Part I

Systems of Linear Equations

Section 1

Introduction to Systems of Linear Equations

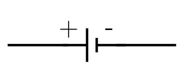
Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a linear equation?
- What is a system of linear equations?
- What is a solution set of a system of linear equations?
- What are equivalent systems of linear equations?
- What operations can we use to solve a system of linear equations?

Application: Electrical Circuits

Linear algebra is concerned with the study of systems of linear equations. There are two important aspects to linear systems. One is to use given information to set up a system of equations that represents the information (this is called *modeling*), and the other is to solve the system. As an example of modeling, we consider the application to the very simple electrical circuit. An electrical circuit consists of

- one or more electrical sources, denoted by 

- one or more resistors, denoted by .

A source is a power supply like a battery, and a resistor is an object that consumes the electricity, like a lamp or a computer. A simple circuit consists of one or more sources connected to resistors, like the one shown in Figure 1.2. The straight lines in the circuit indicate wires through which current flows. The points labeled P and Q are called *junctions* or *nodes*.

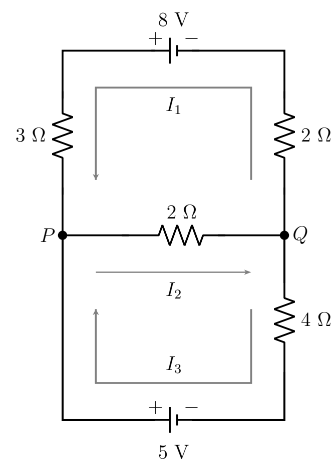


Figure 1.1: A circuit.

The source creates a charge that produces potential energy E measured in volts (V). Current flows out of the positive terminal of a source and runs through each branch of the circuit. Let I_1, I_2, and I_3 be the currents as illustrated in Figure 1.2. The goal is to find the current flowing in each branch of the circuit.

Linear algebra comes into play when analyzing a circuit based on the relationship between current I, resistance R, and voltage E. There are laws governing electrical circuits that state that $E = IR$ across a resistor. Additionally, Kirchoff's Current and Voltage Laws indicate how current behaves within the whole circuit. Using all these laws together, we derive the system

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 5I_1 + 2I_2 &= 8 \\ 2I_2 + 4I_3 &= 5, \end{aligned}$$

where I_1, I_2, and I_3 are the currents at the points indicated in Figure 1.2. To finish analyzing the circuit, we now need to solve this system. In this section we will begin to learn systematic methods for solving systems of linear equations. More details about the derivation of these circuit equations can be found at the end of this section.

Introduction

Systems of linear equations arise in almost every field of study: mathematics, statistics, physics, chemistry, biology, economics, sociology, computer science, engineering, and many, many others. We will study the theory behind solving systems of linear equations, implications of this theory, and applications of linear algebra as we proceed throughout this text.

Preview Activity 1.1.

(1) Consider the following system of two linear equations in two unknowns, x_1, x_2:

$$\begin{aligned} 2x_1 - 3x_2 &= 0 \\ x_1 - x_2 &= 1. \end{aligned}$$

One way to solve such a system of linear equations is the method of substitution (where one equation is solved for one variable and then the resulting expression is substituted into the remaining equations). This method works well for simple systems of two equations in two unknowns, but becomes complicated if the number or complexity of the equations is increased.

Another method is elimination – the method that we will adopt in this book. Recall that the elimination method works by multiplying each equation by a suitable constant so that the coefficients of one of the variables in each equation is the same. Then we subtract corresponding sides of these equations to eliminate that variable.

Use the method of elimination to show that this system has the unique solution $x_1 = 3$ and $x_2 = 2$. Explain the specific steps you perform when using elimination.

(2) Recall that a linear equation in two variables can be represented as a line in $\mathbb{R}^2$, the Cartesian plane, where one variable corresponds to the horizontal axis and the other to the vertical axis. Represent the two equations $2x_1 - 3x_2 = 0$ and $x_1 - x_2 = 1$ in $\mathbb{R}^2$ and illustrate the solution to the system in your picture.

(3) The previous example should be familiar to you as a system of two equations in two unknowns. Now we consider a system of three equations in three unknowns

$$I_1 - I_2 + I_3 = 0 \tag{1.1}$$
$$5I_1 + 2I_2 = 8 \tag{1.2}$$
$$2I_2 + 4I_3 = 5. \tag{1.3}$$

that arises from our electrical circuit in Figure 1.2, with currents I_1, I_2, and I_3 as indicated in the circuit. In the remainder of this preview activity we will apply the method of elimination to solve the system of linear equations (1.1), (1.2), and (1.3).

(a) Replace equation (1.2) with the new equation obtained by multiplying both sides of equation (1.1) by 5 and then subtracting corresponding sides of this equation from the appropriate sides of equation (1.2). Show that the resulting system is

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 7I_2 - 5I_3 &= 8 \\ 2I_2 + 4I_3 &= 5. \end{aligned} \tag{1.4}$$

(b) Now eliminate the variable I_2 from the last two equations in the system in part (a) by using equations (1.3) and (1.4) to show that $I_3 = 0.5$. Explain your process.

(c) Once you know the value for I_3, how can you find I_2? Then how do you find I_1? Use your method to show that the solution to this system is the ordered triple (1,1.5,0.5). Interpret the result in terms of currents.

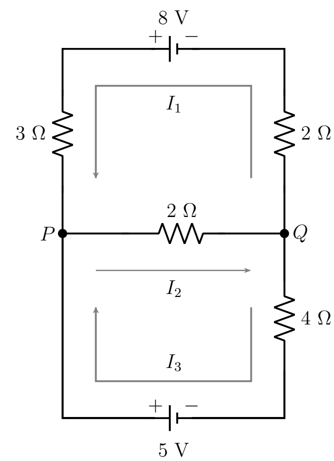


Figure 1.2: A circuit.

Notation and Terminology

To study linear algebra, we will need to agree on some general notation and terminology to represent our systems.

An equation like $4x_1 + x_2 = 8$ is called a linear equation because the variables (x_1 and x_2 in this case) are raised to the first power, and there are no products of variables. The equation $4x_1 + x_2 = 8$ is a linear equation in two variables, but we can make a linear equation with any number of variables we like.

Definition 1.1. A **linear equation** in the variables x_1, x_2, …, x_n is an equation of the form

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

where n is a positive integer and a_1, a_2, …, a_n and b are constants. The constants a_1, a_2, …, a_n are called the **coefficients** of the equation.

We can use any labels for the variables in a linear equation that we like, e.g., I_1, x_1, t_1, and you should become comfortable working with variables in any form. We will usually use subscripts, as in $x_1, x_2, x_3, \ldots$, to represent the variables as this notation allows us to have any number of variables. Other examples of linear equations are

$$x + 2y = 4 \quad \text{and} \quad \sqrt{2}x_1 - 3x_2 = \frac{1}{4}x_3 + \pi\,.$$

On the other hand, the equations

$$\frac{1}{x} + y - z = 0 \quad \text{and} \quad 2x_1 = \sqrt{x_2} - 5$$

are non-linear equations.

Definition 1.2. A **system of linear equations** is a collection of one or more linear equations in the same variables.

For example, the two equations

$$\begin{aligned} x_1 - x_2 &= 1 \\ 2x_1 + x_2 &= 5 \end{aligned} \tag{1.5}$$

form a system of two linear equations in variables x_1, x_2.

Definition 1.3. A **solution** to a system of linear equations is an ordered n-tuple $(s_1, s_2, \ldots, s_n)$ of numbers so that we obtain all true statements in the system when we replace the variable in order with s_1, s_2, $\ldots$, and s_n.

For example, $x_1 = 2, x_2 = 1$, or simply $(2, 1)$, is a solution to the above system of linear equations in (1.5) as can be checked by substituting the variables into each equation. In solving a system of linear equations, we are interested in finding the set of all solutions, which we will call the *solution set of the system*. For the above system in (1.5), the solution set is the set containing the single point $(2, 1)$, denoted $\{(2, 1)\}$, because there is only one solution. If we consider just the equation $x_1 - x_2 = 0$ as our system, the solution set is the line $x_1 = x_2$ in the plane. More generally, a set of solutions is a collection of ordered n-tuples of numbers. We denote the set of all ordered n-tuples of numbers as $\mathbb{R}^n$. So, for example, $\mathbb{R}^2$ is the set of all ordered pairs, or just the standard coordinate plane, and $\mathbb{R}^3$ is the set of all ordered triples, or the three-dimensional space.

Solving Systems of Linear Equations

In Preview Activity 1.1, we were introduced to linear systems and the method of elimination for a system of two or three variables. Our goal now is to come up with a systematic method that will reduce any linear system to one that is easy to solve without changing the solution set of the system. Two linear systems will be called *equivalent* if they have the same solution set.

The operations we used in Preview Activity 1.1 to systematically eliminate variables so that we can solve a linear system are called *elementary operations on a system of linear equations* or just *elementary operations*. In the exercises you will argue that elementary operations do not change the solution set to a system of linear equations, a fact that is summarized in the following theorem.

Theorem 1.4. *The* ***elementary operations*** *on a system of linear equations:*

(1) replacing one equation by the sum of that equation and a scalar multiple of another equation;

(2) interchanging two equations;

(3) replacing an equation by a nonzero scalar multiple of itself;

do not change the solution set to the system of equations.

When we apply these elementary operations our ultimate goal is to produce a system of linear equations in a simplified form with the same solution set, where the number of variables eliminated from the equations increase as we move from top to bottom. This method is called the *elimination* method.

Activity 1.1. For systems of linear equations with a small number of variables, many different methods could be used to find a solution. However, when a system gets large, ad-hoc methods become unwieldy. One of our goals is to develop an algorithmic approach to solving systems of linear equations that can be programmed and applied to any linear system, so we want to work in a very prescribed method as indicated in this activity. Ultimately, once we understand how the algorithm works, we will use calculators/computers to do the work. Apply the elimination method as described to show that the solution set of the following system is $(2, -1, 1)$.

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

(a) Use the first equation to eliminate the variable x_1 in the second and third equations.

(b) Use the new second equation to eliminate the variable x_2 in the third equation and find the value of x_3.

(c) Find values of x_2 and then x_1.

Important Note: Technically, we don't really add two equations or multiply an equation by a scalar. When we refer to a scalar multiple of an equation, we mean the equation obtained by equating the scalar multiple of the expression on the left side of the equation and the same scalar multiple of the expression on the right side of the equation. Similarly, when we refer to a sum of two equations, we don't really add the equations themselves. Instead, we mean the equation obtained by equating the sum of the expressions on the left sides of the equations to the sum of the expressions on the right sides of the equations. We will use the terminology "scalar multiple of an equation" and "sum of two equations" as shorthand to mean what is described here.

Another Important Note: There is an important and subtle point to consider here. When we use these operations to find a solution to a system of equations, we are assuming that the system has a solution. The application of these operations then tells us what a solution must look like. However, there is no guarantee that the outcome is actually a solution – to be safe we should check to make sure that our result is a solution to the system. In the case of linear systems, though, every one of our operations on equations is reversible (if applied correctly), so the result will always be a solution (but this is not true in general for non-linear systems).

Terminology: A system of equations is called *consistent* if the system has at least one solution. If a system has no solutions, then it is said to be *inconsistent*.

The Geometry of Solution Sets of Linear Systems

We are familiar with linear equations in two variables from basic algebra and calculus (through linear approximations). The set of solutions to a system of linear equations in two variables has some geometry connected to it.

Activity 1.2. Recall that we examined the geometry of the system

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

in Preview Activity 1.1 to show that the resulting solution set consists of a single point in the plane.

In this activity we examine the geometry of the system

$$\begin{aligned} 2x_1 - x_2 &= 1 \\ 2x_1 - 2x_2 &= 2. \end{aligned} \tag{1.6}$$

(a) Consider the linear equation $2x_1 - 2x_2 = 2$ (or, equivalently $2x - 2y = 2$). What is the graph of the solution set (the set of points (x_1, x_2) satisfying this equation) of this single equation in the plane? Draw the graph to illustrate.

(b) How can we represent the solution set of the system (1.6) of two equations graphically? How is this solution set related to the solution set of the single equation $2x_1 - 2x_2 = 2$? Why? How many solutions does the system (1.6) have?

(c) There are exactly three possibilities for the number of solutions to a general system of two linear equations in two unknowns. Describe the geometric representations of solution sets for each of the possibilities. Illustrate each with a specific example (of your own) using a system of equations and sketching its geometric representation.

Activity 1.2 shows that there are three options for the solution set of a system: A system can have no solutions, one solution, or infinitely many solutions.

Now we consider systems of three variables. As an example, let us look at the linear equation $x + y + z = 1$ in the three variables x, y, and z. Notice that the points $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ all satisfy this equation. As a linear equation, the graph of $x + y + z = 1$ will be a plane in three dimensions that contains these three points, as shown in Figure 1.3. Hence when we consider a linear system in three unknowns, we are looking for a point in the three dimensional space that lies on all the planes described by the equations.

Activity 1.3. In this activity we examine the geometry of linear systems of three equations in three unknowns. Recall that each linear equation in three variables has a plane as its solution set. Use a piece of paper to represent each plane.

(a) Is it possible for a general system of three linear equations in three unknowns to have no solutions? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.

(b) Is it possible for a general system of three linear equations in three unknowns to have exactly one solution? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.

(c) Is it possible for a general system of three linear equations in three unknowns to have infinitely many solutions? If so, geometrically describe this situation and then illustrate each with a specific example using a system of equations. If not, explain why not.

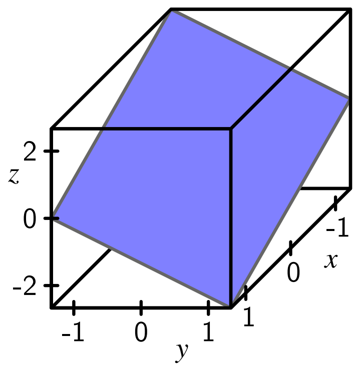


Figure 1.3: The plane $x + y + z = 1$.

Examples

What follows are worked examples that use the concepts from this section.

Example 1.5. Apply the allowable operations on equations to solve the system

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

Example Solution. We begin by eliminating the variable x_1 from all but the first equation. To do so, we replace the third equation with the third equation minus the first equation to obtain the equivalent system

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

Then we replace the fourth equation with the fourth equation minus 2 times the first to obtain the equivalent system

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

To continue the elimination process, we want to eliminate the x_2 variable from our latest third and fourth equations. To do so, we use the second equation so that we do not reinstate an x_1

variable in our new equations. We replace equation three with equation 3 minus 2 times equation 2 to produce the equivalent system

$$\begin{aligned} x_1 + 2x_2 + x_3 - x_4 &= 4 \\ -x_2 - x_3 + 3x_4 &= 6 \\ 3x_3 - 6x_4 &= -15 \\ -7x_2 - x_3 + 3x_4 &= -6. \end{aligned}$$

Then we replace equation four with equation four minus 7 times equation 2, giving us the equivalent system

$$\begin{aligned} x_1 + 2x_2 + x_3 - x_4 &= 4 \\ -x_2 - x_3 + 3x_4 &= 6 \\ 3x_3 - 6x_4 &= -15 \\ 6x_3 - 18x_4 &= -48. \end{aligned}$$

With one more step we can determine the value of x_4. We use the last two equations to eliminate x_3 from the fourth equation by replacing equation four with equation four minus 2 times equation 3. This results in the equivalent system

$$\begin{aligned} x_1 + 2x_2 + x_3 - x_4 &= 4 \\ -x_2 - x_3 + 3x_4 &= 6 \\ 3x_3 - 6x_4 &= -15 \\ -6x_4 &= -18. \end{aligned}$$

The last equation tells us that $-6x_4 = -18$, or $x_4 = 3$. Substituting into the third equation shows that

$$\begin{aligned} 3x_3 - 6(3) &= -15 \\ 3x_3 &= 3 \\ x_3 &= 1. \end{aligned}$$

The second equation shows that

$$\begin{aligned} -x_2 - 1 + 3(3) &= 6 \\ -x_2 &= -2 \\ x_2 &= 2. \end{aligned}$$

Finally, the first equation tells us that

$$\begin{aligned} x_1 + 2(2) + 1 - 3 &= 4 \\ x_1 &= 2. \end{aligned}$$

So the solution to our system is $x_1 = 2$, $x_2 = 2$, $x_3 = 1$, and $x_4 = 3$. It is worth substituting back into our original system to check to make sure that we have not made any arithmetic mistakes.

Example 1.6. A mining company has three mines. One day of operation at the mines produces the following output.

- Mine 1 produces 25 tons of copper, 600 kilograms of silver and 15 tons of manganese.
- Mine 2 produces 30 tons of copper, 500 kilograms of silver and 10 tons of manganese.
- Mine 3 produces 20 tons of copper, 550 kilograms of silver and 12 tons of manganese.

Suppose the company has orders for 550 tons of copper, 11350 kilograms of silver and 250 tons of manganese.

Write a system of equations to answer the question: how many days should the company operate each mine to exactly fill the orders? State clearly what the variables in your system represent. Then find the general solution of your system.

Example Solution. For our system, let x_1 be the number of days mine 1 operates, x_2 be the number of days mine 2 operates, and x_3 be the number of days mine 3 operates. Since mine 1 produces 25 tons of coper each day, in x_1 days mine 1 will produce $25x_1$ tons of copper. Mine 2 produces 30 tons of copper each day, so in x_2 days mine 2 will produce $30x_2$ tons of copper. Also, mine 3 produces 20 tons of copper each day, so in x_3 days mine 3 will produce $20x_3$ tons of copper. Since the company needs to supply a total of 550 tons of copper, we need to have $25x_1 + 30x_2 + 20x_3 = 550$. Similar analyses of silver and manganese give us the system

$$\begin{aligned} 25x_1 + 30x_2 + 20x_3 &= 550 \\ 600x_1 + 500x_2 + 550x_3 &= 11350 \\ 15x_1 + 10x_2 + 12x_3 &= 250. \end{aligned}$$

To solve the system, we eliminate the variable x_2 from the second and third equations by replacing equation two with equation two minus 24 times equation one and replacing equation three with equation three minus $\frac{3}{5}$ times equation one. This produces the equvalent system

$$\begin{aligned} 25x_1 + 30x_2 + 20x_3 &= 550 \\ -220x_2 + 70x_3 &= -1850 \\ -8x_2 &= -80 \quad . \end{aligned}$$

We are fortunate now that we can determine the value of x_2 from the third equation, which tells us that $x_2 = 10$. Substituting into the second equation shows that

$$\begin{aligned} -220(10) + 70x_3 &= -1850 \\ 70x_3 &= 350 \\ x_3 &= 5. \end{aligned}$$

Substituting into the first equation allows us to determine the value for x_1:

$$\begin{aligned} 25x_1 + 30(10) + 20(5) &= 550 \\ 25x_1 &= 150 \\ x_1 &= 6. \end{aligned}$$

So the company should run mine 1 for 6 days, mine 2 for 10 days, and mine 3 for 5 days to meet this demand.

Summary

In this section we introduced linear equations and systems of linear equations.

- Informally, a linear equation is an equation in which each term is either a constant or a constant times a variable. More formally, a linear equation in the variables x_1, x_2, ..., x_n is an equation of the form
$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$
where n is a positive integer and a_1, a_2, ..., a_n and b are constants.
- A system of linear equations is a collection of one or more linear equations in the same variables.
- Informally, a solution to a system of linear equations is a point that satisfies all of the equations in the system. More formally, a solution to a system of linear equation in n variables x_1, x_2, ..., x_n is an ordered n-tuple $(s_1, s_2, \ldots, s_n)$ of numbers so that we obtain all true statements in the system when we replace x_1 with s_1, x_2 with s_2, ..., and x_n with s_n.
- Two linear systems are equivalent if they have the same solution set.
- The following operations on a system of equations do not change the solution set:
  (1) Replace one equation by the sum of that equation and a scalar multiple of another equation.
  (2) Interchange two equations.
  (3) Replace an equation by a nonzero scalar multiple of itself.

Exercises

(1) In the method of elimination there are three operations we can apply to solve a system of linear equations. For this exercise we focus on a system of equations in three unknowns x_1, x_2, and x_3, but the arguments generalize to a system with any number of variables. Consider the general system of three equations in three unknowns

$$\begin{aligned} 4x_1 - 4x_2 + 4x_3 &= 0 \\ 4x_1 + 2x_2 \phantom{{}+ 4x_3} &= 8 \\ 2x_2 + 5x_3 &= 9. \end{aligned}$$

The goal of this exercise is to understand why the three operations on a system do not change the solutions to the system. Recall that a solution to a system with unknowns x_1, x_2, and x_3 is a set of three numbers, one for x_1, one for x_2, and one for x_3 that satisfy all of the equations in the system.

(a) Explain why, if we have a solution to this system, then that solution is also a solution to any constant k times the second equation.

(b) Explain why, if we have a solution to this system, then that solution is also a solution to the sum of the first equation and k times the third equation for any constant k.

(2) Alice stopped by a coffee shop two days in a row at a conference to buy drinks and pastries. On the first day, she bought a cup of coffee and two muffins for which she paid \$6.87. The next day she bought two cups of coffee and three muffins (for herself and a friend). Her bill was \$11.25. Use the method of elimination to determine the price of a cup of coffee, and the price of a muffin. Clearly explain your set-up for the problem (Assume you are explaining your solution to someone who has not solved the problem herself/himself).

(3) Alice stopped by a coffee shop three days in a row at a conference to buy drinks and pastries. On the first day, she bought a cup of coffee, a muffin and a scone for which she paid \$6.15. The next day she bought two cups of coffee, three muffins and a scone (for herself and friends). Her bill was \$12.20. The last day she bought a cup of coffee, two muffins and two scones, and paid \$10.35. Determine the price of a cup of coffee, the price of a muffin and the price of a scone. Clearly explain your set-up for the problem (Assume you are explaining your solution to someone who has not solved the problem herself/himself).

(4) (a) Find an example of a system of two linear equations in variables x, y for each of the following three cases:

   i. where the equations correspond to two non-parallel lines,
   ii. two parallel distinct lines,
   iii. two identical lines (represented with different equations).

   (b) Describe how the relationship between the coefficients of the variables of the two equations in parts (ii) and (iii) are different than the relationship between those coefficients in part (i) (Note: Please make sure your system examples are different than the examples in the activities, and that they are your own examples.)

(5) In a grid of wires in thermal equilibrium, the temperature at interior nodes is the average of the temperatures at adjacent nodes. Consider the grid as shown in Figure 1.4, with x_1, x_2, and x_3 the temperatures (in degrees Centigrade) at the indicated interior nodes, and fixed temperatures at the other nodes as shown. For example, the nodes adjacent to the node with temperature x_1 have temperatures of x_2, 200, 0, and 0, so when the grid is in thermal equilibrium x_1 is the average of these temperatures:

$$x_1 = \frac{x_2 + 200 + 0 + 0}{4}.$$

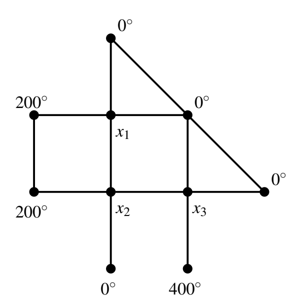


Figure 1.4: A grid of wires.

(a) Determine equations for the temperatures x_2 and x_3 if the grid is in thermal equilibrium to construct a system of three linear equations in x_1, x_2, and x_3 that models node temperatures in the grid in thermal equilibrium.

(b) Use the method of elimination to find a specific solution to the system that makes sense in context.

(6) We have seen that a linear system of two equations in two unknowns can have no solutions, one solution, or infinitely many solutions. Find, if possible, a specific example of each of the following. If not possible, explain why.

(a) A linear system of three equations in two unknowns with no solutions.

(b) A linear system of three equations in two unknowns with exactly one solution.

(c) A linear system of three equations in two unknowns with exactly two solutions.

(d) A linear system of three equations in two unknowns with infinitely many solutions.

(7) We have seen that a linear system of three equations in three unknowns can have no solutions, one solution, or infinitely many solutions. Find, if possible, a specific example of each of the following. If not possible, explain why.

(a) A linear system of two equations in three unknowns with no solutions.

(b) A linear system of two equations in three unknowns with exactly one solution.

(c) A linear system of two equations in three unknowns with exactly two solutions.

(d) A linear system of two equations in three unknowns with infinitely many solutions.

(8) Find a system of three linear equations in two variables u, v whose solution is $u = 2$, $v = 1$.

(9) Consider the system of linear equations

$$\begin{aligned} x_1 + hx_2 &= 2 \\ 3x_1 + 5x_2 &= 1 \end{aligned}$$

where h is an unknown constant.

(a) Determine the solution(s) of this system for all possible h values, if a solution exists. (Note: Your answers for the variables will depend on the h.)

(b) How do your answers change if the second equation in the system above is changed to $3x_1 + 5x_2 = 6$?

(10) Suppose we are given a system of two linear equations

$$x_1 + 2x_2 - x_3 = 1 \tag{1.7}$$

$$3x_1 + x_2 + 2x_3 = -1. \tag{1.8}$$

Find another system of two linear equations E_1 and E_2 in the variables x_1, x_2, and x_3 that are not multiples of each other or of equations (1.7) or (1.8) so that any solution (x_1, x_2, x_3) to the system (1.7) and (1.8) is a solution to the system E_1 and E_2.

True/False Questions

In many sections you will be given True/False questions. In each of the True/False questions, you will be given a statement, such as “If we add corresponding sides of two linear equations, then the resulting equation is a linear equation.” and “One can find a system of two equations in two unknowns that has infinitely many solutions.”. Your task will be to determine the truth value of the statement and to give a brief justification for your choice.

Note that a *general* statement is considered *true* only when it is always true. For example, the first of the above statements, “If we add corresponding sides of two linear equations, then the resulting equation is a linear equation.”, is a general statement. For this statement to be true, the equation we obtain by adding corresponding sides of any two linear equations has to be linear. If we can find two equations that do not give a linear equation when combined in this way, then this statement is false.

Note that an *existential* statement is considered *true* if there is at least one example which makes is true. For example, the latter of the above statements, “One can find a system of two equations in two unknowns that has infinitely many solutions.”, is an existential statement. For this statement to be true, existence of a system of two equations in two unknowns with infinitely many solutions should suffice. If it is impossible to find two such equations, then this statement is false.

To justify that something always happens or never happens, one would need to refer to other statements whose truth is known, such as theorems, definitions. In particular, giving an **example** of two linear equations that produce a linear equation when we add corresponding sides *does not justify* why the sum of **any** two linear equations is also linear. Using the definition of linear equations, however, we can justify why this new equation will always be linear: each side of a linear equation is linear, and adding linear expressions always produces a linear sum.

To justify that there are examples of something happening or not happening, one would need to give a specific example. For example, in justifying the claim that there is a system of two equations in two unknowns with infinitely many solutions, it is not enough to say “An equation in two unknowns is a line in the xy-plane, so there can be two equations with the same line as their solution.”. In general, you should avoid the words “can”, “possibly”, “maybe”, etc., in your justifications. Instead, giving an example such as “The linear system $x + y = 1$ and $2x + 2y = 2$ of two equations in two unknowns has infinitely many solutions since the second equation gives the same line as the first in the xy-plane.” provides complete justification beyond a reasonable doubt.

Each response to a True/False statement should be more than just True or False. It is important that you provide **justification** for your responses.

(1) (a) **True/False** The set of all solutions of a linear equation can be represented graphically as a line.

(b) **True/False** The set of all solutions of a linear equation in two variables can be represented graphically as a line.

(c) **True/False** The set of all solutions of an equation in two variables can be represented graphically as a line.

(d) **True/False** A system of three linear equations in two unknowns cannot have a unique solution.

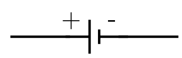
(e) **True/False** A system of three linear equations in three unknowns has a unique solution.

Project: Modeling an Electrical Circuit and the Wheatstone Bridge Circuit

Mathematical modeling, or the act of creating equations to model given information, is an important part of problem solving. In this section we will see how we derived the system of equations

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 5I_1 + 2I_2 &= 8 \\ 2I_2 + 4I_3 &= 5, \end{aligned}$$

to represent the electrical current in the circuit shown in Figure 1.2. Recall that a circuit consists of

- one or more electrical sources (like a battery), denoted by 

- one or more resistors (like any appliance that you plug into a wall outlet), denoted by .

The source creates a charge that produces potential energy E measured in volts (V). No substance conducts electricity perfectly, there is always some price to pay (energy loss) to moving electricity. Electrical current I in amperes (A) is the flow of the electric charge in the circuit. (A current of 1 ampere means that 6.2×10^{18} electrons pass through the circuit per second.) Current flows out of the positive terminal of a source and runs through each branch of the circuit. Let I_1 be the current flowing through the upper branch, I_2 the current through middle branch, and I_3 the current through the lower branch as illustrated in Figure 1.2. The goal is to find the current flowing in each branch of the circuit.

Linear algebra comes into play when analyzing a circuit based on the relationship between current, resistance, and potential. Three basic principles govern current low in a circuit.

(1) Resistance R in ohms (Ω) can be thought of as a measure of how difficult it is to move a charge along a circuit. When a current flows through a resistor, it must expend some energy, called a *voltage drop*. Ohm's Law states that the voltage drop E across a resistor is the product of the current I passing through the resistor and the resistance R. That is,

$$E = IR.$$

(2) Kirchoff's Current Law states that at any point in an electrical circuit, the sum of currents flowing into that point is equal to the sum of currents flowing out of that point.

(3) Kirchoff's Voltage Law says that around any closed loop the sum of the voltage drops is equal to the sum of the voltage rises.

To see how these laws allow us to model the circuit in Figure 1.2, we will need three equations in I_1, I_2, and I_3 to determine the values of these currents. Let us first apply Kirchoff's Current Law to the point P. The currents flowing into point P are I_1 and I_3, and the current flowing out is I_2. This produces the equation $I_1 + I_3 = I_2$, or

$$I_1 - I_2 + I_3 = 0.$$

Project Activity 1.1. Apply Kirchoff's Current Law to the point Q to obtain an equation in I_1, I_2, and I_3. What do you notice?

We have three variables to determine, so we still need two more equations in I_1, I_2, and I_3. Next we apply Kirchoff's Voltage Law to the top loop in the circuit in Figure 1.2. We will assume the following sign conventions:

- A current passing through a resistor produces a voltage drop if it flows in the direction of loop (and a voltage rise if the current passes in the opposite direction of the loop).
- A current passing through a source in the direction of the loop produces a voltage drop if it flows from $+$ to $-$ and a voltage rise if it flows from $-$ to $+$, while a current passing through a source in the opposite direction of the loop produces a voltage rise if it flows from $+$ to $-$ and a voltage drop if it flows from $-$ to $+$.

(The directions chosen in Figure 1.2 for the voltage flow are arbitrary – if we reverse the flow then we just replace voltage drops with voltage rises and obtain the same equations. If a solution shows that a current is negative, then that current flows in the direction opposite of what is shown.)

If we move in the counterclockwise direction around the top loop in the circuit in Figure 1.2, there is a voltage rise through the source of 8 volts. This must equal the voltage drop in this loop. The current I_1 passing though the resistor of resistance 2 Ω produces a voltage drop of $2I_1$ volts. Similarly, the current I_1 passing through the resistor of resistance 3Ω produces a voltage drop of $3I_1$ volts. Finally, the current I_2 passing through the resistor of resistance 2 Ω produces a voltage drop of $2I_2$ volts. So Kirchoff's Voltage Law applied to the top loop in the circuit in Figure 1.2 gives us the equation $2I_1 + 3I_1 + 2I_2 = 8$ or

$$5I_1 + 2I_2 = 8.$$

Project Activity 1.2. Apply Kirchoff's Voltage Law to the bottom loop in the circuit in Figure 1.2 to obtain an equation in I_1, I_2, and I_3. Compare the three equations we have found to those in the introduction.

Project Activity 1.3. Consider the circuit as shown in Figure 1.5, with a single source and five resistors with resistances R_1, R_2, R_3, R_4, and R_5 as labeled.

(a) Assume the following information. The voltage E is 13 volts, $R_1 = R_2 = R_3 = R_5 = 1\Omega$, and $R_4 = 2\Omega$. Follow the directions given to find the currents I_0, I_1, I_2, I_3, I_4, and I_5.

i. Use Kirchoff's Current Law to show that $I_0 = I_1 + I_2$, $I_3 = I_1 - I_5$, and $I_4 = I_2 + I_5$. Thus, we reduce the problem to three variables.

