

1. Systems of Linear Equations

1.1 Solutions and Elementary Operations

Practical problems in many fields of study—such as biology, business, chemistry, computer science, economics, electronics, engineering, physics and the social sciences—can often be reduced to solving a system of linear equations. Linear algebra arose from attempts to find systematic methods for solving these systems, so it is natural to begin this book by studying linear equations.

If a , b , and c are real numbers, the graph of an equation of the form

$$ax + by = c$$

is a straight line (if a and b are not both zero), so such an equation is called a *linear* equation in the variables x and y . However, it is often convenient to write the variables as x_1, x_2, \dots, x_n , particularly when more than two variables are involved. An equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

is called a **linear equation** in the n variables x_1, x_2, \dots, x_n . Here a_1, a_2, \dots, a_n denote real numbers (called the **coefficients** of x_1, x_2, \dots, x_n , respectively) and b is also a number (called the **constant term** of the equation). A finite collection of linear equations in the variables x_1, x_2, \dots, x_n is called a **system of linear equations** in these variables. Hence,

$$2x_1 - 3x_2 + 5x_3 = 7$$

is a linear equation; the coefficients of x_1, x_2 , and x_3 are 2, -3 , and 5, and the constant term is 7. Note that each variable in a linear equation occurs to the first power only.

Given a linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$, a sequence s_1, s_2, \dots, s_n of n numbers is called a **solution** to the equation if

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

that is, if the equation is satisfied when the substitutions $x_1 = s_1, x_2 = s_2, \dots, x_n = s_n$ are made. A sequence of numbers is called a **solution to a system** of equations if it is a solution to every equation in the system.

For example, $x = -2, y = 5, z = 0$ and $x = 0, y = 4, z = -1$ are both solutions to the system

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

A system may have no solution at all, or it may have a unique solution, or it may have an infinite family of solutions. For instance, the system $x + y = 2, x + y = 3$ has no solution because the sum of two numbers cannot be 2 and 3 simultaneously. A system that has no solution is called **inconsistent**; a system with at least one solution is called **consistent**. The system in the following example has infinitely many solutions.

Example 1.1.1

Show that, for arbitrary values of s and t ,

$$\begin{aligned}x_1 &= t - s + 1 \\x_2 &= t + s + 2 \\x_3 &= s \\x_4 &= t\end{aligned}$$

is a solution to the system

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

Solution. Simply substitute these values of x_1 , x_2 , x_3 , and x_4 in each equation.

$$\begin{aligned}x_1 - 2x_2 + 3x_3 + x_4 &= (t - s + 1) - 2(t + s + 2) + 3s + t = -3 \\2x_1 - x_2 + 3x_3 - x_4 &= 2(t - s + 1) - (t + s + 2) + 3s - t = 0\end{aligned}$$

Because both equations are satisfied, it is a solution for all choices of s and t .

The quantities s and t in Example 1.1.1 are called **parameters**, and the set of solutions, described in this way, is said to be given in **parametric form** and is called the **general solution** to the system. It turns out that the solutions to *every* system of equations (if there *are* solutions) can be given in parametric form (that is, the variables x_1, x_2, \dots are given in terms of new independent variables s, t , etc.). The following example shows how this happens in the simplest systems where only one equation is present.

Example 1.1.2

Describe all solutions to $3x - y + 2z = 6$ in parametric form.

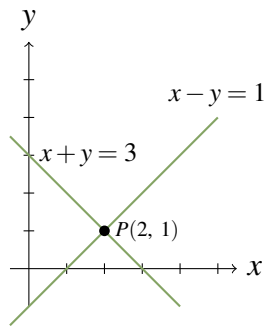
Solution. Solving the equation for y in terms of x and z , we get $y = 3x + 2z - 6$. If s and t are arbitrary then, setting $x = s$, $z = t$, we get solutions

$$\begin{aligned}x &= s \\y &= 3s + 2t - 6 \quad s \text{ and } t \text{ arbitrary} \\z &= t\end{aligned}$$

