Linear Algebra

A work in progress

Linear Algebra

A work in progress

Ryan Higginbottom
Washington and Jefferson College

April 20, 2024

[1] creativecommons.org/licenses/by-sa/4.0

Contents

Chapter 1

Solving Systems of Linear Equations

Linear Algebra is a remarkable area of mathematics. It is wide, deep, and has a huge variety of applications. One of the greatest parts of the subject is that it takes very little in terms of prerequisites to reap great rewards!

1.1 Systems of Linear Equations

We will introduce the first set of big ideas in this book with an example.

Example 1.1.1 Alex reaches into his pocket and pulls out a handful of coins. He tells us that he's holding 90 cents in his hand, consisting of only nickels (worth five cents each) and quarters (worth 25 cents each). How many coins of each type is Alex holding?

We can model this information with a linear equation. Let x be the number of nickels in Alex's hand, and let y be the number of quarters. The equation that captures the information Alex shared is

$$5x + 25y = 90. \tag{1.1}$$

In this case, we don't have enough information to answer Alex's question. It could be that Alex is holding three nickels and three quarters $(15 + 75 = 90)$ or that he is holding 13 nickels and one quarter $(65 + 25 = 90)$. There are quite a few solutions to equation (1.1). (Note that in this example, since it would not make sense to have only part of a coin, we need our values of x and y to be non-negative integers.)

Now imagine that Alex gives us additional information by telling us that he is holding exactly ten coins. We can put this information into a second linear equation, and we now have what is called a **system of equations**. We need to know what values of x and y satisfy the following equations simultaneously:

$$\begin{aligned} 5x + 25y &= 90 \\ x + y &= 10. \end{aligned}$$

A small amount of work shows us that the values $x = 8$ and $y = 2$ satisfy these equations simultaneously. (Though you may have been able to produce those numbers yourself, don't worry right now about where those numbers came from. We'll get there soon.) This means that Alex is holding eight nickels and two quarters. □

The first major task of Linear Algebra is to learn how to handle systems of linear equations like the one given in Example 1.1.1. We will learn a method to analyze such systems and determine their solutions (if they have solutions). We need a number of definitions as we get started.

Note 1.1.2 We will use the symbol $\mathbb{R}$ to denote the set of all real numbers. At the beginning of our work, we will be using $\mathbb{R}$ as our set of numbers for almost everything. However, when we get to Section 2.1, we will move away from $\mathbb{R}$ to a more general description of a set of numbers that "works" for solving linear equations.

Definition 1.1.3 A **linear equation** in the variables $x_1, \dots, x_n$ is one which can be written in the form

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

where the a_i and b are constants. The numbers $a_1, \dots, a_n, b$ all come from $\mathbb{R}$. In the special situation where $b = 0$, this is called a **homogeneous linear equation**. (Note that when there are only two variables in view we may use x and y instead of x_1 and x_2; similarly, when there are only three variables present, we may use x, y, and z.) $\Diamond$

The word "linear" in the phrase "linear equation" should make us think of a single power of a variable. In a linear equation, therefore, we will have no terms involving x_1^4, $\sin(x_2)$, x_2x_3, $\sqrt{x_4}$, or anything other than a single power of a variable.

Example 1.1.4 The following are linear equations:

$$4x + 2y = 98, \qquad -2x_1 - 6x_2 + 5x_3 = 0, \qquad 14x_1 - 15x_3 = -7.$$

These are all linear equations because every appearance of a variable term in these equations contains only a single variable raised to the first power.

The following are not linear equations:

$$4x^2 - y = 9, \qquad \ln(x_1) + 3x_2 = 2, \qquad x_1x_2 - \tan(x_3) = 0.$$

These are not linear equations because each equation has at least one variable term with something other than a single variable to a single power. $\square$

Definition 1.1.5 A **system of linear equations** is a collection of one or more equations involving the same variables. (We sometimes shorten this and refer to a **linear system**.) When all of these linear equations are homogeneous linear equations, this may be called a **homogeneous linear system**. $\Diamond$

Example 1.1.6 Here is a system of linear equations involving the variables x_1, x_2, x_3, and x_4:

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

Note that not all variables need to be present in each equation. When a variable is missing, we consider that variable to have a zero coefficient in that equation. It may be convenient (and preferred) to align the terms with the same variables vertically, but when an equation lacks one of the variables this creates a blank horizontal space. $\square$

Definition 1.1.7 A **solution** to a system of linear equations in n variables is a list of n numbers, $(c_1, \dots, c_n)$, such that when the corresponding variables are assigned these numeric values (plug c_1 in everywhere for x_1, plug c_2 in everywhere for x_2, and so on), all the equations are true statements. The set of all solutions is called the **solution set** of the linear system. Two linear systems are said to be **equivalent** if their solution sets are equal. $\diamond$

The language of *solution sets* and *equivalent linear systems* may seem unnecessarily complex. However, the method (see Algorithm 1.3.9) for solving a linear system is much easier to describe with these terms firmly in hand.