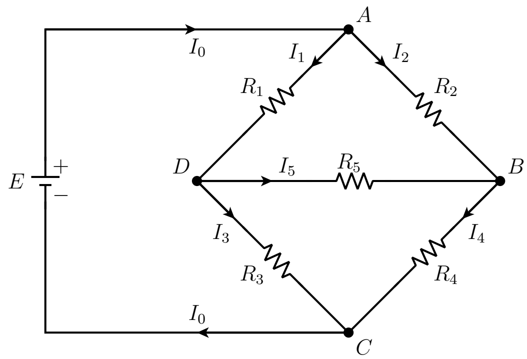


Figure 1.5: A Wheatstone bridge circuit.

ii. Apply Kirchoff's Voltage Law to three loops to show that the currents must satisfy the linear system

$$\begin{aligned} 2I_1 \quad\quad - I_5 &= 13 && (1.9)\\ 3I_2 + 2I_5 &= 13 && (1.10)\\ I_1 - I_2 + I_5 &= 0. && (1.11) \end{aligned}$$

iii. Solve the system to find the unknown currents.

(b) The circuit pictured in Figure 1.5 is called a *Wheatstone bridge* (invented by Samuel Hunter Christie in 1833 and popularized by Sir Charles Wheatstone in 1843). The Wheatstone bridge is a circuit designed to determine an unknown resistance by balancing two paths in a circuit. It is set up so that the resistances of resistors R_1 and R_2 are known, R_3 is a variable resistor and we want to find the resistance of R_4. The resistor R_5 is replaced with a voltmeter, and the resistance of R_3 is varied until the voltmeter reads 0. This balances the circuit and tells the resistance of resistor R_4. Show that if the current I_5 in Figure 1.5 is 0 (so the circuit is balanced), then $R_4 = \frac{R_2 R_3}{R_1}$ (which is how we calculate the unknown resistance R_4). Do this in general and do not use any specific values for the resistances or the voltage.

Section 2

The Matrix Representation of a Linear System

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a matrix?
- How do we associate a matrix to a system of linear equations?
- What row operations can we perform on an augmented matrix of a linear system to solve the system of linear equations?
- What are pivots, basic variables, and free variables?
- How many solutions can a system of linear equations have?
- When is a linear system consistent?
- When does a linear system have infinitely many solutions? A unique solution?
- How can we represent the set of solutions to a consistent system if the system has infinitely many solutions?

Application: Simpson's Rule

You may recall that Simpson's Rule from calculus ($\frac{2}{3}$ of the midpoint approximation plus $\frac{1}{3}$ of the trapezoid approximation) is a formula that can be used to approximate definite integrals. One the one hand, Simpson's Rule is a weighted average of the midpoint and trapezoid sum, but that does not completely explain why Simpson's Rule is so much better than either the midpoint or trapezoid sum. While the midpoint and trapezoid sums use line segments to approximate a function on an interval, Simpson's Rule uses parabolas. In order to use Simpsons Rule, we need to know how to exactly fit a quadratic function to three points. More details about this process can be found at the end of this section. This idea of fitting a polynomial to a set of data points has uses in other areas as well. For example, two common applications of Bézier curves are font design and drawing tools. When fitting a polynomial to a large set of data points, our systems of equations can become quite large, and can be difficult to solve by hand. In this section we will see how to use matrices to more conveniently represent systems of equations of any size. We also consider how the elimination process works on the matrix representation of a linear system and how we can determine the existence of solutions and the form of solutions of a linear system.

Introduction

When working with a linear system, the labels for the variables are irrelevant to the solution – the only thing that matters is the coefficients of the variables in the equations and the constants on the other side of the equations. For example, given a linear system of the form

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5, \end{aligned} \tag{2.1}$$

the important information in the system can be represented as

$$\begin{matrix} 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 6 \\ 4 & 2 & 1 & 5 \end{matrix}$$

where we interpret the first three numbers in each horizontal row to represent the coefficients of the variables a, b and c, respectively, and the last number to be the constant on the right hand side of the equation. This tells us that we can record all the necessary information about our system in a rectangular array of numbers. Such an array is called a *matrix*.

Definition 2.1. A **matrix** is a rectangular array of quantities or expressions.

We usually delineate a matrix by enclosing its entries in square brackets $[*]$. For the system in (2.1), there are two corresponding matrices:

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

The matrix on the left is the matrix of the coefficients of the system, and is called the *coefficient matrix* of the system. The matrix on the right is the matrix of coefficients and the constants, and is called the *augmented matrix* of the system (where we say we augment the coefficient matrix with the additional column of constants). We will separate the augmented column from the coefficient matrix with a vertical line to keep it clear that the last column is an augmented column of constants and not a column of coefficients.[1]

Terminology. There is some important terminology related to matrices.

- Any number in a matrix is called an *entry* of the matrix.
- The collection of entries in an augmented matrix that corresponds to a given equation (that is reading the entries from left to right, or a horizontal set of entries) is called a *row* of the matrix. We number the rows from top to bottom in a matrix. For example, $\begin{bmatrix} 1 & -1 & 1 \end{bmatrix}$ is the first row and $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ is the second row of the coefficient matrix of the system (2.1).
- The set of entries as we read from top to bottom (or a vertical set of entries that correspond to one fixed variable or the constants on the right hand sides of the equations) is called a *column* of the matrix. We number the columns from left to right in a matrix. For example, $\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ is the first column and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is the third column of the coefficient matrix of the system (2.1).
- The *size* of a matrix is given as $m \times n$ where m is the number of rows and n is the number of columns. The coefficient matrix above is a 3×3 matrix since it has 3 rows and 3 columns, while the augmented matrix is a 3×4 matrix as it has 4 columns.

Preview Activity 2.1.

(1) Write the augmented matrix for the following linear system. If needed, rearrange an equation to ensure that the variables appear in the same order on the left side in each equation with the constants being on the right hand side of each equation.

$$\begin{aligned} -x_3 + 3 + 2x_2 &= -x_1 \\ -3 + 2x_3 &= -x_2 \\ -2x_2 + x_1 &= 3x_3 - 7 \end{aligned} \tag{2.2}$$

(2) Write the linear system in variables x_1, x_2 and x_3, appearing in the natural order that corresponds to the following augmented matrix. Then solve the linear system using the elimination method.

$$\left[\begin{array}{ccr|r} 1 & 1 & -1 & 4 \\ 1 & 2 & 2 & 3 \\ 2 & 3 & -3 & 11 \end{array}\right]$$

(3) Consider the three types of elementary operations on systems of equations introduced in Section 1. Each row of an augmented matrix of a system corresponds to an equation, so each elementary operation on equations corresponds to an operation on rows (called row operations).

[1] You should note that not every author uses this convention – when they do not, it is important that you be careful to understand if the matrix has an augmented column or not.

(a) Describe the row operation that corresponds to interchanging two equations.

(b) Describe the row operation that corresponds to multiplying an equation by a nonzero scalar.

(c) Describe the row operation that corresponds to replacing one equation by the sum of that equation and a scalar multiple of another equation.

Simplifying Linear Systems Represented in Matrix Form

Once we have stored the information about a linear system in an augmented matrix, we can perform the elementary operations directly on the augmented matrix.

Recall that the allowable operations on a system of equations are the following:

(1) Replacing one equation by the sum of that equation and a scalar multiple of another equation.

(2) Interchanging the positions of two equations.

(3) Replacing an equation by a nonzero scalar multiple of itself.

Recall that we use these elementary operations to transform a system, with the ultimate goal of finding a simpler, equivalent system that we can solve. Since each row of an augmented matrix corresponds to an equation, we can translate these operations on equations to corresponding operations on rows (called *row operations* or *elementary row operations*):

(1) Replacing one row by the sum of that row and a scalar multiple of another row.

(2) Interchanging two rows.

(3) Replacing a row by a nonzero scalar multiple of itself.

Activity 2.1. Consider the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5 \end{aligned}$$

with corresponding augmented matrix

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

(a) As a first step in solving our system, we might eliminate a_2 from the second equation. This means that the corresponding entry in the second row and first column of the augmented matrix will become 0. Find a row operation that adds a multiple of the first row to the second row to achieve this goal. Then write the system of equations that corresponds to this new augmented matrix.

(b) Now that we have eliminated the a_2 terms from the second equation, we eliminate the a_2 term from the third equation. Find an appropriate row operation that does that, and write the corresponding system of linear equations that corresponds to the new augmented matrix.

(c) Now you should have a system in which the last two rows correspond to a system of 2 linear equations in two unknowns. Use a row operation that adds a multiple of the second row to the third row to turn the coefficient of a_1 in the third row to 0. Then write the corresponding system of linear equations.

(d) Your simplified system and its augment matrix are in *row echelon form* and this system is solvable using *back-substitution* (substituting the known variable values into the previous equation to find the value of another variable). Solve the system.

Reflection 1. Do you see how this standard elimination process can be generalized to any linear system with any number of variables to produce a simplified system? Do you see why the process does not change the solutions of the system? If needed, can you modify the standard elimination process to obtain a simplified system in which the last equation contains only the variable a_2, the next to last equation contains only the variables a_1, a_2, etc.? Understanding the standard process will enable you to be able to modify it, if needed, in a problem.

Activity 2.1 illustrates how we can perform all of the operations on equations with operations on the rows of augmented matrices to reduce a system to a solvable form. Each time we perform an operation on the system of equations (or on the rows of an augmented matrix) we obtain an equivalent system (or an augmented matrix corresponding to an equivalent system). For completeness, we list the operations on equations and the corresponding row operations below that can be used to solve our polynomial fitting system. Throughout the process we will let E_1, E_2, and E_3 be the first, second, and third equations in the system and R_1, R_2, and R_3 the first, second, and third rows of the augmented matrices. The notation $E_1 + E_2$ placed next to equation E_2 means means that we replace the second equation in the system with the sum of the first two equations. We start with the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ a_2 + a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5 \end{aligned}$$

On the left we demonstrate the operations on equations and on the right the corresponding operations on rows of the augmented matrix.

$$E_2 - E_1 \to E_2 \quad \begin{aligned} a_2 - a_1 + a_0 &= 2 \\ 2a_1 &= 4 \\ 4a_2 + 2a_1 + a_0 &= 5 \end{aligned} \qquad R_2 - R_1 \to R_2 \left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 4 \\ 4 & 2 & 1 & 5 \end{array}\right]$$

$$E_3 - 4E_1 \to E_3 \quad \begin{aligned} a_2 - a_1 + a_0 &= 2 \\ 2a_1 &= 4 \\ 6a_1 - 3a_0 &= -3 \end{aligned} \qquad R_3 - 4R_1 \to R_3 \left[\begin{array}{ccc|c} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 4 \\ 0 & 6 & -3 & -3 \end{array}\right]$$

$$\begin{array}{rrrrl} & a_2 - & a_1 + & a_0 = & 2 \\ & & 2a_1 & = & 4 \\ E_3 - 3E_2 \to E_3 & & & -3a_0 = & -15 \end{array} \qquad \begin{array}{r} \\ \\ R_3 - 3R_2 \to R_3 \end{array} \left[\begin{array}{rrr|r} 1 & -1 & 1 & 2 \\ 0 & 2 & 0 & 4 \\ 0 & 0 & -3 & -15 \end{array}\right]$$

Now we can solve the last equation for a_0 to find that $a_0 = 5$. The second equation gives us $a_1 = 2$.[2] Finally, using the first equation with the already determined values of a_0 and a_1 gives us $a_2 = -1$. Thus we have found the solution to the polynomial fitting system to be $a_2 = -1$, $a_1 = 2$, and $a_0 = 5$.

We summarize the steps of the (partial) elimination on matrices we used above to solve a general linear system in the variables $x_1, x_2, \ldots, x_n$.

(1) Interchange equations if needed to ensure that the coefficient of x_1 (or, more generally, the first non-zero variable) in the first equation is non-zero.

(2) Use the first equation to eliminate x_1 (or, the first non-zero variable) from other equations by adding a multiple of the first equation to the others.

(3) After x_1 is eliminated from all equations but the first equation, focus on the rest of the equations. Repeat the process of elimination on these equations to eliminate x_2 (or, the next non-zero variable) all but the second equation.

(4) Once the process of eliminating variables recursively is finished, solve for the variables in a backwards fashion starting with the last equation and substituting known values in the equations above as they become known.

This elimination method where the variables are eliminated from lower equations is called the *forward elimination phase* as it eliminates variables in the forward direction. Solving for variables using substitution into upper equations is called *back substitution*. The matrix representation of a linear system after the forward elimination process is said to be in *row echelon form*. We will define this form and the elimination process on the matrices more precisely in the next section.

Linear Systems with Infinitely Many Solutions

Each of the systems that we solved so far have had a unique (exactly one) solution. The geometric representation of linear systems with two equations in two variables shows that this does not always have to be the case. We also have linear systems with no solution and systems with infinitely many solutions. We now consider the problem of how to represent the set of solutions of a linear system that has infinitely many solutions. (Systems with infinitely many solutions will also be of special interest to us a bit later when we study eigenspaces of a matrix.)

Activity 2.2. Consider the system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ x_1 + x_2 - 3x_3 &= 0 \\ 2x_1 + 3x_2 - 4x_3 &= 1. \end{aligned}$$

[2] If there had been an a_0 term in the second equation, we could have substituted $a_0 = 5$ and solved for a_1

(a) Without explicitly solving the system, check that $(-1, 1, 0)$ and $(4, -1, 1)$ are solutions to this system.

(b) Without explicitly solving the system, show that $x_1 = -1 + 5t$, $x_2 = 1 - 2t$, and $x_3 = t$ is a solution to this system for any value of t. What values of t yield the solutions $(-1, 1, 0)$ and $(4, -1, 1)$ from part (a)? The equations $x_1 = -1 + 5t$, $x_2 = 1 - 2t$, and $x_3 = t$ form what is called a *parametric solution* to the system with *parameter* t.

(c) Part (b) shows that our system has infinitely many solutions. We were given solutions in part (b) – but how do we find these solutions and how do we know that these are all of the solutions? We address those questions now.

If we apply row operations to the augmented matrix

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

of this system, we can reduce this system to one with augmented matrix

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

i. What is it about this reduced form of the augmented matrix that indicates that the system has infinitely many solutions?

ii. Since the system has infinitely many solutions, we will not be able to explicitly determine values for each of the variables. Instead, at least one of the variables can be chosen arbitrarily. What is it about the reduced form of the augmented matrix that indicates that x_3 is convenient to choose as the arbitrary variable?

iii. Letting x_3 be arbitrary (we call x_3 a *free* variable), use the second row to show that $x_2 = 1 - 2x_3$ (so that we can write x_2 in terms of the arbitrary variable x_3).

iv. Use the first row to show that $x_1 = 5x_3 - 1$ (and we can write x_1 in terms of the arbitrary variable x_3). Compare this to the solutions from part (b).

After using the elimination method, the first non-zero coefficient (from the left) of each equation in the linear system is in a different position. We call each such coefficient a *pivot* and a variable corresponding to a pivot a *basic variable*. In the system

$$\begin{aligned} a_2 - a_1 + a_0 &= 2 \\ 2a_1 \qquad &= 4 \\ -3a_0 &= -15 \end{aligned}$$

the basic variables are a_2, a_1, a_0 for the first, second, and third equations, respectively. In the system,

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

the basic variables are x_1 and x_2 for the first and second equations, respectively, while the third equation does not have a basic variable. Through back-substitution, we can solve for each variable in a unique way if each appears as the basic variable in an equation. If, however, a variable is *free*, meaning that it is not the basic variable of an equation, we cannot solve for that variable explicitly. We instead assign a distinct parameter to each such free variable and solve for the basic variables in terms of these parameters.

Definition 2.2. The first non-zero coefficient (from the left) in an equation in a linear system after elimination is called a **pivot**. A variable corresponding to a pivot is a **basic variable** and while a variable not corresponding to a pivot is a **free variable**.

Activity 2.3. Each matrix is an augmented matrix for a linear system after elimination. Identify the basic variables (if any) and free variables (if any). Then write the general solution (if there is a solution) expressing all variables in terms of the free variables. Use any symbols you like for the variables.

(a) $\left[\begin{array}{ccc|c} 1 & 0 & 2 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

(b) $\left[\begin{array}{ccr|c} 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$

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

Reflection 2. Does the existence of a row of 0's always mean a free variable? Can you think of an example where there is a row of 0's but none of the variables is free? How do the numbers of equations and the variables compare in that case?

Linear Systems with No Solutions

We saw in the previous section that geometrically two parallel and distinct lines represent a linear system with two equations in two unknowns which has no solution. Similarly, two parallel and distinct planes in three dimensions represent a linear system with two equations in three unknowns which has no solution. We can have at least four different geometric configurations of three planes in three dimensions representing a system with no solution. But how do these geometrical configurations manifest themselves algebraically?

Activity 2.4. Consider the linear system

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

(a) Apply the elimination process to the augmented matrix of this system. Write the system of equations that corresponds to the final reduced matrix.

(b) Discuss which feature in the final simplified system makes it easy to determine that the system has no solution. Similarly, what features in the matrix representation makes is easy to see the system has no solution?

We summarize our observations about when a system has a solution, and which of those cases has a unique solution.

Theorem 2.3. *A linear system is consistent if after the elimination process there is no equation of the form $0 = b$ where b is a non-zero number. If a linear system is consistent and has a free variable, then it has infinitely many solutions. If it is consistent and has no free variables, then there is a unique solution.*

Examples

What follows are worked examples that use the concepts from this section.

Example 2.4. Consider the linear system

$$\begin{aligned} x_1 - \ \ x_2 \qquad\quad + 2x_4 &= 1 \\ 2x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ x_1 - \ \ x_2 + \ \ x_3 - \ \ x_4 &= 0 \\ 4x_1 + \ \ x_2 - \ \ x_3 + 6x_4 &= 5. \end{aligned}$$

(a) Set up the augmented matrix for this linear system.

(b) Find all solutions to the system using forward elimination.

(c) Suppose, after forward elimination, the augmented matrix of the system

$$\begin{aligned} x_1 - \ \ x_2 \qquad\quad + 2x_4 &= 1 \\ 2x_1 + 3x_2 - 2x_3 + 5x_4 &= 4 \\ x_1 - \ \ x_2 + \ \ x_3 - \ \ x_4 &= 0 \\ 4x_1 + \ \ x_2 - \ \ x_3 + 6x_4 &= h. \end{aligned}$$

has the form

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

For which values of h does this system have:

i. No solutions?
ii. A unique solution? Find the solution.
iii. Infinitely many solution? Determine all solutions?

Example Solution.

(a) The augmented matrix for this system is

$$\left[\begin{array}{rrrr|r} 1 & -1 & 0 & 2 & 1 \\ 2 & 3 & -2 & 5 & 4 \\ 1 & -1 & 1 & -1 & 0 \\ 4 & 1 & -1 & 6 & 5 \end{array}\right].$$

(b) We apply forward elimination, first making the entries below the 1 in the upper left all 0. We do this by replacing row two with row two minus 2 times row 1, row three with row three minus row 1, and row four with row four minus 4 row one. This produces the augmented matrix

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

Now we eliminate the leading 5 in the fourth row by replacing row four with row four minus row two to obtain the augmented matrix

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

When we replace row four with row four minus row three, we wind up with a row of zeros:

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

We see that there is no pivot in column four, so x_4 is a free variable. We can solve for the other variables in terms of x_4. The third row shows us that

$$\begin{aligned} x_3 - 3x_4 &= -1 \\ x_3 &= 3x_4 - 1. \end{aligned}$$

The second row tells us that

$$\begin{aligned} 5x_2 - 2x_3 + x_4 &= 2 \\ 5x_2 &= 2x_3 - x_4 + 2 \\ 5x_2 &= 2(3x_4 - 1) - x_4 + 2 \\ 5x_2 &= 5x_4 \\ x_2 &= x_4. \end{aligned}$$

Finally, the first row gives us

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

So this system has infinitely many solutions, with $x_1 = -x_4 + 1$, $x_2 = x_4$, $x_3 = 3x_4 - 1$, and x_4 is arbitrary. As a check, notice that

$$(-x_4 + 1) - x_4 + 2x_4 = 1$$

and so this solution satisfies the first equation in our system. You should check to verify that it also satisfies the other three equations.

(c) i. The system has no solutions when there is an equation of the form $0 = b$ for some nonzero number b. The last row will correspond to an equation of the form $0 = h - 5$. So our system will have no solutions when $h \neq 5$.

ii. When $h \neq 5$, the system has no solutions. When $h = 5$, the variable x_4 is a free variable and the system has infinitely many solutions. So there are no values of h for which the system has exactly one solution.

iii. When $h = 5$, the variable x_4 is a free variable and the system has infinitely many solutions. The solutions were already found in part (a).

Example 2.5. After applying row operations to the augmented matrix of a system of linear equations, each of which describes a plane in 3-space, the following augmented matrix was obtained:

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

(a) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 0$ and $b = 2$.

(b) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 0$ and $b = 6$.

(c) Describe, algebraically and geometrically, all solutions (if any), to this system when $a = 1$ and $b = 12$.

Example Solution. Throughout, we will let the variables x, y, and z correspond to the first, second, and third columns, respectively, of our augmented matrix.

(a) When $a = 0$ and $b = 2$ our augmented matrix has the form

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

This matrix corresponds to the system

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

There are no equations of the form $0 = b$ for a nonzero constant b, so the system is consistent. There are no free variables, so the system has a unique solution. Algebraically, the solution is $x = 2$, $z = \frac{1}{2}$, and $y = -\frac{5}{2}$. Geometrically, this tells us that the three planes given by the original system intersect in a single point.

(b) When $a = 0$ and $b = 6$ our augmented matrix has the form

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

The last row corresponds to the equation $0 = 1$, so our system is inconsistent and has no solution. Geometrically, this tells us that the three planes given by the original system do not all intersect at any common points.

(c) When $a = 1$ and $b = 12$ our augmented matrix reduces to

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

There are no rows that correspond to equations of the form $0 = c$ for a nonzero constant c, so the system is consistent. The variable y is a free variable, so the system has infinitely many solutions. Algebraically, the solutions are y is free, is $z = -\frac{1}{3}$, and $x = 2 - y$. Geometrically, this tells us that the three planes given by the original system intersect in the line with $z = -\frac{1}{3}$, and $x = 2 - y$.

Summary

- A matrix is just a rectangular array of numbers or objects.
- Given a system of linear equations, with the variables listed in the same order in each equation, we represent the system by writing the coefficients of the first equation as the first row of a matrix, the coefficients of the second equation as the second row, and so on. This creates the coefficient matrix of the system. We then augment the coefficient matrix with a column of the constants that appear in the equations. This gives us the augmented matrix of the system.
- The operations that we can perform on equations translate exactly to row operations that we can perform on an augmented matrix:
  (1) Replacing one row by the sum of that row and a scalar multiple of another row.
  (2) Interchanging two rows.
  (3) Replacing a row by a nonzero scalar multiple of itself.
- The forward elimination phase of the elimination method recursively eliminates the variables in a linear system to reach an equivalent but simplified system.
- The first non-zero entry in an equation in a linear system after elimination is called a pivot.
- A basic variable in a linear system corresponds to a pivot of the system. A free variable is a variable that is not basic.
- A linear system can be inconsistent (no solutions), have a unique solution (if consistent and every variable is a basic variable), or have infinitely many solutions (if consistent and there is a free variable).
- A linear system has no solutions if, after elimination, there is an equation of the form $0 = b$ where b is a nonzero number.
- A linear system after the elimination method can be solved using back-substitution. The free variables can be chosen arbitrarily and the basic variables can be solved in terms of the free variables through the back-substitution process.

Exercises

(1) Consider the system of linear equations whose augmented matrix is

$$\left[\begin{array}{cc|r} 1 & 3 & -1 \\ 2 & h & k \end{array}\right]$$

where h and k are unknown constants. For which values of h and k does this system have

(a) a unique solution,

(b) infinitely many solutions,

(c) no solution?

(2) Consider the following system:

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

Check that when $h = -3$ the system has infinitely many solutions, while when $h \neq -3$ the system has a unique solution.

(3) If possible, find a system of three equations (not in reduced form) in three variables whose solution set consists only of the point $x_1 = 2, x_2 = -1, x_3 = 0$.

(4) What are the possible geometrical descriptions of the solution set of two linear equations in $\mathbb{R}^3$? (Recall that $\mathbb{R}^3$ is the three-dimensional xyz-space – that is, the set of all ordered triples of the form (x, y, z)).

(5) Two students are talking about when a linear system has infinitely many solutions.

> Student 1: So, if we have a linear system whose augmented matrix has a row of zeros, then the system has infinitely many solutions, doesn't it?
>
> Student 2: Well, but what if there is a row of the form $[\,0\,0\,\ldots\,0\,|\,b\,]$ with a nonzero b right above the row of 0's?
>
> Student 1: OK, maybe I should ask "If we have a consistent linear system whose augmented matrix has a row of zeros, then the system has infinitely many solutions, doesn't it?"
>
> Student 2: I don't know. It still doesn't sound enough to me, but I'm not sure why.

Is Student 1 right? Or is Student 2's hunch correct? Justify your answer with a specific example if possible.

(6) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** A system of linear equations in two unknowns can have exactly five solutions.

(b) **True/False** A system of equations with all the right hand sides equal to 0 has at least one solution.

(c) **True/False** A system of equations where there are fewer equations than the number of unknowns (known as an underdetermined system) cannot have a unique solution.

(d) **True/False** A system of equations where there are more equations than the number of unknowns (known as an overdetermined system) cannot have a unique solution.

(e) **True/False** A consistent system of two equations in three unknowns cannot have a unique solution.

(f) **True/False** If a system with three equations and three unknowns has a solution, then the solution is unique.

(g) **True/False** If a system of equations has two different solutions, then it has infinitely many solutions.

(h) **True/False** If there is a row of zeros in the row echelon form of the augmented matrix of a system of equations, the system has infinitely many solutions.

(i) **True/False** If there is a row of zeros in the row echelon form of the augmented matrix of a system of n equations in n variables, the system has infinitely many solutions.

(j) **True/False** If a system has no free variables, then the system has a unique solution.

(k) **True/False** If a system has a free variable, then the system has infinitely many solutions.

Project: A Polynomial Fitting Application: Simpson's Rule

As discussed in the introduction, Simpson's Rule for approximating a definite integral models the integrand with a quadratic polynomial on each interval. To better understand this method, we consider how to fit a quadratic to three points.

Suppose we are given a collection of three points in the plane: (x_1, y_1), (x_2, y_2) and (x_3, y_3). There is exactly one quadratic polynomial $p(x)$ which goes through these points, i.e. there is exactly one quadratic $p(x)$ such that for each x_i, $p(x_i) = y_i$. This is an example of *polynomial curve fitting*.

Suppose our given points are $(-1, 2)$, $(1, 6)$, $(2, 5)$. To fit a quadratic to these points, consider a general quadratic of the form $p(x) = a_2x^2 + a_1x + a_0$. By substituting the x value of each of the given points and setting that equal to the y value of that point, we find three equations

$$(-1)^2a_2 - a_1 + a_0 = 2\ , \qquad a_2 + a_1 + a_0 = 6\ , \qquad (2)^2a_2 + 2a_1 + a_0 = 5$$

that give us a system of three equations in the three unknowns a_2, a_1, and a_0:

$$\begin{aligned} a_2 - \ \ a_1 + a_0 &= 2 \\ a_2 + \ \ a_1 + a_0 &= 6 \\ 4a_2 + 2a_1 + a_0 &= 5. \end{aligned}$$

This system is the example we considered in Preview Activity 2.1, whose solution is $a_2 = -1$, $a_1 = 2$, and $a_0 = 5$. A graph of $q(x) = -x^2 + 2x + 5$ along with the three points $(-1, 2)$, $(1, 6)$, $(2, 5)$ is shown in Figure 2.1.

Project Activity 2.1. In this activity we model the sine function on the interval $[a, b]$, where $a = -\frac{\pi}{2}$ and $b = \pi$ with a collection of quadratics. Let $f(x) = \sin(x)$. We partition the interval $[a, b]$ using 6 partition points. Let $x_0 = -\frac{\pi}{2}$, $x_1 = -\frac{\pi}{4}$, $x_2 = 0$, $x_3 = \frac{\pi}{4}$, $x_4 = \frac{\pi}{2}$, $x_5 = \frac{3\pi}{4}$, and $x_6 = \pi$. We need 3 points to determine a quadratic, so the interval $[a, b]$ will be partitioned into 3 subintervals: $[x_0, x_2]$, $[x_2, x_4]$, and $[x_4, x_6]$.

(a) Set up a system of linear equations to fit a quadratic $q_1(x) = r_1x^2 + s_1x + t_1$ to the 3 points $(x_0, f(x_0))$, $(x_1, f(x_1))$, and $(x_2, f(x_2))$. (The solution to this system to 3 decimal places is $r_1 \approx 0.336$, $s_1 \approx 1.164$, and $t_1 = 0$.)

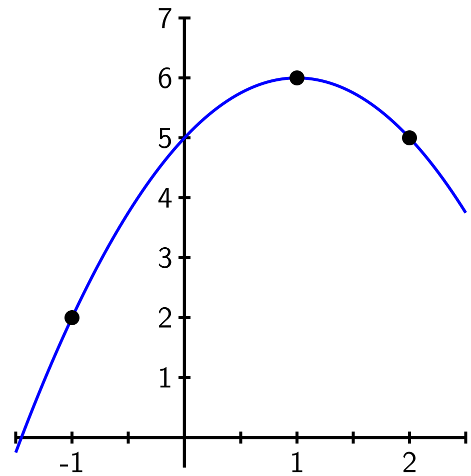


Figure 2.1: A quadratic fit to the points $(-1, 2)$, $(1, 6)$, $(2, 5)$.

(b) Set up a system of linear equations to fit a quadratic $q_2(x) = r_2x^2 + s_2x + t_2$ to the 3 points $(x_2, f(x_2))$, $(x_3, f(x_3))$, and $(x_4, f(x_4))$. (The solution to this system to 3 decimal places is $r_2 \approx -0.336$, $s_2 \approx 1.164$, and $t_2 = 0$.)

(c) Set up a system of linear equations to fit a quadratic $q_3(x) = r_3x^2 + s_3x + t_3$ to the 3 points $(x_4, f(x_4))$, $(x_5, f(x_5))$, and $(x_6, f(x_6))$. (The solution to this system to 3 decimal places is $r_3 \approx -0.336$, $s_3 \approx 0.946$, and $t_3 \approx 0.343$.)

(d) Plot the three quadratics on their intervals against the graph of f. Explain what you see.

Project Activity 2.1 illustrates how we can model a function on an interval using a sequence of quadratic functions. Now we apply this polynomial curve fitting technique to derive the general formula for Simpson's Rule for approximating definite integrals. The Simpson sum $S(n)$ is found by using parabolic arcs to approximate the graph of f on each subinterval rather than line segments. This allows Simpsons's Rule to more closely approximate the value of the definite integral with a smaller number of subintervals, although Simpson's Rule requires more calculations. Recall that to approximate a definite integral of a function f on an interval $[a, b]$, we partition $[a, b]$ into equal length subintervals. For Simpson's Rule, we partition $[a, b]$ into $n = 2m$ subintervals of equal length $\Delta x = \frac{b-a}{n}$. (Note that we need an even number of subintervals since we have to use three points for each parabola.) For each k we let $x_k = a + k\Delta x$ and $y_k = f(x_k)$. We approximate f on each subinterval using a quadratic. So we need to find the quadratic $Q(x) = c_2x^2 + c_1x + c_0$ that passes through two consecutive end points as well as the midpoint of a subinterval. That is, we need to find the coefficients of Q so that Q passes through the points (x_k, y_k), (x_{k+2}, y_{k+2}), and the midpoint (x_{k+1}, y_{k+1}) on the interval $[x_k, x_{k+2}]$ (so that we have three points to which to fit a parabola). Note that the length of the interval $[x_k, x_{k+2}]$ is $2\Delta x$. To make the calculations easier, we will translate our function so that our leftmost point is $(-r, y_k)$. Then the middle point is $(0, y_{k+1})$ and the rightmost point is (r, y_{k+2}), where $r = \Delta x$.

Project Activity 2.2.

(a) Set up a linear system that will determine the coefficients c_2, c_1, and c_0 so that the polynomial $Q(x) = c_2x^2 + c_1x + c_0$ passes through the points $(-r, y_k)$, $(0, y_{k+1})$, and (r, y_{k+2}) with $r \neq 0$. Remember that the unknowns in this system are c_2, c_1, and c_0.

(b) We apply row operations to the matrix $\begin{bmatrix} r^2 & -r & 1 & y_k \\ 0 & 0 & 1 & y_{k+1} \\ r^2 & r & 1 & y_{k+2} \end{bmatrix}$ and obtain the matrix $\begin{bmatrix} r^2 & -r & 1 & y_k \\ 0 & 2r & 0 & y_{k+2} - y_k \\ 0 & 0 & 1 & y_{k+1} \end{bmatrix}$. Use these matrices to show that $c_2 = \frac{y_k - 2y_{k+1} + y_{k+2}}{2r^2}$, $c_1 = \frac{y_{k+2} - y_k}{2r}$, and $c_0 = y_{k+1}$.