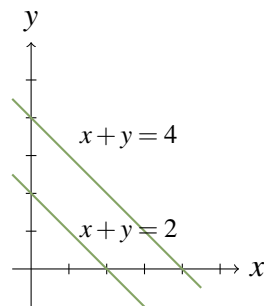
Of course we could have solved for x : $x = \frac{1}{3}(y - 2z + 6)$. Then, if we take $y = p$, $z = q$, the solutions are represented as follows:

$$\begin{aligned}x &= \frac{1}{3}(p - 2q + 6) \\y &= p \\z &= q\end{aligned} \quad p \text{ and } q \text{ arbitrary}$$

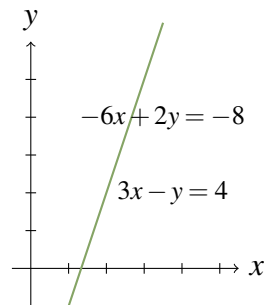
The same family of solutions can “look” quite different!



(a) Unique Solution
($x = 2, y = 1$)



(b) No Solution



(c) Infinitely many solutions
($x = t, y = 3t - 4$)

Figure 1.1.1

When only two variables are involved, the solutions to systems of linear equations can be described geometrically because the graph of a linear equation $ax + by = c$ is a straight line if a and b are not both zero. Moreover, a point $P(s, t)$ with coordinates s and t lies on the line if and only if $as + bt = c$ —that is when $x = s, y = t$ is a solution to the equation. Hence the solutions to a *system* of linear equations correspond to the points $P(s, t)$ that lie on *all* the lines in question.

In particular, if the system consists of just one equation, there must be infinitely many solutions because there are infinitely many points on a line. If the system has two equations, there are three possibilities for the corresponding straight lines:

1. *The lines intersect at a single point. Then the system has a unique solution corresponding to that point.*
2. *The lines are parallel (and distinct) and so do not intersect. Then the system has no solution.*
3. *The lines are identical. Then the system has infinitely many solutions—one for each point on the (common) line.*

These three situations are illustrated in Figure 1.1.1. In each case the graphs of two specific lines are plotted and the corresponding equations are indicated. In the last case, the equations are $3x - y = 4$ and $-6x + 2y = -8$, which have identical graphs.

With three variables, the graph of an equation $ax + by + cz = d$ can be shown to be a plane (see Section 4.2) and so again provides a “picture” of the set of solutions. However, this graphical method has its limitations: When more than three variables are involved, no physical image of the graphs (called hyperplanes) is possible. It is necessary to turn to a more “algebraic” method of solution.

Before describing the method, we introduce a concept that simplifies the computations involved. Consider the following system

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

of three equations in four variables. The array of numbers¹

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

occurring in the system is called the **augmented matrix** of the system. Each row of the matrix consists of the coefficients of the variables (in order) from the corresponding equation, together with the constant

¹A rectangular array of numbers is called a **matrix**. Matrices will be discussed in more detail in Chapter 2.

term. For clarity, the constants are separated by a vertical line. The augmented matrix is just a different way of describing the system of equations. The array of coefficients of the variables

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

is called the **coefficient matrix** of the system and $\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$ is called the **constant matrix** of the system.

Elementary Operations

The algebraic method for solving systems of linear equations is described as follows. Two such systems are said to be **equivalent** if they have the same set of solutions. A system is solved by writing a series of systems, one after the other, each equivalent to the previous system. Each of these systems has the same set of solutions as the original one; the aim is to end up with a system that is easy to solve. Each system in the series is obtained from the preceding system by a simple manipulation chosen so that it does not change the set of solutions.