Example 1.1.8 Consider the following linear system:

$$\begin{aligned} 2x + y &= -1 \\ x - 3y &= 17. \end{aligned}$$

A solution to this system is $(2, -5)$. (We use this notation to mean $x = 2$ and $y = -5$.) We verify this claim by plugging these numbers in for the variables and checking that both equations turn out to be true. (In fact, this is the *only* solution to this system.) $\square$

The fact that we can write down a system of linear equations does not mean that it has a solution. Many linear systems have no solution at all. Others have one or an infinite number of solutions. This can be illustrated by some examples of lines graphed in the familiar setting of $\mathbb{R}^2$.

Example 1.1.9 We first consider the linear system consisting of two simple equations where the solution to each equation is a line in $\mathbb{R}^2$:

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

Readers will likely realize that the solution to this system happens when the graphs of these lines intersect. We see from the graph below that the intersection occurs at the point $(2, 1)$. Since this is the only intersection point, there is only one solution to this system.

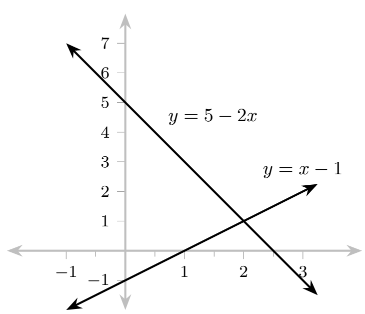

Figure 1.1.10

We now consider a second linear system:

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

The graphs of these two lines appear below, but in this case the lines do not intersect at all because they are parallel. (The slope of both lines is -2.)

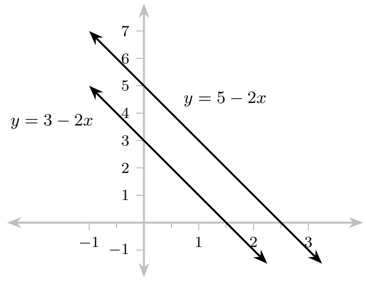

Figure 1.1.11

Here is a third linear system in two variables:

$$\begin{aligned} 2x + y &= 5 \\ 6x + 3y &= 15. \end{aligned}$$

Each of these equations has a solution set whose graph is a line. In this particular case, we obtain *the same line* for both equations, which means that the linear system has an infinite number of solutions. Each one of the infinite solutions to the first equation is a solution to the second equation, and vice versa.

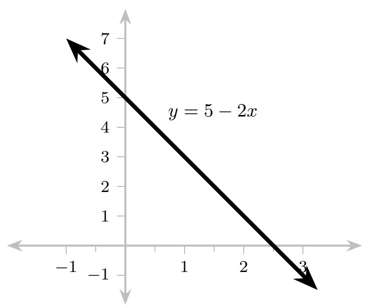

Figure 1.1.12

□

What we saw in Example 1.1.9 is no accident. This simple collection of examples in $\mathbb{R}^2$ exposes our need for the following definitions.

Definition 1.1.13 A system of linear equations is **consistent** if it has at least one solution. A system is **inconsistent** if it has no solutions. When a linear system has only one solution we say that the solution is **unique**. ◊

This definition and the principle illustrated in Example 1.1.9 mean that whenever we encounter a system of linear equations we have two questions to ask. *Is this system consistent? If the system is consistent, is the solution unique?*

Reading Questions

1. This question is designed to help you understand linear equations.

(a) Write down an example of a linear equation involving the variables x, y, and z.

(b) Write down an example of an equation involving the variables x, y, and z which is *not* linear.

2. Find the solution set to the following linear system:

$$\begin{aligned} 3x + 7y &= 9 \\ -2x + 6y &= -2. \end{aligned}$$

Exercises

1. Carlos has ten coins totalling \$1.10. Each coin is either a nickel, a dime, or a quarter. He has twice as many nickels as he has quarters. How many of each coin does Carlos have?

Model this problem with a linear system. (You do not need to find the solution set of this linear system.)

2. Suppose that $f(x) = ax^2 + bx + c$ is a function whose graph passes through the points $(-1, 6)$, $(1, 0)$, and $(3, -2)$. What are the values of a, b, and c?

Model this problem with a linear system. (You do not need to find the solution set of this linear system.)