(c) Our goal is to ultimately approximate $\int_a^b f(x)\,dx$ by approximating f with quadratics on each subinterval. Use the fact that

$$\int_{x_k}^{x_{k+2}} f(x)\,dx \approx \int_{x_k}^{x_{k+2}} Q(x)\,dx = \int_{-r}^{r} Q(x)\,dx,$$

to show that

$$\int_{x_k}^{x_{k+2}} f(x)\,dx = \frac{1}{3}\left(y_k + 4y_{k+1} + y_{k+2}\right)\Delta x.$$

(d) Now we can derive Simpson's Rule. Use an additive property of the definite integral to show that

$$\int_a^b f(x)\,dx \approx S(n),$$

where

$$S(n) = \left(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots 2y_{n-2} + 4y_{n-1} + y_n\right)\frac{\Delta x}{3}$$

is the Simpson's Rule approximation.

Notice that we can rewrite the Simpson's Rule approximation as

$$\begin{aligned}
&\frac{1}{3}\left(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 2y_{n-2} + 4y_{n-1} + y_n\right)\Delta x \\
&\quad = \frac{2}{3}\left[2\left(y_1 + y_3 + \cdots + y_{n-1}\right)\right]\Delta x \\
&\qquad + \frac{2}{3}\left(\frac{y_0 + y_2}{2} + \frac{y_2 + y_4}{2} + \cdots + \frac{y_{n-2} + y_n}{2}\right)\Delta x \\
&\quad = \frac{1}{3}\left[\left(y_1 + y_3 + \cdots + y_{n-1}\right)(2\Delta x)\right] \\
&\qquad + \frac{1}{3}\left(\frac{y_0 + y_2}{2} + \frac{y_2 + y_4}{2} + \cdots + \frac{y_{n-2} + y_n}{2}\right)(2\Delta x) \\
&\quad = \frac{2M(n) + T(n)}{3},
\end{aligned}$$

where $M(n)$ is the midpoint sum and $T(n)$ is the trapezoid sum using n subdivisions of the interval $[a, b]$. Therefore, the weighted average $\frac{2M(n)+T(n)}{3}$ of the midpoint and trapezoid sums gives an approximation using quadratic functions.

Section 3

Row Echelon Forms

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is the row echelon form of a matrix?
- What is the procedure to obtain the row echelon form of any matrix?
- What is the reduced row echelon form of a matrix?
- What is the procedure to obtain the reduced row echelon form of any matrix?
- What do the echelon forms of the augmented matrix for a linear system tell us about the solutions to the system?

Application: Balancing Chemical Reactions

Linear systems have applications in chemistry when balancing chemical equations. When a chemical reaction occurs, molecules of different substances combine to create molecules of other substances. Chemists represent such reactions with chemical equations. To balance a chemical equation means to find the number of atoms of each element involved that will preserve the number of atoms in the reaction. As an example, consider the chemical equation

$$C_2H_6 + O_2 \to CO_2 + H_2O. \tag{3.1}$$

This equation asks about what will happen when the chemicals ethane (C_2H_6) and oxygen (O_2), called the *reactants* of the reaction, combine to produce carbon dioxide (CO_2) and water (H_2O), called the *products* of the reaction (note that oxygen gas is *diatomic*, so that oxygen atoms are paired). The arrow indicates that it is the reactants that combine to form the products. Any chemical reaction has to obey the Law of Conservation of Mass that says that mass can neither be created nor destroyed in a chemical reaction. Consequently, a chemical reaction requires the same number of atoms on both sides of the reaction. In other words, the total mass of the reactants must equal the total mass of the products. In reaction (3.1) the chemicals involved are made up of carbon (C), hydrogen (H), and oxygen (O) atoms. To balance the equation, we need to know how many molecules of each chemical are combined to preserve the number of atoms of C, H, and O. This can be done by setting up a linear system of equations of the form

$$\begin{array}{rcrcrcrcl} 2x_1 & & & - & x_3 & & & = & 0 \\ 6x_1 & & & & & - & 2x_4 & = & 0 \\ & & 2x_2 & - & 2x_3 & - & x_4 & = & 0, \end{array}$$

where x_1, x_2, x_3, and x_4 represent the number of molecules of C_2H_6, O_2, CO_2, and H_2O, respectively, in the reaction and then solving the system. Specific details can be found at the end of this section.

Introduction

In the previous sections, we identified operations on a given linear system with corresponding equivalent operations on the matrix representations which simplify the system and its matrix representation without changing the solutions of the system. Our end goal was to obtain a system which could be solved using back substitution, such as

$$\begin{array}{rcrcrcl} x_1 & - & x_2 & + & x_3 & = & 0 \\ & & 6x_2 & - & x_3 & = & 8 \\ & & & & x_3 & = & 1. \end{array}$$

The augmented matrix for this system is

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 0 & 6 & -1 & 8 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

The matrices of linear systems which can be solved via back substitution are said to be in *row echelon form* (or simply *echelon form*). We will define the properties of matrices in this form precisely in this section. Our goal will be to prescribe a precise procedure for converting any matrix to an equivalent one in row echelon form without having to convert back to the system representation.

Preview Activity 3.1. We want to determine a suitable form for an augmented matrix that can be obtained from row operations so that it is straightforward to find the solutions to the system. We begin with some examples.

(1) Write the linear system corresponding to each of the following augmented matrices. Use the linear system to determine which systems have their variables eliminated completely in the forward direction, or equivalently determine for which systems the next step in the solution

process is back substitution (possibly using free variables). Explain your reasoning. You do not need to solve the systems.

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

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

(2) Shown below are two row reduced forms of the system

$$\begin{aligned} 2x_1 \quad - x_3 \quad &= 0 \\ 6x_1 \quad\quad - 2x_4 &= 0 \\ 2x_2 - 2x_3 - x_4 &= 0. \end{aligned}$$

Of the systems that correspond to these augmented matrices, which is easier to solve and why?

$$\left[\begin{array}{rrrr|r} 2 & 0 & -1 & 0 & 0 \\ 0 & 2 & -2 & -1 & 0 \\ 0 & 0 & 3 & -2 & 0 \end{array}\right] \qquad \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -\frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{7}{6} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right]$$

The Echelon Forms of a Matrix

In the previous sections we saw how to simplify a linear system and its matrix representation via the elimination method without changing the solution set. This process is more efficient when performed on the matrix representation rather than on the system itself. Furthermore, the process of applying row operations to any augmented matrix is one that can be automated. In order to write an algorithm that can be used with any size augmented matrix to the extent that it can be applied even by a computer program, it is necessary to have a consistent procedure and a stopping point for the simplification process. The two main properties that we want the simplified augmented matrix to satisfy are that it should be easy to see if the system has solutions from the simplified matrix, and in cases when there are solutions, the general form of the solutions can be easily found. Hence the topic of this section is to define the process of elimination completely and generally.

We begin by discussing the *row echelon* or, simply, *echelon* form of a matrix. We know that the forward phase of the elimination on a linear system produces a system which can be solved by back substitution. The matrix representation of such a simplified system is said to be in *row echelon* or simply *echelon* form. Note that matrices in this form have the first nonzero entry in each row to the right of and below the first nonzero entry in the preceding row. Our next step is to formally describe this form – one that you tried to explain in problem 3 of Preview Activity 3.1.

Definition 3.1. A rectangular matrix is in **row echelon form** (or simply **echelon form**) if it has the following properties:

(1) All nonzero rows are above any rows of all zeros.

(2) Each **pivot** (the first non-zero entry reading from left to right) in a row is in a column to the right of the pivot of the row above it.

A pivot is also called a *leading entry* of a row. Note that properties (1) and (2) above imply that all entries in a column below a pivot are zeros. It can be shown that the positions of these pivots, called **pivot positions**, are unique and tell us quite a bit about a matrix and the solutions of the linear system it corresponds to. The columns that the pivots are in, called **pivot columns**, will also have useful properties as we will see soon.

Reflection 3. Compare the row echelon form of an augmented matrix to the corresponding system. Do you clearly see the correspondence between the requirements of the row echelon form and the properly eliminated variables in the system? Can you quickly come up with a system which will be in row echelon form when represented in augmented matrix form? Can you modify the standard row echelon form definition to cover cases where the elimination process eliminates the variables from last to first? For example, in a system with three equations in three unknowns, the last variable, say x_3, can be eliminated from the second equation, and the last two variables, say x_2, x_3 can be eliminated from the last equation. How would you define this modified row echelon form for a general system with this modified elimination process?

Once an augmented matrix is in row echelon form, we can use back substitution to solve the corresponding system. However, we can make solving much easier with just a little more elimination work.

Row operations are easy to apply, so if we are inclined, there is no reason to stop at the row echelon form. For example, starting with the following matrix

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

in row echelon form, we could take the row operations even farther and avoid the process of back substitution altogether. First, we multiply the last row by $1/2$ to simplify that row:

$$\begin{array}{r} \\ \\ \frac{1}{2}R_3 \to R_3 \end{array} \left[\begin{array}{rrrr|r} 2 & -1 & 2 & 2 & 7 \\ 0 & 1 & 3 & -1 & -1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right].$$

Then we use the third row to eliminate entries above the third pivot:

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

We can continue in this manner (we call this process *backward elimination*) to make 0 all of the entries above the pivots (one in the second column, and one in the fourth) with the pivots being 1, to ultimately obtain the equivalent augmented matrix

$$\left[\begin{array}{rrrr|r} 1 & 0 & 1 & 0 & 2 \\ 0 & 1 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right].$$

The system corresponding to this augmented matrix is

$$\begin{array}{rcrcrcl} x_1 & & + & x_3 & & = & 2 \\ & x_2 & + & 3x_3 & & = & 1 \\ & & & & x_4 & = & 2 \end{array}$$

so we can just directly read off the solution to the system: x_3 free and $x_1 = 2 - x_3, x_2 = 1 - 3x_3, x_4 = 2$. This final row reduced form makes solving the system very easy, and this form is called the *reduced row echelon* form.

Definition 3.2. A rectangular matrix is in **reduced row echelon form** (or **reduced echelon form**) if the matrix is in row echelon form and

(3) The pivot in each nonzero row is 1.

(4) Each pivot is the only nonzero entry in its column.

In short, the reduced row echelon form of a matrix is a row echelon form in which all the pivots are 1 and any entries below and above the pivots are 0.

If we use either of these two row echelon forms, solving the original system becomes straightforward and, as a result, these matrix forms are stopping points for the row operation algorithm to solve a system. It is also very easy to write a computer program to perform row operations to obtain and row echelon or reduced row echelon form of the matrix, making hand computations unnecessary. We will discuss this shortly.

Reflection 4. Compare the reduced row echelon form of an augmented matrix to the corresponding system. Do you clearly see the correspondence between the requirements of the reduced row echelon form and the way the variables appear in the equations in the system? Can you quickly come up with a system which will be in reduced row echelon form when represented in augmented matrix form?

Note. We have used the elimination method on augmented matrices so far. However, the elimination method can be applied on just the coefficient matrix, or other matrices that will arise in other contexts, and will provide useful information in each of those cases. Therefore, the row echelon form and reduced row echelon form is defined for any matrix, and from now on, a matrix will be a general matrix unless explicitly specified to be an augmented matrix.

Activity 3.1. Identify which of the following matrices is in row echelon form (REF) and/or reduced row echelon form (RREF). For those in row and/or reduced row echelon form, identify the pivots clearly by circling them. For those that are not in a given form, state which properties the matrix fails to satisfy.

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

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

Determining the Number of Solutions of a Linear System

Consider the system

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

The augmented matrix for this system is

$$\left[\begin{array}{rrrr|r} 1 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 & 4 \end{array}\right].$$

Note that this matrix is already in row echelon form. The reduced row echelon form of this augmented matrix is

$$\left[\begin{array}{rrrr|r} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 2 \\ 0 & 0 & 1 & -2 & 4 \end{array}\right]. \tag{3.2}$$

Since there are leading 1s in the first three columns, we can use those entries to write x_1, x_2, and x_3 in terms of x_4. We then choose x_4 to be arbitrary and write the remaining variables in terms of x_4. Let $x_4 = t$. Solving the third equation for x_3 gives us $x_3 = 4 + 2t$. The second equation shows that $x_2 = 2 + t$, and the first that $x_1 = 0$. Each value of t provides a solution to the system, so our system has infinitely many solutions. These solutions are

$$x_1 = 0, \; x_2 == 2 + t, \; x_3 = 4 + 2t, \; \text{ and } \; x_4 = t,$$

where t can have any value.

Activity 3.2. We have seen examples of systems with no solutions, one solution, and infinitely many solutions. As we will see in this activity, we can recognize the number of solutions to a system by analyzing the pivot positions in the augmented matrix of the system.

(a) Write an example of an augmented matrix in row echelon form so that the last column of the (whole) matrix is a pivot column. What is the system of equations corresponding to your augmented matrix? How many solutions does your system have? Why?

(b) Consider the reduced row echelon form (3.2). Based on the columns of this matrix, explain how we know that the system it represents is consistent.

(c) The system with reduced row echelon form (3.2) is consistent. What is it about the columns of the coefficient matrix that tells us that this system has infinitely many solutions?

(d) Suppose that a linear system is consistent and that the coefficient matrix has m rows and n columns.

i. If every column of the coefficient matrix is a pivot column, how many solutions must the system have? Why? What relationship must exist between m and n? Explain.

ii. If the coefficient matrix has at least one non-pivot column, how many solutions must the system have? Why?

When solving a linear system of equations, the free variables can be chosen arbitrarily and we can write the basic variables in terms of the free variables. Therefore, the existence of a free variable leads to infinitely many solutions for consistent systems. However, it is possible to have a system with free variables which is inconsistent. (Can you think of an example?)

Producing the Echelon Forms

In this part, we consider the formal process of creating the row and reduced row echelon forms of matrices. The process of creating the row echelon form is the equivalent of the elimination method on systems of linear equations.

Activity 3.3. Each of the following matrices is at most a few steps away from being in the requested echelon form. Determine what row operations need to be completed to turn the matrix into the required form.

(a) Turn into REF: $\begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$

(b) Turn into REF: $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}$

(c) Turn into RREF: $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(d) Turn into RREF: $\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

(e) Turn into RREF: $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$

(f) Turn into RREF: $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{bmatrix}$

The complete process of applying row operations to reduce an augmented matrix to a row or reduced row echelon form can be expressed as a recursive process in an algorithmic fashion, making it possible to program computers to solve linear systems. Here are the steps to do so:

Step 1: Begin with the leftmost nonzero column (if there is one). This will be a pivot column.

Step 2: Select a nonzero entry in this pivot column as a pivot. If necessary, interchange rows to move this entry to the first row (this entry will be a pivot).

Step 3: Use row operations to create zeros in all positions below the pivot.

Step 4: Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

To obtain the reduced row echelon form we need one more step.

Step 5: Beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by an appropriate row multiplication.

The algorithm described in steps 1-4 will produce the row echelon form of the matrix. This algorithm is called *Gaussian elimination*. When we add step 5 to produce the reduced row echelon form, the algorithm is called *Gauss-Jordan elimination*.

Activity 3.4. Consider the matrix $\begin{bmatrix} 0 & 2 & 4 & 1 \\ -1 & 3 & 0 & 6 \\ 0 & 4 & 8 & 2 \\ 1 & -3 & 0 & -2 \end{bmatrix}$.

(a) Perform Gaussian elimination to reduce the matrix to row echelon form. Clearly identify each step used. Compare your row echelon form to that of another group. Do your results agree? If not, who is right?

(b) Now continue applying row operations to obtain the reduced row echelon form of the matrix. Clearly identify each step. Compare your row echelon form to that of another group. Do your results agree? If not, who is right?

If we compare row echelon forms from Activity 3.4, it is likely that different groups or individuals produced different row echelon forms. That is because the row echelon form of a matrix is not unique. (Is the row echelon form ever unique?)

However, if row operations are applied correctly, then we will all arrive at the same reduced row echelon form in Activity 3.4:

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

It turns out that the reduced row echelon form of a matrix is unique.

Two matrices who are connected by row operations are said to be *row equivalent*.

Definition 3.3. A matrix B is **row equivalent** to a matrix A if B can be obtained by applying elementary row operations to A.

Since every elementary row operation is reversible, if B is row equivalent to A, then A is also row equivalent to B. Thus, we just say that A and B are row equivalent. While the row echelon form of a matrix is not unique, it is the case that the reduced row echelon form of a matrix is unique.

Theorem 3.4. *Every matrix is row equivalent to a unique matrix in reduced row echelon form.*

The reduced row echelon form of a matrix that corresponds to a system of linear equations provides us with a equivalent system whose solutions are easy to find. As an example, consider the system

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

with augmented matrix

$$\left[\begin{array}{rrrr|r} 0 & 2 & 4 & 1 & 0 \\ -1 & 3 & 0 & 6 & 0 \\ 0 & 4 & 8 & 2 & 0 \\ 1 & -3 & 0 & -2 & 0 \end{array}\right].$$

Notice that the coefficient matrix (the left hand side portion of the augmented matrix) of this system is same as the matrix we considered in Activity 3.4. Since we are augmenting with a column of zeros, no row operations will change those zeros in the augmented column. So the row operations applied in Activity 3.4 will give us the reduced row echelon form of this augmented matrix as

$$\left[\begin{array}{cccc|c} 1 & 0 & 6 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Note that the third column is not a pivot column. That means that the variable x_3 is a free variable. There are pivots in the other three columns of the coefficient matrix, so we can solve for x_1, x_2, and x_4 in terms of x_3. These variables are the basic variables. The third row of the augmented matrix tells us that $x_4 = 0$. The second row corresponds to the equation $x_2 + 2x_3 = 0$, and solving for x_2 shows that $x_2 = -2x_3$. Finally, the first row tells us that $x_1 + 6x_3 = 0$, so $x_1 = -6x_3$. Therefore, the general solution to this system of equations is

$$x_1 = -6x_3, \quad x_2 = -2x_3, \quad x_3 \text{ is free}, \quad x_4 = 0.$$

The fact that x_3 is free means that we can choose any value for x_3 that we like and obtain a specific solution to the system. For example, if $x_3 = -1$, then we have the solution $x_1 = 6$, $x_2 = 2$, $x_3 = -1$, and $x_4 = 0$. Check this to be sure.

Activity 3.5. Each matrix below is an augmented matrix for a linear system after elimination with variables $x_1, x_2, \ldots$ in that order. Identify the basic variables (if any) and free variables (if any). Then find the general solution (if there is a solution) expressing all variables in terms of the free variables.

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

(d) $\left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right]$ (e) $\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

Recall that in the previous section, we determined the criteria for when a system has a unique solution, or infinitely many solutions, or no solution. With the use of the row echelon form of the augmented matrix, we can rewrite these criteria as follows:

Theorem 3.5.

(1) A linear system is consistent if in the row echelon form of the augmented matrix representing the system no pivot is in the rightmost column.

(2) If a linear system is consistent and the row echelon form of the coefficient matrix does not have a pivot in every column, then the system has infinitely many solutions.

(3) If a linear system is consistent and there is a pivot in every column of the row echelon form of the coefficient matrix, then the system has a unique solution.

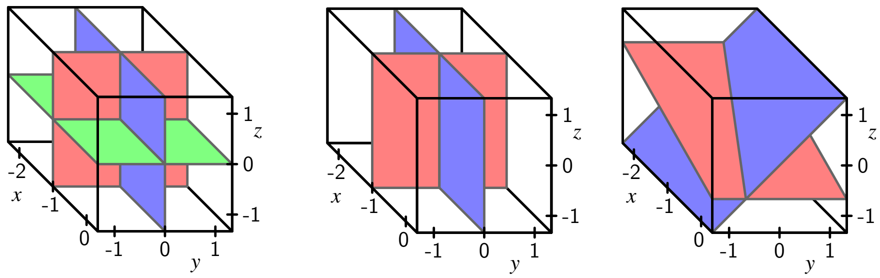


Figure 3.1: Figures for Activity 3.6.

Activity 3.6.

(a) For each part, the reduced row echelon form of the augmented matrix of a system of equations in variables x, y, and z (in that order) is given. Use the reduced row echelon form to find the solution set to the original system of equations.

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

iv. Each of the three systems above is represented as one of the graphs in Figure 3.1. Match each figure with a system.

(b) The reduced row echelon form of the augmented matrix of a system of equations in variables x, y, z, and t (in that order) is given. Use the reduced row echelon form to find the solution set to the original system of equations:

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

Examples

What follows are worked examples that use the concepts from this section.

Example 3.6. Consider the linear system

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

(a) Find the augmented matrix for this system.

(b) Use row operations to find a row echelon form of the augmented matrix of this system.

(c) Use row operations to find the reduced row echelon form of the augmented matrix of this system.

(d) Find the solution(s), if any, to the system.

Example Solution. Before we can find the augmented matrix of this system, we need to rewrite the system so that the variables are all on one side and the constant terms are on the other side of the equations. Doing so yields the equivalent system

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

Note that this is not the only way to rearrange the system. For example, for the second equation, could be written instead as $4x_1 + 2x_2 - 2x_3 = 0$ to minimize the number of negative signs in the equation.

(a) The augmented matrix for this system is

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

(b) Our first steps to row echelon form are to eliminate the entries below the leading entry in the first row. To do this we replace row two with row two plus 2 times row 1 and we replace row three with row three plus row one. This produces the row equivalent matrix

$$\left[\begin{array}{rrr|r} 2 & -1 & 6 & 2 \\ 0 & -4 & 14 & 4 \\ 0 & 0 & 10 & 10 \end{array}\right].$$

This matrix is now in row echelon form.

(c) To continue to find the reduced row echelon form, we replace row two with row two times $-\frac{1}{4}$ to get a leading 1 in the second row, and we replace row three with row three times $\frac{1}{10}$ to get a leading 1 in the third row and obtain the row equivalent matrix

$$\left[\begin{array}{rrr|r} 2 & -1 & 6 & 2 \\ 0 & 1 & -\frac{7}{2} & -1 \\ 0 & 0 & 1 & 1 \end{array}\right].$$

Now we perform backwards elimination to make the entries above the leading 1s equal to 0, starting with the third column and working backwards. Replace row one with row one minus 6 times row three and replace row two with row two plus $\frac{7}{2}$ row three to obtain the row equivalent matrix

$$\left[\begin{array}{rrr|r} 2 & -1 & 0 & -4 \\ 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array}\right].$$

For the second column, we replace row one with row one plus row two to obtain the row equivalent matrix

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

Since the leading entry in row one is not a one, we have one more step before we have the reduced row echelon form. Finally, we replace row one with row one times $\frac{1}{2}$. This gives us the reduced row echelon form

$$\left[\begin{array}{rrr|r} 1 & 0 & 0 & -\frac{3}{4} \\ 0 & 1 & 0 & \frac{5}{2} \\ 0 & 0 & 1 & 1 \end{array}\right].$$

(d) We can read off the solution to the system from the reduced row echelon form: $x_1 = -\frac{3}{4}$, $x_2 = \frac{5}{2}$, and $x_3 = 1$. You should check in the original equations to make sure we have the correct solution.

Example 3.7. In this example, a and b are unknown scalars. Consider the system with augmented matrix

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

Find all values of a and b so that the system has:

(a) Exactly one solution (and find the solution)

(b) No solutions

(c) Infinitely many solutions (and find all solutions)

Example Solution. Let x_1, x_2, and x_3 be the variables corresponding to the first, second, and third columns, respectively, of the augmented matrix. To answer these questions, we row reduce the augmented matrix. We interchange rows one and two and then also rows two and three to obtain the matrix

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

Now we replace row three with row three minus row one to produce the row equivalent matrix

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

Next, replace row three with row three minus 2 times row two. This yields the row equivalent matrix

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

We now have a row echelon form.

(a) The system will have exactly one solution when the last row has the form $[0\ 0\ u\ v]$ where u is not zero. Thus, the system has exactly one solution when $a - 2 \neq 0$, or when $a \neq 2$. In this case, the solution is

$$\begin{aligned} x_3 &= \frac{3-b}{a-2}, \\ x_2 &= -x_3 = \frac{b-2}{a-2} \\ x_1 &= b. \end{aligned}$$

You should check to ensure that this solution is correct. The other cases occur when $a = 2$.

(b) When $a = 2$ and $3 - b \neq 0$ (or $b \neq 3$), then we have a row of the form $[0\ 0\ 0\ t]$, where t is not 0. In these cases there are no solutions.

(c) When $a = 2$ and $b = 3$, then the last row is a row of all zeros. In this case, the system is consistent and x_3 is a free variable, so the system has infinitely many solutions. The solutions are

$$\begin{aligned} &x_1 = b \\ &x_2 = -x_3 \\ &x_3 \text{ is free.} \end{aligned}$$

You should check to ensure that this solution is correct.

Summary

In this section we learned about the row echelon and reduced row echelon forms of a matrix and some of the things these forms tell us about solutions to systems of linear equations.

- A matrix is in row echelon form if

  (1) All nonzero rows are above any rows of all zeros.

  (2) Each **pivot** (the first nonzero entry) of a row is in a column to the right of the pivot of the row above it.

- Once an augmented matrix is in row echelon form, we can use back substitution to solve the corresponding linear system.

- To reduce a matrix to row echelon form we do the following:

  - Begin with the leftmost nonzero column (if there is one). This will be a pivot column.
  - Select a nonzero entry in this pivot column as a pivot. If necessary, interchange rows to move this entry to the first row (this entry will be a pivot).
  - Use row operations to create zeros in all positions below the pivot.
  - Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. Apply the preceding steps to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

- A matrix is in reduced row echelon form if it is in row echelon form and

  (3) The pivot in each nonzero row is 1.

  (4) Each pivot is the only nonzero entry in its column.

- To obtain the reduced row echelon form from the row echelon form, beginning with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by an appropriate row multiplication.

- Both row echelon forms of an augmented matrix tell us about the number of solutions to the corresponding linear system.

  - A linear system is inconsistent if and only if a row echelon form of the augmented matrix of the system contains a row of the form

    $$[0\,0\,0\;\cdots\;0\,*],$$

    where $*$ is not zero. Another way to say this is that a linear system is inconsistent if and only if the last column of the augmented matrix of the system is a pivot column.
  - A consistent linear system will have a unique solution if and only if each column but the last in the augmented matrix of the system is a pivot column. This is equivalent to saying that a consistent linear system will have a unique solution if and only if the consistent system has no free variables.
  - A consistent linear system will have infinitely many solutions if and only if the coefficient matrix of the system contains a non-pivot column. In that case, the free variables corresponding to the non-pivot columns can be chosen arbitrarily and the basic variables corresponding to pivot columns can be written in terms of the free variables.
  - A linear system can have no solutions, exactly one solution, or infinitely many solutions.

Exercises

(1) Represent the following linear system in variables x_1, x_2, x_3 in augmented matrix form and use row reduction to find the general solution of the system.

$$\begin{aligned} x_1 + x_2 - x_3 &= 4 \\ x_1 + 2x_2 + 2x_3 &= 3 \\ 2x_1 + 3x_2 - 3x_3 &= 11. \end{aligned}$$

(2) Represent the following linear system in variables x_1, x_2, x_3 in augmented matrix form after rearranging the terms and use row reduction to find all solutions to the system.

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

(3) Check that the reduced row echelon form of the matrix

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

is

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

(4) Consider the following system:

$$\begin{aligned} x - 2y + z &= -1 \\ 2y - 4z &= 6 \\ hy - 2z &= 1. \end{aligned}$$

(a) Find a row echelon form of the augmented matrix for this system.

(b) For which values of h, if any, does the system have (i.) no solutions, (ii.) exactly one solution, (iii.) infinitely many solutions? Find the solutions in each case.

(5) Find the general solution of the linear system corresponding to the following augmented matrix:

$$\left[\begin{array}{cccc|c} 1 & -1 & 2 & 1 & 2 \\ -1 & 2 & 2 & -1 & -5 \\ 1 & 1 & 10 & 2 & -1 \end{array}\right].$$

(6) What are the conditions, if any, on the a, b, c values so that the following augmented matrix corresponds to a consistent linear system? How many solutions will the consistent system have? Explain.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & a \\ 2 & 3 & 7 & b \\ -1 & -4 & -1 & c \end{array}\right].$$

(7) In this exercise the symbol ■ denotes a non-zero number and the symbol * denotes any real number (including 0).

(a) Is the augmented matrix

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

in a form to which back substitution will easily give the solutions to the system? Explain your reasoning. (Hint: In order to help see what happens in the general case, substitute some numbers in place of the ■'s and *'s and answer the question for that specific system first. Then determine if your answer generalizes.)

(b) The above matrix is a possible form of an augmented matrix with 2 rows and 3 columns corresponding to a linear system after forward elimination, i.e., a linear system for which back substitution will easily give the solutions. Determine the other possible such forms of the nonzero augmented matrices with 2 rows and 3 columns. As in part (a), use the symbol ■ to denote a non-zero number and * to denote any real number.

(8) Give an example of a linear system with a unique solution for which a row echelon form of the augmented matrix of the system has a row of 0's.

(9) Come up with an example of an augmented matrix with 0's in the rightmost column corresponding to an inconsistent system, if possible. If not, explain why not.

(10) Find two different row echelon forms which are equivalent to the same matrix not given in row echelon form.

(11) Determine all possible row echelon forms of a 2×2 matrix. Use the symbol ■ to denote a non-zero number and * to denote a real number with no condition on being 0 or not to represent entries.

(12) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** The number of pivots of an $m \times n$ matrix cannot exceed m. (Note: Here m, n are some unknown numbers.)

(b) **True/False** The row echelon form of a matrix is unique.

(c) **True/False** The reduced row echelon form of a matrix is unique.

(d) **True/False** A system of equations where there are fewer equations than the number of unknowns (known as an underdetermined system) cannot have a unique solution.

(e) **True/False** A system of equations where there are more equations than the number of unknowns (known as an overdetermined system) cannot have a unique solution.

(f) **True/False** If a row echelon form of the *augmented matrix* of a system of three equations in two unknowns has three pivots, then the system is inconsistent.

(g) **True/False** If the coefficient matrix of a system has pivots in every row, then the system is consistent.

(h) **True/False** If there is a row of zeros in a row echelon form of the augmented matrix of a system of equations, the system has infinitely many solutions.

(i) **True/False** If there is a row of zeros in a row echelon form of the augmented matrix of a system of n equations in n variables, the system has infinitely many solutions.

(j) **True/False** If a linear system has no free variables, then the system has a unique solution.

(k) **True/False** If a linear system has a free variable, then the system has infinitely many solutions.

Project: Modeling a Chemical Reaction

Recall the chemical equation

$$C_2H_6 + O_2 \rightarrow CO_2 + H_2O$$

from the beginning of this section. This equation illustrates the reaction between ethane (C_2H_6) and oxygen (O_2),called the *reactants*, to produce carbon dioxide (CO_2) and water (H_2O), called the *products* of the reaction. In any chemical reaction, the total mass of the reactants must equal the total mass of the products. In our reaction the chemicals involved are made up of carbon (C), hydrogen (H), and oxygen (O) atoms. To balance the equation, we need to know how many molecules of each chemical are combined to preserve the number of atoms of C, H, and O.

Let x_1 be the number of molecules of C_2H_6, x_2 the number of molecules of O_2, x_3 the number of molecules of CO_2, and x_4 the number of molecules of H_2O in the reaction. We can then represent this reaction as

$$x_1C_2H_6 + x_2O_2 \rightarrow x_3CO_2 + x_4H_2O.$$

In each molecule (e.g., ethane C_2H_6), the subscripts indicate the number of atoms of each element in the molecule. So 1 molecule of ethane contains 2 atoms of carbon and 6 atoms of hydrogen. Thus, there are 2 atoms of carbon in C_2H_6 and 0 atoms of carbon in O_2, giving us $2x_1$ carbon atoms in x_1 molecules of C_2H_6 and 0 carbon atoms in x_2 molecules of O_2. On the product side of the reaction there is 1 carbon atom in CO_2 and 0 carbon atoms in H_2O. To balance the reaction, we know that the number of carbon atoms in the products must equal the number of carbon atoms in the reactants.

Project Activity 3.1.

(a) Set up an equation that balances the number of carbon atoms on both sides of the reaction.

