzero* vector $v$ in $\mathbb{R}^n$ such that

$$Av = \lambda v.$$

The vector $v$ is called an *eigenvector* corresponding to $\lambda$. 

235
In Example 1, the numbers 4, 2, −1 are eigenvalues of $A$ and $v = (8, 5, 2)$, $u = (2, 3, −2)$, $w = (1, 0, −1)$ are corresponding eigenvectors.

Eigenvalues (vectors) are also called characteristic values (vectors) or proper values (vectors).

Calculating the Eigenvalues of a Matrix

Let $A$ be an $n \times n$ matrix and suppose that $\lambda$ is an eigenvalue of $A$ with corresponding eigenvector $v$. Then

$$Av = \lambda v \quad \text{implies} \quad Av - \lambda v = 0.$$ 

The latter equation can be written

$$(A - \lambda I_n)v = 0.$$ 

This equation says that the homogeneous system of linear equations

$$(A - \lambda I_n)x = 0$$

has the nontrivial solution $v$. As we saw in Section 5.6, a homogeneous system of $n$ linear equations in $n$ unknowns has a nontrivial solution if and only if the determinant of the matrix of coefficients is zero. Thus, $\lambda$ is an eigenvalue of $A$ if and only if

$$\det(A - \lambda I_n) = 0.$$ 

The equation $\det(A - \lambda I_n) = 0$ is called the characteristic equation of the matrix $A$. The roots of the characteristic equation are the eigenvalues of $A$.

We’ll start with a special case which we’ll illustrate with a $3 \times 3$ matrix.

**Example 2.** Let $A$ be an upper triangular matrix:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ 0 & b_2 & b_3 \\ 0 & 0 & c_3 \end{pmatrix}.\$$

Then the eigenvalues of $A$ are the entries on the main diagonal.

$$\det(A - \lambda I_3) = \begin{vmatrix} a_1 - \lambda & a_2 & a_3 \\ 0 & b_2 - \lambda & b_3 \\ 0 & 0 & c_3 - \lambda \end{vmatrix} = (a_1 - \lambda)(b_2 - \lambda)(c_3 - \lambda).$$

Thus, the eigenvalues of $A$ are: $\lambda_1 = a_1$, $\lambda_2 = b_2$, $\lambda_3 = c_3$.

The general result is this: If $A$ is either an upper triangular matrix or a lower triangular matrix, then the eigenvalues of $A$ are the entries on the main diagonal of $A$.

Now we’ll look at arbitrary matrices; that is, no special form.
Example 3. Find the eigenvalues of

\[ A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}. \]

**SOLUTION** We need to find the numbers \( \lambda \) that satisfy the equation \( \det(A - \lambda I_2) = 0 \):

\[ A - \lambda I_2 = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \lambda & -3 \\ -2 & 2 - \lambda \end{pmatrix} \]

and

\[ \begin{vmatrix} 1 - \lambda & -3 \\ -2 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda) - 6 = \lambda^2 - 3\lambda - 4 = (\lambda + 1)(\lambda - 4). \]

Therefore the eigenvalues of \( A \) are \( \lambda_1 = -1 \) and \( \lambda_2 = 4 \).

Example 4. Find the eigenvalues of

\[ A = \begin{pmatrix} 1 & -3 & 1 \\ -1 & 1 & 1 \\ 3 & -3 & -1 \end{pmatrix}. \]

**SOLUTION** We need to find the numbers \( \lambda \) that satisfy the equation \( \det(A - \lambda I_3) = 0 \):

\[ A - \lambda I_3 = \begin{pmatrix} 1 & -3 & 1 \\ -1 & 1 & 1 \\ 3 & -3 & -1 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 - \lambda & -3 & 1 \\ -1 & 1 - \lambda & 1 \\ 3 & -3 & -1 - \lambda \end{pmatrix} \]

and

\[ \det A = \begin{vmatrix} 1 - \lambda & -3 & 1 \\ -1 & 1 - \lambda & 1 \\ 3 & -3 & -1 - \lambda \end{vmatrix}. \]