3. For each part, determine whether the equations form a linear system.

(a)

$$\begin{aligned} 3y - 4z + x &= \sqrt{2} \\ -z + 2w + 7x &= 14 \end{aligned}$$

(b)

$$\begin{aligned} 2x - 10y + \frac{3}{z} &= 16 \\ 7x + 12z &= -1 \end{aligned}$$

(c)

$$\begin{aligned} x - e^4 z &= 9 \\ \ln(8)x + 13y &= -3 \end{aligned}$$

4. Determine whether or not each given list is a solution to this linear system.

$$\begin{aligned} x_1 - 2x_2 + 8x_3 &= -5 \\ -x_1 + 3x_2 - 10x_3 &= 6 \\ 2x_1 - 3x_2 + 14x_3 &= -9. \end{aligned}$$

(a) $(-7, 3, 1)$

(b) $(-5, 0, 0)$

(c) $(-4, 4, 1)$

(d) $(5, -3, -2)$

5. Determine whether or not each given list is a solution to this linear system.

$$\begin{aligned} x - y + 3z &= -6 \\ -3x + 4y - 10z &= 22 \\ -2x + 4y - 8z &= 21. \end{aligned}$$

(a) $(1, 1, -2)$

(b) $(-2, 3, 4)$

(c) $(0, 3, -1)$

6. For the following linear systems, graph the solution set of each equation. Then graph the solution set of the linear system.

(a)

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

(b)

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

(c)

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

(d)

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

(e)

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

Writing Exercises

7. For each of the following, write an example of a 2×2 linear system with the given property. (The system must be one you've not yet seen!) Explain why your example has the property.

(a) The system has no solutions.

(b) The system has exactly one solution.

(c) The system has infinitely many solutions.

8. Consider an $m \times n$ linear system and suppose that $(c_1, \ldots, c_n)$ and $(d_1, \ldots, d_n)$ are both solutions to this system. Under what conditions is $(c_1 + d_1, \ldots, c_n + d_n)$ also a solution to the same system? Explain.

9. Consider an $m \times n$ linear system and suppose that $(c_1, \ldots, c_n)$ and $(d_1, \ldots, d_n)$ are both solutions to this system. Let t be a real number. Prove that $(tc_1 + (1-t)d_1, \ldots, tc_n + (1-t)d_n)$ is also a solution to this same linear system.

1.2 Matrices

In this section we will introduce matrices, one of the main computational tools in linear algebra. We will learn how to translate the information about a linear system to a matrix and then manipulate that matrix to solve the original system.

Some of the matrix manipulations later in this section may seem to come from nowhere. We intend this first example to motivate these upcoming operations.

Example 1.2.1 Consider the following system of two linear equations in two variables:

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

We first multiply both sides of the top equation by $\frac{1}{3}$ in order to get "easier" coefficients on our variables. After taking this action, we have the following equivalent linear system:

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

We can now use the first equation to eliminate one of the variables in the second equation. If we add -2 times the first equation to the second equation, we'll only have the y variable left. The second equation is transformed, resulting in this equivalent linear system:

$$x + y = -1 \tag{1.2}$$
$$-3y = 9. \tag{1.3}$$

We can now multiply both sides of Equation (1.3) by $-\frac{1}{3}$ to find that $y = -3$. Plugging this value in for y in Equation (1.2) quickly gives us $x = 2$. We have solved the linear system at this point, and we have determined that the only solution to the system is $(2, -3)$. (Readers should check that this is in fact a solution by plugging these values into the original linear equations.) □

Example 1.2.1 is much longer than it needs to be, and at this point some readers may wonder what they've gotten themselves into—the mathematics so far (such as it is) is far from advanced. The real point of this example is to make explicit the operations used to solve a simple linear system. Once we have better notation and terminology, solving linear systems will be much faster (especially with the aid of technology). This is where matrices come in so handy.

Definition 1.2.2 A **matrix** is a rectangular array of numbers. If m and n are natural numbers, then an $m \times n$ matrix is one with m rows and n columns. The entries in a matrix are referred to by their row and column numbers, so entry (i, j) is the number in the ith row and jth column. (Row numbers increase from the top of the matrix down, and column numbers increase from the left of the matrix to the right.)

The **main diagonal** of a matrix refers to those entries on the (i, i)-diagonal of the matrix—starting at the upper left and going down to the right. In other words, an entry is on the main diagonal if and only if its row and column numbers are the same.

A **submatrix** of a matrix refers to the matrix that remains after removing one or more rows and/or columns from a matrix. ◊

Every system of linear equations generates two important matrices—the **coefficient matrix** and the **augmented matrix**.

Definition 1.2.3 Given a system of m linear equations in n variables (hereafter, we will call this an $m \times n$ linear system),

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \quad & \quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

the **coefficient matrix** of the system is

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

and the **augmented matrix** of the system is

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right].$$

We form the augmented matrix by "augmenting" the coefficient matrix with the column of constants. ◊

Note that the number of equations in the linear system corresponds to the number of rows of both the coefficient and augmented matrices, and the number of variables in the linear system corresponds to the number of columns in the coefficient matrix. (The augmented matrix of a linear system has one more column than the number of variables.)