As an illustration, we solve the system $x + 2y = -2$, $2x + y = 7$ in this manner. At each stage, the corresponding augmented matrix is displayed. The original system is

$$\begin{array}{l} x + 2y = -2 \\ 2x + y = 7 \end{array} \quad \left[\begin{array}{cc|c} 1 & 2 & -2 \\ 2 & 1 & 7 \end{array} \right]$$

First, subtract twice the first equation from the second. The resulting system is

$$\begin{array}{l} x + 2y = -2 \\ -3y = 11 \end{array} \quad \left[\begin{array}{cc|c} 1 & 2 & -2 \\ 0 & -3 & 11 \end{array} \right]$$

which is equivalent to the original (see Theorem 1.1.1). At this stage we obtain $y = -\frac{11}{3}$ by multiplying the second equation by $-\frac{1}{3}$. The result is the equivalent system

$$\begin{array}{l} x + 2y = -2 \\ y = -\frac{11}{3} \end{array} \quad \left[\begin{array}{cc|c} 1 & 2 & -2 \\ 0 & 1 & -\frac{11}{3} \end{array} \right]$$

Finally, we subtract twice the second equation from the first to get another equivalent system.

$$\begin{array}{l} x = \frac{16}{3} \\ y = -\frac{11}{3} \end{array} \quad \left[\begin{array}{cc|c} 1 & 0 & \frac{16}{3} \\ 0 & 1 & -\frac{11}{3} \end{array} \right]$$

Now *this* system is easy to solve! And because it is equivalent to the original system, it provides the solution to that system.

Observe that, at each stage, a certain operation is performed on the system (and thus on the augmented matrix) to produce an equivalent system.

Definition 1.1 Elementary Operations

The following operations, called **elementary operations**, can routinely be performed on systems of linear equations to produce equivalent systems.

- I. Interchange two equations.
- II. Multiply one equation by a nonzero number.
- III. Add a multiple of one equation to a different equation.

Theorem 1.1.1

Suppose that a sequence of elementary operations is performed on a system of linear equations. Then the resulting system has the same set of solutions as the original, so the two systems are equivalent.

The proof is given at the end of this section.

Elementary operations performed on a system of equations produce corresponding manipulations of the *rows* of the augmented matrix. Thus, multiplying a row of a matrix by a number k means multiplying *every entry* of the row by k . Adding one row to another row means adding *each entry* of that row to the corresponding entry of the other row. Subtracting two rows is done similarly. Note that we regard two rows as equal when corresponding entries are the same.

In hand calculations (and in computer programs) we manipulate the rows of the augmented matrix rather than the equations. For this reason we restate these elementary operations for matrices.

Definition 1.2 Elementary Row Operations

The following are called **elementary row operations** on a matrix.

- I. Interchange two rows.
- II. Multiply one row by a nonzero number.
- III. Add a multiple of one row to a different row.

In the illustration above, a series of such operations led to a matrix of the form

$$\left[\begin{array}{cc|c} 1 & 0 & * \\ 0 & 1 & * \end{array} \right]$$

where the asterisks represent arbitrary numbers. In the case of three equations in three variables, the goal is to produce a matrix of the form

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

This does not always happen, as we will see in the next section. Here is an example in which it does happen.

Example 1.1.3

Find all solutions to the following system of equations.

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

Solution. The augmented matrix of the original system is

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

To create a 1 in the upper left corner we could multiply row 1 through by $\frac{1}{3}$. However, the 1 can be obtained without introducing fractions by subtracting row 2 from row 1. The result is

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

The upper left 1 is now used to “clean up” the first column, that is create zeros in the other positions in that column. First subtract 2 times row 1 from row 2 to obtain

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

Next subtract 4 times row 1 from row 3. The result is

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

This completes the work on column 1. We now use the 1 in the second position of the second row to clean up the second column by subtracting row 2 from row 1 and then adding row 2 to row 3. For convenience, both row operations are done in one step. The result is

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & -7 & -8 \end{array} \right]$$

Note that the last two manipulations *did not affect* the first column (the second row has a zero there), so our previous effort there has not been undermined. Finally we clean up the third column. Begin by multiplying row 3 by $-\frac{1}{7}$ to obtain

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & 1 & \frac{8}{7} \end{array} \right]$$

Now subtract 3 times row 3 from row 1, and then add 2 times row 3 to row 2 to get

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{3}{7} \\ 0 & 1 & 0 & \frac{2}{7} \\ 0 & 0 & 1 & \frac{8}{7} \end{array} \right]$$

The corresponding equations are $x = -\frac{3}{7}$, $y = \frac{2}{7}$, and $z = \frac{8}{7}$, which give the (unique) solution.