Expanding across the first row (remember, you can go across any row or down any column), we have

\[ \begin{vmatrix} 1 - \lambda & -3 & 1 \\ -1 & 1 - \lambda & 1 \\ 3 & -3 & -1 - \lambda \end{vmatrix} = (1 - \lambda)[(1 - \lambda)(-1 - \lambda) + 3] + 3[1 + \lambda - 3] + [3 - 3(1 - \lambda)] \]

\[ = -\lambda^3 + \lambda^2 + 4\lambda - 4 = -(\lambda + 2)(\lambda - 1)(\lambda - 2). \]

Therefore the eigenvalues of \( A \) are \( \lambda_1 = -2 \), \( \lambda_2 = 1 \), \( \lambda_3 = 2 \).

Note that the characteristic equation of our \( 2 \times 2 \) matrix is a polynomial equation of degree 2, a quadratic equation; the characteristic equation of our \( 3 \times 3 \) matrix is a
polynomial equation of degree 3; a cubic equation. This is true in general; the characteristic equation of an \( n \times n \) matrix \( A \) is

\[
\begin{vmatrix}
  a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\
  a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{n1} & a_{2n} & \cdots & a_{nn} - \lambda \\
\end{vmatrix} = p(\lambda) = 0,
\]

where \( p \) is a polynomial of degree \( n \) in \( \lambda \). The polynomial \( p \) is called the characteristic polynomial of \( A \).

Recall the following facts about a polynomial \( p \) of degree \( n \) with real coefficients.

1. \( p \) has exactly \( n \) roots, counting multiplicities.
2. \( p \) may have complex roots, but if \( a + bi \) is a root, then its conjugate \( a - bi \) is also a root; the complex roots of \( p \) occur in conjugate pairs, counting multiplicities.
3. \( p \) can be factored into a product of linear and quadratic factors — the linear factors corresponding to the real roots of \( p \) and the quadratic factors corresponding to the complex roots.

**Example 5.** Calculate the eigenvalues of

\[
A = \begin{pmatrix}
  1 & -1 \\
  4 & 1 \\
\end{pmatrix}.
\]

**SOLUTION**

\[
\begin{vmatrix}
  1 - \lambda & -1 \\
  4 & 1 - \lambda \\
\end{vmatrix} = (1 - \lambda)^2 + 4 = \lambda^2 - 2\lambda + 5.
\]

The roots of \( \lambda^2 - 2\lambda + 5 = 0 \) are \( \lambda_1 = 1 + 2i, \lambda_2 = 1 - 2i; \) \( A \) has complex eigenvalues.

**Example 6.** Calculate the eigenvalues of

\[
A = \begin{pmatrix}
  1 & -3 & 3 \\
  3 & -5 & 3 \\
  6 & -6 & 4 \\
\end{pmatrix}.
\]

**SOLUTION**

\[
\begin{vmatrix}
  1 - \lambda & -3 & 3 \\
  3 & -5 - \lambda & 3 \\
  6 & -6 & 4 - \lambda \\
\end{vmatrix} = -\lambda^3 + 12\lambda + 16 = -\lambda(\lambda + 2)^2(\lambda - 4).
\]

The eigenvalues of \( A \) are \( \lambda_1 = \lambda_2 = -2, \lambda_3 = 4; \) \(-2\) is an eigenvalue of multiplicity 2.
Calculating the Eigenvectors of a Matrix

We note first that if \( \lambda \) is an eigenvalue of an \( n \times n \) matrix \( A \) and \( v \) is a corresponding eigenvector, then any nonzero multiple of \( v \) is also an eigenvector of \( A \) corresponding to \( \lambda \): if \( Av = \lambda v \) and \( r \) is any nonzero number, then

\[
A(rv) = rAv = r(\lambda v) = \lambda rv.
\]

Thus, each eigenvalue has infinitely many corresponding eigenvectors.