(b) Balance the numbers of hydrogen and oxygen atoms in the reaction to explain why

$$\begin{aligned} 6x_1 &= 2x_4 \\ 2x_2 &= 2x_3 + x_4. \end{aligned}$$

(c) So the system of linear equations that models this chemical reaction is

$$\begin{array}{rcrcrcrcl} 2x_1 & & & - & x_3 & & & = & 0 \\ 6x_1 & & & & & - & 2x_4 & = & 0 \\ & & 2x_2 & - & 2x_3 & - & x_4 & = & 0. \end{array}$$

Find all solutions to this system and then balance the reaction. Note that we cannot have a fraction of a molecule in our reaction. (Hint: Some of the work needed is done in Preview Activity 3.1.)

Project Activity 3.2. Chemical reactions can be very interesting.

(a) Carbon dioxide, CO_2, is a familiar product of combustion. For example, when we burn glucose, $C_6H_{12}O_6$, the products of the reaction are carbon dioxide and water:

$$C_6H_{12}O_6 + O_2 \to CO_2 + H_2O. \tag{3.3}$$

Use the techniques developed in this project to balance this reaction.

(b) To burn glucose, we need to add oxygen to make the combustion happen. Carbon dioxide is different in that it can burn without the presence of oxygen. For example, when we mix magnesium (Mg) with dry ice (CO_2), the products are magnesium oxide (Mg) and carbon (C). This is an interesting reaction to watch: you can see it at many websites, e.g., http://www.ebaumsworld.com/video/watch/404311/ or https://www.youtube.com/watch?v=-6dfi8LyRLA.

Use the method determined above to balance the chemical reaction

$$Mg + CO_2 \to MgO + C. \tag{3.4}$$

Section 4

Vector Representation

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a vector?
- How do we define operations on vectors?
- What is a linear combination of vectors?
- How do we determine if one vector is a linear combination of a given set of vectors?
- How do we represent a linear system as a vector equation?
- What is the span of a set of vectors?
- What are possible geometric representations of the span of a vector, or the span of two vectors?

Application: The Knight's Tour

Chess is a game played on an 8×8 grid which utilizes a variety of different pieces. One piece, the knight, is different from the other pieces in that it can jump over other pieces. However, the knight is limited in how far it can move in a given turn. For these reasons, the knight is a powerful, but often under-utilized, piece.

A knight can move two units either horizontally or vertically, and one unit perpendicular to that. Four knight moves are as illustrated in Figure 4.1, and the other four moves are the opposites of these.

The knight's tour problem is the mathematical problem of finding a knight's tour, that is a se-

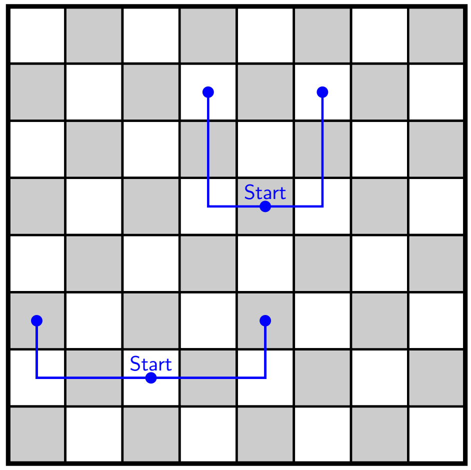


Figure 4.1: Moves a knight can make.

quence of knight moves so the the knight visits each square exactly once. While we won't consider a knight's tour in this text, we will see using linear combinations of vectors that a knight can move from its initial position to any other position on the board, and that it is possible to determine an sequence of moves to make that happen.

Introduction

So far we learned of a convenient method to represent a linear system using matrices. We now consider another representation of a linear system using *vectors*. Vectors can represent concepts in the physical world like velocity, acceleration, and force – but we will be interested in vectors as algebraic objects in this class. Vectors will form the foundation for everything we will do in linear algebra. For now, the following definition will suffice.

Definition 4.1. A (real) **vector** is a finite list of real numbers in a specified order. Each number in the list is referred to as an **entry** or **component** of the vector.

Note: For the majority of this text, we will work with real vectors. However, A vector does not need to be restricted to have real entries. At times we will use complex vectors and even vectors in other types of sets. The types of sets we use will be ones that have structure just like the real numbers. Recall that a real number is a number that has a decimal representation, either finite or repeating (rational numbers) or otherwise (irrational numbers). We can add and multiply real numbers as we have done throughout our mathematical careers, and the real numbers have a certain structure given in the following theorem that we will treat as an axiom – that is, we assume these properties without proof. We denote the set of real numbers with the symbol $\mathbb{R}$.

Theorem 4.2. *Let x, y, and z be real numbers. Then*

- $x + y \in \mathbb{R}$ *and* $xy \in \mathbb{R}$ *(*The name given to this property is closure. That is, the set $\mathbb{R}$ is closed under addition and multiplication.*)*

- $x + y = y + x$ *and* $xy = yx$ *(*The name given to this property is commutativity. That is addition and multiplication are commutative operations in $\mathbb{R}$.*)*

- $(x+y)+z = x+(y+z)$ *and* $(xy)z = x(yz)$ *(*The name given to this property is associativity. That is, addition and multiplication is associative operations in $\mathbb{R}$.*)*

- *There is an element* 0 *in* $\mathbb{R}$ *such that* $x + 0 = x$ *(*The element 0 is called the additive identity in $\mathbb{R}$.*)*

- *There is an element* 1 *in* $\mathbb{R}$ *such that* $(1)x = x$ *(*The element 1 is called the multiplicative identity in $\mathbb{R}$.*)*

- *There is an element* $-x$ *in* $\mathbb{R}$ *such that* $x + (-x) = 0$ *(*The element $-x$ is the additive inverse of x in $\mathbb{R}$.*)*

- *If* $x \neq 0$*, there is an element* $\frac{1}{x}$ *in* $\mathbb{R}$ *such that* $x\left(\frac{1}{x}\right) = 1$ *(*The element $\frac{1}{x}$ is the multiplicative inverse of the nonzero element x in $\mathbb{R}$.*)*

- $x(y + z) = (xy) + (xz)$ *(*The is the distributive property. That is, multiplication distributes over addition in $\mathbb{R}$.*)*

Any set that satisfies the properties listed in Theorem 4.2 is called a *field*. We our vectors are made from elements of a field, we call those elements of the field *scalars*.

We will algebraically represent a vector as a matrix with one column. For example, $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a vector with 2 entries, and we say that $\mathbf{v}$ is a vector in 2-space. By 2-space we mean $\mathbb{R}^2$, which can be geometrically modeled as the plane. Here the symbol $\mathbb{R}$ indicates that the entries of $\mathbf{v}$ are real numbers and the superscript 2 tells us that $\mathbf{v}$ has two entries. Similarly, vectors in $\mathbb{R}^3$ have three entries, e.g., $\begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. The collection of column vectors with three entries can be geometrically modeled as three-dimensional space. If a vector $\mathbf{v}$ has n entries we say that $\mathbf{v}$ is a vector in $\mathbb{R}^n$ (or n-space). Vectors are also often indicated with arrows, so we might also see a vector $\mathbf{v}$ written as $\vec{v}$. It is important when writing to differentiate between a vector $\mathbf{v}$ and a scalar v. These are quite different objects and it is up to us to make sure we are clear what a symbol represents. We will use boldface letters to represent vectors.

A vector like $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is called a *column vector* of *size* 2×1 (two rows, one column). We can define an addition operation on two vectors of the same size by adding corresponding components, such as

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

Similarly, we can define scalar multiplication of a vector by multiplying each component of the vector by the scalar. For example,

$$3\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Since we can add vectors and multiply vectors by scalars, we can then add together scalar multiples of vectors. For completeness, we define vector subtraction as adding a scalar multiple:

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-1)\mathbf{u}.$$

This definition is equivalent to defining subtraction of $\mathbf{u}$ from $\mathbf{v}$ by subtracting components of $\mathbf{u}$ from the corresponding components of $\mathbf{v}$.

Preview Activity 4.1.

(1) Given vectors

$$\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix},$$

determine the components of the vector $3\mathbf{v} + \mathbf{u} - 2\mathbf{w}$ using the operations defined above.

(2) In mathematics, any time we define operations on objects, such as addition of vectors, we ask which properties the operation has. For example, one might wonder if $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$ for any two vectors $\mathbf{u}, \mathbf{v}$ of the same size. If this property holds, we say that the *addition of vectors is a commutative operation*. However, to verify this property we cannot use examples since the property must hold for any two vectors. For simplicity, we focus on two-dimensional vectors $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Using these arbitrary vectors, can we say that $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$? If so, justify. If not, give a counterexample. (Note: Giving a counterexample is the best way to justify why a general statement is not true.)

(3) One way to geometrically represent vectors with two components uses a point in the plane to correspond to a vector. Specifically, the vector $\begin{bmatrix} x \\ y \end{bmatrix}$ corresponds to the point (x, y) in the plane. As a specific example, the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ corresponds to the point $(1, 2)$ in the plane. This representation will be especially handy when we consider infinite collections of vectors as we will do in this problem.

   (a) On the same set of axes, plot the points that correspond to 5-6 scalar multiples of the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Make sure to use variety of scalar multiples covering possibilities with $c > 0, c < 0, c > 1, 0 < c < 1, -1 < c < 0$. If we consider the collection of all possible scalar multiples of this vector, what do we obtain?

   (b) What would the collection of all scalar multiples of the vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ form in the plane?

   (c) What would the collection of all scalar multiples of the vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ form in the three-dimensional space?

(4) Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ in $\mathbb{R}^2$. We are interested in finding all vectors that can be formed as a sum of scalar multiples of $\mathbf{u}$ and $\mathbf{v}$.

(a) On the same set of axes, plot the points that correspond to the vectors $\mathbf{u}, \mathbf{v}, \mathbf{u}+\mathbf{v}, 1.5\mathbf{u}, 2\mathbf{v}, -\mathbf{u}, -\mathbf{v}, -\mathbf{u}+2\mathbf{v}$. Plot other random sums of scalar multiples of $\mathbf{u}$ and $\mathbf{v}$ using several scalar multiples (including those less than 1 or negative) (that is, find other vectors of the form $a\mathbf{u}+b\mathbf{v}$ where a and b are any scalars.).

(b) If we considered sums of all scalar multiples of $\mathbf{u}, \mathbf{v}$, which vectors will we obtain? Can we obtain any vector in $\mathbb{R}^2$ in this form?

Vectors and Vector Operations

As discussed in Preview Activity 4.1, a vector is simply a list of numbers. We can add vectors of like size and multiply vectors by scalars. These operations define a structure on the set of all vectors with the same number of components that will be our major object of study in linear algebra. Ultimately we will expand our idea of vectors to a more general context and study what we will call *vector spaces*.

In Preview Activity 4.1 we saw how to add vectors and multiply vectors by scalars in $\mathbb{R}^2$, and this idea extends to $\mathbb{R}^n$ for any n. Before we do so, one thing we didn't address in Preview Activity 4.1 is what it means for two vectors to be equal. It should seem reasonable that two vectors are equal if and only if they have the same corresponding components. More formally, if we let

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \text{ and } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

be vectors in $\mathbb{R}^n$, then $\mathbf{u} = \mathbf{v}$ if $u_i = v_i$ for every i between 1 and n. Note that this statement implies that a vector in $\mathbb{R}^2$ cannot equal a vector in $\mathbb{R}^3$ because they don't have the same number of components. With this in mind we can now define the sum $\mathbf{u}+\mathbf{v}$ of the vectors $\mathbf{u}$ and $\mathbf{v}$ to be the vector in $\mathbb{R}^n$ defined by

$$\mathbf{u}+\mathbf{v} = \begin{bmatrix} u_1+v_1 \\ u_2+v_2 \\ \vdots \\ u_n+v_n \end{bmatrix}.$$

In other words, to add two vectors of the same size, we add corresponding components.

Similarly, we can define scalar multiplication of a vector. If c is a scalar, then the scalar multiple $c\mathbf{v}$ of the vector $\mathbf{v}$ is the vector in $\mathbb{R}^n$ defined by

$$c\mathbf{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

In other words, the scalar multiple $c\mathbf{v}$ of the vector $\mathbf{v}$ is the vector obtained by multiplying each component of the vector $\mathbf{v}$ by the scalar c. Since we can add vectors and multiply vectors by scalars, we can then add together scalar multiples of vectors. For completeness, we define vector subtraction as adding a scalar multiple:

$$\mathbf{v} - \mathbf{u} = \mathbf{v} + (-1)\mathbf{u}.$$

This definition is equivalent to defining subtraction of $\mathbf{u}$ from $\mathbf{v}$ by subtracting components of $\mathbf{u}$ from the corresponding components of $\mathbf{v}$.

After defining operations on objects, we should wonder what kinds of properties these operations have. For example, with the operation of addition of real numbers we know that $1+2$ is equal to $2+1$. This is called the *commutative* property of scalar addition and says that order does not matter when we add real numbers. It is natural for us to ask if similar properties hold for the vector operations, addition and scalar multiplication, we defined. You showed in Preview Activity 4.1 that the addition operation is also commutative on vectors in $\mathbb{R}^2$.

In the activity below we consider how the two operations, addition and scalar multiplication, interact with each other. In real numbers, we know that multiplication is distributive over addition. Is that true with vectors as well?

Activity 4.1. We work with vectors in $\mathbb{R}^2$ to make the notation easier.

Let a be an arbitrary scalar, and $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ be two *arbitrary* vectors in $\mathbb{R}^2$. Is $a(\mathbf{u}+\mathbf{v})$ equal to $a\mathbf{u} + a\mathbf{v}$? What property does this imply about the scalar multiplication and addition operations on vectors?

Similar arguments can be used to show the following properties of vector addition and multiplication by scalars.

Theorem 4.3. *Let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let a and b be scalars. Then*

(1) $\mathbf{v} + \mathbf{u} = \mathbf{u} + \mathbf{v}$

(2) $(\mathbf{v} + \mathbf{u}) + \mathbf{w} = \mathbf{v} + (\mathbf{u} + \mathbf{w})$

(3) The vector $\mathbf{z} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ *has the property that* $\mathbf{v} + \mathbf{z} = \mathbf{v}$. *The vector* $\mathbf{z}$ *is called the* ***zero vector****.*

(4) $(-1)\mathbf{v} + \mathbf{v} = \mathbf{z}$. *The vector* $(-1)\mathbf{v} = -\mathbf{v}$ *is called the* ***additive inverse*** *of the vector* $\mathbf{v}$.

(5) $(a+b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

(6) $a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$

(7) $(ab)\mathbf{v} = a(b\mathbf{v})$

(8) $1\mathbf{v} = \mathbf{v}$.

We will later see that the above properties make the set $\mathbb{R}^n$ a *vector space*. These properties just say that, for the most part, we can manipulate vectors just as we do real numbers. Please note, though, that there is no multiplication or division of vectors.

Geometric Representation of Vectors and Vector Operations

We can geometrically represent a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in $\mathbb{R}^2$ as the point (v_1, v_2) in the plane as we did in Preview Activity 4.1. We can similarly represent a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ in $\mathbb{R}^3$ as the point (v_1, v_2, v_3) in the three-dimensional space. This geometric representation will be handy when we consider collections of infinitely many vectors, as we will do when we consider the span of a collection of vectors later in this section.

We can also represent the vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ in $\mathbb{R}^2$ as the directed line segment (or arrow) from the origin to the point (v_1, v_2) as shown in Figure 4.2 to aid in the visualization.

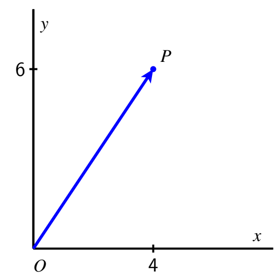


Figure 4.2: The vector $\left[\begin{smallmatrix} 4 \\ 6 \end{smallmatrix}\right]$ in $\mathbb{R}^2$.

The fact that the vector in Figure 4.2 is represented by the directed line segment from the origin to the point (4,6) means that this vector is the vector $\mathbf{v} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$. If O is the origin and P is the point $(4, 6)$, we will also denote this vector as $\overrightarrow{OP}$ – so

$$\overrightarrow{OP} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

In this way we can think of vectors as having direction and length. With the Pythagorean Theorem, we can see that the length of a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is $\sqrt{v_1^2 + v_2^2}$. This idea can be applied to vectors in any space. If $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}$ is a vector in $\mathbb{R}^n$, then the **length** of $\mathbf{v}$, denoted $|\mathbf{v}|$ is the scalar

$$||\mathbf{v}|| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

Thinking of vectors having direction and length is especially useful in visualizing the addition of vectors. The geometric interpretation of the sum of two vectors can be seen in Figures 4.3 and 4.4.

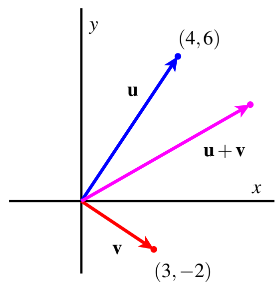


Figure 4.3: A vector sum.

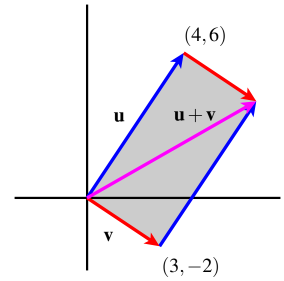


Figure 4.4: Geometric interpretation.

Let $\mathbf{u} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Then $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$ as shown in Figure 4.3. Figure 4.4 provides a context to interpret this vector sum geometrically. Using the parallelogram imposed on the three vectors, we see that if vectors $\mathbf{u}$ and $\mathbf{v}$ are both placed to start at the origin, then the vector sum $\mathbf{u}+\mathbf{v}$ can be visualized geometrically as the directed line segment from the origin to the fourth corner of the parallelogram.

In Preview Activity 4.1 we considered scalar multiples of a vector in $\mathbb{R}^2$. The arrow representation helps in visualizing scalar multiples as well. Geometrically, a scalar multiple $c\mathbf{v}$ of a nonzero vector $\mathbf{v}$ is a vector in the same direction as $\mathbf{v}$ if $c > 0$ and in the opposite direction as $\mathbf{v}$ if $c < 0$. If $c > 1$, scalar multiplication stretches the vector, while $0 < c < 1$ shrinks the vector. We also saw that the collection of all scalar multiples of a vector $\mathbf{v}$ in $\mathbb{R}^2$ gives us a line through the origin and $\mathbf{v}$, except when $\mathbf{v} = \mathbf{0}$ in which case we only obtain $\mathbf{0}$. In other words, for a nonzero vector $\mathbf{v}$, the set $S = \{c\mathbf{v} : c \text{ is a scalar}\}$ is the line through the origin and $\mathbf{v}$ in $\mathbb{R}^2$.

All of these properties generalize to vectors in $\mathbb{R}^3$. Specifically, the scalar multiple $c\mathbf{v}$ is a vector in the same or opposite direction as $\mathbf{v}$ based on the sign of c, and is a stretched or shrunken version of $\mathbf{v}$ based on whether $|c| > 1$ or $|c| < 1$. Also, the collection of all multiples of a non-zero vector $\mathbf{v}$ in $\mathbb{R}^3$ form a line through the origin.

Linear Combinations of Vectors

The concept of linear combinations is one of the fundamental ideas in linear algebra. We will use linear combinations to describe almost every important concept in linear algebra – the span of a set of vectors, the range of a linear transformation, bases, the dimension of a vector space – to name just a few.

In Preview Activity 4.1, we considered the sets of all scalar multiples of a single nonzero vector in $\mathbb{R}^2$ and in $\mathbb{R}^3$. We also considered the set of all sums of scalar multiples of two nonzero vectors. These results so far gives us an idea of geometrical descriptions of sets of vectors generated by one or two vectors. Oftentimes we are interested in what vectors can be made from a given collection of vectors. For example, suppose we have two different water-benzene-acetic acid chemical solutions, one with 40% water, 50% benzene and 10% acetic acid, the other with 52% water, 42% benzene and 6% acid. An experiment we want to conduct requires a chemical solution with 43% water, 48% benzene and 9% acid. We would like to know if we make this new chemical solution by mixing the first two chemical solutions, or do we have to run to the chemical solutions market to get the chemical solution we want.

We can set up a system of equations for each ingredient and find the answer. But we can also consider each chemical solution as a vector, where the components represent the water, benzene and acid percentages. So the two chemical solutions we have are represented by the vectors $\mathbf{v}_1 = \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$. If we mix the two chemical solutions with varying amounts of each ingredient, then the question of whether we can make the desired chemical solution becomes the question of whether the equation

$$c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix} = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}$$

has a solution. (You will determine if this equation has a solution in Exercise 5.)

We might also be interested in what other chemical solutions we can make from the two given solutions. This amounts to determining which vectors can be written in the form $c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ for scalars c_1 and c_2. Vectors that are created from sums of scalar multiples of given vectors are called linear combinations of those vectors. More formally,

Definition 4.4. A **linear combination** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ in $\mathbb{R}^n$ is any vector of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m, \tag{4.1}$$

where $c_1, c_2, \ldots, c_m$ are scalars that we will refer to as the **weights**.

In the chemical solutions example, the vector $c_1 \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix} + c_2 \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ for scalars c_1 and c_2 is a linear combination of the vectors $\begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$ and $\begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$ with weights c_1 and c_2, and the set of linear combinations of the given chemical solution vectors tells us exactly which chemical solutions we can make from the given ones. This is one example of how linear combinations can arise in applications.

The set of all linear combinations of a fixed collection of vectors has a very nice algebraic structure and, in small dimensions, allows us to use a geometrical description to aid our understanding. In the above example, this collection gives us the type of chemical solutions we can make by combining the first two solutions in varying amounts.

Activity 4.2. Our chemical solution example illustrates that it can be of interest to determine whether certain vectors can be written as a linear combination of given vectors. We explore that idea in more depth in this activity. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

(a) Calculate the linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ with corresponding weights (scalar multiples) 1 and 2. The resulting vector is a vector which can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(b) Can $\mathbf{w} = \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}$ be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$? If so, which linear combination? If not, explain why not.

(c) Can $\mathbf{w} = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$ be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$? If so, which linear combination? If not, explain why not.

(d) Let $\mathbf{w} = \begin{bmatrix} 0 \\ 6 \\ -2 \end{bmatrix}$. The problem of determining if $\mathbf{w}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ is equivalent to the problem of finding scalars x_1 and x_2 so that

$$\mathbf{w} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2. \tag{4.2}$$

i. Combine the vectors on the right hand side of equation (4.2) into one vector, and then set the components of the vectors on both sides equal to each other to convert the vector equation (4.2) to a linear system of three equations in two variables.

ii. Use row operations to find a solution, if it exists, to the system you found in the previous part of this activity. If you find a solution, verify in (4.2) that you have found appropriate weights to produce the vector $\mathbf{w}$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

Note that to find the weights that make $\mathbf{w}$ a linear combination of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, we simply solved the linear system corresponding to the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \mid \mathbf{w}],$$

where the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{w}$ form the columns of an augmented matrix, and the solution of the system gave us the weights of the linear combination. In general, if we want to find weights c_1, c_2, ..., c_m so that a vector $\mathbf{w}$ in $\mathbb{R}^n$ is a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ in $\mathbb{R}^n$, we solve the system corresponding to the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_m | \ \mathbf{w}].$$

Any solution to this system will gives us the weights. If this system has no solutions, then $\mathbf{w}$ cannot be written as a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$. This shows us the equivalence of the linear system and its vector equation representation. Specifically, we have the following result.

Theorem 4.5. *The vector equation*

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + \cdots + x_m\mathbf{v}_m = \mathbf{w}$$

has the same solution set as the linear system represented by the augmented matrix

$$[\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \ \cdots \ \mathbf{v}_m | \ \mathbf{w}].$$

In particular, the system has a solution if and only if $\mathbf{w}$ *is a linear combination of the vectors* $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_m$.

Activity 4.3.

(a) Represent the following linear system as a vector equation. After finding the vector equation, compare your vector equation to the matrix representation you found in Preview Activity 4.1. (Note that this is the same linear system from Preview Activity 3.1.)

$$\begin{aligned} -x_3 + 3 + 2x_2 &= -x_1 \\ -3 + 2x_3 &= -x_2 \\ -2x_2 + x_1 &= 3x_3 - 7 \end{aligned}$$

(b) Represent the following vector equation as a linear system and solve the linear system.

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 11 \end{bmatrix}$$

The Span of a Set of Vectors

As we saw in the previous section, the question of whether a system of linear equations has a solution is equivalent to the question of whether the vector obtained by the non-coefficient constants in the system is a linear combination of the vectors obtained from the columns of the coefficient matrix of the system. So if we were interested in finding for which constants the system has a solution, we would look for the collection of all linear combinations of the columns. We call this collection the *span* of these vectors. In this section we investigate the concept of span both algebraically and geometrically.

Our work in Preview Activity 4.1 seems to indicate that the span of a set of vectors, i.e., the collection of all linear combinations of this set of vectors, has a nice structure. As we mentioned above, the span of a set of vectors represents the collection of all constant vectors for which a linear system has a solution, but we will also see that other important objects in linear algebra can be represented as the span of a set of vectors.

Definition 4.6. The **span** of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ in $\mathbb{R}^n$ is the collection of all linear combinations of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$.

Notation: We denote the span of a set of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ as

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}.$$

So

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\} = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m : c_1, c_2, \ldots, c_m \text{ are scalars}\}.$$

The curly braces, { }, are used in denoting sets. They represent the whole set formed by the objects included between them. So $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ represents the collection of the vectors formed by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ for an arbitrary number m. Note that m can be 1, meaning that the collection can contain only one vector $\mathbf{v}_1$.

We now investigate what the span of a set of one or two vectors is, both from an algebraic and geometric perspective, and consider what happens for more general spanning sets.

Activity 4.4.

(a) By definition, $\text{Span}\left\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$ is the collection of all vectors which are scalar multiples of $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Determine which vectors are in this collection. If we plot all these vectors with each vector being represented as a point in the plane, what do they form?

(b) Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ in $\mathbb{R}^3$. By definition,

$$\text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$$

is the collection of all linear combinations of the form

$$x_1\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + x_2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

where x_1 and x_2 are any scalars.

i. Find four different vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ and indicate the weights (the values of x_1 and x_2) for each linear combination. (Hint: It is really easy to find 3 vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ for any $\mathbf{v}_1, \mathbf{v}_2$.)

ii. Are there any vectors in $\mathbb{R}^3$ that are not in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$? Explain. Verify your result.

iii. Set up a linear system to determine which vectors $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ are in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Specifically, which $\mathbf{w}$ can be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$?

iv. Geometrically, what shape do the vectors in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ form inside $\mathbb{R}^3$?

(c) Is it possible for $\text{Span}\{\mathbf{z}_1, \mathbf{z}_2\}$ to be a line for two vectors $\mathbf{z}_1, \mathbf{z}_2$ in $\mathbb{R}^3$?

(d) What do you think are the possible geometric descriptions of a span of a set of vectors in $\mathbb{R}^2$? Explain.

(e) What do you think are the possible spans of a set of vectors in $\mathbb{R}^3$? Explain.

Examples

What follows are worked examples that use the concepts from this section.

Example 4.7. For each of the following systems,

- express an arbitrary solution to the system algebraically as a linear combination of vectors,
- find a set of vectors that spans the solution set,
- describe the solution set geometrically.

(a)

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

(b)

$$\begin{aligned} x_1 + 2x_2 + \ \ 3x_3 &= 0 \\ 2x_1 + 4x_2 + \ \ 6x_3 &= 0 \\ 4x_1 + 8x_2 + 12x_3 &= 0. \end{aligned}$$

Example Solution. In each example, we use technology to find the reduced row echelon form of the augmented matrix.

(a) The reduced row echelon form of the augmented matrix

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

is

$$\left[\begin{array}{rrr|r} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

- There is no pivot in the x_3 column, so x_3 is a free variable. Since the system is consistent, it has infinitely many solutions. We can write both x_1 and x_2 in terms of x_3 as $x_2 = -x_3$ and $x_1 = -x_3$. So the general solution to the system has the algebraic form
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_3 \\ -x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}.$$
So every solution to this system is a scalar multiple (linear combination) of the vector $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

- Since every solution to the system is a scalar multiple of the vector $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, the solution set to the system is Span $\left\{ \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \right\}$.
- As the set of scalar multiples of a single vector, the solution set to this system is a line in $\mathbb{R}^3$ through the origin and the point $(-1, -1, 1)$.

(b) The reduced row echelon form of the augmented matrix

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 2 & 4 & 6 & 0 \\ 4 & 8 & 12 & 0 \end{array}\right]$$

is

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

- There are no pivots in the x_2 and x_3 columns, so x_2 and x_3 are free variables. Since the system is consistent, it has infinitely many solutions. We can write x_1 in terms of x_2 and x_3 as $x_1 = -2x_2 - 3x_3$. So the general solution to the system has the algebraic form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 - 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}.$$

So every solution to this system is a linear combination of the vectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$.

- Since every solution to the system is a linear combination of the vectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$, the solution set to the system is

$$\text{Span} \left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

- As the set of linear combinations of two vectors, the solution set to this system is a plane in $\mathbb{R}^3$ through the origin and the points $(-2, 1, 0)$ and $(-3, 0, 1)$.

Example 4.8. Let $W = \left\{ \begin{bmatrix} s+t \\ r+2s \\ r-3t \\ r+s+t \end{bmatrix} : r, s, t \in \mathbb{R} \right\}$.

(a) Find three vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ such that $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(b) Can $\mathbf{w} = \begin{bmatrix} -2 \\ -4 \\ -1 \\ 0 \end{bmatrix}$ be written as a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$? If so, find such a linear combination. If not, justify your response. What does your result tell us about the relationship between $\mathbf{w}$ and W? Explain.

(c) Can $\mathbf{u} = \begin{bmatrix} 3 \\ -4 \\ 1 \\ -1 \end{bmatrix}$ be written as a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$? If so, find such a linear combination. If not, justify your response. What does your result tell us about the relationship between $\mathbf{w}$ and W? Explain.

(d) What relationship, if any, exists between $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and Span W? Explain.

Example Solution.

(a) Every vector in W has the form

$$\begin{bmatrix} s+t \\ r+2s \\ r-3t \\ r+s+t \end{bmatrix} = \begin{bmatrix} 0 \\ r \\ r \\ r \end{bmatrix} + \begin{bmatrix} s \\ 2s \\ 0 \\ s \end{bmatrix} + \begin{bmatrix} t \\ 0 \\ -3t \\ t \end{bmatrix}$$
$$= r \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \end{bmatrix}$$

for some real numbers r, s, and t. Thus, $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ -3 \\ 1 \end{bmatrix}$.

(b) To determine if $\mathbf{w}$ is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, we row reduced the augmented matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3 \mid \mathbf{w}]$. The reduced row echelon form of the matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3 \mid \mathbf{w}]$ is

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

The system with this as augmented matrix is consistent. If we let x_1, x_2, and x_3 be the variables corresponding to the first three columns, respectively, of this augmented matrix, then we see that $x_1 = 2$, $x_2 = -3$, and $x_3 = 1$. So $\mathbf{w}$ can be written as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ as

$$\mathbf{w} = 2\mathbf{v}_1 - 3\mathbf{v}_2 + \mathbf{v}_3.$$

Since $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, it follows that $\mathbf{w} \in W$.

(c) To determine if $\mathbf{u}$ is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, we row reduced the augmented matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3 \mid \mathbf{u}]$. The reduced row echelon form of the matrix $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3 \mid \mathbf{u}]$ is

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

The last row shows that the system with this as augmented matrix is inconsistent. So $\mathbf{u}$ cannot be written as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Since $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, it follows that $\mathbf{u} \notin W$.

(d) We know that $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = W$. Now Span W contains the linear combinations of vectors in W, which are all linear combinations of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. Thus, Span W is just the set of linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. We conclude that Span $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = W$.

Summary

- A vector is a list of numbers in a specified order.
- We add two vectors of the same size by adding corresponding components. In other words, if $\mathbf{u}$ and $\mathbf{v}$ are vectors of the same size and u_i and v_i are the i components of $\mathbf{u}$ and $\mathbf{v}$, respectively, then $\mathbf{u}+\mathbf{v}$ is the vector whose ith component is u_i+v_i for each i. Geometrically, we represent the sum of two vectors using the Parallelogram Rule: The vector $\mathbf{u} + \mathbf{v}$ is the directed line segment from the origin to the 4th point of the parallelogram formed by the origin and the vectors $\mathbf{u}, \mathbf{v}$.
- A scalar multiple of a vector is found by multiplying each component of the vector by that scalar. In other words, if v_i is the i component of the vector $\mathbf{v}$ and c is any scalar, then $c\mathbf{v}$ is the vector whose i component is cv_i for each i. Geometrically, a scalar multiple of a nonzero vector $\mathbf{v}$ is a vector in the same direction as $\mathbf{v}$ if $c > 0$ and in the opposite direction if $c < 0$. If $|c| > 1$, the vector is stretched, and if $|c| < 1$, the vector is shrunk.