Every elementary row operation can be **reversed** by another elementary row operation of the same type (called its **inverse**). To see how, we look at types I, II, and III separately:

Type I Interchanging two rows is reversed by interchanging them again.

Type II Multiplying a row by a nonzero number k is reversed by multiplying by $1/k$.

Type III Adding k times row p to a different row q is reversed by adding $-k$ times row p to row q (in the new matrix). Note that $p \neq q$ is essential here.

To illustrate the Type III situation, suppose there are four rows in the original matrix, denoted R_1 , R_2 , R_3 , and R_4 , and that k times R_2 is added to R_3 . Then the reverse operation adds $-k$ times R_2 to R_3 . The following diagram illustrates the effect of doing the operation first and then the reverse:

$$\begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{bmatrix} \rightarrow \begin{bmatrix} R_1 \\ R_2 \\ R_3 + kR_2 \\ R_4 \end{bmatrix} \rightarrow \begin{bmatrix} R_1 \\ R_2 \\ (R_3 + kR_2) - kR_2 \\ R_4 \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{bmatrix}$$

The existence of inverses for elementary row operations and hence for elementary operations on a system of equations, gives:

Proof of Theorem 1.1.1. Suppose that a system of linear equations is transformed into a new system by a sequence of elementary operations. Then every solution of the original system is automatically a solution of the new system because adding equations, or multiplying an equation by a nonzero number, always results in a valid equation. In the same way, each solution of the new system must be a solution to the original system because the original system can be obtained from the new one by another series of elementary operations (the inverses of the originals). It follows that the original and new systems have the same solutions. This proves Theorem 1.1.1. \square

Exercises for 1.1

Exercise 1.1.1 In each case verify that the following are solutions for all values of s and t .

a. $x = 19t - 35$
 $y = 25 - 13t$
 $z = t$

is a solution of

$$\begin{aligned} 2x + 3y + z &= 5 \\ 5x + 7y - 4z &= 0 \end{aligned}$$

b. $x_1 = 2s + 12t + 13$
 $x_2 = s$
 $x_3 = -s - 3t - 3$
 $x_4 = t$

is a solution of

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

Exercise 1.1.2 Find all solutions to the following in parametric form in two ways.

a. $3x + y = 2$ b. $2x + 3y = 1$
 c. $3x - y + 2z = 5$ d. $x - 2y + 5z = 1$

Exercise 1.1.3 Regarding $2x = 5$ as the equation $2x + 0y = 5$ in two variables, find all solutions in parametric form.

Exercise 1.1.4 Regarding $4x - 2y = 3$ as the equation $4x - 2y + 0z = 3$ in three variables, find all solutions in parametric form.

Exercise 1.1.5 Find all solutions to the general system $ax = b$ of one equation in one variable (a) when $a = 0$ and (b) when $a \neq 0$.

Exercise 1.1.6 Show that a system consisting of exactly one linear equation can have no solution, one solution, or infinitely many solutions. Give examples.

Exercise 1.1.7 Write the augmented matrix for each of the following systems of linear equations.

a. $x - 3y = 5$ b. $x + 2y = 0$
 $2x + y = 1$ $y = 1$
 c. $x - y + z = 2$ d. $x + y = 1$
 $x - z = 1$ $y + z = 0$
 $y + 2x = 0$ $z - x = 2$

Exercise 1.1.8 Write a system of linear equations that has each of the following augmented matrices.

a. $\left[\begin{array}{ccc|c} 1 & -1 & 6 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & -1 & 0 & 1 \end{array} \right]$ b. $\left[\begin{array}{ccc|c} 2 & -1 & 0 & -1 \\ -3 & 2 & 1 & 0 \\ 0 & 1 & 1 & 3 \end{array} \right]$