Let \( A \) be an \( n \times n \) matrix. If \( \lambda \) is an eigenvalue of \( A \) with corresponding eigenvector \( v \), then \( v \) is a solution of the homogeneous system of equations

\[
(A - \lambda I_n)x = 0. \tag{1}
\]

Therefore, to find an eigenvector of \( A \) corresponding to the eigenvalue \( \lambda \), we need to find a nontrivial solution of (1). This takes us back to the solution methods of Sections 5.3 and 5.4.

**Example 7.** Find the eigenvalues and corresponding eigenvectors of

\[
\begin{pmatrix}
2 & 4 \\
1 & -1
\end{pmatrix}.
\]

**SOLUTION** The first step is to find the eigenvalues.

\[
\det(A - \lambda I_2) = \begin{vmatrix}
2 - \lambda & 4 \\
1 & -1 - \lambda
\end{vmatrix} = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2).
\]

The eigenvalues of \( A \) are \( \lambda_1 = 3, \lambda_2 = -2 \).

Next we find an eigenvector for each eigenvalue. Equation (1) for this problem is

\[
\begin{pmatrix}
2 - \lambda & 4 \\
1 & -1 - \lambda
\end{pmatrix}x = \begin{pmatrix}
2 - \lambda & 4 \\
1 & -1 - \lambda
\end{pmatrix} \begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} = \begin{pmatrix} 0, 0 \end{pmatrix}. \tag{2}
\]

We have to deal with each eigenvalue separately. We set \( \lambda = 3 \) in (2) to get

\[
\begin{pmatrix}
-1 & 4 \\
1 & -4
\end{pmatrix} \begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} = 0.
\]

The augmented matrix for this system of equations is

\[
\begin{pmatrix}
-1 & 4 & 0 \\
1 & -4 & 0
\end{pmatrix}
\]

which row reduces to

\[
\begin{pmatrix}
1 & -4 & 0 \\
0 & 0 & 0
\end{pmatrix}.
\]
The solution set is \( x_1 = 4a, \ x_2 = a, \ a \) any real number.

We get an eigenvector by choosing a value for \( a \). Since an eigenvector is, by definition, a nonzero vector, we must choose \( a \neq 0 \). Any such \( a \) will do; we’ll let \( a = 1 \). Then, an eigenvector corresponding to the eigenvalue \( \lambda_1 = 3 \) is \( \mathbf{v}_1 = (4, 1) \). Here is a verification:

\[
\begin{pmatrix}
2 & 4 \\
1 & -1
\end{pmatrix}
\begin{pmatrix}
4 \\
1
\end{pmatrix} =
\begin{pmatrix}
12 \\
3
\end{pmatrix} = 3 \begin{pmatrix}
4 \\
1
\end{pmatrix}.
\]

Now we set \( \lambda = -2 \) in (2) to get

\[
\begin{pmatrix}
4 & 4 \\
1 & 1
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2
\end{pmatrix} = 0.
\]

The augmented matrix for this system of equations is

\[
\begin{pmatrix}
4 & 4 & 0 \\
1 & 1 & 0
\end{pmatrix}
\]

which row reduces to

\[
\begin{pmatrix}
1 & 1 & 0 \\
0 & 0 & 0
\end{pmatrix}.
\]

The solution set is \( x_1 = -a, \ x_2 = a, \ a \) any real number. Again, to get an eigenvector corresponding to \( \lambda = -2 \), we can choose any nonzero number for \( a \); we’ll choose \( a = -1 \). Then, an eigenvector corresponding to the eigenvalue \( \lambda_2 = -2 \) is \( \mathbf{v}_2 = (1, -1) \). We leave it to you to verify that \( A\mathbf{v}_2 = -2\mathbf{v}_2 \).