- An important concept is that of a linear combination of vectors. In words, a linear combination of a collection of vectors is a sum of scalar multiples of the vectors. More formally, we defined a linear combination of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ in $\mathbb{R}^n$ is any vector of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m$, where c_1, c_2, ..., c_m are scalars.

- To find weights c_1, c_2, ..., c_m so that a vector $\mathbf{w}$ in $\mathbb{R}^n$ is a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ in $\mathbb{R}^n$, we simply solve the system corresponding to the augmented matrix

$$[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_m|\ \mathbf{w}].$$

- The collection of all linear combinations of a set of vectors is called the span of the set of vectors. More formally, the span of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_m$ in $\mathbb{R}^n$ is the set

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m : c_1, c_2, \ldots, c_m \text{ are scalars}\},$$

which we denote as Span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$. Geometrically, the span of a single nonzero vector $\mathbf{v}$ in any dimension is the line through the origin and the vector $\mathbf{v}$. The span of two vectors $\mathbf{v}_1, \mathbf{v}_2$ in any dimension neither of which is a multiple of the other is a plane through the origin containing both vectors.

Exercises

(1) Given vectors $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ in $\mathbb{R}^2$, determine if $\mathbf{w} = \begin{bmatrix} -4 \\ -1 \end{bmatrix}$ can be written as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. If so, determine the weights of $\mathbf{u}$ and $\mathbf{v}$ which produce $\mathbf{w}$.

(2) Given vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix} -1 \\ 3 \\ 3 \end{bmatrix}$ in $\mathbb{R}^3$, determine if $\mathbf{w} = \begin{bmatrix} 5 \\ 5 \\ 1 \end{bmatrix}$ can be written as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$. If so, determine the weights of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ which produce $\mathbf{w}$. Reflect on the result. Is there anything special about the given vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$?

(3) Let $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$ in $\mathbb{R}^3$. Determine which vectors $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$ in $\mathbb{R}^3$ can be written as a linear combination of $\mathbf{u}$ and $\mathbf{v}$. Does the set of $\mathbf{w}$'s include the 0 vector? If so, determine which weights in the linear combination produce the 0 vector. If not, explain why not.

(4) Consider vectors $\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ in $\mathbb{R}^3$.

   (a) Find four specific linear combinations of the vectors $\mathbf{u}$ and $\mathbf{v}$.

   (b) Explain why the zero vector must be a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

(c) What kind of geometric shape does the set of all linear combinations of $\mathbf{u}$ and $\mathbf{v}$ have in $\mathbb{R}^3$?

(d) Can we obtain any vector in $\mathbb{R}^3$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$? Explain.

(5) Suppose we have two different water-benzene-acetic acid solutions, one with 40% water, 50% benzene and 10% acetic acid, the other with 52% water, 42% benzene and 6% acid.

(a) An experiment we want to conduct requires a solution with 43% water, 48% benzene and 9% acid. Representing each acid solution as a vector, determine if we can we make this new acid solution by mixing the first two solutions, or do we have to run to the chemical solutions market to get the solution we want?

(b) Using the water-benzene-acetic acid solutions in the previous problem, can we obtain an acid solution which contains 50% water, 43% benzene and 7% acid?

(c) Determine the relationship between the percentages of water, benzene, and acid in solutions which can be obtained by mixing the two given water-benzene-acetic acid solutions above.

(6) Is the vector $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ in Span $\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$? Justify your answer.

(7) Describe geometrically each of the following sets.

(a) Span $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right\}$ in $\mathbb{R}^2$

(b) Span $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}$ in $\mathbb{R}^3$

(8) Consider the linear system

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

(a) Find the general solution to this system.

(b) Find two specific vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ so that the solution set to this system is Span$\{\mathbf{v}_1, \mathbf{v}_2\}$.

(9) Answer the following question as yes or no. Verify your answer. If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, then $\mathbf{v}$ is in Span$\{\mathbf{u}, \mathbf{u} - \mathbf{v}\}$.

(10) Let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let a and b be scalars. Verify Theorem 4.3. That is, show that

(a) $\mathbf{v} + \mathbf{u} = \mathbf{u} + \mathbf{v}$

(b) $(\mathbf{v} + \mathbf{u}) + \mathbf{w} = \mathbf{v} + (\mathbf{u} + \mathbf{w})$

(c) The vector $\mathbf{z} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$ has the property that $\mathbf{v} + \mathbf{z} = \mathbf{v}$.

(d) $(-1)\mathbf{v} + \mathbf{v} = \mathbf{z}$.

(e) $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$

(f) $a(\mathbf{v} + \mathbf{u}) = a\mathbf{v} + a\mathbf{u}$

(g) $(ab)\mathbf{v} = a(b\mathbf{v})$

(h) $1\mathbf{v} = \mathbf{v}$.

(11) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** A vector in $\mathbb{R}^2$, i.e. a two-dimensional vector, is also a vector in $\mathbb{R}^3$.

(b) **True/False** Any vector in $\mathbb{R}^2$ can be visualized as a vector in $\mathbb{R}^3$ by adding a 0 as the last coordinate.

(c) **True/False** The zero vector is a scalar multiple of any other vector (of the same size).

(d) **True/False** The zero vector cannot be a linear combination of two non-zero vectors.

(e) **True/False** Given two vectors $\mathbf{u}$ and $\mathbf{v}$, the vector $\frac{1}{2}\mathbf{u}$ is a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

(f) **True/False** Given any two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, we can obtain any vector in $\mathbb{R}^2$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

(g) **True/False** Given any two distinct vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^2$, we can obtain any vector in $\mathbb{R}^2$ as a linear combination of $\mathbf{u}$ and $\mathbf{v}$.

(h) **True/False** If $\mathbf{u}$ can be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, then $2\mathbf{u}$ can also be expressed as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(i) **True/False** The span of any two vectors neither of which is a multiple of the other can be visualized as a plane through the origin.

(j) **True/False** Given any vector, the collection of all linear combinations of this vector can be visualized as a line through the origin.

(k) **True/False** The span of any collection of vectors includes the $\mathbf{0}$ vector.

(l) **True/False** If the span of $\mathbf{v}_1$ and $\mathbf{v}_2$ is all of $\mathbb{R}^2$, then so is the span of $\mathbf{v}_1$ and $\mathbf{v}_1 + \mathbf{v}_2$.

(m) **True/False** If the span of $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ is all of $\mathbb{R}^3$, then so is the span of $\mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{v}_2 + \mathbf{v}_3$.

Project: Analyzing Knight Moves

To understand where a knight can move in a chess game, we need to know the initial setup. A chess board is an 8×8 grid. To be able to refer to the individual positions on the board, we will place the board so that its lower left corner is at the origin, make each square in the grid have side length 1, and label each square with the point at the lower left corner. This is illustrated at left in Figure 4.5.

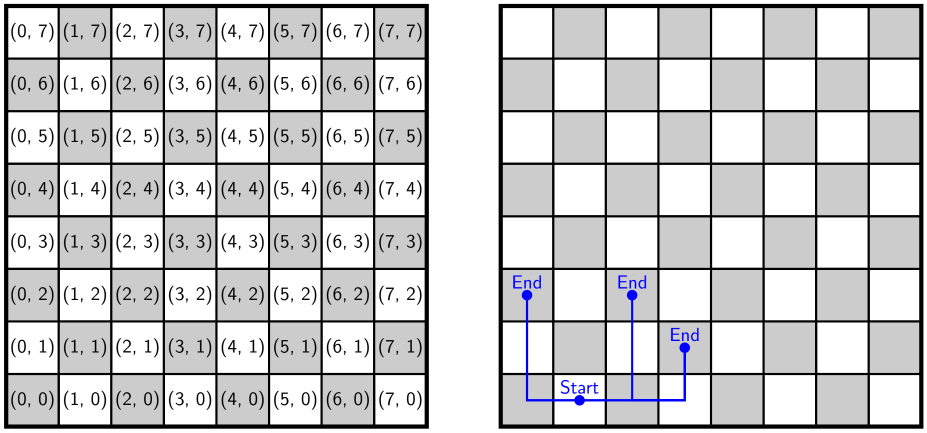


Figure 4.5: Initial knight placement and moves.

Each player has two knights to start the game, for one player the knights would begin in positions $(1, 0)$ and $(6, 0)$. Because of the symmetry of the knight's moves, we will only analyze the moves of the knight that begins at position $(1, 0)$. This knight has only three allowable moves from its starting point (assuming that the board is empty), as shown at right in Figure 4.5. The questions we will ask are: given any position on the board, can the knight move from its start position to that position using only knight moves and, what sequence of moves will make that happen. To answer these questions we will use linear combinations of knight moves described as vectors.

Each knight move can be described by a vector. A move one position to the right and two up can be represented as $\mathbf{n}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Three other moves are $\mathbf{n}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{n}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $\mathbf{n}_4 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. The other four knight moves are the additive inverses of these four. Any sequence of moves by the knight is given by the linear combination

$$x_1\mathbf{n}_1 + x_2\mathbf{n}_2 + x_3\mathbf{n}_3 + x_4\mathbf{n}_4.$$

A word of caution: the knight can only make complete moves, so we are restricted to integer (either positive or negative) values for x_1, x_2, x_3, and x_4. You can use the GeoGebra app at https://www.geogebra.org/m/dfwtskrj to see the effects the weights have on the knight moves. We should note here that since addition of vectors is commutative, the order in which we apply our moves does not matter. However, we may need to be careful with the order so that our knight does not leave the chess board.

Project Activity 4.1.

(a) Explain why the vector equation

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_1\mathbf{n}_1 + x_2\mathbf{n}_2 + x_3\mathbf{n}_3 + x_4\mathbf{n}_4 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$$

will tell us if it is possible for the knight to move from its initial position at $(1, 0)$ to the position (5,2).

(b) Find all solutions, if any, to the system from part (a). If it is possible to find a sequence of moves that take the knight from its initial position to position $(5, 2)$, find weights x_1, x_2, x_3, and x_4 to accomplish this move. (Be careful – we must have solutions in which x_1, x_2, x_3, and x_4 are integers.) Is there more than one sequence of possible moves? You can check your solution with the GeoGebra app at https://www.geogebra.org/m/dfwtskrj.

Project Activity 4.1 shows that it is possible for our knight to move to position $(5, 2)$ on the board. We would like to know if it is possible to move to any position on the board. That is, we would like to know if the integer span of the four moves $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}_3$, and $\mathbf{n}_4$ will allow our knight to cover the entire board. This takes a bit more work.

Project Activity 4.2. Given any position (a, b), we want to know if our knight can move from its start position $(1, 0)$ to position (a, b).

(a) Write a vector equation whose solution will tell us if it is possible for our knight to move from its start position $(1, 0)$ to position (a, b).

(b) Show that the solution to part (a) can be written in the form

$$x_1 = \frac{1}{4}\left(-5x_3 + 3x_4 + b + 2(a-1)\right) \tag{4.3}$$

$$x_2 = \frac{1}{4}\left(3x_3 - 5x_4 + b - 2(a-1)\right) \tag{4.4}$$

x_3 is free

x_4 is free.

To answer our question if our knight can reach any position, we now need to determine if we can always find integer values of x_3 and x_4 to make equations (4.3) and (4.4) have integer solutions. In other words, we need to find values of x_3 and x_4 so that $-5x_3 + 3x_4 + b + 2(a - 1)$ and $3x_3 - 5x_4 + b - 2(a - 1)$ are multiples of 4. How we do this could depend on the parity (even or odd) of a and b. For example, if a is odd and b is even, say $a = 2r + 1$ and $b = 2s$ for some integers r and s, then

$$x_1 = \frac{1}{4}\left(-5x_3 + 3x_4 + 2s + 4r\right)$$

$$x_2 = \frac{1}{4}\left(3x_3 - 5x_4 + 2s - 4r\right).$$

With a little trial and error we can see that if we let $x_3 = x_4 = s$, then $x_1 = r$ and $x_2 = -r$ is a solution with integer weights. For example, when $a = 5$ and $b = 2$ we have $r = 2$ and $s = 1$. This makes $x_1 = 2$, $x_2 = -2$, $x_3 = 1 = x_4$. Compare this to the solution(s) you found in Project Activity 4.1. This analysis shows us how to move our knight to any position (a, b) where a is odd and b is even.

Project Activity 4.3. Complete the analysis as above to determine if there are integer solutions to our knight's move system in the following cases.

(a) a odd and b odd

(b) a even and b even

(c) a even and b odd.

Project Activity 4.3 shows that for any position on the chess board, using linear combinations of move vectors, we can find a sequence of moves that takes our knight to that position. (We actually haven't shown that these moves can be made so that our knight always stays on the board – we leave that question to you.)

Section 5

The Matrix-Vector Form of a Linear System

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- How and when is the matrix-vector product $A\mathbf{x}$ defined?
- How can a system of linear equations be written in matrix-vector form?
- How can we tell if the system $A\mathbf{x} = \mathbf{b}$ is consistent for a given vector $\mathbf{b}$?
- How can we tell if the system $A\mathbf{x} = \mathbf{b}$ is consistent for every vector $\mathbf{b}$?
- What is a homogeneous system? What can we say about the solution set to a homogeneous system?
- What must be true about pivots in the coefficient matrix A in order for the homogeneous system $A\mathbf{x} = \mathbf{0}$ to have a unique solution?
- How are the solutions to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ related to the solutions of the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$?

Application: Modeling an Economy

An economy is a very complex system. An economy is not a well-defined object, there are many factors that influence an economy, and it is often unclear how the factors influence each other. Mathematical modeling plays an important role in attempting to understand an economy.

In 1941 Wassily Leontief developed the first empirical model of a national economy. Around 1949 Leontief used data from the U.S. Bureau of Labor Statistics to divide the U.S. economy into

500 sectors. He then set up linear equations for each sector. This system was too large for the computers at the time to solve, so he then aggregated the information into 42 sectors. The Harvard Mark II computer was used to solve this system, one of the first significant uses of computers for mathematical modeling. Leontief won the 1973 Nobel Prize in economics for his work.

With such large models (Leontief's models are called *input-output* models) it is important to find a shorthand way to represent the resulting systems. In this section we will see how to represent any size system of linear equations in a very convenient way. Later, we will analyze a small economy using input-output models.

Introduction

There is another useful way to represent a system of linear equations using a matrix-vector product that we investigate in this section. To understand how this product comes about, recall that we can represent the linear system

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

as a vector equation as

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}. \tag{5.1}$$

We can view the left hand side of Equation (5.1) as a **matrix-vector product**. Specifically, if $A = \begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$, then we define the *matrix-vector product* $A\mathbf{x}$ as the left hand side Equation (5.1). So the matrix vector product $A\mathbf{x}$ is the linear combination of the columns of A with weights from the vector $\mathbf{x}$ in order.

With this definition, the vector equation in (5.1) can be expressed as a matrix-vector equation as

$$\begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

We call this representation the **matrix-vector form** of the system. Note that the matrix A in this expression is the same as the coefficient matrix that appears in the augmented matrix representation of the system.

We can use the above definition of the matrix-vector product as a linear combination with any matrix and any vector, as long as it is meaningful to use the entries in the vector as weights for the columns of the matrix. For example, for $A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, then we can define $A\mathbf{v}$ to be the linear combination of the columns of A with weights 3 and 4:

$$A\mathbf{v} = 3\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 4\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 13 \\ 7 \end{bmatrix}.$$

However, note that if $\mathbf{v}$ had three entries, this definition would not make sense since we do not have three columns in A. In those cases, we say $A\mathbf{v}$ is not defined. We will later see that this definition can be generalized to matrix-matrix products, by treating the vector as a special case of a matrix with one column.

Preview Activity 5.1.

(1) Write the vector equation

$$x_1\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_3\begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 11 \end{bmatrix}$$

in matrix-vector form. Note that this is the vector equation whose augmented matrix representation was given in Problem 2 in Preview Activity 2.1. Compare your matrix A and the right hand side vector to the augmented matrix. Do not solve the system.

(2) Given the matrix-vector equation

$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 2 \\ 1 & -2 & -3 \end{bmatrix}\mathbf{x} = \begin{bmatrix} -3 \\ 3 \\ -7 \end{bmatrix}$$

represent the system corresponding to this equation. Note that this should correspond to the system (or an equivalent system where an equation might be multiplied by (-1)) in Problem 1 of Preview Activity 2.1.

(3) Find the indicated matrix-vector products, if possible. Express as one vector.

(a) $\begin{bmatrix} 2 & -2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

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

(c) $\begin{bmatrix} -6 & -2 & 1 \\ 2 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix}$

(4) As you might have noticed, systems with all the constants being 0 are special in that they always have a solution. (Why?) So we might consider grouping systems into two types: Those of the form $A\mathbf{x} = \mathbf{b}$, where not all of the entries of the vector $\mathbf{b}$ are 0, and those of the form $A\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the vector of all zeros. Systems like $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ contains

at least one non-zero entry, are called *nonhomogeneous* systems, and systems of the form $A\mathbf{x} = \mathbf{0}$ are called *homogeneous* systems. For every nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ there is a corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$, and there is a useful connection between the solutions to the nonhomogeneous system and the corresponding homogeneous system. For example, consider the nonhomogeneous system

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

with

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \text{and } \mathbf{b} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}. \tag{5.2}$$

The augmented matrix representation of this system is $[A \mid \mathbf{b}]$. If we reduce this augmented matrix, we find

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

From this RREF, we immediately see that the general solution is that x_3 is free, $x_2 = x_3 - 2$, and $x_1 = 2 - 3x_3$. In vector form, we can represent this general solution as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 - 3x_3 \\ x_3 - 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}. \tag{5.3}$$

The rightmost expression above is called the *parametric vector form* of the solution.

If we had a system where the general solution involved more than one free variable, then we would write the parametric vector form to include one vector multiplying each free variable. For example, if the general solution of a system were that x_2 and x_3 are free and $x_1 = 2 + x_2 + 3x_3$, then the parametric vector form would be

$$\mathbf{x} = \begin{bmatrix} 2 + x_2 + 3x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

Note that the parametric vector form expresses the solutions as a linear combination of a number of vectors, depending on the number of free variables, with an added constant vector. This expression helps us in interpreting the solution set geometrically, as we will see in this section.

(a) Find the general solution to the homogeneous system

$$A\mathbf{x} = \mathbf{0}$$

with A and $\mathbf{x}$ as in (5.2) and compare it to the solution to the nonhomogeneous system in (5.3). What do you notice?

(b) Find the general solution to the nonhomogeneous system

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

with

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \text{and} \ \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

and express it in parametric vector form. Then find the general solution to the corresponding homogeneous system and express it in parametric vector form. How are the two solution sets related?

(c) Make a conjecture about the relationship between the solutions to a consistent non-homogeneous system $A\mathbf{x} = \mathbf{b}$ and the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. Be as specific as possible.

The Matrix-Vector Product

The matrix-vector product we defined in Preview Activity 5.1 for a specific example generalizes in a very straightforward manner, and provides a convenient way to represent a system of linear equations of any size using matrices and vectors. In addition to providing us with an algebraic approach to solving systems via matrices and vectors – leading to a powerful geometric relationship between solution sets of homogeneous and non-homogeneous systems – this representation allows us to think of a linear system from a dynamic perspective, as we will see later in the section on matrix transformations.

The matrix-vector product $A\mathbf{x}$ is a linear combination of the columns of A with weights from $\mathbf{x}$. To define this product in general, we will need a little notation. Recall that a matrix is made of rows and columns – the entries reading from left to right form the *rows* of the matrix and the entries reading from top to bottom form the *columns*. For example, the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}.$$

has three rows and four columns. The number of rows and columns of a matrix is called the *size* of the matrix, so A is a 3 by 4 matrix (also written as 3×4). We often need to have a way to reference the individual entries of a matrix A, and to do so we typically give a label, say a_{ij} to the entry in the ith row and jth column of A. So in our example we have $a_{23} = 7$. We also write $A = [a_{ij}]$ to indicate a matrix whose i, jth entry is a_{ij}. At times it is convenient to write a matrix in terms of its rows or columns. If $A = [a_{ij}]$ is an $m \times n$ matrix, then we will write

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

or, if we let $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ denote the rows of the matrix A, then we can write A as[1]

$$A = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \vdots \\ \mathbf{r}_m \end{bmatrix}.$$

We can also write A in terms of its columns, $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$, as

$$A = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n].$$

In general, the product of a matrix with a vector is defined as follows.

Definition 5.1. Let A be an $m \times n$ matrix with columns $\mathbf{c}_1$, $\mathbf{c}_2$, ..., $\mathbf{c}_n$, and let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be a vector in $\mathbb{R}^n$. The **matrix-vector product** $A\mathbf{x}$ is

$$A\mathbf{x} = x_1\mathbf{c}_2 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n.$$

Important Note: The matrix-vector product $A\mathbf{x}$ is defined only when the number of entries of the vector $\mathbf{x}$ is equal to the number of columns of the matrix A. That is, if A is an $m \times n$ matrix, then $A\mathbf{x}$ is defined only if $\mathbf{x}$ is a column vector with n entries.

The Matrix-Vector Form of a Linear System

As we saw in Preview Activity 5.1, the matrix-vector product provides us with a short hand way of representing a system of linear equations. In general, every linear system can be written in matrix-vector form as follows.

The linear system

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

of m equations in n unknowns can be written in matrix-vector form as $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

[1] Technically, the rows of A are made from the entries of the row vectors, but we use this notation as a shorthand.

This general system can also be written in the vector form

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

With this last representation, we now have four different ways to represent a system of linear equations (as a system of linear equations, as an augmented matrix, in vector equation form, and in matrix-vector equation form), and it is important to be able to translate between them. As an example, the system

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

from the introduction to this section has corresponding augmented matrix

$$\left[\begin{array}{rrrr|r} 1 & 4 & 2 & 4 & 1 \\ 2 & -1 & -5 & -1 & 2 \\ 3 & 7 & 1 & 7 & 3 \end{array}\right],$$

is expressed in vector form as

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ -5 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} 4 \\ -1 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

and has matrix-vector form

$$\begin{bmatrix} 1 & 4 & 2 & 4 \\ 2 & -1 & -5 & -1 \\ 3 & 7 & 1 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}.$$

Activity 5.1. In this activity, we will use the equivalence of the different representations of a system to make useful observations about when a system represented as $A\mathbf{x} = \mathbf{b}$ has a solution.

(a) Consider the system

$$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

Write the matrix-vector product on the left side of this equation as a linear combination of the columns of the coefficient matrix. Find weights that make the vector $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ a linear combination of the columns of the coefficient matrix.

(b) From this point on we consider the general case where A is an $m \times n$ matrix. Use the vector equation representation to explain why the system $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A. (Note that 'if and only if' is an expression to mean that if one side of the expression is true, then the other side must also be true.) (Hint: Compare to what you did in part (a).)

(c) Use part (b) and the definition of span to explain why the system $A\mathbf{x} = \mathbf{b}$ has a solution if and only if the vector $\mathbf{b}$ is in the span of the columns of A.

(d) Use part (c) to explain why the system $A\mathbf{x} = \mathbf{b}$ always has a solution for any vector $\mathbf{b}$ in $\mathbb{R}^m$ if and only if the span of the columns of A is all of $\mathbb{R}^m$.

(e) Use the augmented matrix representation and the criterion for a consistent system to explain why the system $A\mathbf{x} = \mathbf{b}$ is consistent for all vectors $\mathbf{b}$ if and only if A has a pivot position in every row.

We summarize our observations from the above activity in the following theorem.

Theorem 5.2. *Let A be an $m \times n$ matrix. The following statements are equivalent:*

(1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution for every vector $\mathbf{b}$ in $\mathbb{R}^m$.

(2) Every vector $\mathbf{b}$ in $\mathbb{R}^m$ can be written as a linear combination of the columns of A.

(3) The span of the columns of A is $\mathbb{R}^m$.

(4) The matrix A has a pivot position in each row.

In the future, if we need to determine whether a system has a solution for every $\mathbf{b}$, we can refer to this theorem without having to argue our reasoning from scratch.

Properties of the Matrix Vector Product

As we have done before, we have a new operation (the matrix-vector product), so we should wonder what properties it has.

Activity 5.2. In this activity, we consider whether the matrix-vector product distributes vector addition. In other words: Is $A(\mathbf{u} + \mathbf{v})$ equal to $A\mathbf{u} + A\mathbf{v}$?

We work with arbitrary vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$ and an arbitrary matrix A with 3 columns (so that $A\mathbf{u}$ and $A\mathbf{v}$ are defined) to simplify notation. Let $A = [\mathbf{c}_1 \ \mathbf{c}_2 \ \mathbf{c}_3]$ (note that each $\mathbf{c}_i$ represents a column of A), $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$. Use the definition of the matrix-vector product along with the properties of vector operations to show that

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}.$$

Similar arguments using the definition of matrix-vector product along with the properties of vector operations can be used to show the following theorem:

Theorem 5.3. *Let A be an $m \times n$ matrix, $\mathbf{u}$ and $\mathbf{v}$ $n \times 1$ vectors, and c a scalar. Then*

(1) $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$

(2) $c(A\mathbf{v}) = A(c\mathbf{v})$

Homogeneous and Nonhomogeneous Systems

As we saw before, the systems with all the right hand side constants being 0 are special in that they always have a solution. (Why?) So we might consider grouping systems into two types: Those of the form $A\mathbf{x} = \mathbf{b}$, where not all of the entries of the vector $\mathbf{b}$ are 0, and those of the form $A\mathbf{x} = \mathbf{0}$, where $\mathbf{0}$ is the vector of all zeros. Systems like $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ contains at least one non-zero entry, are called **nonhomogeneous** systems, and systems of the form $A\mathbf{x} = \mathbf{0}$ are called **homogeneous** systems. For every nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ there is a corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$. We now investigate the connection between the solutions to the nonhomogeneous system and the corresponding homogeneous system.

Activity 5.3. In this activity we will consider the relationship between the solution sets of nonhomogeneous systems and those of the corresponding homogeneous systems.

(a) Find the solution sets of the system
$$A\mathbf{x} = \mathbf{b}$$
where
$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \text{and} \ \mathbf{b} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$
and the corresponding homogeneous system (i.e. where we replace $\mathbf{b}$ with $\mathbf{0}$.)

(b) Find the solution sets of the system
$$A\mathbf{x} = \mathbf{b}$$
where
$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \text{and} \ \mathbf{b} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$
and the corresponding homogeneous system.

(c) What are the similarities/differences between solutions of the nonhomogeneous system and its homogeneous counterpart?

As we saw in the above activity, there is a relationship between solutions of a nonhomogeneous and the corresponding homogeneous system. Let us formalize this relationship. If the general solution of a system involves free variables, we can represent the solutions in **parametric vector form** to have a better idea about the geometric representation of the solution set. Suppose the solution is that x_3 is free, $x_2 = -2 + x_3$, and $x_1 = 2 - 3x_3$. In vector form, we can represent this general solution as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 - 3x_3 \\ x_3 - 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}. \tag{5.4}$$

From this representation, we see that the solution set is a line through the origin (formed by multiples of $\begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$) shifted by the added vector $\begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$. The solution to the homogeneous system on the other does not have the shift.

Algebraically, we see that every solution to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ can be written in the form $\mathbf{p} + \mathbf{v}_h$, where $\mathbf{p}$ is a particular solution to $A\mathbf{x} = \mathbf{b}$ and $\mathbf{v}_h$ is a solution to the corresponding homogeneous system $A\mathbf{x} = \mathbf{0}$.

To understand why this always happens, we will verify the result algebraically for an arbitrary A and $\mathbf{b}$. Assuming that $\mathbf{p}$ is a particular solution to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$, we need to show that:

- if $\mathbf{v}$ is an arbitrary solution to the nonhomogeneous system, then $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$, where $\mathbf{v}_h$ is some solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, and
- if $\mathbf{v}_h$ is an arbitrary solution to the homogeneous system, then $\mathbf{p} + \mathbf{v}_h$ is a solution to the nonhomogeneous system.

To verify the first condition, suppose that $\mathbf{v}$ is a solution to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$. Since we want $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$, we need to verify that $\mathbf{v} - \mathbf{p}$ is a solution for the homogeneous system so that we can assign $v_h = \mathbf{v} - \mathbf{p}$. Note that

$$A(\mathbf{v} - \mathbf{p}) = A\mathbf{v} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}\,,$$

using the distributive property of matrix-vector product over vector addition. Hence $\mathbf{v}$ is of the form $\mathbf{p} + \mathbf{v}_h$ with $\mathbf{v}_h = \mathbf{0}$.

To verify the second condition, consider a vector of the form $\mathbf{p}+\mathbf{v}_h$, where $\mathbf{v}_h$ is a homogeneous solution. We have

$$A(\mathbf{p} + \mathbf{v}_h) = A\mathbf{p} + A\mathbf{v}_h = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

and so $\mathbf{p} + \mathbf{v}_h$ is a solution to $A\mathbf{x} = \mathbf{b}$.

Our work above proves the following theorem.

Theorem 5.4. *Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some $\mathbf{b}$ and $\mathbf{p}$ is a solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ consists of all vectors of the form $\mathbf{v} = \mathbf{p} + \mathbf{v}_h$ where $\mathbf{v}_h$ is a solution to $A\mathbf{x} = \mathbf{0}$.*

The Geometry of Solutions to the Homogeneous System

There is a simple geometric interpretation to the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ based on the number of free variables that imposes a geometry on the solution set of the corresponding nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ (when consistent) due to Theorem 5.4.

Activity 5.4. In this activity we consider geometric interpretations of the solution sets of homogeneous and nonhomogeneous systems.

(a) Consider the system $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \\ -1 & 3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 2 \\ -6 \\ -2 \end{bmatrix}$. The general solution to this system has the form $\begin{bmatrix} 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, where x_2 is any real number.

i. Let $\mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$. What does the set of all vectors of the form $x_2\mathbf{v}$ look like geometrically? Draw a picture in $\mathbb{R}^2$ to illustrate. (Recall that we refer to all the vectors of the form $x_2\mathbf{v}$ simply as Span$\{\mathbf{v}\}$.)

ii. Let $\mathbf{p} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. What effect does adding the vector $\mathbf{p}$ to each vector in Span$\{\mathbf{v}\}$ have on the geometry of Span$\{\mathbf{v}\}$? Finally, what does this mean about the geometry of the solution set to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$?

(b) Consider the system $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & -3 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$. The general solution to this system has the form $\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, where x_2, x_3 are any real numbers.

i. Let $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$. Use our results from Section 4 to determine the geometric shape of Span$\{\mathbf{u}, \mathbf{v}\}$, the set of all vectors of the form $x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, where x_2, x_3 are any real numbers.

ii. Let $\mathbf{p} = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix}$. What's the geometric effect of adding the vector $\mathbf{p}$ to each vector in Span$\{\mathbf{u}, \mathbf{v}\}$? Finally, what does this mean about the geometry of the solution set to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$?

Our work in the above activity shows the geometric shape of the solution set of a consistent nonhomogeneous system is the same as the geometric shape of the solution set of the corresponding homogeneous system. The only difference between the two solution sets is that one is a shifted version of the other.

Examples

What follows are worked examples that use the concepts from this section.

Example 5.5. We now have several different ways to represent a system of linear equations. Rewrite the system in an equivalent form

$$\begin{aligned} 11x_1 + 4x_2 - 5x_3 - 2x_4 &= 63 \\ 15x_1 + 5x_2 + 2x_3 - 2x_4 &= 68 \\ 6x_1 + 2x_2 + \ \ x_3 - \ \ x_4 &= 26 \\ 9x_1 + 3x_2 + 2x_3 - \ \ x_4 &= 40. \end{aligned}$$

(a) as an augmented matrix

(b) as an equation involving a linear combination of vectors

(c) using a matrix-vector product

Then solve the system.

Example Solution.

(a) The augmented matrix for this system is

$$\left[\begin{array}{cccr|c} 11 & 4 & -5 & -2 & 63 \\ 15 & 5 & 2 & -2 & 68 \\ 6 & 2 & 1 & -1 & 26 \\ 9 & 3 & 2 & -1 & 40 \end{array}\right].$$

(b) If we make vectors from the columns of the augmented matrix, we can write this system in vector form as

$$x_1 \begin{bmatrix} 11 \\ 15 \\ 6 \\ 9 \end{bmatrix} + x_2 \begin{bmatrix} 4 \\ 5 \\ 2 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 2 \\ 1 \\ 2 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -2 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 63 \\ 68 \\ 26 \\ 40 \end{bmatrix}.$$

(c) The coefficient matrix for this system is $\begin{bmatrix} 11 & 4 & -5 & -2 \\ 15 & 5 & 2 & -2 \\ 6 & 2 & 1 & -1 \\ 9 & 3 & 2 & -1 \end{bmatrix}$, and the matrix-vector form of the system is

$$\begin{bmatrix} 11 & 4 & -5 & -2 \\ 15 & 5 & 2 & -2 \\ 6 & 2 & 1 & -1 \\ 9 & 3 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 63 \\ 68 \\ 26 \\ 40 \end{bmatrix}.$$

Using technology, we find that the reduced row echelon form of the augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 7 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right].$$

So the solution to this system is $x_1 = 3$, $x_2 = 7$, $x_3 = -2$, and $x_4 = 4$.