Exercise 1.1.9 Find the solution of each of the following systems of linear equations using augmented matrices.

a. $x - 3y = 1$ b. $x + 2y = 1$
 $2x - 7y = 3$ $3x + 4y = -1$
 c. $2x + 3y = -1$ d. $3x + 4y = 1$
 $3x + 4y = 2$ $4x + 5y = -3$

Exercise 1.1.10 Find the solution of each of the following systems of linear equations using augmented matrices.

a. $x + y + 2z = -1$ b. $2x + y + z = -1$
 $2x + y + 3z = 0$ $x + 2y + z = 0$
 $-2y + z = 2$ $3x - 2z = 5$

Exercise 1.1.11 Find all solutions (if any) of the following systems of linear equations.

a. $3x - 2y = 5$ b. $3x - 2y = 5$
 $-12x + 8y = -20$ $-12x + 8y = 16$

Exercise 1.1.12 Show that the system

$$\begin{cases} x + 2y - z = a \\ 2x + y + 3z = b \\ x - 4y + 9z = c \end{cases}$$

is inconsistent unless $c = 2b - 3a$.

Exercise 1.1.13 By examining the possible positions of lines in the plane, show that two equations in two variables can have zero, one, or infinitely many solutions.

Exercise 1.1.14 In each case either show that the statement is true, or give an example² showing it is false.

- If a linear system has n variables and m equations, then the augmented matrix has n rows.
- A consistent linear system must have infinitely many solutions.
- If a row operation is done to a consistent linear system, the resulting system must be consistent.
- If a series of row operations on a linear system results in an inconsistent system, the original system is inconsistent.

Exercise 1.1.15 Find a quadratic $a + bx + cx^2$ such that the graph of $y = a + bx + cx^2$ contains each of the points $(-1, 6)$, $(2, 0)$, and $(3, 2)$.

Exercise 1.1.16 Solve the system $\begin{cases} 3x + 2y = 5 \\ 7x + 5y = 1 \end{cases}$ by changing variables $\begin{cases} x = 5x' - 2y' \\ y = -7x' + 3y' \end{cases}$ and solving the resulting equations for x' and y' .

Exercise 1.1.17 Find a , b , and c such that

$$\frac{x^2 - x + 3}{(x^2 + 2)(2x - 1)} = \frac{ax + b}{x^2 + 2} + \frac{c}{2x - 1}$$

[Hint: Multiply through by $(x^2 + 2)(2x - 1)$ and equate coefficients of powers of x .]

Exercise 1.1.18 A zookeeper wants to give an animal 42 mg of vitamin A and 65 mg of vitamin D per day. He has two supplements: the first contains 10% vitamin A and 25% vitamin D; the second contains 20% vitamin A and 25% vitamin D. How much of each supplement should he give the animal each day?

Exercise 1.1.19 Workmen John and Joe earn a total of \$24.60 when John works 2 hours and Joe works 3 hours. If John works 3 hours and Joe works 2 hours, they get \$23.90. Find their hourly rates.

Exercise 1.1.20 A biologist wants to create a diet from fish and meal containing 183 grams of protein and 93 grams of carbohydrate per day. If fish contains 70% protein and 10% carbohydrate, and meal contains 30% protein and 60% carbohydrate, how much of each food is required each day?

1.2 Gaussian Elimination

The algebraic method introduced in the preceding section can be summarized as follows: Given a system of linear equations, use a sequence of elementary row operations to carry the augmented matrix to a “nice” matrix (meaning that the corresponding equations are easy to solve). In Example 1.1.3, this nice matrix took the form

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

The following definitions identify the nice matrices that arise in this process.

²Such an example is called a **counterexample**. For example, if the statement is that “all philosophers have beards”, the existence of a non-bearded philosopher would be a counterexample proving that the statement is false. This is discussed again in Appendix B.