**NOTE:** It is important to understand that in finding eigenvectors we can assign any nonzero value to the parameter in the solution set of the system of equations \( (A - \lambda I)\mathbf{x} = \mathbf{0} \). Typically we’ll choose values which will avoid fractions in the eigenvector and, just because it reads better, we like to have the first component of an eigenvector be non-negative. Such choices are certainly not required.

**Example 8.** Find the eigenvalues and corresponding eigenvectors of

\[
\begin{pmatrix}
2 & 2 & 2 \\
-1 & 2 & 1 \\
1 & -2 & -1
\end{pmatrix}.
\]

**SOLUTION** First we find the eigenvalues.

\[
\det(A - \lambda I) =
\begin{vmatrix}
2 - \lambda & 2 & 2 \\
-1 & 2 - \lambda & 1 \\
1 & -2 & -1 - \lambda
\end{vmatrix} = \lambda^3 - 3\lambda^2 + 2\lambda = \lambda(\lambda - 1)(\lambda - 2).
\]

The eigenvalues of \( A \) are \( \lambda_1 = 0, \ \lambda_2 = 1, \ \lambda_3 = 2 \).
Next we find an eigenvector for each eigenvalue. Equation (1) for this problem is

\[
\begin{pmatrix}
2 - \lambda & 2 & 2 \\
-1 & 2 - \lambda & 1 \\
1 & -2 & -1 - \lambda
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
= 0.
\tag{3}
\]

We have to deal with each eigenvalue separately. We set \( \lambda = 0 \) in (3) to get

\[
\begin{pmatrix}
2 & 2 & 2 \\
-1 & 2 & 1 \\
1 & -2 & -1
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
= 0.
\]

The augmented matrix for this system of equations is

\[
\begin{pmatrix}
2 & 2 & 2 & 0 \\
-1 & 2 & 1 & 0 \\
1 & -2 & -1 & 0
\end{pmatrix}
\]

which row reduces to

\[
\begin{pmatrix}
1 & 1 & 1 & 0 \\
0 & 1 & 2/3 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

The solution set is \( x_1 = -\frac{1}{3}a, \ x_2 = -\frac{2}{3}a, \ x_3 = a, \ a \) any real number.

We get an eigenvector by choosing a value for \( a \). We’ll choose \( a = -3 \) (this avoids having fractions as components, a convenience). Thus, an eigenvector corresponding to the eigenvalue \( \lambda_1 = 0 \) is \( \mathbf{v}_1 = (1, 2, -3) \). Here is a verification:

\[
\begin{pmatrix}
2 & 2 & 2 \\
-1 & 2 & 1 \\
1 & -2 & -1
\end{pmatrix}
\begin{pmatrix}
1 \\
2 \\
-3
\end{pmatrix}
= 0
\]

Next we set \( \lambda = 1 \) in (3) to get

\[
\begin{pmatrix}
1 & 2 & 2 \\
-1 & 1 & 1 \\
1 & -2 & -2
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix}
= 0.
\]

The augmented matrix for this system of equations is

\[
\begin{pmatrix}
1 & 2 & 2 & 0 \\
-1 & 1 & 1 & 0 \\
1 & -2 & -2 & 0
\end{pmatrix}
\]

which row reduces to

\[
\begin{pmatrix}
1 & 2 & 2 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]
The solution set is \( x_1 = 0, \ x_2 = -a, \ x_3 = a, \ a \) any real number. If we let \( a = -1 \), we get the eigenvector \( v_2 = (0, \ 1, \ -1) \). You can verify that \( Av_2 = v_2 \).