Example 1.2.4 For the linear system

$$\begin{aligned} 2x_1 - 7x_2 + x_3 &= -8 \\ -x_1 + 4x_3 &= -2, \end{aligned}$$

the coefficient and augmented matrices are, respectively,

$$\begin{bmatrix} 2 & -7 & 1 \\ -1 & 0 & 4 \end{bmatrix} \qquad \text{and} \qquad \left[\begin{array}{ccc|c} 2 & -7 & 1 & -8 \\ -1 & 0 & 4 & -2 \end{array}\right].$$

Note the 0 in position $(2, 2)$ as it corresponds to the absence of an x_2 term in the second equation of the linear system. □

We must get comfortable switching between linear systems and their associated matrices. In particular, we need to understand why specific forms of matrices are especially useful when solving linear systems.

Example 1.2.5 Consider the following matrix as the augmented matrix for a linear system:

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

In one sense, this is the best possible augmented matrix we could have for a

3×3 linear system, because the solutions to the system are obvious. Translating this matrix back to equation form gives us $x = 7$, $y = -2$, and $z = \frac{2}{3}$. □

Almost no linear systems will come to us with an augmented matrix as simple as the one in Example 1.2.5. (This is mostly because such a linear system is, well, boring. It takes no work to solve a system like this.) Our goal, however, is to take any given linear system and manipulate its augmented matrix to be *as close to* this sort of matrix as we can get.

As we work with augmented matrices, we are restricted in the arithmetic we perform on them because, most of all, we want to preserve the solution sets of the corresponding linear systems. In our next definition, we describe the three "legal" ways we have to manipulate a matrix in this fashion.

Definition 1.2.6 The following operations on a matrix are called **elementary row operations**.

1. Add a multiple of one row to another row, replacing that second row with the result. (We will call this the **replace** row operation.)
2. Multiply every entry in a row by a nonzero constant. (We will call this the **scale** row operation.)
3. Switch the location of any two rows in the matrix. (We will call this the **switch** row operation.)

Two matrices are called **row equivalent** if there is a sequence of elementary row operations that transforms one matrix into the other. ◊

When defining "row equivalent" in the previous definition, careful readers will note one implied fact that must be checked. The word "equivalent" brings to mind an *equivalence relation*, which includes the property of the relation being symmetric. Therefore, the definition implies that all of the elementary row operations are "reversible"—that is, for each elementary row operation, there is an elementary row operation that reverses the change that was just made. This is something to prove! (See Exercise 1.2.9.)

Theorem 1.2.7 *Suppose that A and B are augmented matrices corresponding to systems of linear equations. Then if A and B are row equivalent, the linear systems to which they correspond are also equivalent.*

Proof. We will prove this statement directly. Because elementary row operations only involve one or two rows of a matrix at a time, it is sufficient to prove this theorem for systems of two linear equations.
We suppose that we have the following $2 \times n$ linear system:

$$\begin{aligned} a_{11}x_1 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots a_{2n}x_n &= b_2, \end{aligned}$$

which produces the following augmented matrix:

$$A = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \end{array}\right].$$

Further, we suppose that $(c_1, \ldots, c_n)$ is a solution to the linear system. If we apply the *switch* row operation to A, this corresponds to writing the second equation above the first in the linear system. It is immediate that $(c_1, \ldots, c_n)$ is still a solution to this system.
If we apply the *scale* row operation to A, multiplying row 1 (without loss of

generality) by a nonzero constant d, then we have the matrix

$$B = \left[\begin{array}{ccc|c} da_{11} & \cdots & da_{1n} & db_1 \\ a_{21} & \cdots & a_{2n} & b_2 \end{array}\right].$$

We must show that $(c_1, \ldots, c_n)$ is a solution to the corresponding linear system. It is obvious that $(c_1, \ldots, c_n)$ satisfies the second equation in this new linear system since that equation is unchanged. If $(c_1, \ldots, c_n)$ satisfied the first equation of the original system, then

$$a_{11}c_1 + \cdots + a_{1n}c_n = b_1.$$

We can now show that $(c_1, \ldots, c_n)$ satisfies the first equation of the second system by substitution:

$$\begin{aligned} da_{11}c_1 + \cdots + da_{1n}c_n &= d(a_{11}c_1 + \cdots + a_{1n}c_n) \\ &= d(b_1). \end{aligned}$$

We must now show that the *replace* row operation preserves solutions. We let k be a nonzero constant and we replace (without loss of generality) the second row of A with the old second row plus k times the first row. Here is the resulting matrix C:

$$C = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ (ka_{11} + a_{21}) & \cdots & (ka_{1n} + a_{2n}) & (kb_1 + b_2) \end{array}\right].$$