Example 5.6. Consider the homogeneous system

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

(a) Find the general solution to this homogeneous system and express the system in parametric vector form.

(b) Let $A = \left[\begin{array}{rrr} 1 & 8 & -1 \\ 1 & -7 & 2 \\ 3 & 4 & 1 \end{array}\right]$, and let $\mathbf{b} = \left[\begin{array}{r} -6 \\ 9 \\ 2 \end{array}\right]$. Show that $\left[\begin{array}{r} -1 \\ 0 \\ 5 \end{array}\right]$ is a solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$.

(c) Use the results from part (a) and (b) to write the parametric vector form of the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$. (Do this without directly solving the system $A\mathbf{x} = \mathbf{b}$.)

(d) Describe what the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ and the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ look like geometrically.

Example Solution.

(a) The augmented matrix of the homogeneous system is

$$\left[\begin{array}{rrr|r} 1 & 8 & -1 & 0 \\ 1 & -7 & 2 & 0 \\ 3 & 4 & 1 & 0 \end{array}\right],$$

and the reduced row echelon form of this augmented matrix is

$$\left[\begin{array}{ccr|c} 1 & 0 & \frac{3}{5} & 0 \\ & 1 & -\frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Since there is no corresponding equation of the form $0 = b$ for a nonzero constant b, this system is consistent. The third column contains no pivot, so the variable x_3 is free, $x_2 =$

$\frac{1}{5}x_3$ and $x_1 = -\frac{3}{5}x_3$. In parametric vector form the general solution to the homogeneous system is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\frac{3}{5}x_3 \\ \frac{1}{5}x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix}.$$

(b) Since

$$A\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix} = (-1)\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} + (0)\begin{bmatrix} 8 \\ -7 \\ 4 \end{bmatrix} + (5)\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} -1-5 \\ -1+10 \\ -3+5 \end{bmatrix} = \begin{bmatrix} -6 \\ 9 \\ 2 \end{bmatrix},$$

we conclude that $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$ is a solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$.

(c) We know that every solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ has the form of the general solution to the homogeneous system plus a particular solution to the non-homogeneous system. Combining the results of (a) and (b) we see that the general solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix} + x_3 \begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix},$$

where x_3 can be any real number.

(d) The solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$ is the span of the vector $\begin{bmatrix} -\frac{3}{5} \\ \frac{1}{5} \\ 1 \end{bmatrix}$. Geometrically, this set of points is a line through the origin and the point $(-3, 1, 5)$ in $\mathbb{R}^3$. The solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is the translation of the line through the origin and $(-3, 1, 5)$ by the vector $\begin{bmatrix} -1 \\ 0 \\ 5 \end{bmatrix}$. In other words, the solution to the non-homogeneous system $A\mathbf{x} = \mathbf{b}$ is the line in $\mathbb{R}^3$ through the points $(-1, 0, 5)$ and $(-4, 1, 10)$.

Summary

- If $A = [\mathbf{c}_1\ \mathbf{c}_2\ \cdots\ \mathbf{c}_n]$ is an $m \times n$ matrix with columns $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$, and if $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ is a vector in $\mathbb{R}^n$, then the matrix-vector product $A\mathbf{x}$ is defined to be the linear combination of the columns of A with corresponding weights from $\mathbf{x}$ – that is
$$A\mathbf{x} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n.$$

- A linear system
$$\begin{aligned} a_{11}x_1 + \ a_{12}x_2 + \cdots + \ a_{1n}x_n &= \ b_1 \\ a_{21}x_1 + \ a_{22}x_2 + \cdots + \ a_{2n}x_n &= \ b_2 \\ &\vdots \\ a_{m1}x_1 + a_{2m}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$
can be written in matrix form as
$$A\mathbf{x} = \mathbf{b},$$
where
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

- The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A.

- The system $A\mathbf{x} = \mathbf{b}$ is consistent for every vector $\mathbf{b}$ if every row of A contains a pivot.

- A homogeneous system is a system of the form $A\mathbf{x} = \mathbf{0}$ for some $m \times n$ matrix A. Since the zero vector in $\mathbb{R}^n$ satisfies $A\mathbf{x} = \mathbf{0}$, a homogeneous system is always consistent.

- A homogeneous system can have one or infinitely many different solutions. The homogeneous system $A\mathbf{x} = \mathbf{0}$ has exactly one solution if and only if each column of A is a pivot column.

- The solutions to the consistent nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ have the form $\mathbf{p} + \mathbf{v}_h$, where $\mathbf{p}$ is a particular solution to the nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ and $\mathbf{v}_h$ is a solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$. In other words, the solution space to a consistent nonhomogeneous system $A\mathbf{x} = \mathbf{b}$ is a translation of the solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$ by a particular solution to the nonhomogeneous system.

Finally, we argued an important theorem.

Theorem 5.7. *Let A be an $m \times n$ matrix. The following statements are equivalent.*

(1) *The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution for every vector $\mathbf{b}$ in $\mathbb{R}^m$.*

(2) *Every vector $\mathbf{b}$ in $\mathbb{R}^m$ can be written as a linear combination of the columns of A.*

(3) *The span of the columns of A is $\mathbb{R}^m$.*

(4) *The matrix A has a pivot position in each row.*

We will continue to add to this theorem, so it is a good idea for you to begin now to remember the equivalent conditions of this theorem.

Exercises

(1) Write the system

$$\begin{aligned} x_1 + 2x_2 + 2x_3 + x_4 &= -1 \\ 4x_1 - 8x_2 + 3x_3 - 9x_4 &= 2 \\ x_1 + 6x_2 - 4x_3 + 12x_4 &= -1 \end{aligned}$$

in matrix-vector form. Explicitly identify the coefficient matrix and the vector of constants.

(2) Write the linear combination

$$x_1 \begin{bmatrix} 1 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 10 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

as a matrix-vector product.

(3) Represent the following matrix-vector equation as a linear system and find its solution.

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$$

(4) Represent the following matrix-vector equation as a linear system and find its solution.

$$\begin{bmatrix} 1 & -2 & -1 \\ 2 & 2 & -2 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -4 \\ 8 \end{bmatrix}$$

(5) Another way of defining the matrix-vector product uses the concept of the *scalar product* of vectors.[2] Given a $1 \times n$ matrix $\mathbf{u} = [u_1\ u_2\ \dots\ u_n]$[3] and an $n \times 1$ vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, we define the scalar product $\mathbf{u} \cdot \mathbf{v}$ as

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 + \cdots + u_nv_n.$$

We then define the matrix-vector product $A\mathbf{x}$ as the vector whose entries are the scalar products of the rows of A with $\mathbf{x}$. As an example, if $A = \begin{bmatrix} 2 & 3 & 4 \\ 1 & -2 & 3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then

$$A\mathbf{x} = \begin{bmatrix} 2x_1 + 3x_2 + 4x_3 \\ x_1 + (-2)x_2 + 3x_3 \end{bmatrix}.$$

Calculate the matrix-vector product $A\mathbf{x}$ where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ using both methods of finding the matrix-vector product to show that the two definitions are equivalent for size 2×2 matrices.

(6) Find the value of a such that

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

where $*$'s represent unknown values.

(7) Suppose we have

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ -1 & 2 & 3 & 1 \\ 2 & 3 & 1 & a \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

where b_i's represent unknown values.

(a) In order to find the value of a, which of the b_i's do we need to know? Why?

(b) Suppose the b_i(s) that we need to know is(are) equal to 9. What is the value of a?

(8) Suppose we are given

$$A\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad A\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

for an unknown A and two unknown vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$. Using matrix-vector product properties, evaluate $A\mathbf{w}$ where $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$.

[2] Note that some authors refer to the scalar product as the *dot product*.

[3] We can identify a $1 \times n$ matrix $\mathbf{u} = [u_1\ u_2\ \dots\ u_n]$ with the $n \times 1$ vector $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$, so we ofter refer to $[u_1\ u_2\ \dots\ u_n]$ as a vector.

(9) Suppose we are given

$$A\begin{bmatrix}1\\2\\1\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix} \text{ and } A\begin{bmatrix}1\\0\\2\end{bmatrix}=\begin{bmatrix}0\\2\end{bmatrix}.$$

After expressing $\begin{bmatrix}-1\\6\\-5\end{bmatrix}$ as a linear combination of $\begin{bmatrix}1\\2\\1\end{bmatrix}$ and $\begin{bmatrix}1\\0\\2\end{bmatrix}$, use the matrix-vector product properties to determine $A\begin{bmatrix}-1\\6\\-5\end{bmatrix}$.

(10) (a) The non-homogeneous system (with unknown constants a and b)

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

has a solution which lies on the x-axis (i.e. $y=z=0$). Find this solution.

(b) If the corresponding homogeneous system

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

has its general solution expressed in parametric vector form as $z\cdot\begin{bmatrix}0\\1\\1\end{bmatrix}$, find the general solution for the non-homogeneous system using your answer to part (a).

(c) Find the conditions on a and b that make the system from (a) have the general solution you found in (b).

(11) Find the general solution to the non-homogeneous system

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

Using the parametric vector form of the solutions, determine what the solution set to this non-homogeneous system looks like geometrically. Be as specific as possible. (Include information such as whether the solution set is a point, a line, or a plane, etc.; whether the solution set passes through the origin or is shifted from the origin in a specific direction by a specific number of units; and how the solution is related to the corresponding homogeneous system.)

(12) Come up with an example of a 3×3 matrix A for which the solution set of $A\mathbf{x}=\mathbf{0}$ is a line, and a 3×3 matrix A for which the solution set of $A\mathbf{x}=\mathbf{0}$ is a plane.

(13) Suppose we have three vectors $\mathbf{v}_1,\mathbf{v}_2$ and $\mathbf{v}_3$ satisfying $\mathbf{v}_3=2\mathbf{v}_1-\mathbf{v}_2$. Let A be the matrix with vectors $\mathbf{v}_1,\mathbf{v}_2$ and $\mathbf{v}_3$ as the columns in that order. Find a non-zero $\mathbf{x}$ such that $A\mathbf{x}=\mathbf{0}$ using this information.

(14) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** If the system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, then so does the system $A\mathbf{x} = \mathbf{b}$ for **any** right-hand-side $\mathbf{b}$.

(b) **True/False** If $\mathbf{x}_1$ is a solution for $A\mathbf{x} = \mathbf{b}_1$ and $\mathbf{x}_2$ is a solution for $A\mathbf{x} = \mathbf{b}_2$, then $\mathbf{x}_1 + \mathbf{x}_2$ is a solution for $A\mathbf{x} = \mathbf{b}_1 + \mathbf{b}_2$.

(c) **True/False** If an $m \times n$ matrix A has a pivot in every row, then the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.

(d) **True/False** If an $m \times n$ matrix A has a pivot in every row, then the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$.

(e) **True/False** If A and B are row equivalent matrices and the columns of A span $\mathbb{R}^m$, then so do the columns of B.

(f) **True/False** All homogeneous systems have either a unique solution or infinitely many solutions.

(g) **True/False** If a linear system is not homogeneous, then the solution set does not include the origin.

(h) **True/False** If a solution set of a linear system does not include the origin, the system is not homogeneous.

(i) **True/False** If the system $A\mathbf{x} = \mathbf{b}$ has a unique solution for some $\mathbf{b}$, then the homogeneous system has only the trivial solution.

(j) **True/False** If A is a 3×4 matrix, then the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has non-trivial solutions.

(k) **True/False** If A is a 3×2 matrix, then the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has non-trivial solutions.

Project: Input-Output Models

There are two basic types of input-output models: closed and open. The closed model assumes that all goods produced are consumed within the economy – no trading takes place with outside entities. In the open model, goods produced within the economy can be traded outside the economy.

To work with a closed model, we use an example (from *Input-Output Economics* by Wassily Leontief). Assume a simple three-sector economy consisting of agriculture (growing wheat), manufacturing (producing cloth), and households (supplying labor). Each sector of the economy relies on goods from the other sectors to operate (e.g., people must eat to work and need to be clothed). To model the interactions between the sectors, we consider how many units of product is needed as input from one sector to another to produce one unit of product in the second sector. For example, assume the following:

- to produce one unit (say dollars worth) of agricultural goods requires 25% of a unit of agricultural output, 28% of a unit of manufacturing output, and 27% of a unit of household output;

- to produce one unit of manufactured goods requires 20% of a unit of agricultural output, 60% of a unit of manufacturing output, and 60% of a unit of household output;
- to produce one unit of household goods requires 55% of a unit of agricultural output, 12% of a unit of manufacturing output, and 13% of a unit of household output.

These assumptions are summarized in Table 5.1.

into\from	Agriculture	Manufacture	Households
Agriculture	0.25	0.28	0.27
Manufacture	0.20	0.60	0.60
Households	0.55	0.12	0.13

Table 5.1: Summary of simple three sector economy.

This model is said to be *closed* because all good produced are used up within the economy. If there are goods that are not used within the economy the model is said to be *open*. Open models will be examined later.

The economist's goal is to determine what level of production in each section meets the following requirements:

- the production from each sector meets the needs of all of the sectors and
- there is no overproduction.

Project Activity 5.1. We can use techniques from linear algebra to determine the levels of production that precisely meet the two goals of the economist.

(a) Suppose that the agricultural output is x_1 units, the manufacturing output is x_2 units, and the household output is x_3 units. We represent this data as a *production vector* $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. To produce a unit of agriculture requires 0.25 units of agriculture, 0.28 units of manufacturing, and 0.27 units of household. If x_1 units of agriculture, x_2 units of manufacturing, and x_3 units of household products are are produced, then agriculture can produce

$$0.25x_1 + 0.28x_2 + 0.27x_3$$

units. In order to meet the needs of agriculture and for there to be no overproduction, we must then have

$$0.25x_1 + 0.28x_2 + 0.27x_3 = x_1.$$

Write similar equations for the manufacturing and household sectors of the economy.

(b) Find the augmented matrix for the system of linear equations that represent production of the three sectors from part (a), and then solve the system to find the production levels that meet the economist's two goals.

(c) Suppose the production level of the household sector is 200 million units (dollars). Find the production levels of the agricultural and manufacturing sectors that meet the economist's two goals.

In general, a matrix derived from a table like Table 5.1 is called a *consumption* matrix, which we will denote as C. (In the example discussed here $C = \begin{bmatrix} 0.25 & 0.28 & 0.27 \\ 0.20 & 0.60 & 0.60 \\ 0.55 & 0.12 & 0.13 \end{bmatrix}$.) A consumption matrix $C = [c_{ij}]$, where c_{ij} represents the proportion of the output of sector j that is consumed by sector i, satisfies two important properties.

- Since no sector can consume a negative amount or an amount that exceeds the output of another sector, we must have $0 \leq c_{ij} \leq 1$ for all i and j.
- If there are n sectors in the economy, the fact that all output is consumed within the economy implies that $c_{1j} + c_{2j} + \cdots + c_{nj} = 1$. In other words, the column sums of C are all 1.

In our example, if we let $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then we can write the equations that guarantee that the production levels satisfy the two economists' goal in matrix form as

$$\mathbf{x} = C\mathbf{x}. \tag{5.5}$$

Now we can rephrase the question to be answered as which production vectors $\mathbf{x}$ satisfy equation (5.5). When $C\mathbf{x} = \mathbf{x}$, then the system is in equilibrium, that is output exactly meets needs. Any solution $\mathbf{x}$ that satisfies (5.5) is called a *steady state* solution.

Project Activity 5.2. Is there a steady state solution for the closed system of Agriculture, Manufacturing, and Households? If so, find the general steady state solution. If no, explain why.

So far, we considered the case where the economic system was *closed*. This means that the industries that were part of the system sold products only to each other. However, if we want to represent the demand from other countries, from households, capital building, etc., we need an *open model*. In an article in the *Scientific American* Leontief organized the 1958 American economy into 81 sectors. The production of each of these sectors relied on production from the all of the sectors. Here we present a small sample from Leontief's 81 sectors, using Petroleum, Textiles, Transportation, and Chemicals as our sectors of the economy. Leontief's model assumed that the production of 1 unit of output of

- petroleum requires 0.1 unit of petroleum, 0.2 units of transportation, and 0.4 units of chemicals;
- textiles requires 0.4 units of petroleum, 0.1 unit of textiles, 0.15 units of transportation, and 0.3 units of chemicals;
- transportation requires 0.6 units of petroleum, 0.1 unit of transportation, and 0.25 units of chemicals;

- chemicals requires 0.2 units of petroleum, 0.1 unit of textiles, 0.3 units of transportation, and 0.2 units of chemicals.

A summary of this information is in Table 5.2. Assume the units are measured in dollars.

into\from	Petroleum	Textiles	Transportation	Chemicals
Petroleum	0.10	0.00	0.20	0.40
Textiles	0.40	0.10	0.15	0.30
Transportation	0.60	0.00	0.10	0.25
Chemicals	0.20	0.10	0.30	0.20

Table 5.2: Summary of four sector economy.

In the open model, there is another part of the economy, called the *open sector*, that does not produce goods or services but only consumes them. If this sector (think end consumers, for example) demands/consumes d_1 units of Petroleum, d_2 units of Textiles, d_3 units of Transportation, and d_4 units of Chemicals, we put this into a *final demand vector* $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$.

An economist would want to find the production level where the demand from the good/service producing sectors of the economy plus the final demand from the open sector exactly matches the output in each of the sectors. Let x_1 represent the number of units of petroleum output, x_2 the number of units of textiles output, x_3 the number of units of transportation output, and x_4 the number of units of chemical output during any time period. Then the production vector is $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$. So an economist wants to find the production vectors $\mathbf{x}$ such that

$$\begin{aligned}
0.10x_1 \qquad\qquad\; &+ 0.20x_3 + 0.40x_4 + d_1 = x_1 \\
0.40x_1 + 0.10x_2 &+ 0.15x_3 + 0.30x_4 + d_2 = x_2 \\
0.60x_1 \qquad\qquad\; &+ 0.10x_3 + 0.25x_4 + d_3 = x_3 \\
0.20x_1 + 0.10x_2 &+ 0.30x_3 + 0.20x_4 + d_4 = x_4,
\end{aligned}$$

where $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix}$ is the demand vector from the open market. The matrix

$$E = \begin{bmatrix} 0.10 & 0.00 & 0.20 & 0.40 \\ 0.40 & 0.10 & 0.15 & 0.30 \\ 0.60 & 0.00 & 0.10 & 0.25 \\ 0.20 & 0.10 & 0.30 & 0.20 \end{bmatrix}$$

derived from Table 5.2, is called the *exchange* matrix.

Project Activity 5.3.

(a) Suppose the final demand vector in our four sector economy is $\begin{bmatrix} 500 \\ 200 \\ 400 \\ 100 \end{bmatrix}$. Find the production levels that satisfy our system.

(b) Does this economy defined by the exchange matrix E have production levels that exactly meet internal and external demands regardless of the external demands? That is, does the system of equations

$$\begin{aligned}
0.10x_1 + 0.20x_3 + 0.40x_4 + d_1 &= x_1 \\
0.40x_1 + 0.10x_2 + 0.15x_3 + 0.30x_4 + d_2 &= x_2 \\
0.60x_1 + 0.10x_3 + 0.25x_4 + d_3 &= x_3 \\
0.20x_1 + 0.10x_2 + 0.30x_3 + 0.20x_4 + d_4 &= x_4
\end{aligned}$$

have a solution regardless of the values of d_1, d_2, d_3, and d_4? Explain.

Section 6

Linear Dependence and Independence

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What are two ways to describe what it means for a set of vectors in $\mathbb{R}^n$ to be linearly independent?
- What are two ways to describe what it means for a set of vectors in $\mathbb{R}^n$ to be linearly dependent?
- If S is a set of vectors, what do we mean by a basis for Span S?
- Given a nonzero set S of vectors, how can we find a linearly independent subset of S that has the same span as S?
- How do we recognize if the columns of a matrix A are linearly independent?
- How can we use a matrix to determine if a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors is linearly independent?
- How can we use a matrix to find a minimal spanning set for a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$?

Application: Bézier Curves

Bézier curves are simple curves that were first developed in 1959 by French mathematician Paul de Casteljau, who was working at the French automaker Citroën. The curves were made public in 1962 by Pierre Bézier who used them in his work designing automobiles at the French car maker Renault. In addition to automobile design, Bézier curves have many other uses. Two of the most common applications of Bézier curves are font design and drawing tools. As an example, the letter

"S" in Palatino font is shown using Bézier curves in Figure 6.1. If you've used Adobe Illustrator, Photoshop, Macromedia Freehand, Fontographer, or any other of a number of drawing programs, then you've used Bézier curves. At the end of this section we will see how Bézier curves can be defined using linearly independent vectors and linear combinations of vectors.

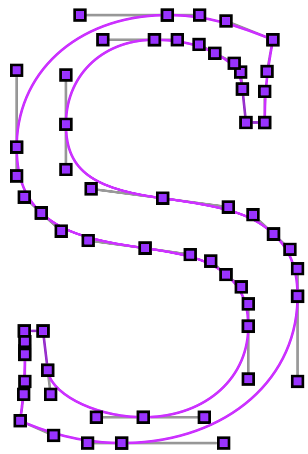


Figure 6.1: A letter S.

Introduction

In Section 4 we saw how to represent water-benzene-acetic acid chemical solutions with vectors, where the components represent the water, benzene and acid percentages. We then considered a problem of determining if a given chemical solution could be made by mixing other chemical solutions. Suppose we now have three different water-benzene-acetic acid chemical solutions, one with 40% water, 50% benzene and 10% acetic acid, the second with 52% water, 42% benzene and 6% acid, and a third with 46% water, 46% benzene and 8% acid. We represent the first chemical solution with the vector $\mathbf{v}_1 = \begin{bmatrix} 40 \\ 50 \\ 10 \end{bmatrix}$, the second with the vector $\mathbf{v}_2 = \begin{bmatrix} 52 \\ 42 \\ 6 \end{bmatrix}$, and the third with the vector $\mathbf{v}_3 = \begin{bmatrix} 46 \\ 46 \\ 8 \end{bmatrix}$. By combining these three chemical solutions we can make a chemical solution with 43% water, 48% benzene and 9% acid as follows

$$\frac{7}{12}\mathbf{v}_1 + \frac{1}{12}\mathbf{v}_2 + \frac{1}{3}\mathbf{v}_3 = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}.$$

However, if we had noticed that the third chemical solution can actually be made from the first two, that is,

$$\frac{1}{2}\mathbf{v}_1 + \frac{1}{2}\mathbf{v}_2 = \mathbf{v}_3,$$

we might have realized that we don't need the third chemical solution to make the 43% water, 48% benzene and 9% acid chemical solution. In fact,

$$\frac{3}{4}\mathbf{v}_1 + \frac{1}{4}\mathbf{v}_2 = \begin{bmatrix} 43 \\ 48 \\ 9 \end{bmatrix}.$$

(See Exercise 5 of Section 4.) Using the third chemical solution (represented by $\mathbf{v}_3$) uses more information than we actually need to make the desired 43% water, 48% benzene and 9% acid chemical solution because the vector $\mathbf{v}_3$ is redundant – all of the material we need to make $\mathbf{v}_3$ is contained in $\mathbf{v}_1$ and $\mathbf{v}_2$. This is the basic idea behind linear independence – representing information in the most efficient way.

Information is often contained in and conveyed through vectors – especially linear combinations of vectors. In this section we will investigate the concepts of linear dependence and independence of a set of vectors. Our goal is to be able to efficiently determine when a given set of vectors forms a *minimal spanning set*. A minimal spanning set is a spanning set that contains the smallest number of vectors to obtain all of the vectors in the span. An important aspect of a minimal spanning set is that every vector in the span can be written in one and only one way as a linear combination of the vectors in the minimal spanning set. This will allow us to define the important notion of the dimension of a vector space.

Review of useful information: Recall that a linear combination of vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$ in $\mathbb{R}^n$ is a sum of scalar multiples of $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$. That is, a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$ is a vector of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k,$$

where c_1, c_2, ..., c_k are scalars.

Recall also that the collection of all linear combinations of a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is called the span of the set of vectors. That is, the span $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of the set $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$ of vectors in $\mathbb{R}^n$ is the set

$$\{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k : \text{ where } c_1, c_2, \ldots, c_k \text{ are scalars}\}.$$

For example, a linear combination of vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$ is $2\mathbf{v}_1 - 3\mathbf{v}_2 = \begin{bmatrix} 2 \\ 8 \\ 1 \end{bmatrix}$. All linear combinations of these two vectors can be expressed as the collection of vectors of the form $\begin{bmatrix} c_1 \\ c_1 - 2c_2 \\ 2c_1 + c_2 \end{bmatrix}$ where c_1, c_2 are scalars. Suppose we want to determine whether $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is in the span, in other words if $\mathbf{w}$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2$. This means we are looking for c_1, c_2 such that

$$\begin{bmatrix} c_1 \\ c_1 - 2c_2 \\ 2c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

we solve for the system represented with the augmented matrix

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

By reducing this matrix, we find that there are no solutions of the system, which implies that $\mathbf{w}$ is not a linear combination of $\mathbf{v}_1, \mathbf{v}_2$. Note that we can use any names we please for the scalars, say x_1, x_2, if we prefer.

Preview Activity 6.1. Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ -6 \end{bmatrix}$, and let $\mathbf{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$. If $\mathbf{b}$ is in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, we are interested in the most efficient way to represent $\mathbf{b}$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

(1) The vector $\mathbf{b}$ is in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ if there exist x_1, x_2, and x_3 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 = \mathbf{b}.$$

(Recall that we can use any letters we want for the scalars. They are simply unknown scalars we want to solve for.)

(a) Explain why $\mathbf{b}$ is in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$. (Hint: What is the matrix we need to reduce?)

(b) Write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. In how many ways can $\mathbf{b}$ be written as a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$? Explain.

(2) In problem 1 we saw that the vector $\mathbf{b}$ could be written in infinitely many different ways as linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. We now ask the question if we really need all of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ to make $\mathbf{b}$ as a linear combination in a unique way.

(a) Can the vector $\mathbf{b}$ be written as a linear combination of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$? If not, why not? If so, in how many ways can $\mathbf{b}$ be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$? Explain.

(b) If possible, write $\mathbf{b}$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(3) In problem 1 we saw that $\mathbf{b}$ could be written in infinitely many different ways as a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. However, the vector $\mathbf{b}$ could only be written in one way as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. So $\mathbf{b}$ is in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathbf{b}$ is also in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$. This raises a question – is *any* vector in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ also in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$. If so, then the vector $\mathbf{v}_3$ is redundant in terms of forming the span of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. For the sake of efficiency, we want to recognize and eliminate this redundancy.

(a) Can $\mathbf{v}_3$ be written as a linear combination of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$? If not, why not? If so, write $\mathbf{v}_3$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.

(b) Use the result of part (a) to decide if *any* vector in Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$.

Linear Independence

In this section we will investigate the concepts of linear dependence and independence of a set of vectors. Our goal is to be able to efficiently determine when a given set of vectors forms a *minimal spanning set*. This will involve the concepts of span and linear independence. Minimal spanning sets are important in that they provide the most efficient way to represent vectors in a space, and will later allow us to define the dimension of a vector space.

In Preview Activity 6.1 we considered the case where we had a set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of three vectors, and the vector $\mathbf{v}_3$ was in the span of $\{\mathbf{v}_1, \mathbf{v}_2\}$. So the the vector $\mathbf{v}_3$ did not add anything to the span of $\{\mathbf{v}_1, \mathbf{v}_2\}$. In other words, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ was larger than it needed to be in order to generate the vectors in its span – that is, $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. However, neither of the vectors in the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ could be removed without changing its span. In this case, the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ is what we will call a *minimal spanning set* or a *basis* for Span S. There are two important properties that make $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for Span S. The first is that every vector in Span S can be written as linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$ (we also use the terminology that the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ span Span S), and the second is that every vector in Span S can be written in exactly one way as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. This second property is the property of linear independence, and it is the property that makes the spanning set *minimal*.

To make a spanning set minimal, we want to be able to write every vector in the span in a unique way in terms of the spanning vectors. Notice that the zero vector can always be written as a linear combination of any set of vectors using 0 for all of the weights. So to have a *minimal* or *linearly independent* spanning set, that is, to have a unique representation for each vector in the span, it will need to be the case that the *only* way we can write the zero vector as a linear combination of a set of vectors is if all of the weights are 0. This leads us to the definition of a linearly independent set of vectors.

Definition 6.1. A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is **linearly independent** if the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

for scalars $x_1, x_2, \ldots, x_k$ has only the trivial solution

$$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

If a set of vectors is not linearly independent, then the set is **linearly dependent**.

Alternatively, we say that the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent (or dependent) if the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent (or dependent).

Note that the definition tells us that a set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly dependent if there are scalars x_1, x_2, ..., x_n, not all of which are 0 so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

Activity 6.1. Which of the following sets in $\mathbb{R}^2$ or $\mathbb{R}^3$ is linearly independent and which is linearly dependent? Why? For the linearly dependent sets, write one of the vectors as a linear combination of the others, if possible.

(a) $S_1 = \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 8 \\ 1 \end{bmatrix}, \begin{bmatrix} -4 \\ 8 \\ 0 \end{bmatrix} \right\}$.

(b) $S_2 = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \right\}$. (Hint: What relationship must exist between two vectors if they are linearly dependent?)

(c) The vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ as shown in Figure 6.2.

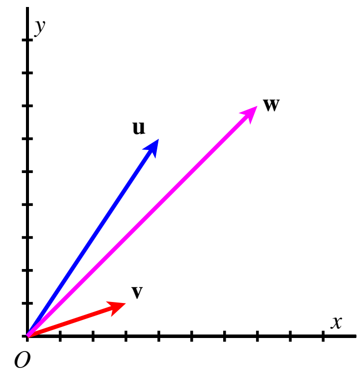


Figure 6.2: Vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

Activity 6.1 (a) and (c) illustrate how we can write one of the vectors in a linearly dependent set as a linear combination of the others. This would allow us to write at least one of the vectors in the span of the set in more than one way as a linear combination of vectors in this set. We prove this result in general in the following theorem.

Theorem 6.2. *A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly dependent if and only if at least one of the vectors in the set can be written as a linear combination of the remaining vectors in the set.*

The next activity is intended to help set the stage for the proof of Theorem 6.2.

Activity 6.2. The statement of Theorem 6.2 is a bi-conditional statement (an if and only if statement). To prove this statement about the set S we need to show two things about S. One: we must demonstrate that if S is a linearly dependent set, then at least one vector in S is a linear combination of the other vectors (this is the "only if" part of the biconditional statement) and Two: if at least one vector in S is a linear combination of the others, then S is linearly dependent (this is the "if" part of the biconditional statement). We illustrate the main idea of the proof using a three vector set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

(a) First let us assume that S is a linearly dependent set and show that at least one vector in S is a linear combination of the other vectors. Since S is linearly dependent we can write the zero vector as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ with at least one nonzero weight. For example, suppose

$$2\mathbf{v}_1 + 3\mathbf{v}_2 + 4\mathbf{v}_3 = \mathbf{0}. \quad (6.1)$$

Solve Equation (6.1) for the vector $\mathbf{v}_2$ to show that $\mathbf{v}_2$ can be written as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_3$. Conclude that $\mathbf{v}_2$ is a linear combination of the other vectors in the set S.

(b) Now we assume that at least one of the vectors in S is a linear combination of the others. For example, suppose that

$$\mathbf{v}_3 = \mathbf{v}_1 + 5\mathbf{v}_2. \tag{6.2}$$

Use vector algebra to rewrite Equation 6.2 so that $\mathbf{0}$ is expressed as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ such that the weight on $\mathbf{v}_3$ is not zero. Conclude that the set S is linearly dependent.

Now we provide a formal proof of Theorem 6.2, using the ideas from Activity 6.2.

Proof of Theorem 6.2. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. We will begin by verifying the first statement.