Finally, we set \( \lambda = 2 \) in (3) to get

\[
\begin{pmatrix}
1 & 2 & 2 \\
-1 & 1 & 1 \\
1 & -2 & -2
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = 0.
\]

The augmented matrix for this system of equations is

\[
\begin{pmatrix}
0 & 2 & 2 & 0 \\
-1 & 0 & 1 & 0 \\
1 & -2 & -3 & 0
\end{pmatrix}
\]

which row reduces to

\[
\begin{pmatrix}
1 & 0 & -1 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

The solution set is \( x_1 = a, \ x_2 = -a, \ x_3 = a, \ a \) any real number. If we let \( a = 1 \), we get the eigenvector \( v_3 = (1, \ -1, \ 1) \). You can verify that \( Av_3 = v_3 \).

In Example 6 we found the eigenvectors \( v_1 = (4, \ 1) \) and \( v_2 = (1, \ -1) \) corresponding to the eigenvalues \( \lambda_1 = 3 \) and \( \lambda_2 = -2 \), respectively. Notice that \( v_1 \) and \( v_2 \) are linearly independent.

In Example 7 we found the eigenvectors \( v_1 = (1, \ 2, \ -3), \ v_2 = (0, \ 1, \ -1), \ v_3 = (1, \ -1, \ 1) \) corresponding to the eigenvalues \( \lambda_1 = 0, \ \lambda_2 = 1, \ \lambda_3 = 2 \). As you can check, these vectors are linearly independent.

The general result is this:

**THEOREM 1.** If \( v_1, \ v_2, \ldots, \ v_k \) are eigenvectors of a matrix \( A \) corresponding to distinct eigenvalues \( \lambda_1, \ \lambda_2, \ldots, \ \lambda_k \), then \( v_1, \ v_2, \ldots, \ v_k \) are linearly independent.

We’ll prove this result for two vectors. The method can be extended by induction to a set of \( k \) vectors.

Let \( v_1 \) and \( v_2 \) be eigenvectors of a matrix \( A \) corresponding to eigenvalues \( \lambda_1 \) and \( \lambda_2 \), \( \lambda_1 \neq \lambda_2 \). Suppose that \( v_1, \ v_2 \) are linearly dependent. Then there exist two numbers \( c_1, \ c_2 \), not both zero, such that

\[
c_1v_1 + c_2v_2 = 0. \quad (a)
\]

Since \( \lambda_1 \) and \( \lambda_2 \) are distinct, at least one of them is nonzero; assume \( \lambda_2 \neq 0 \) and multiply (a) by \( \lambda_2 \). We get

\[
c_1\lambda_2v_1 + c_2\lambda_2v_2 = 0. \quad (b)
\]

242
Now “multiply” (a) by $A$. This gives

$$A(c_1v_1 + c_2v_2) = c_1Av_1 + c_2Av_2 = c_1\lambda_1v_1 + c_2\lambda_2v_2 = A0 = 0$$

and we have

$$c_1\lambda_1v_1 + c_2\lambda_2v_2 = 0.$$  \hspace{1cm} (c)

Subtracting (c) from (b), we get

$$c_1(\lambda_2 - \lambda_1)v_1 = 0.$$  

Since $\lambda_2 - \lambda_1 \neq 0$ and $v_1 \neq 0$, it follows that $c_1 = 0$. Substituting $c_1 = 0$ into (a) gives $c_2 = 0$. Thus, $v_1, v_2$ cannot be linearly dependent; they must be linearly independent.

This is an important result for the next chapter.

Even though we have restricted our treatment of differential equations and linear algebra to real-valued functions and real numbers, the need to consider complex numbers does arise. We saw this in our treatment of second order linear differential equations with constant coefficients. The need to consider complex numbers arises here as well — the characteristic polynomial of an $n \times n$ matrix $A$ may have complex roots. Of course, since the characteristic polynomial has real coefficients, complex roots occur in conjugate pairs. Also, we have to expect that an eigenvector corresponding to a complex eigenvalue will be complex; that is, have complex number components. Here is the main theorem in this regard.

**Theorem 2.** Let $A$ be a (real) $n \times n$ matrix. If the complex number $\lambda = a + bi$ is an eigenvalue of $A$ with corresponding (complex) eigenvector $u + iv$, then $\lambda = a - bi$, the conjugate of $a + bi$, is also an eigenvalue of $A$ and $u - iv$ is a corresponding eigenvector.