In the linear system which corresponds to C, the first equation is unchanged from the first equation in the linear system that corresponds to A, so we only need to be concerned with the second equation. By virtue of the fact that $(c_1, \ldots, c_n)$ satisfied both equations of the first system, we know that

$$a_{11}c_1 + \cdots + a_{1n}c_n = b_1$$

and

$$a_{21}c_1 + \cdots + a_{2n}c_n = b_2.$$

We can now use this in the following calculation:

$$\begin{aligned} &(ka_{11} + a_{21})c_1 + \cdots + (ka_{1n} + a_{2n})c_n \\ &\qquad = (ka_{11}c_1 + \cdots + ka_{1n}c_n) + (a_{21}c_1 + \cdots + a_{2n}c_n) \\ &\qquad = k(a_{11}c_1 + \cdots + a_{1n}c_n) + (a_{21}c_1 + \cdots + a_{2n}c_n) \\ &\qquad = kb_1 + b_2. \end{aligned}$$

There is technically one more collection of facts to verify before this proof is complete. We have shown that, for all three elementary row operations, if $(c_1, \ldots, c_n)$ is a solution to the linear system corresponding to A then it will also be a solution to the linear system corresponding to B (where B is the result of applying one elementary row operation to A). But "equivalent" linear systems means that the solution sets are equal *as sets*. This means that we must assume we have a solution for the linear system corresponding to the matrix B and show that it is a solution for the linear system corresponding to A. We claim that this concern can be dispensed with by invoking another result in this section. Connecting the last dots of this proof is left to the reader in Exercise 1.2.12. ■

Roughly speaking, we want to use elementary row operations to transform the augmented matrix for a linear system into a matrix which has 1s along the main diagonal and 0s as the other entries in those columns. This is not always possible, and we will describe the situation more precisely below, but here is an example to illustrate the process.

Example 1.2.8 We start by considering the following matrix as the augmented matrix of a linear system:

$$\left[\begin{array}{ccc|c} 2 & 2 & -1 & 8 \\ -3 & -2 & 2 & -12 \\ 5 & 0 & 4 & 11 \end{array}\right].$$

We first *scale* the first row by $\frac{1}{2}$ to produce a 1 in the $(1,1)$ entry:

$$\left[\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 4 \\ -3 & -2 & 2 & -12 \\ 5 & 0 & 4 & 11 \end{array}\right].$$

We then *replace* the second row with the sum of the second row and three times the first row:

$$\left[\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 4 \\ 0 & 1 & \frac{1}{2} & 0 \\ 5 & 0 & 4 & 11 \end{array}\right].$$

We will again use the 1 in the $(1,1)$ entry to "eliminate" the 5 in the $(3,1)$ entry. We *replace* the third row with -5 times the first row plus the third:

$$\left[\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 4 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & -5 & \frac{13}{2} & -9 \end{array}\right].$$

Now that we have "cleared out" the entries under the $(1,1)$ entry, we can do the same for the 1 in the $(2,2)$ entry. (In future examples we may need to *scale* first to have a 1 here.) We *replace* the third row with 5 times the second row plus the third:

$$\left[\begin{array}{ccc|c} 1 & 1 & -\frac{1}{2} & 4 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 9 & -9 \end{array}\right].$$

We can now *scale* the third row by $\frac{1}{9}$ to produce 1s along the main diagonal:

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

We are halfway done, as we have 0s below the main diagonal. We now need to use elementary row operations to produce 0s above the main diagonal. We first *replace* the second row with $-\frac{1}{2}$ times the third plus the second:

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

We can again use the 1 in the $(3,3)$ position to produce a 0 in the $(1,3)$ position. We *replace* the first row with $\frac{1}{2}$ times the third row plus the first:

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

Finally, we use the 1 in the $(2,2)$ entry to produce a 0 in the $(1,2)$ entry. We *replace* the first row with -1 times the second row plus the first:

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

We now have the matrix in the form we wanted, because we can read off the solution: $x = 3$, $y = \frac{1}{2}$, and $z = -1$. □

We will now define and standardize this form of the augmented matrix that is so helpful in solving the related linear system. In what follows, by "nonzero row (column)" we mean a row (column) with at least one nonzero entry, and by the "leading entry" of a row we mean the row's leftmost nonzero entry.

Definition 1.2.9 A matrix A is in **row-echelon form** (REF) if all of the following hold:

- all rows of all zeros are collected at the bottom of the matrix;
- each leading entry of a row is in a column to the right of the column of the leading entry for the row above it; and
- all entries in a column below a leading entry are zeros.