We assume that S is a linearly dependent set and show that at least one vector in S is a linear combination of the others. Since S is linearly dependent, there are scalars x_1, x_2, …, x_n, not all of which are 0, so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}. \tag{6.3}$$

We don't know which scalar(s) are not zero, but there is at least one. So let us assume that x_i is not zero for some i between 1 and k. We can then subtract $x_i\mathbf{v}_i$ from both sides of Equation (6.3) and divide by x_i to obtain

$$\mathbf{v}_i = \frac{x_1}{x_i}\mathbf{v}_1 + \frac{x_2}{x_i}\mathbf{v}_2 + \cdots + \frac{x_{i-1}}{x_i}\mathbf{v}_{i-1} + \frac{x_{i+1}}{x_i}\mathbf{v}_{i+1} + \frac{x_{i+2}}{x_i}\mathbf{v}_{i+2} + \cdots + \frac{x_k}{x_i}\mathbf{v}_k.$$

Thus, the vector $\mathbf{v}_i$ is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{i-1}$, $\mathbf{v}_{i+1}$, …, $\mathbf{v}_k$, and at least one of the vectors in S is a linear combination of the other vectors in S.

To verify the second statement, we assume that at least one of the vectors in S can be written as a linear combination of the others and show that S is then a linearly dependent set. We don't know which vector(s) in S can be written as a linear combination of the others, but there is at least one. Let us suppose that $\mathbf{v}_i$ is a linear combination of the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{i-1}$, $\mathbf{v}_{i+1}$, …, $\mathbf{v}_k$ for some i between 1 and k. Then there exist scalars x_1, x_2, …, x_{-1}, x_{i+1}, …, x_n so that

$$\mathbf{v}_i = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_{i+1}\mathbf{v}_{i+1} + x_{i+2}\mathbf{v}_{i+2} + \cdots + x_k\mathbf{v}_k.$$

It follows that

$$\mathbf{0} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + (-1)\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + x_{i+2}\mathbf{v}_{i+2} + \cdots + x_k\mathbf{v}_k.$$

So there are scalars there are scalars x_1, x_2, …, x_n (with $x_i = -1$), not all of which are 0, so that

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

This makes S a linearly dependent set. ■

With a linearly dependent set, at least one of the vectors in the set is a linear combination of the others. With a linearly independent set, this cannot happen – no vector in the set can be written as a linear combination of the others. This result is given in the next theorem. You may be able to see how Theorems 6.2 and 6.3 are logically equivalent.

Theorem 6.3. *A set* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *of vectors in* $\mathbb{R}^n$ *is linearly independent if and only if no vector in the set can be written as a linear combination of the remaining vectors in the set.*

Activity 6.3. As was hinted at in Preview Activity 6.1, an important consequence of a linearly independent set is that every vector in the span of the set can be written in one and only one way as a linear combination of vectors in the set. It is this uniqueness that makes linearly independent sets so useful. We explore this idea in this activity for a linearly independent set of three vectors. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a linearly independent set of vectors in $\mathbb{R}^n$ for some n, and let $\mathbf{b}$ be a vector in Span S. To show that $\mathbf{b}$ can be written in exactly one way as a linear combination of vectors in S, we assume that

$$\mathbf{b} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 \quad \text{and} \quad \mathbf{b} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + y_3\mathbf{v}_3$$

for some scalars x_1, x_2, x_3, y_1, y_2, and y_3. We need to demonstrate that $x_1 = y_1$, $x_2 = y_2$, and $x_3 = y_3$.

(a) Use the two different ways of writing $\mathbf{b}$ as a linear combination of $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ to come up with a linear combination expressing $\mathbf{0}$ as a linear combination of these vectors.

(b) Use the linear independence of the vectors $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ to explain why $x_1 = y_1, x_2 = y_2$, and $x_3 = y_3$.

Activity 6.3 contains the general ideas to show that any vector in the span of a linearly independent set can be written in one and only one way as a linear combination of the vectors in the set. The weights of such a linear combination provide us a *coordinate system* for the vectors in terms of the basis. Two familiar examples of coordinate systems are the Cartesian coordinates in the xy-plane, and xyz-space. We will revisit the coordinate system idea in a later chapter.

In the next theorem we state and prove the general case of any number of linearly independent vectors producing unique representations as linear combinations.

Theorem 6.4. *Let* $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *be a linearly independent set of vectors in* $\mathbb{R}^n$*. Any vector in Span* S *can be written in one and only one way as a linear combination of the vectors* $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$.

Proof. Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a linearly independent set of vectors in $\mathbb{R}^n$, and let $\mathbf{b}$ be a vector in Span S. By definition, it follows that $\mathbf{b}$ can be written as a linear combination of the vectors in S. It remains for us to show that this representation is unique. So assume that

$$\mathbf{b} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k \quad \text{and} \quad \mathbf{b} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_k\mathbf{v}_k \tag{6.4}$$

for some scalars x_1, x_2, …, x_k, and y_1, y_2, …, y_k. Then

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \cdots + y_k\mathbf{v}_k.$$

Subtracting all terms from the right side and using a little vector algebra gives us

$$(x_1 - y_1)\mathbf{v}_1 + (x_2 - y_2)\mathbf{v}_2 + \cdots + (x_k - y_k)\mathbf{v}_k = \mathbf{0}.$$

The fact that S is a linearly independent set implies that

$$x_1 - y_1 = 0, \; x_2 - y_2 = 0, \; \ldots, \; x_k - y_k = 0,$$

showing that $x_i = y_i$ for every i between 1 and k. We conclude that the representation of $\mathbf{b}$ as a linear combination of the linearly independent vectors in S is unique.

■

Determining Linear Independence

The definition and our previous work give us a straightforward method for determining when a set of vectors in $\mathbb{R}^n$ is linearly independent or dependent.

Activity 6.4. In this activity we learn how to use a matrix to determine in general if a set of vectors in $\mathbb{R}^n$ is linearly independent or dependent. Suppose we have k vectors $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$ in $\mathbb{R}^n$. To see if these vectors are linearly independent, we need to find the solutions to the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}. \tag{6.5}$$

If we let $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_k]$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$, then we can write the vector equation (6.5) in matrix form $A\mathbf{x} = \mathbf{0}$. Let B be the reduced row echelon form of A.

(a) What can we say about the pivots of B in order for $A\mathbf{x} = \mathbf{0}$ to have exactly one solution? Under these conditions, are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$ linearly independent or dependent?

(b) What can we say about the rows or columns of B in order for $A\mathbf{x} = \mathbf{0}$ to have infinitely many solutions? Under these conditions, are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$ linearly independent or dependent?

(c) Use the result of parts (a) and (b) to determine if the vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 3 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}$ in $\mathbb{R}^4$ are linearly independent or dependent. If dependent, write one of the vectors as a linear combination of the others. You may use the fact that the matrix $\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 2 & 2 & 2 \\ 0 & 3 & 1 \end{bmatrix}$ is row equivalent to $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

Minimal Spanning Sets

It is important to note the differences and connections between linear independence, span, and minimal spanning set.

- The set $S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ is not a minimal spanning set for $\mathbb{R}^3$ even though S is a linearly independent set. Note that S does not span $\mathbb{R}^3$ since the vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is not in Span S.

- The set $T = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is not a minimal spanning set for $\mathbb{R}^3$ even though Span $T = \mathbb{R}^3$. Note that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

so T is not a linearly independent set.

- The set $U = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a minimal spanning set for $\mathbb{R}^3$ since it satisfies both characteristics of a minimal spanning set: Span $U = \mathbb{R}^3$ AND U is linearly independent.

The three concepts – linear independence, span, and minimal spanning set – are different. The important point to note is that minimal spanning set must be both linearly independent and span the space.

To find a minimal spanning set we will often need to find a smallest subset of a given set of vectors that has the same span as the original set of vectors. In this section we determine a method for doing so.

Activity 6.5. Let $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 0 \\ -4 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} -3 \\ 4 \\ 18 \end{bmatrix}$ in $\mathbb{R}^3$. Assume that the reduced row echelon form of the matrix $A = \begin{bmatrix} -1 & 2 & 0 & -3 \\ 0 & 0 & 1 & 4 \\ 2 & -4 & 3 & 18 \end{bmatrix}$ is $\begin{bmatrix} 1 & -2 & 0 & 3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(a) Write the general solution to the homogeneous system $A\mathbf{x} = \mathbf{0}$, where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$. Write all linear combinations of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ that are equal to $\mathbf{0}$, using weights that only involve x_2 and x_4.

(b) Explain how we can conveniently choose the weights in the general solution to $A\mathbf{x} = \mathbf{0}$ to show that the vector $\mathbf{v}_4$ is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$. What does this tell us about Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$?

(c) Explain how we can conveniently choose the weights in the general solution to $A\mathbf{x} = \mathbf{0}$ to show why the vector $\mathbf{v}_2$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_3$. What does this tell us about Span$\{\mathbf{v}_1, \mathbf{v}_3\}$ and Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$?

(d) Is $\{\mathbf{v}_1, \mathbf{v}_3\}$ a minimal spanning set for Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$? Explain your response.

Activity 6.5 illustrates how we can use a matrix to determine a minimal spanning set for a given set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in $\mathbb{R}^n$.

- Form the matrix $A = [\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_k]$.
- Find the reduced row echelon form $[B \mid \mathbf{0}]$ of $[A \mid \mathbf{0}]$. If B contains non-pivot columns, say for example that the ith column is a non-pivot column, then we can choose the weight x_i corresponding to the ith column to be 1 and all weights corresponding to the other non-pivot columns to be 0 to make a linear combination of the columns of A that is equal to $\mathbf{0}$. This allows us to write $\mathbf{v}_i$ as a linear combination of the vectors corresponding to the pivot columns of A as we did in the proof of Theorem 6.3. So every vector corresponding to a non-pivot column is in the span of the set of vectors corresponding to the pivot columns. The vectors corresponding to the pivot columns are linearly independent, since the matrix with those columns has every column as a pivot column. Thus, the set of vectors corresponding to the pivot columns of A forms a minimal spanning set for $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

IMPORTANT NOTE! The set of pivot columns of the reduced row echelon form of A will normally not have the same span as the set of columns of A, so it is critical that we use columns of A, NOT B in our minimal spanning set.

Activity 6.6. Find a minimal spanning set for the span of the set

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

Activity 6.5 also illustrates a general process by which we can find a minimal spanning set – that is the smallest subset of vectors that has the same span. This process will be useful later when we consider vectors in arbitrary vector spaces. The idea is that if we can write one of the vectors in a set S as a linear combination of the remaining vectors, then we can remove that vector from the set and maintain the same span. In other words, begin with the span of a set S and follow these steps:

Step 1. If S is a linearly independent set, we already have a minimal spanning set.

Step 2. If S is not a linearly independent set, then one of the vectors in S is a linear combination of the others. Remove that vector from S to obtain a new set T. It will be the case that Span T = Span S.

Step 3. If T is a linearly independent set, then T is a minimal spanning set. If not, repeat steps 2 and 3 for the set T until you arrive at a linearly independent set.

This process is guaranteed to stop as long as the set contains at least one nonzero vector. A verification of the statement in Step 2 that Span T = Span S is given in the next theorem.

Theorem 6.5. *Let* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ *be a set of vectors in* $\mathbb{R}^n$ *so that for some* i *between 1 and* k, $\mathbf{v}_i$ *is in Span*$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}$. *Then*

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}.$$

Proof. Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$ so that $\mathbf{v}_i$ is in the span of $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{i-1}$, $\mathbf{v}_{i+1}$, …, and $\mathbf{v}_k$ for some i between 1 and k. To show that

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\},$$

we need to show that

(1) every vector in Span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is in Span$\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}$, and

(2) every vector in Span$\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}$ is in Span$\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

Let us consider the second containment. Let $\mathbf{x}$ be a vector in the span of $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{i-1}$, $\mathbf{v}_{i+1}$, …, and $\mathbf{v}_k$. Then

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k$$

for some scalars x_1, x_2, …, x_{i-1}, x_{i+1}, …, x_k. Note that

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + (0)\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k$$

as well, so $\mathbf{x}$ is in Span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$. Thus,

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}.$$

(This same argument shows a more general statement that if S is a subset of T, then Span $S \subseteq$ Span T.)

Now we demonstrate the first containment. Here we need the assumption that $\mathbf{v}_i$ is in Span$\{\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_{i-1}$, $\mathbf{v}_{i+1}$, …, $\mathbf{v}_k\}$ for some i between 1 and k. That assumption gives us

$$\mathbf{v}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_k\mathbf{v}_k \tag{6.6}$$

for some scalars c_1, c_2, …, c_{i-1}, c_{i+1}, …, c_k. Now let $\mathbf{x}$ be a vector in the span of $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_k$. Then

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k$$

for some scalars x_1, x_2, …, x_k. Substituting from (6.6) shows that

$$\begin{aligned}\mathbf{x} &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k \\ &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} + x_i\mathbf{v}_i + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k \\ &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_{i-1}\mathbf{v}_{i-1} \\ &\quad\quad + x_i[c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_k\mathbf{v}_k] \\ &\quad\quad + x_{i+1}\mathbf{v}_{i+1} + \cdots + x_k\mathbf{v}_k \\ &= (x_1 + x_ic_1)\mathbf{v}_1 + (x_2 + x_ic_2)\mathbf{v}_2 + \cdots + (x_{i-1} + x_ic_{i-1})\mathbf{v}_{i-1} \\ &\quad\quad + (x_{i+1} + x_ic_{i+1})\mathbf{v}_{i+1} \cdots + (x_k + x_ic_k)\mathbf{v}_k.\end{aligned}$$

So $\mathbf{x}$ is in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}$ and

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}.$$

Since the two sets are subsets of each other, they must be equal sets. We conclude that

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}.$$

■

The result of Theorem 6.5 is that if we have a finite set S of vectors in $\mathbb{R}^n$, we can eliminate those vectors that are linear combinations of others until we obtain a smallest set of vectors that still has the same span. As mentioned earlier, we call such a minimal spanning set a basis.

Definition 6.6. Let S be a set of vectors in $\mathbb{R}^n$. A subset B of S is a **basis** for Span S if B is linearly independent and Span B = Span S.

IMPORTANT NOTE: A basis is defined by two characteristics. A basis must span the space in question and a basis must be a linearly independent set. It is the linear independence that makes a basis a *minimal* spanning set.

We have worked with a familiar basis in $\mathbb{R}^2$ throughout our mathematical careers. A vector $\begin{bmatrix} a \\ b \end{bmatrix}$ in $\mathbb{R}^2$ can be written as

$$\begin{bmatrix} a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

So the set $\{\mathbf{e}_1, \mathbf{e}_2\}$, where $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ spans $\mathbb{R}^2$. Since the columns of $[\mathbf{e}_1 \ \mathbf{e}_2]$ are linearly independent, so is the set $\{\mathbf{e}_1, \mathbf{e}_2\}$. Therefore, the set $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis for $\mathbb{R}^2$. The vector $\mathbf{e}_1$ is in the direction of the positive x-axis and the vector $\mathbf{e}_2$ is in the direction of the positive y-axis, so decomposing a vector $\begin{bmatrix} a \\ b \end{bmatrix}$ as a linear combination of $\mathbf{e}_1$ and $\mathbf{e}_2$ is akin to identifying the vector with the point (a, b) as we discussed earlier. The set $\{\mathbf{e}_1, \mathbf{e}_2\}$ is called the *standard basis* for $\mathbb{R}^2$.

This idea is not restricted to $\mathbb{R}^2$. Consider the vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ \cdots, \ \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

in $\mathbb{R}^n$. That is, the vector $\mathbf{e}_i$ is the vector with a 1 in the ith position and 0s everywhere else. Since the matrix $[\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n]$ is the identity matrix, the set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{R}^n$. The set $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is called the *standard basis* for $\mathbb{R}^n$.

As we will see later, bases[1] are of fundamental importance in linear algebra in that bases will allow us to define the dimension of a vector space and will provide us with coordinate systems.

We conclude this section with an important theorem that is similar Theorem 5.3.

Theorem 6.7. *Let A be an $m \times n$ matrix. The following statements are equivalent.*

(1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector $\mathbf{b}$ in the span of the columns of A.

(2) The matrix equation $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.

(3) The columns of A are linearly independent.

(4) The matrix A has a pivot position in each column.

Examples

What follows are worked examples that use the concepts from this section.

Example 6.8. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 6 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{v}_3 = \begin{bmatrix} 3 \\ -6 \\ 2 \\ -7 \end{bmatrix}$, and $\mathbf{v}_4 = \begin{bmatrix} 5 \\ -2 \\ 2 \\ -5 \end{bmatrix}$.

(a) Is the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ linearly independent or dependent. If independent, explain why. If dependent, write one of the vectors in S as a linear combination of the other vectors in S.

(b) Find a subset B of S that is a basis for Span S. Explain how you know you have a basis.

Example Solution.

(a) We need to know the solutions to the vector equation

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 = \mathbf{0}.$$

If the equation has as its only solution $x_1 = x_2 = x_3 = x_4 = 0$ (the trivial solution), then the set S is linearly independent. Otherwise the set S is linearly dependent.

To find the solutions to this system, we row reduce the augmented matrix

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

[1]The plural of basis is bases.

(Note that we really don't need the augmented column of zeros – row operations won't change that column at all. We just need to know that the column of zeros is there.) Technology shows that the reduced row echelon form of this augmented matrix is

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

The reduced row echelon form tells us that the vector equation is consistent, and the fact that there is no pivot in the fourth column shows that the system has a free variable and more than just the trivial solution. We conclude that S is linearly dependent.

Moreover, the general solution to our vector equation is

$$\begin{aligned} x_1 &= -3x_3 - 5x_4 \\ x_2 &= 2x_3 + 2x_4 \\ x_3 &\text{ is free} \\ x_4 &\text{ is free.} \end{aligned}$$

Letting $x_4 = 0$ and $x_3 = 1$ shows that one non-trivial solution to our vector equation is

$$x_1 = -3,\ x_2 = 2,\ x_3 = 1, \text{ and } x_4 = 0.$$

Thus,

$$-3\mathbf{v}_1 + 2\mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0},$$

or

$$\mathbf{v}_3 = 3\mathbf{v}_1 - 2\mathbf{v}_2$$

and we have written one vector in S as a linear combination of the other vectors in S.

(b) We have seen that the pivot columns in a matrix A form a minimal spanning set (or basis) for the span of the columns of A. From part (a) we see that the pivot columns in the reduced row echelon form of $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \mathbf{v}_4]$ are the first and second columns. So a basis for the span of the columns of A is $\{\mathbf{v}_1, \mathbf{v}_2\}$. Since the elements of S are the columns of A, we conclude that the set $B = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a subset of S that is a basis for Span S.

Example 6.9. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 3 \\ -7 \\ 2 \end{bmatrix}$, and $\mathbf{v}_3 = \begin{bmatrix} -5 \\ 6 \\ 10 \end{bmatrix}$.

(a) Is the set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ a basis for $\mathbb{R}^3$? Explain.

(b) Let $\mathbf{v}_4 = \begin{bmatrix} -5 \\ 6 \\ h \end{bmatrix}$, where h is a scalar. Are there any values of h for which the set $S' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ is not a basis for $\mathbb{R}^3$? If so, find all such values of h and explain why S' is not a basis for $\mathbb{R}^3$ for those values of h.

Example Solution.

(a) We need to know if the vectors in S are linearly independent and span $\mathbb{R}^3$. Technology shows that the reduced row echelon form of

$$A = \begin{bmatrix} 1 & 3 & -5 \\ 1 & -7 & 6 \\ 0 & 2 & 10 \end{bmatrix}$$

is

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

Since every column of $[\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ is a pivot column, the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent. The fact that there is a pivot in every row of the matrix A means that the equation $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$ in $\mathbb{R}^3$. Since $A\mathbf{x}$ is a linear combination of the columns of A with weights from $\mathbf{x}$, tt follows that the columns of A span $\mathbb{R}^3$. We conclude that the set S is a basis for $\mathbb{R}^3$.

(b) Technology shows that a row echelon form of $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_4]$ is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -10 & 11 \\ 0 & 0 & h + \frac{11}{5} \end{bmatrix}.$$

The columns of A are all pivot columns (hence linearly independent) as long as $h \neq -\frac{11}{5}$, and are linearly dependent when $h = -\frac{11}{5}$. So the only value of h for which S' is not a basis for $\mathbb{R}^3$ is $h = -\frac{11}{5}$.

Summary

- A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly independent if the vector equation

  $$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

  for scalars $x_1, x_2, \ldots, x_k$ has only the trivial solution

  $$x_1 = x_2 = x_3 = \cdots = x_k = 0.$$

  Another way to think about this is that a set of vectors is linearly independent if no vector in the set can be written as a linear combination of the other vectors in the set.

- A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly dependent if the vector equation

  $$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$$

  has a nontrivial solution. That is, we can find scalars $x_1, x_2, \ldots, x_k$ that are not all 0 so that

  $$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}.$$

  Another way to think about this is that a set of vectors is linearly dependent if at least one vector in the set can be written as a linear combination of the other vectors in the set.

- If S is a set of vectors, a subset B of S is a basis for Span S if B is a linearly independent set and Span B = Span S.

- Given a nonzero set S of vectors, we can remove vectors from S that are linear combinations of remaining vectors in S to obtain a linearly independent subset of S that has the same span as S.

- The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}$.

- The set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent if and only if every column of the matrix $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_k]$, is a pivot column.

- If $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \cdots\ \mathbf{v}_k]$, then the vectors in the pivot columns of A form a minimal spanning set for Span$\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$.

Exercises

(1) Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$$

Is the set consisting of these vectors linearly independent? If so, explain why. If not, make a single change in one of the vectors so that the set is linearly independent.

(2) Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \ \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ c \end{bmatrix}$$

For which values of c is the set consisting of these vectors linearly independent?

(3) In a lab, there are three different water-benzene-acetic acid solutions: The first one with 36% water, 50% benzene and 14% acetic acid; the second one with 44% water, 46% benzene and 10% acetic acid; and the last one with 38% water, 49% benzene and 13% acid. Since the lab needs space, the lab coordinator wants to determine whether all solutions are needed, or if it is possible to create one of the solutions using the other two. Can you help the lab coordinator?

(4) Given vectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, find a vector $\mathbf{v}_3$ in $\mathbb{R}^3$ so that the set consisting of $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ is linearly independent.

(5) Consider the span of $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 4 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 3 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 3 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 6 \end{bmatrix}.$$

(a) Is the set S a minimal spanning set of Span S? If not, determine a minimal spanning set, i.e. a basis, of Span S.

(b) Check that the vector $\mathbf{u} = \begin{bmatrix} 6 \\ 5 \\ -2 \\ 1 \end{bmatrix}$ is in Span S. Find the unique representation of $\mathbf{u}$ in terms of the basis vectors.

(6) Come up with a 4×3 matrix with linearly independent columns, if possible. If not, explain why not.

(7) Come up with a 3×4 matrix with linearly independent columns, if possible. If not, explain why not.

(8) Give an example of vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ such that a minimal spanning set for Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is equal to that of Span$\{\mathbf{v}_1, \mathbf{v}_2\}$; and an example of three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ such that a minimal spanning set for Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is equal to that of Span$\{\mathbf{v}_1, \mathbf{v}_3\}$.

(9) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** If $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are three vectors none of which is a multiple of another, then these vectors form a linearly independent set.

(b) **True/False** If $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ in $\mathbb{R}^n$ are linearly independent vectors, then so are $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ for any $\mathbf{v}_4$ in $\mathbb{R}^n$.

(c) **True/False** If $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ and $\mathbf{v}_4$ in $\mathbb{R}^n$ are linearly independent vectors, then so are $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$.

(d) **True/False** A 3×4 matrix cannot have linearly independent columns.

(e) **True/False** If two vectors span $\mathbb{R}^2$, then they are linearly independent.

(f) **True/False** The space $\mathbb{R}^3$ cannot contain four linearly independent vectors.

(g) **True/False** If two vectors are linearly dependent, then one is a scalar multiple of the other.

(h) **True/False** If a set of vectors in $\mathbb{R}^n$ is linearly dependent, then the set contains more than n vectors.

(i) **True/False** The columns of a matrix A are linearly independent if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.

(j) **True/False** Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a minimal spanning set for W, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$ cannot also be a minimal spanning set for W.

(k) **True/False** Let $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$. If $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a minimal spanning set for W, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ cannot also be a minimal spanning set for W.

(l) **True/False** If $\mathbf{v}_3 = 2\mathbf{v}_1 - 3\mathbf{v}_2$, then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a minimal spanning set for $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$.

Project: Generating Bézier Curves

Bézier curves can be created as linear combinations of vectors. In this section we will investigate how cubic Bézier curves (the ones used for fonts) can be realized through linear and quadratic Bézier curves. We begin with linear Bézier curves.

Project Activity 6.1. Start with two vectors $\mathbf{p}_0$ and $\mathbf{p}_1$. Linear Bézier curves are linear combinations

$$\mathbf{q} = (1-t)\mathbf{p}_0 + t\mathbf{p}_1$$

of the vectors $\mathbf{p}_0$ and $\mathbf{p}_1$ for scalars t between 0 and 1. (You can visualize these linear combinations using the GeoGebra file `Linear Bezier` at https://www.geogebra.org/m/HvrPhh86. With this file you can draw the vectors $\mathbf{q}$ for varying values of t. You can move the points $\mathbf{p}_0$ and $\mathbf{p}_1$ in the GeoGebra file, and the slider controls the values of t. The point identified with $\mathbf{q}$ is traced as t is changed.) For this activity, we will see what the curve $\mathbf{q}$ corresponds to by evaluating certain points on the curve in a specific example. Let $\mathbf{p}_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{p}_1 = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$.

(a) What are the components of the vector $(1-t)\mathbf{p}_0 + t\mathbf{p}_1$ if $t = \frac{1}{2}$? Where is this vector in relation to $\mathbf{p}_0$ and $\mathbf{p}_1$? Explain.

(b) What are the components of the vector $(1-t)\mathbf{p}_0 + t\mathbf{p}_1$ if $t = \frac{1}{3}$? Where is this vector in relation to $\mathbf{p}_0$ and $\mathbf{p}_1$? Explain.

(c) What are the components of the vector $(1-t)\mathbf{p}_0 + t\mathbf{p}_1$ for an arbitrary t? Where is this vector in relation to $\mathbf{p}_0$ and $\mathbf{p}_1$? Explain.

For each value of t, the vector $\mathbf{q} = (1-t)\mathbf{p}_0 + t\mathbf{p}_1$ is a linear combination of the vectors $\mathbf{p}_0$ and $\mathbf{p}_1$. Note that when $t = 0$, we have $\mathbf{q} = \mathbf{p}_0$ and when $t = 1$ we have $\mathbf{q} = \mathbf{p}_1$, and for $0 \leq t \leq 1$ Project Activity 6.1 shows that the vectors $\mathbf{q}$ trace out the line segment from $\mathbf{p}_0$ to $\mathbf{p}_1$. The span $\{(1-t)\mathbf{p}_0 + t\mathbf{p}_1\}$ of the vectors $\mathbf{p}_0$ and $\mathbf{p}_1$ for $0 \leq t \leq 1$ is a linear Bézier curve. Once we have a construction like this, it is natural in mathematics to extend it and see what happens. We do that in the next activity to construct quadratic Bézier curves.

Project Activity 6.2. Let $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$ be vectors in the plane. We can then let

$$\mathbf{q}_0 = (1-t)\mathbf{p}_0 + t\mathbf{p}_1 \quad \text{and} \quad \mathbf{q}_1 = (1-t)\mathbf{p}_1 + t\mathbf{p}_2$$

be the linear Bézier curves as defined in Project Activity 6.1. Since $\mathbf{q}_0$ and $\mathbf{q}_1$ are vectors, we can define $\mathbf{r}$ as

$$\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1.$$

(You can visualize these linear combinations using the GeoGebra file `Quadraric Bezier` at https://www.geogebra.org/m/VWCZZBXz. With this file you can draw the vectors $\mathbf{r}$ for varying values of t. You can move the points $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$ in the GeoGebra file, and the slider controls the values of t. The point identified with $\mathbf{r}$ is traced as t is changed.) In this activity we investigate how the vectors $\mathbf{r}$ change as t changes. For the remainder of this activity, let $\mathbf{p}_0 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$, and $\mathbf{p}_2 = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$.

(a) At what point (in terms of $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$) is the vector $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ when $t = 0$? Explain using the definition of $\mathbf{r}$.

(b) At what point (in terms of $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$) is the vector $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ when $t = 1$? Explain using the definition of $\mathbf{r}$.

(c) Find by hand the components of the vector $(1-t)\mathbf{q}_0 + t\mathbf{q}_1$ with $t = \frac{1}{4}$. Compare with the result of the GeoGebra file.

The span $\{(1-t)\mathbf{q}_0 + t\mathbf{q}_1\}$ of the vectors $\mathbf{q}_0$ and $\mathbf{q}_1$, or the set of points traced out by the vectors $\mathbf{r}$ for $0 \leq t \leq 1$, is a quadratic Bézier curve. To understand why this curve is called quadratic, we examine the situation in a general context in the following activity.

Project Activity 6.3. Let $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$ be arbitrary vectors in the plane. Write $\mathbf{r} = (1-t)\mathbf{q}_0 + t\mathbf{q}_1$ as a linear combination of $\mathbf{p}_0$, $\mathbf{p}_1$, and $\mathbf{p}_2$. That is, write $\mathbf{r}$ in the form $a_0\mathbf{p}_0 + a_1\mathbf{p}_1 + a_2\mathbf{p}_2$ for some scalars (that may depend on t) a_0, a_1, and a_2. Explain why the result leads us to call these vectors *quadratic* Bézier curves.

Notice that if any one of the $\mathbf{p}_i$ lies on the line determined by the other two vectors, then the quadratic Bézier curve is just a line segment. So to obtain something non-linear we need to choose our vectors so that that doesn't happen.

Quadratic Bézier curves are limited, because their graphs are parabolas. For applications we need higher order Bézier curves. In the next activity we consider cubic Bézier curves.

Project Activity 6.4. Start with four vectors $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ – the points defined by these vectors are called *control points* for the curve. As with the linear and quadratic Bézier curves, we let

$$\mathbf{q}_0 = (1-t)\mathbf{p}_0 + t\mathbf{p}_1, \quad \mathbf{q}_1 = (1-t)\mathbf{p}_1 + t\mathbf{p}_2, \quad \text{and} \quad \mathbf{q}_2 = (1-t)\mathbf{p}_2 + t\mathbf{p}_3.$$

Then let

$$\mathbf{r}_0 = (1-t)\mathbf{q}_0 + t\mathbf{q}_1 \quad \text{and} \quad \mathbf{r}_1 = (1-t)\mathbf{q}_1 + t\mathbf{q}_2.$$

We take this one step further to generate the cubic Bézier curves by letting

$$\mathbf{s} = (1-t)\mathbf{r}_0 + t\mathbf{r}_1.$$

(You can visualize these linear combinations using the GeoGebra file `Cubic Bezier` at https://www.geogebra.org/m/EDAhudy9. With this file you can draw the vectors $\mathbf{s}$ for varying values of t. You can move the points $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ in the GeoGebra file, and the slider controls the values of t. The point identified with $\mathbf{s}$ is traced as t is changed.) In this activity we investigate how the vectors $\mathbf{s}$ change as t changes. For the remainder of this activity, let $\mathbf{p}_0 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\mathbf{p}_1 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\mathbf{p}_2 = \begin{bmatrix} 9 \\ -3 \end{bmatrix}$, and $\mathbf{p}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

(a) At what point (in terms of $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$) is the vector $\mathbf{s} = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$ when $t = 0$? Explain using the definition of $\mathbf{s}$.

(b) At what point (in terms of $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$) is the vector $\mathbf{s} = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$ when $t = 1$? Explain using the definition of $\mathbf{s}$.

(c) Find by hand the components of the vector $(1-t)\mathbf{r}_0 + t\mathbf{r}_1$ with $t = \frac{3}{4}$. Compare with the result of the GeoGebra file.

The span $\{(1-t)\mathbf{r}_0 + t\mathbf{r}_1\}$ of the vectors $\mathbf{r}_0$ and $\mathbf{r}_1$, or the set of points traced out by the vectors $\mathbf{s}$ for $0 \leq t \leq 1$, is a cubic Bézier curve. To understand why this curve is called cubic, we examine the situation in a general context in the following activity.

Project Activity 6.5. Let $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ be arbitrary vectors in the plane. Write $\mathbf{s} = (1-t)\mathbf{r}_0 + t\mathbf{r}_1$ as a linear combination of $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$. That is, write $\mathbf{s}$ in the form $b_0\mathbf{p}_0 + b_1\mathbf{p}_1 + b_2\mathbf{p}_2 + b_3\mathbf{p}_3$ for some scalars (that may depend on t) b_0, b_1, b_2, and b_3. Explain why the result leads us to call these vectors *cubic* Bézier curves.