The proof is a simple application of complex arithmetic:

\[
\begin{align*}
A(u + iv) &= (a + bi)(u + iv) \\
\overline{A(u + iv)} &= \overline{(a + bi)(u + iv)} \hspace{1cm} \text{(the overline denotes complex conjugate)} \\
A(u + iv) &= \overline{(a + bi)} \overline{(u + iv)} \\
A(u - iv) &= (a - bi)(u - iv)
\end{align*}
\]

**Example 9.** As we saw in Example 4, the matrix

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

has the complex eigenvalues $\lambda_1 = 1 + 2i$, $\lambda_2 = 1 - 2i$. We’ll find the eigenvectors. The nice thing about a pair of complex eigenvalues is that we only have to calculate one eigenvector. Equation (1) for this problem is

$$\begin{pmatrix} 1 - \lambda & -1 \\ 4 & 1 - \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0.$$  

243
Substituting \( \lambda_1 = 1 + 2i \) in this equation gives
\[
\begin{pmatrix} -2i & -1 \\ 4 & -2i \end{pmatrix} \mathbf{x} = \begin{pmatrix} -2i & -1 \\ 4 & -2i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\]

The augmented matrix for this system of equations is
\[
\begin{pmatrix} -2i & -1 & 0 \\ 4 & -2i & 0 \end{pmatrix}
\]
which row reduces to
\[
\begin{pmatrix} 1 & -\frac{1}{2}i & 0 \\ 0 & 0 & 0 \end{pmatrix}
\]
The solution set is \( x_1 = \frac{1}{2}ia, \ x_2 = a, \ a \) any number; in this case, either real or complex.

If we set \( x_2 = -2i \) we get the eigenvector \( \mathbf{v}_1 = (1, -2i) = (1, 0) + i(0, -2) \):
\[
\begin{pmatrix} 1 & -1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2i \end{pmatrix} = \begin{pmatrix} 1 + 2i \\ 4 - 2i \end{pmatrix} = (1 + 2i) \begin{pmatrix} 1 \\ -2i \end{pmatrix}.
\]

Now, by Theorem 2, an eigenvector corresponding to \( \lambda_2 = 1 - 2i \) is \( \mathbf{v}_2 = (1, 0) - i(0, -2) \).

As we will see in the next chapter, eigenvalues of multiplicity greater than 1 can cause complications in solving linear differential systems. In the next two examples we'll illustrate what can happen when an eigenvalue has multiplicity 2.

**Example 10.** Find the eigenvalues and eigenvectors of
\[
A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}
\]

**SOLUTION** First we find the eigenvalues:
\[
\det(A - \lambda I_3) = \begin{vmatrix} 1 - \lambda & -3 & 3 \\ 3 & -5\lambda & 3 \\ 6 & -6 & 4 - \lambda \end{vmatrix} = 16 + 12\lambda - \lambda^3 = -(\lambda - 4)(\lambda + 2)^2.
\]
The eigenvalues of \( A \) are \( \lambda_1 = 4, \ \lambda_2 = \lambda_3 = -2; \ -2 \) is an eigenvalue of multiplicity 2.

You can verify that \( \mathbf{v}_1 = (1, 1, 2) \) is an eigenvector corresponding to \( \lambda_1 = 4 \).

Now we investigate what happens with the eigenvalue \(-2\):
\[
(A - (-2)I_3)\mathbf{x} = \begin{pmatrix} 3 & -3 & 3 \\ 3 & -3 & 3 \\ 6 & -6 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.
\]
The augmented matrix for this system of equations is
\[
\begin{pmatrix}
3 & -3 & 3 & 0 \\
3 & -3 & 3 & 0 \\
6 & -6 & 6 & 0 \\
\end{pmatrix}
\]
which row reduces to
\[
\begin{pmatrix}
1 & -1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}.
\]
The solution set of the corresponding system of equations is \(x_1 = a - b, a, b\) any real numbers. We can assign any values we want to \(a\) and \(b\), except \(a = b = 0\) (an eigenvector is a non-zero vector). Setting \(a = 1, b = 0\) gives the eigenvector \(v_2 = (1, 1, 0)\); setting \(a = 0, b = -1\) gives the eigenvector \(v_3 = (1, 0, -1)\).