If a matrix A is in row-echelon form and also satisfies the following two conditions, it is in **reduced row-echelon form** (RREF):

- each leading entry in a nonzero row is 1; and
- each leading 1 is the only nonzero entry in its column.

◊

Though that definition is a mouthful, it is useful. Here is an example showing some matrices that do and do not meet these criteria.

Example 1.2.10 The following two matrices are in REF but not RREF:

$$\begin{bmatrix} 2 & -1 & 3 & 0 \\ 0 & 5 & -3 & 7 \\ 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} -4 & 9 & 6 \\ 0 & 12 & -10 \end{bmatrix}.$$

The following two matrices are in RREF:

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

Neither of the following matrices are in REF or RREF:

$$\begin{bmatrix} 0 & 2 & 5 & -8 \\ 0 & -1 & 7 & 7 \\ 4 & 2 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 2 & 3 & -2 & 5 \\ 0 & -1 & 3 & 6 \\ -4 & -3 & 0 & 12 \end{bmatrix}.$$

□

Definition 1.2.11 When a matrix A is in row-echelon form, then the leading entry in each row is called a **pivot**. The location of this entry within the matrix is called a **pivot position**. Finally, any column containing a pivot is called a

pivot column. ◊

One of the reasons we have Definition 1.2.9 is because (we will prove this below) *every* matrix can be put into RREF. What follows here is a description of the algorithm used to transform every matrix into RREF. This process is usually called "Gaussian elimination" or "Gauss-Jordan elimination."

Algorithm 1.2.12 The Row Reduction Algorithm. *This row reduction algorithm consists of five steps. The first four (collectively) produce a matrix in row-echelon form; after the fifth step the matrix will be in reduced row-echelon form.*

1. *Start with the leftmost nonzero column. This will be a pivot column with the pivot position at the top.*
2. *Select a nonzero entry in this column and use the* switch *operation to move it to the top of the column (if necessary).*
3. *Use elementary row operations to create zeros below this pivot.*
4. *Ignore the row with the pivot just considered. Apply steps 1–3 to the submatrix that remains. Repeat this process until all nonzero rows have been handled.*
5. *Select the rightmost pivot and use the* scale *operation to make it a 1. Now use elementary row operations to create 0s above this pivot. Move upward and to the left, repeating this process for all remaining pivots.*

Example 1.2.13 We consider the following matrix A and put it into reduced row-echelon form:

$$A = \left[\begin{array}{cccc|c} -2 & -2 & 6 & 14 & 4 \\ 2 & 3 & -4 & -4 & -1 \\ -3 & -5 & 4 & -2 & -2 \end{array}\right].$$

The first column is nonzero, and we will leave the current top row in its place. Though it is not *necessary* to create a 1 in the pivot position at this point in the process, it is often useful to do so. (We are not violating Algorithm 1.2.12 by producing a leading 1 this early in the process, but this is a step of row reduction that many, including this author, prefer to perform now to make future steps less painful.) We multiply the first row by $-\frac{1}{2}$ to achieve this:

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & -7 & -2 \\ 2 & 3 & -4 & -4 & -1 \\ -3 & -5 & 4 & -2 & -2 \end{array}\right].$$

We now use the 1 in the $(1,1)$ position to create zeros in the column below it. We add -2 times row 1 to row 2 and we add 3 times row 1 to row 3. Since there is no "interaction" between these operations, we will perform them at the same time, though the reader should certainly take one operation at a time if this combination raises one's blood pressure:

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & -7 & -2 \\ 0 & 1 & 2 & 10 & 3 \\ 0 & -2 & -5 & -23 & -8 \end{array}\right].$$

According to the algorithm, we now ignore row 1 and repeat the process for the remaining matrix. There is already a 1 at the "top" of (this portion of) the second column, so we use that entry to create zeros below it. We add twice the second row to the third row:

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & -7 & -2 \\ 0 & 1 & 2 & 10 & 3 \\ 0 & 0 & -1 & -3 & -2 \end{array}\right].$$

The matrix is in row-echelon form now, so we proceed to step 5 of the algorithm. We multiply the third row by -1 to produce a 1 in the $(3,3)$ position:

$$\left[\begin{array}{cccc|c} 1 & 1 & -3 & -7 & -2 \\ 0 & 1 & 2 & 10 & 3 \\ 0 & 0 & 1 & 3 & 2 \end{array}\right].$$

We now use the 1 we just created in order to produce zeros in the column above it. We add -2 times the third row to the second, and we add 3 times the third row to the first:

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

The final step is to use the pivot in position $(2,2)$ in order to create a 0 above it. We add -1 times the second row to the first:

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 5 \\ 0 & 1 & 0 & 4 & -1 \\ 0 & 0 & 1 & 3 & 2 \end{array}\right]. \tag{1.4}$$

The process is now complete, so the matrix in (1.4) is the result of reducing A to RREF. □

We note that this example is different from Example 1.2.8 in an important way. The matrix in Example 1.2.13 corresponds to a system with three equations and four variables, while the matrix in Example 1.2.8 corresponds to a system with three equations and three variables. We only mention this to point out that the size of the original matrix puts some restrictions on the possibilities for its RREF, and the reader should be aware of this when completing the reading questions and the exercises at the end of this section.