Just as with the quadratic case, we need certain subsets of the set of control vectors to be linearly independent so that the cubic Bézier curve does not degenerate to a quadratic or linear Bézier curve.

More complicated and realistic shapes can be represented by piecing together two or more Bézier curves as illustrated with the letter "S" in Figure 6.1. Suppose we have two cubic Bézier curves, the first with control points $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ and the second with control points $\mathbf{p}'_0$, $\mathbf{p}'_1$, $\mathbf{p}'_2$, and $\mathbf{p}'_3$. You may have noticed that $\mathbf{p}_1$ lies on the tangent line to the first Bézier curve at $\mathbf{p}_0$ and that $\mathbf{p}_2$ lies on the tangent line to the first Bézier curve at $\mathbf{p}_3$. (Play around with the program `Cubic Bezier` to convince yourself of these statements. This can be proved in a straightforward manner using vector calculus.) So if we want to make a smooth curve from these two Bézier curves, the curves will need to join together smoothly at $\mathbf{p}_3$ and $\mathbf{p}'_0$. This will force $\mathbf{p}_3 = \mathbf{p}'_0$ and the tangents at $\mathbf{p}_3 = \mathbf{p}'_0$ will have to match. This implies that $\mathbf{p}_2$, $\mathbf{p}_3$, and $\mathbf{p}'_1$ all have to lie on this common tangent line. Keeping this idea in mind, use the GeoGebra file `Cubic Bezier Pair` at https://www.geogebra.org/m/UwxQ6RPk to find control points for the pair of Bézier curves that create your own letter S.

Section 7

Matrix Transformations

Focus Questions

By the end of this section, you should be able to give precise and thorough answers to the questions listed below. You may want to keep these questions in mind to focus your thoughts as you complete the section.

- What is a matrix transformation?
- What properties do matrix transformations have? (In particular, what properties make matrix transformations *linear*?)
- What is the domain of a matrix transformation defined by an $m \times n$ matrix? Why?
- What are the range and codomain of a matrix transformation defined by an $m \times n$ matrix? Why?
- What does it mean for a matrix transformation to be one-to-one? If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, what are the conditions on A that make T a one-to-one transformation?
- What does it mean for a matrix transformation to be onto? If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, what are the conditions on A that make T an onto transformation?

Application: Computer Graphics

As we will discuss, left multiplication by an $m \times n$ matrix defines a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Such a function defined by matrix multiplication is called a matrix transformation. In this section we study some of the properties of matrix transformations and understand how, using the pivots of the matrix, to determine when the output of a matrix transformation covers the whole space R^m or when a transformation maps distinct vectors to distinct outputs.

Matrix transformations are used extensively in computer graphics to produce animations as seen in video games and movies. For example, consider the dancing figure at left in Figure 7.1. We can identify certain control points (e.g., the point at the neck, where the arms join the torso, etc.) to mark the locations of important points. Using just the control points we can reconstruct the figure. Each control point can be represented as a vector, and so we can manipulate the figure by manipulating the control points with matrix transformations. We will explore this idea in more detail later in this section.

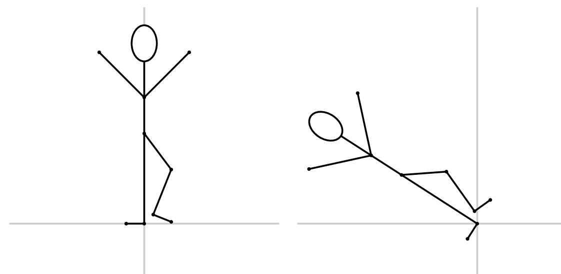


Figure 7.1: A dancing figure and a rotated dancing figure.

Introduction

In this section we will consider special functions which take vectors as inputs and produce vectors as outputs. We will use matrix multiplication to produce the output vectors.

If A is an $m \times n$ matrix and $\mathbf{x}$ is a vector in $\mathbb{R}^n$, then the matrix-vector product $A\mathbf{x}$ is a vector in $\mathbb{R}^m$. (Pick some specific n, m values to understand this statement better.) Therefore, left multiplication by the matrix A takes an input vector $\mathbf{x}$ in $\mathbb{R}^n$ and produces an output vector $A\mathbf{x}$ in $\mathbb{R}^m$, which we will refer to as the *image* of $\mathbf{x}$ under the transformation. This defines a function T from $\mathbb{R}^n$ to $\mathbb{R}^m$ where

$$T(\mathbf{x}) = A\mathbf{x}\,.$$

These functions are the matrix transformations.

Definition 7.1. A **matrix transformation** is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$T(\mathbf{x}) = A\mathbf{x}$$

for some $m \times n$ matrix A.

Many of the transformations we consider in this section are from $\mathbb{R}^2$ to $\mathbb{R}^2$ so that we can visualize the transformations. As an example, let us consider the transformation T defined by

$$T\left(\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right]\right) = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{c} x_1 \\ x_2 \end{array}\right].$$

If we plot the input vectors $\mathbf{u}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$, $\mathbf{u}_2 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$, $\mathbf{u}_3 = \left[\begin{array}{c} 1 \\ 2 \end{array}\right]$, and $\mathbf{u}_4 = \left[\begin{array}{r} -1 \\ 1 \end{array}\right]$ (as (blue) circles) and their images $T(\mathbf{u}_1) = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]\left[\begin{array}{c} 1 \\ 0 \end{array}\right] = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$, $T(\mathbf{u}_2) = \left[\begin{array}{r} 0 \\ -1 \end{array}\right]$,

$T(\mathbf{u}_3) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$, and $T(\mathbf{u}_4) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (as (red) $\times$'s) on the same set of axes as shown in Figure 7.2, we see that this transformation reflects the input vectors across the x-axis. We can also see this algebraically since the reflection of the point (x_1, x_2) around the x-axis is the point $(x_1, -x_2)$, and

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix}.$$

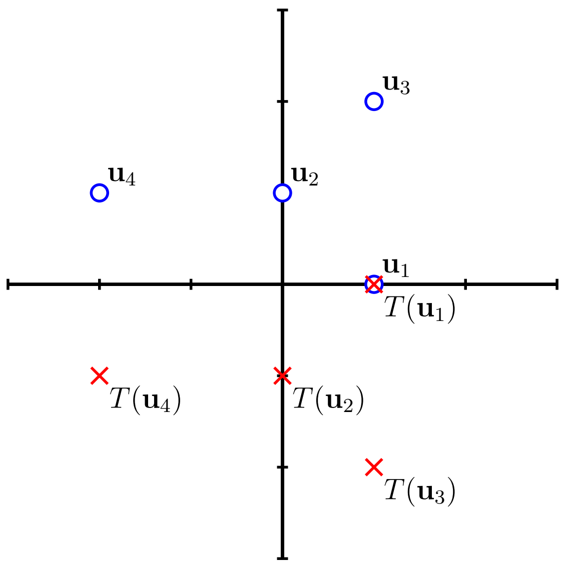


Figure 7.2: Inputs and outputs of the transformation T.

Preview Activity 7.1. We now consider other transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$.

(1) Suppose a transformation T is defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(a) Find $T(\mathbf{u}_i)$ for each of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. (In other words, substitute $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ into the formula above to see what output is obtained.)

(b) Plot all input vectors and their images on the same axes in $\mathbb{R}^2$. Clearly identify which image corresponds to which input vector. Then give a geometric description of what this transformation does.

(2) The transformation in the introduction performs a reflection across the x-axis. Find a matrix transformation that performs a reflection across the y-axis.

(3) Suppose a transformation T is defined by

$$T(\mathbf{x}) = A\mathbf{x},$$

where

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

(a) Find $T(\mathbf{u}_i)$ for each of $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

(b) Plot all input vectors and their images on the same axes in $\mathbb{R}^2$. Give a geometric description of this transformation.

(c) Is there an input vector which produces $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as an output vector?

(d) Find all input vectors that produce the output vector $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Is there a unique input vector, or multiple input vectors?

Properties of Matrix Transformations

A matrix transformation is a function. When dealing with functions in previous mathematics courses we have used the terms domain and range with our functions. Recall that the domain of a function is the set of all allowable inputs into the function and the range of a function is the set of all outputs of the function. We do the same with transformations. If T is the matrix transformation $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A, then T maps vectors from $\mathbb{R}^n$ into $\mathbb{R}^m$. So $\mathbb{R}^n$ is the *domain* of T – the set of all input vectors. However, the set $\mathbb{R}^m$ is only the target set for T and not necessarily the range of T. We call $\mathbb{R}^m$ the *codomain* of T, while the *range* of T is the set of all output vectors. The range is always a subset of the codomain, but the two sets do not have to be equal. In addition, if a vector $\mathbf{b}$ in $\mathbb{R}^m$ satisfies $\mathbf{b} = T(\mathbf{x})$ for some $\mathbf{x}$ in $\mathbb{R}^n$, then we say that $\mathbf{b}$ is the *image* of $\mathbf{x}$ under the transformation T.

Because of the properties of the matrix-vector product, if the matrix transformation T is defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A, then

$$T(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$$

and

$$T(c\mathbf{u}) = A(c\mathbf{u}) = cA\mathbf{u} = cT(\mathbf{u})$$

for any vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ and any scalar c. So every matrix transformation T satisfies the following two important properties:

(1) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ and

(2) $T(c\mathbf{u}) = cT(\mathbf{v})$.

The first property says that a matrix transformation T preserves sums of vectors and the second that T preserves scalar multiples of vectors.

Activity 7.1. Let T be a matrix transformation, and let $\mathbf{u}$ and $\mathbf{v}$ be vectors in the domain of T so that $T(\mathbf{u}) = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $T(\mathbf{v}) = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}$.

(a) Exactly which vector is $T(2\mathbf{u} - 3\mathbf{v})$? Explain.

(b) If a and b are any scalars, what is the vector $T(a\mathbf{u} + b\mathbf{v})$? Why?

As we saw in Activity 7.1, we can combine the two properties of a matrix transformation T into one: for any scalars a and b and any vectors $\mathbf{u}$ and $\mathbf{v}$ in the domain of T we have

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v}). \tag{7.1}$$

We can then extend equation (7.1) (by mathematical induction) to any finite linear combination of vectors. That is, if $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k$ are any vectors in the domain of a matrix transformation T and if x_1, x_2, ..., x_k are any scalars, then

$$T(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k) = x_1T(\mathbf{v}_1) + x_2T(\mathbf{v}_2) + \cdots + x_kT(\mathbf{v}_k). \tag{7.2}$$

In other words, a matrix transformation preserves linear combinations. For this reason matrix transformations are examples of a larger set of transformation that are called *linear* transformations. We will discuss general linear transformations in a later section.

There is one other important property of a matrix transformation for us to consider. The functions we encountered in earlier mathematics courses, e.g., $f(x) = 2x + 1$, could send the input 0 to any output. However, as a consequence of the definition, any matrix transformation T maps the zero vector to the zero vector because

$$T(\mathbf{0}) = A\mathbf{0} = \mathbf{0}\,.$$

Note that the two vectors $\mathbf{0}$ in the last equation may not be the same vector – if $T : \mathbb{R}^n \to \mathbb{R}^m$, then the first $\mathbf{0}$ is in $\mathbb{R}^n$ and the second in $\mathbb{R}^m$. It should be clear from the context which vector $\mathbf{0}$ is meant.

Onto and One-to-One Transformations

The problems we have been asking about solutions to systems of linear equations can be rephrased in terms of matrix transformations. The question about whether a system $A\mathbf{x} = \mathbf{b}$ is consistent for any vector $\mathbf{b}$ is also a question about the existence of a vector $\mathbf{x}$ so that $T(\mathbf{x}) = \mathbf{b}$, where T is the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$.

Activity 7.2. Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ where A is

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

(a) Find $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$. If it is not possible to find one or both of the output vectors, indicate why.

(b) What are the domain and codomain of T? Why? (Recall that the domain is the space of all input vectors, while the codomain is the space in which the output vectors are contained.)

(c) Can you find a vector $\mathbf{x}$ for which $T(\mathbf{x}) = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}$? Can you find a vector $\mathbf{x}$ for which $T(\mathbf{x}) = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$?

(d) Which $\mathbf{b} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ are the image vectors for this transformation? Is the range of T equal to the codomain of T? Explain.

(e) The previous question can be rephrased as a matrix equation question. We are asking whether $A\mathbf{x} = \mathbf{b}$ is consistent for every $\mathbf{b}$. How is the answer to this question related to the pivots of A?

If T is a matrix transformation, Activity 7.2 illustrates that the range of a matrix transformation T may not equal its codomain. In other words, there may be vectors $\mathbf{b}$ in the codomain of T that are not the image of any vector in the domain of T. If it is the case for a matrix transformation T that there is always a vector $\mathbf{x}$ in the domain of T such that $T(\mathbf{x}) = \mathbf{b}$ for any vector $\mathbf{b}$ in the codomain of T, then T is given a special name.

Definition 7.2. A matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is **onto** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$.

So the matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each vector $\mathbf{b}$ in $\mathbb{R}^m$. Since the vectors $A\mathbf{x}$ are linear combinations of the columns of A, T is onto exactly when the span of the columns of A is all of $\mathbb{R}^m$. Activity 7.2 shows us that T is onto if every row of A contains a pivot.

Another question to ask about matrix transformations is how many vectors there can be that map onto a given output vector.

Activity 7.3. Let T be the matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$ where A is

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

(a) Find $T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$ and $T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$. If it is not possible to find one or both of the output vectors, indicate why.

(b) What are the domain and codomain of T? Why?

(c) Find $T\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right)$. Are there any other $\mathbf{x}$'s for which $T(\mathbf{x})$ is this same output vector? (Hint: Set up an equation to solve for such $\mathbf{x}$'s.)

(d) Assume more generally that for some vector $\mathbf{b}$, there is a vector $\mathbf{x}$ so that $T(\mathbf{x}) = \mathbf{b}$. Write this as a matrix equation to determine how many solutions this equation has. Explain. How is the answer to this question related to the pivots of A?

The uniqueness of a solution to $A\mathbf{x} = \mathbf{b}$ is the same as saying that the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ maps exactly one vector to $\mathbf{b}$. A matrix transformation T that has the property that every image vector is an image in exactly one way is also a special type of transformation.

Definition 7.3. A matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most* one $\mathbf{x}$ in $\mathbb{R}^n$.

So the matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one if the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution whenever $A\mathbf{x} = \mathbf{b}$ is consistent. Since the vectors $A\mathbf{x}$ are linear combinations of the columns of A, the unique solution requirement indicates that any output vector can be written in exactly one way as a linear combination of the columns of A. This implies that the columns of A are linearly independent. Activity 7.3 indicates that this happens when every column of A is a pivot column.

To summarize, if T is a matrix transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then T is onto if every row of A contains a pivot, and T is one-to-one if every column of A is a pivot column. It is important to note the difference: being one-to-one depends on the rows of A and being onto depends on the columns of A.

Having a matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ can tell us things about m and n. For example, when a matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is one-to-one, it means that there is a unique input vector for every output vector. Since a matrix transformation preserves the algebraic structure of $\mathbb{R}^n$, this implies that the collection of the images of the vectors in the domain of T form a copy of $\mathbb{R}^n$ inside of $\mathbb{R}^m$. If we think of T as a one-to-one matrix transformation with $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix, then every column of A will have to be a pivot column. It follows that if there is a one-to-one matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, we must have $m \geq n$. Similarly, if a matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is onto, then for each $\mathbf{b}$ in $\mathbb{R}^m$, if we select one vector in the domain of T whose image is $\mathbf{b}$, then the collection of these vectors in the domain of T is a copy of $\mathbb{R}^m$ inside of $\mathbb{R}^n$. So if there is an onto matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then $n \geq m$. As a consequence, the only way a matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is both one-to-one and onto is if $n = m$.

We conclude this section by adding new equivalent conditions to Theorems 5.3 and 6.7 from Sections 5 and 6.

Theorem 7.4. *Let A be an $m \times n$ matrix. The following statements are equivalent.*

(1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a solution for every vector $\mathbf{b}$ in $\mathbb{R}^m$.

(2) Every vector $\mathbf{b}$ in $\mathbb{R}^m$ can be written as a linear combination of the columns of A.

(3) The span of the columns of A is $\mathbb{R}^m$.

(4) The matrix A has a pivot position in each row.

(5) The matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.

Theorem 7.5. *Let A be an $m \times n$ matrix. The following statements are equivalent.*

(1) The matrix equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector $\mathbf{b}$ in the span of the columns of A.

(2) The matrix equation $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.

(3) The columns of A are linearly independent.

(4) The matrix A has a pivot position in each column.

(5) The matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.

We will continue to add to these theorems, which will eventually give us many different but equivalent perspectives to look at a linear algebra problem. Please keep these equivalent criteria in mind when considering the best possible approach to a problem.

Examples

What follows are worked examples that use the concepts from this section.

Example 7.6. Let $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 6 & 0 & 3 \\ 2 & -1 & -5 & -8 \end{bmatrix}$ and let $T(\mathbf{x}) = A\mathbf{x}$.

(a) Identify the domain of T. Explain your reasoning.

(b) Is T one-to-one. Explain.

(c) Is T onto? If yes, explain why. If no, describe the range of T as best you can, both algebraically and graphically.

Example Solution.

(a) Since A is a 3×4 matrix, A has four columns. Now $A\mathbf{x}$ is a linear combination of the columns of A with weights from $\mathbf{x}$, so $\mathbf{x}$ must have four entries to correspond to the columns of A. We conclude that the domain of T is $\mathbb{R}^4$.

(b) Technology shows that the reduced row echelon form of A is

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

Since A contains non-pivot columns, the homogeneous system $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions. So T is not one-to-one. In other words, if there is a column of A that is a non-pivot column, then A is not one-to-one.

(c) Since the reduced row echelon form of A has rows of zeros, there will be vectors $\mathbf{b}$ in $\mathbb{R}^3$ such that the reduced row echelon form of $[A\ \mathbf{b}]$ will have a row of the form $[0\ 0 0\ 0\ c]$ for some nonzero scalar c. This means that $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ will have no solution and T is not onto. In other words, if there is a row of A that does not contain a pivot, then T is not onto.

(d) To determine the vectors $\mathbf{b} = \begin{bmatrix} r \\ s \\ t \end{bmatrix}$ so that $T(\mathbf{x}) = A\mathbf{x} = \mathbf{b}$ is consistent, we row reduce the augmented matrix $[A \mid \mathbf{b}]$. Technology shows that an echelon form of $[A\ \mathbf{b}]$ is

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & r \\ 0 & 3 & 3 & 6 & s-3r \\ 0 & 0 & 0 & 0 & t-5r+s \end{array}\right].$$

Thus, the system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $-5r + s + t = 0$. We can then write the general output vector to this system as

$$\mathbf{b} = \begin{bmatrix} r \\ s \\ 5r-s \end{bmatrix} = r\begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} + s\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

with r and s any scalars. Since there are two free variables, the vectors $\mathbf{b}$ in $\mathbb{R}^3$ define a plane through the origin. Letting $r = 0$ and $s = 1$ and $r = 1$ and $s = 0$, we see that two points that lie on this plane are $(0, 1, -1)$ and $(1, 0, 5)$. So the range of T is the plane through the origin and the points $(0, 1, -1)$ and $(1, 0, 5)$.

Example 7.7. A matrix transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} cx \\ y \end{bmatrix}$$

is a contraction in the x direction if $0 < c < 1$ and a dilation in the x direction if $c > 1$.

(a) Find a matrix A such that $T(\mathbf{x}) = A\mathbf{x}$.

(b) Sketch the square S with vertices $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $\mathbf{u}_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Determine and sketch the image of S under T if $c = 2$.

Example Solution.

(a) Since

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ y \end{bmatrix},$$

the matrix $A = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$ has the property that $T(\mathbf{x}) = A\mathbf{x}$.

(b) We can determine the image of S under T by calculating what T does to the vertices of S. Notice that

$$T(\mathbf{u}_1) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$T(\mathbf{u}_2) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$
$$T(\mathbf{u}_3) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$
$$T(\mathbf{u}_4) = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Since T is a linear map, the image of S under T is the polygon with vertices $(0, 0)$, $(1, 0)$, $(2, 1)$, and $(0, 1)$ as shown in Figure 7.3. From Figure 7.3 we can see that T stretches the figure in the x direction only by a factor of 2.

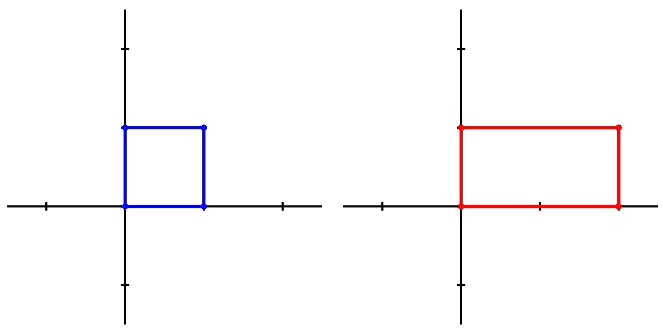


Figure 7.3: The input square S and the output $T(S)$.

Summary

In this section we determined how to represent any matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ as a matrix transformation, and what it means for a matrix transformation to be one-to-one and onto.

- A matrix transformation is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = A\mathbf{x}$ for some $m \times n$ matrix A.
- A matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ satisfies

$$T(a\mathbf{u} + b\mathbf{v}) = aT(\mathbf{u}) + bT(\mathbf{v})$$

  for any scalars a and b and any vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. The fact that T preserves linear combinations is why we say that T is a linear transformation.
- An $m \times n$ matrix A defines the matrix transformation T via

$$T(\mathbf{x}) = A\mathbf{x}.$$

The domain of this transformation is $\mathbb{R}^n$ because the matrix-vector product $A\mathbf{x}$ is only defined if $\mathbf{x}$ is an $n \times 1$ vector.

- If A is an $m \times n$ matrix, then the codomain of the matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is $\mathbb{R}^m$. This is because the matrix-vector product $A\mathbf{x}$ with $\mathbf{x}$ an $n \times 1$ vector is an $m \times 1$ vector. The range of T is the subset of the codomain of T consisting of all vectors of the form $T(\mathbf{x})$ for vectors $\mathbf{x}$ in the domain of T.

- A matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is **one-to-one** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at most* one $\mathbf{x}$ in $\mathbb{R}^n$. If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, then T is one-to-one if each column of A is a pivot column, or if the columns of A are linearly independent.

- A matrix transformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is **onto** if each $\mathbf{b}$ in $\mathbb{R}^m$ is the image of *at least one* $\mathbf{x}$ in $\mathbb{R}^n$. If T is a matrix transformation represented as $T(\mathbf{x}) = A\mathbf{x}$, then T is onto if each row of A contains a pivot position, or if the span of the columns of A is all of $\mathbb{R}^m$.

Exercises

(1) Given matrix $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & -3 \end{bmatrix}$, write the coordinate form of the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$. (Note: Writing a transformation in coordinate form refers to writing the transformation in terms of the entries of the input and output vectors.)

(2) Suppose the transformation T is defined by $T(\mathbf{x}) = A\mathbf{x}$ where

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

Determine if $\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is in the range of T. If so, find all $\mathbf{x}$'s which map to $\mathbf{b}$.

(3) Suppose T is a matrix transformation and

$$T(\mathbf{v}_1) = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ T(\mathbf{v}_2) = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Find $T(2\mathbf{v}_1 - 5\mathbf{v}_2)$.

(4) Given a matrix transformation defined as

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 - x_3 \\ -x_1 + 2x_2 + x_3 \\ 3x_2 - 4x_3 \end{bmatrix}$$

determine the matrix A for which $T(\mathbf{x}) = A\mathbf{x}$.

(5) Suppose a matrix transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ for some unknown A matrix satisfies

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ and } T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

Use the matrix transformation properties to determine $T(\mathbf{x})$ where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Use the expression for $T(\mathbf{x})$ to determine the matrix A.

(6) For each of the following matrices, determine if the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto and if T is one-to-one.

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

(b) $A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 2 & 4 \end{bmatrix}$

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

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

(7) Come up with an example of a one-to-one transformation from $\mathbb{R}^3$ to $\mathbb{R}^4$, if possible. If not, explain why not.

(8) Come up with an example of an onto transformation from $\mathbb{R}^3$ to $\mathbb{R}^4$, if possible. If not, explain why not.

(9) Come up with an example of a one-to-one but not onto transformation from $\mathbb{R}^4$ to $\mathbb{R}^4$, if possible. If not, explain why not.

(10) Two students are talking about when a matrix transformation is one-to-one.

Student 1: If we have a matrix transformation, then we need to check that $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ for which $A\mathbf{x} = \mathbf{b}$ has a solution, right?

Student 2: Well, that's the definition. Each $\mathbf{b}$ in the codomain has to be the image of at most one $\mathbf{x}$ in the domain. So when $\mathbf{b}$ is in the range, corresponding to $A\mathbf{x} = \mathbf{b}$ having a solution, then there is exactly one solution $\mathbf{x}$.

Student 1: But wouldn't it be enough to check that $A\mathbf{x} = \mathbf{0}$ has a unique solution? Doesn't that translate to the other $\mathbf{b}$ vectors? If there is a unique solution for one $\mathbf{b}_1$, then there can't be infinitely many solutions for another $\mathbf{b}_2$.

Student 2: I don't know. It feels to me as if changing the right hand side could change whether there is a unique solution, or infinitely many solutions, or no solution.

Which part of the above conversation do you agree with? Which parts need fixing?

(11) Show that if T is a matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ and L is a line in $\mathbb{R}^n$, then $T(L)$, the image of L, is a line or a single vector. (Note that a line in $\mathbb{R}^n$ is the set of all vectors of the form $\mathbf{v} + c\mathbf{w}$ where c is a scalar, and $\mathbf{v}, \mathbf{w}$ are two fixed vectors in $\mathbb{R}^n$.)

(12) Label each of the following statements as True or False. Provide justification for your response.

(a) **True/False** The range of a transformation is the same as the codomain of the transformation.

(b) **True/False** The codomain of a transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is the span of the columns of A.

(c) **True/False** A one-to-one transformation is a transformation where each input has a unique output.

(d) **True/False** A one-to-one transformation is a transformation where each output can only come from a unique input.

(e) **True/False** If a matrix transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ is one-to-one, then it is also onto.

(f) **True/False** A matrix transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$ cannot be onto.

(g) **True/False** A matrix transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ cannot be onto.

(h) **True/False** A matrix transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$ cannot be one-to-one.

(i) **True/False** If the columns of a matrix A are linearly independent, then the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is onto.

(j) **True/False** If the columns of a matrix A are linearly independent, then the transformation T defined by $T(\mathbf{x}) = A\mathbf{x}$ is one-to-one.

(k) **True/False** If A is an $m \times n$ matrix with n pivots, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.

(l) **True/False** If A is an $m \times n$ matrix with n pivots, then the transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.

(m) **True/False** If $\mathbf{u}$ is in the range of a matrix transformation T, then there is an $\mathbf{x}$ in the domain of T such that $T(\mathbf{x}) = \mathbf{u}$.

(n) **True/False** If T is a one-to-one matrix transformation, then $T(\mathbf{x}) = \mathbf{0}$ has a non-trivial solution.

(o) **True/False** If the transformations $T_1 : \mathbb{R}^m \to \mathbb{R}^n$ and $T_2 : \mathbb{R}^n \to \mathbb{R}^p$ are onto, then the transformation $T_2 \circ T_1$ defined by $T_2 \circ T_1(\mathbf{x}) = T_2(T_1(\mathbf{x}))$ is also onto.

(p) **True/False** If the transformations $T_1 : \mathbb{R}^m \to \mathbb{R}^n$ and $T_2 : \mathbb{R}^n \to \mathbb{R}^p$ are one-to-one, then the transformation $T_2 \circ T_1$ defined by $T_2 \circ T_1(\mathbf{x}) = T_2(T_1(\mathbf{x}))$ is also one-to-one.

Project: The Geometry of Matrix Transformations

In this section we will consider certain types of matrix transformations and analyze their geometry. Much more would be needed for real computer graphics, but the essential ideas are contained in our examples. A GeoGebra applet is available at https://www.geogebra.org/m/rh4bzxee for you to use to visualize the transformations in this project.

Project Activity 7.1. We begin with transformations that produce the rotated dancing image in Figure 7.1. Let R be the matrix transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by

$$R\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = \left[\begin{array}{cc} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right].$$

These matrices are the rotation matrices.

(a) Suppose $\theta = \frac{\pi}{2}$. Then

$$R\left(\left[\begin{array}{c} x \\ y \end{array}\right]\right) = \left[\begin{array}{cc} 0 & -1 \\ 1 & 0 \end{array}\right]\left[\begin{array}{c} x \\ y \end{array}\right].$$

i. Find the images of $\mathbf{u}_1 = \left[\begin{array}{c} 1 \\ 0 \end{array}\right]$, $\mathbf{u}_2 = \left[\begin{array}{c} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{array}\right]$, and $\mathbf{u}_3 = \left[\begin{array}{c} 0 \\ 1 \end{array}\right]$ under R.

ii. Plot the points determined by the vectors from part i. The matrix transformation R performs a rotation. Based on this small amount of data, what would you say the angle of rotation is for this transformation R?

(b) Now let R be the general matrix transformation defined by the matrix

$$\left[\begin{array}{cc} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{array}\right].$$

Follow the steps indicated to show that R performs a counterclockwise rotation of an angle θ around the origin. Let P be the point defined by the vector $\left[\begin{array}{c} x \\ y \end{array}\right] = \left[\begin{array}{c} \cos(\alpha) \\ \sin(\alpha) \end{array}\right]$ and Q the point defined by the vector $\left[\begin{array}{c} w \\ z \end{array}\right] = \left[\begin{array}{c} \cos(\alpha+\theta) \\ \sin(\alpha+\theta) \end{array}\right]$ as illustrated in Figure 7.4.

i. Use the angle sum trigonometric identities

$$\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B)$$
$$\sin(A+B) = \cos(A)\sin(B) + \cos(B)\sin(A)$$

to show that

$$w = \cos(\theta)x - \sin(\theta)y$$
$$z = \sin(\theta)x + \cos(\theta)y.$$

ii. Now explain why the counterclockwise rotation around the origin by an angle θ can be represented by left multiplication by the matrix

$$\left[\begin{array}{cc} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{array}\right].$$

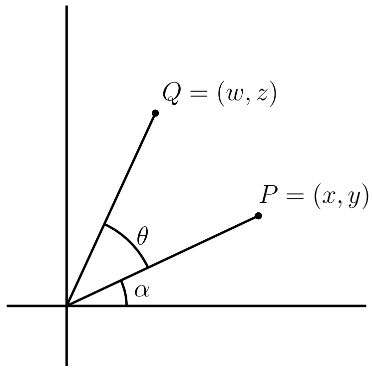


Figure 7.4: A rotation in the plane.

Project Activity 7.1 presented the rotation matrices. Other matrices have different effects.

Project Activity 7.2. Different matrix transformations

(a) Let S be the matrix transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by

$$S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Determine the entries of the output vector $S\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ and explain the action of the transformation S on the dancing figure as illustrated in Figure 7.5. (The transformation S is called a *shear* in the x direction.)

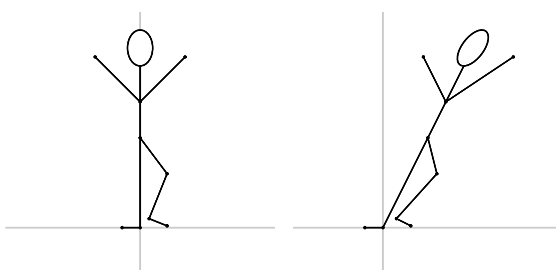


Figure 7.5: A dancing figure and a sheared dancing figure.

(b) Let C be the matrix transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by

$$C\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0.65 & 0 \\ 0 & 0.65 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Determine the entries of the output vector $C\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$ and explain the action of the transformation C on the dancing figure as illustrated in Figure 7.6. (The transformation C is called a *contraction*.) How would your response change if each 0.65 was changed to 2 in the matrix C?

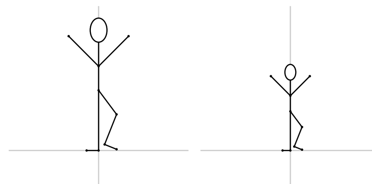


Figure 7.6: A dancing figure and a contracted dancing figure.

So far we have seen specific matrix transformations perform a rotations, shears, and contractions. We can combine these, and other, matrix transformations by composition to change figures in different ways, and to created animations of geometric figures. (As we will see later, combining transformations needs to be done carefully in order to obtain the result we want. For example, if we want to first rotate then translate, in what order should the matrices be applied?)