Note that our two choices of \(a\) and \(b\) produced two linearly independent eigenvectors. The fact that the solution set of the system of equations (a) had two independent parameters guarantees that we can obtain two independent eigenvectors.

**Example 11.** Find the eigenvalues and eigenvectors of
\[
A = \begin{pmatrix}
5 & 6 & 2 \\
0 & -1 & -8 \\
1 & 0 & -2 \\
\end{pmatrix}.
\]

**SOLUTION** First we find the eigenvalues:
\[
\det(A - \lambda I_3) = \begin{vmatrix}
5 - \lambda & 6 & 2 \\
0 & -1 - \lambda & -8 \\
1 & 0 & -2 - \lambda \\
\end{vmatrix} = -36 + 15\lambda + 2\lambda^2 - \lambda^3 = -(\lambda + 4)(\lambda - 3)^2.
\]
The eigenvalues of \(A\) are \(\lambda_1 = -4, \lambda_2 = \lambda_3 = 3\); \(3\) is an eigenvalue of multiplicity 2.

You can verify that \(v_1 = (6, -8, -3)\) is an eigenvector corresponding to \(\lambda_1 = -4\).

We now find the eigenvectors for \(3\):
\[
(A - 3I_3)x = \begin{pmatrix}
2 & 6 & 2 \\
0 & -4 & -8 \\
1 & 0 & -5 \\
\end{pmatrix} \begin{pmatrix}
x_1 \\
x_2 \\
x_3 \\
\end{pmatrix} = \begin{pmatrix}
0 \\
0 \\
0 \\
\end{pmatrix}.
\]
The augmented matrix for this system of equations is
\[
\begin{pmatrix}
2 & 6 & 2 & 0 \\
0 & -4 & -8 & 0 \\
1 & 0 & -5 & 0 \\
\end{pmatrix}
\]
which row reduces to
\[
\begin{pmatrix}
1 & 3 & 1 & 0 \\
0 & 1 & 2 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}.
\]
The solution set of the corresponding system of equations is \( x_1 = 5a, \ x_2 = -2a, \ x_3 = a, \ a \) any real number. Setting \( a = 1 \) gives the eigenvector \( v_2 = (5, -2, 1) \)

In this case, the eigenvalue of multiplicity two yielded only one (independent) eigenvector.

In general, if the matrix \( A \) has an eigenvalue \( \lambda \) of multiplicity \( k \), then \( \lambda \) may have only one (independent) eigenvector, it may have two independent eigenvectors, it may have three independent eigenvectors, and so on, up to \( k \) independent eigenvectors. It can be shown using Theorem 2 that \( \lambda \) cannot have more than \( k \) linearly independent eigenvectors.

**Eigenvalues, Determinant, Inverse, Rank**

There is a relationship between the eigenvalues of an \( n \times n \) matrix \( A \), the determinant of \( A \), the existence of \( A^{-1} \), and the rank of \( A \).

**THEOREM 3.** Let \( A \) be an \( n \times n \) matrix with eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \). (Note: the \( \lambda \)'s here are not necessarily distinct, one or more of the eigenvalues may have multiplicity greater than 1, and they are not necessarily real.) Then

\[
\det A = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdots \lambda_n.
\]

That is, \( \det A \) is the product of the eigenvalues of \( A \).