Reading Questions

1. Write down the coefficient matrix and the augmented matrix for the following linear system:

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

2. Although the reduced row-echelon form of a matrix is unique (we will prove this soon), the row-echelon form of a matrix is not unique. For the following matrix A, write two distinct row-echelon forms:

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

3. Put the following matrix into reduced row-echelon form. Record each of your steps along the way (both the elementary row operations you used and the matrices you obtained as a result):

$$A = \left[\begin{array}{cc|c} 2 & -3 & 5 \\ -1 & 2 & -4 \\ 5 & -9 & 8 \end{array}\right].$$

Exercises

1. Write the augmented matrix that corresponds to the following linear system:

$$\begin{aligned} 2x_2 - 4x_1 + \sqrt{2}x_4 &= 8 \\ 3x_3 - 19x_2 &= -1 \\ 0.5x_4 - 3x_3 + x_1 &= 0. \end{aligned}$$

2. Write the linear system that corresponds to the following matrix, assuming this is the augmented matrix of the system:

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

3. For each of the following, describe the elementary row operation that was used to transform the matrix on the left into the matrix on the right. Be specific in your description of the operation. (You should not just answer "scale" or "switch", etc.)

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

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

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

4. For each of the following matrices, determine if the matrix is in row-echelon form, reduced row-echelon form, both, or neither.

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

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

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

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

5. Consider the following matrix A:

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

Find two distinct matrices B and C which are row equivalent to A and are in row echelon form. (There are many correct answers!)

6. Use Algorithm 1.2.12 to put each of these matrices into RREF.

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

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

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

(d) $\begin{bmatrix} -3 & -8 & -1 \\ -4 & -3 & -3 \\ 2 & 13 & -1 \end{bmatrix}$

7. For each of the following matrix sizes, list all of the possible RREFs for matrices of that size. Use the symbols ■ for a pivot, $*$ for an unspecified number, and 0 for a zero entry.

(a) 2×2

(b) 2×3

(c) 3×2

Writing Exercises

8.

(a) Suppose A is a 3×4 matrix. What is the maximum number of pivots in any RREF of A? Explain.

(b) Suppose B is a 6×4 matrix. What is the minimum number of rows of zeros in any RREF of B? Explain.

9. Prove that each of the elementary row operations is reversible. In other words, if an elementary row operation was used to transform a matrix A into a matrix B, prove that there is another elementary row operation (of the same type) which will transform B back into A.

10. Recall that an equivalence relation is reflexive, symmetric, and transitive. Now, fix the integers m and n and consider row equivalence as a relation on all $m \times n$ matrices. Prove that this is an equivalence relation. (Hint: another exercise in this section may be helpful in your argument.)

11. Prove or disprove: The following two matrices are row equivalent. (Hint: another exercise in this section may be helpful in your argument.)

$$A = \begin{bmatrix} 9 & -9 & -10 \\ -1 & -1 & 9 \\ -5 & -2 & -1 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & 3 & -2 \\ 7 & -2 & 1 \\ 7 & 6 & -8 \end{bmatrix}$$

12. Complete the proof of Theorem 1.2.7. This can be accomplished by proving the statement in the final paragraph of the given proof.

1.3 Results

In the previous section, we learned how to write the augmented matrix of a linear system and Algorithm 1.2.12 provided a process to reduce any matrix to RREF. In this section, we will learn how to use the RREF of a matrix to solve the corresponding linear system. We will also prove some important results related to these solutions.

We first introduce some terminology. These terms relate what we saw in Definition 1.2.11 back to the corresponding linear systems.

Definition 1.3.1 Suppose that A is the coefficient matrix corresponding to a system of linear equations and that A is in REF. Then a variable corresponding to a pivot column in A is called a **basic variable** (or **pivot variable**), and a variable corresponding to a non-pivot column in A is called a **free variable**. $\diamondsuit$

Note 1.3.2 Note that in this definition A is the *coefficient matrix* (not the augmented matrix) of the linear system. We are using the coefficient matrix because we are making a definition concerning variables, and the rightmost column in the augmented matrix does not correspond to a variable in the linear system.

The augmented matrix of an $m \times n$ linear system is of size $m \times (n+1)$. This puts some limitations on the different reduced row-echelon forms that we could see in this context. In the following two examples, we will consider specific reduced row-echelon forms and what they say about the linear systems to which they correspond.