**Proof:** The eigenvalues of \( A \) are the roots of the characteristic polynomial
\[
\det(A - \lambda I) = p(\lambda).
\]

Writing \( p(\lambda) \) in factored form, we have
\[
\det(A - \lambda I) = p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)(\lambda_3 - \lambda) \cdots (\lambda_n - \lambda).
\]

Setting \( \lambda = 0 \), we get
\[
\det A = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdots \lambda_n.
\]

At the end of Section 5.6 we listed equivalences between the determinant, existence of an inverse and rank. With Theorem 3, we can add eigenvalues to the list of equivalences.

Let \( A \) be an \( n \times n \) matrix. The following are equivalent:

1. The system of equations \( Ax = b \) has a unique solution.
2. \( A \) has an inverse.
3. \( \det A \neq 0 \).
4. \( A \) has rank \( n \).
5. 0 is not an eigenvalue of \( A \).

**Exercises 1.8**

Determine the eigenvalues and the eigenvectors.

1. \( A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} \).
2. \( A = \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} \).
3. \( A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \).
4. \( A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \).
5. \( A = \begin{pmatrix} 6 & 5 \\ -5 & -4 \end{pmatrix} \).
6. \( A = \begin{pmatrix} -1 & 1 \\ 4 & 2 \end{pmatrix} \).
7. \( A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \).
8. \( A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \).
9. \( A = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \).
10. \( A = \begin{pmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{pmatrix} \). Hint: 1 is an eigenvalue.
11. \( A = \begin{pmatrix} 15 & 7 & -7 \\ -1 & 1 & 1 \\ 13 & 7 & -5 \end{pmatrix} \). Hint: 2 is an eigenvalue.
12. \( A = \begin{pmatrix} 2 & -2 & 1 \\ 1 & -1 & 1 \\ -3 & 2 & -2 \end{pmatrix} \). Hint: 1 is an eigenvalue.

13. \( A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ -2 & -2 & -1 \end{pmatrix} \). Hint: 1 is an eigenvalue.

14. \( A = \begin{pmatrix} 1 & -3 & 1 \\ 0 & 1 & 0 \\ 3 & 0 & -1 \end{pmatrix} \). Hint: 1 is an eigenvalue.

15. \( A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \).

16. \( A = \begin{pmatrix} 1 & -4 & -1 \\ 3 & 2 & 3 \\ 1 & 1 & 3 \end{pmatrix} \). Hint: 2 is an eigenvalue.

17. \( A = \begin{pmatrix} -1 & 1 & 2 \\ -1 & 1 & 1 \\ -2 & 1 & 3 \end{pmatrix} \).

18. \( A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix} \). Hint: 5 is an eigenvalue.

19. \( A = \begin{pmatrix} 2 & -4 & 2 \\ -4 & 2 & -2 \\ 2 & -2 & -1 \end{pmatrix} \). Hint: -2 is an eigenvalue.

20. \( A = \begin{pmatrix} 4 & 1 & -1 \\ 2 & 5 & -2 \\ 1 & 1 & 2 \end{pmatrix} \). Hint: 3 is an eigenvalue.

21. \( A = \begin{pmatrix} 4 & 2 & -2 & 2 \\ 1 & 3 & 1 & -1 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -3 & 5 \end{pmatrix} \). Hint: 2 is an eigenvalue of multiplicity 2.

22. \( A = \begin{pmatrix} 3 & 5 & -5 & 5 \\ 3 & 1 & 3 & -3 \\ -2 & 2 & 0 & 2 \\ 0 & 4 & -6 & 8 \end{pmatrix} \). Hint: 2 and -2 are eigenvalues.
23. Prove that if $\lambda$ is an eigenvalue of $A$, then for every positive integer $k$, $\lambda^k$ is an eigenvalue of $A^k$.

24. Suppose that $\lambda$ is a nonzero eigenvalue of $A$ with corresponding eigenvector $v$. Prove that if $A$ has an inverse, then $1/\lambda$ is an eigenvalue of $A^{-1}$ with corresponding vector $v$. 