Example 1.3.3 Consider the following as the augmented matrix corresponding to a system of linear equations:

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

This is a 3×4 matrix, so the original linear system has three equations and three variables. Hopefully the reader can see that this matrix is indeed in RREF. (Consult Definition 1.2.9 for a refresher.)

We will now write the equations which correspond to each row of the matrix, and we will pay special attention to the final row:

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

Usually we omit terms in linear equations when the coefficient is 0, but we are including those terms here to make a point. When the coefficient of a variable is 0, the entire linear term disappears since no value of the variable could make the product with the coefficient anything other than 0. This means that each of these equations can be written in a more simple form. In particular, the third equation can be written as $0 = 1$.

This may feel rather elementary, but there are no possible variable values to make $0 = 1$ a true statement. It is false *all of the time*. Since we are searching for values of the variables which satisfy all the equations simultaneously, and since one of the equations has no solution, the linear system has no solution. This is an *inconsistent* linear system.

We proved in Theorem 1.2.7 that row-equivalent matrices correspond to equivalent linear systems. Therefore, we can say that the original linear system for this example is inconsistent. □

What we saw in Example 1.3.3 we will be able to generalize (see Theorem 1.3.5) in our effort to categorize inconsistent linear systems. Before we do that, let's look at an augmented matrix which has a different RREF.

Example 1.3.4 Consider the following matrix as the augmented matrix of a linear system:

$$A = \left[\begin{array}{ccc|c} 1 & 0 & -3 & 5 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

(Again, the reader should verify that this matrix is in reduced row-echelon form.)

The final row of this matrix corresponds to the equation $0 = 0$. Since this equation is true all the time, for all values of the involved variables, we won't consider it any longer as it places no further restrictions on the solutions.

The first two rows of A correspond to the following two linear equations:

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

From Definition 1.3.1, we see that in this system x_1 and x_2 are basic variables and x_3 is a free variable. What does that mean for the solutions of this system? We call x_3 "free" because any element of $\mathbb{R}$ that we put into x_3 will produce a solution for this system. The variable x_3 is "free" to take on any value, and then the values of the basic variables x_1 and x_2 are determined by that value and the linear equations.

For example, in this system of equations, if $x_3 = 2$, then $x_1 = 11$ and $x_2 = 3$, giving $(11, 3, 2)$ as a solution to the linear system. If $x_3 = -1$, then $x_1 = 2$ and $x_2 = 9$, giving $(2, 9, -1)$ as a solution to the system. Since x_3 can take on any value in $\mathbb{R}$, and since we have a solution to the system for each value of x_3, this means we have a solution to the system for each element of $\mathbb{R}$. We conclude that there are an infinite number of solutions to this system.

The solutions to this linear system can be written in a number of ways. We will prefer the following form:

$$\begin{cases} x_1 = 5 + 3x_3 \\ x_2 = 7 - 2x_3 \\ x_3 \text{ is free.} \end{cases}$$

This is called a **parametric description** of the solution set. Sometimes solutions like this are written with the letter t or s in place of x_3 to better match the usage of the word "parameter" elsewhere. We will follow the convention of using the free variables as parameters in our solutions. □

In Example 1.3.4, we saw that having a free variable corresponded to having an infinite number of solutions. But we need to be careful about our conclusions, because there was also a free variable (x_3) in Example 1.3.3, and in that case there were *no solutions* to the system. A free variable only indicates an infinite number of solutions when the system is consistent.

The two previous examples, combined with Example 1.2.5, give us a sense of what solutions to linear systems can look like given certain characteristics of the augmented matrices. We can now state in theorem form what we observed to be true in these examples.

Theorem 1.3.5 *Suppose A is the augmented matrix of a linear system, and suppose that A is in reduced row-echelon form. Then the linear system is consistent if and only if there is no pivot in the final column of A.*

Proof. We note that the pivot columns do not change when a matrix goes from row-echelon form to reduced row-echelon form (see Algorithm 1.2.12), so we are not losing any generality with our assumption that A is in RREF.
This theorem is a biconditional statement, and we will prove one implication directly. We assume there is no pivot in the final column of A. Then when we consider the linear equations which correspond to the rows of A, we see that each of the basic variables can be written in terms of the free variables, if any free variables exist. If no free variables exist, then all basic variables have an assigned value and the system is consistent. In the case that there is at least one free variable, we can pick an element of $\mathbb{R}$—let's say, 0—to assign to each of the free variables. This produces a solution to the linear system, and our system is consistent.
We will prove the contrapositive of the other implication. If there is a pivot in the final column of A, then the corresponding linear equation reduces to $0 = 1$. This means that there is no solution to the linear system, so the system is inconsistent. ∎

Given that two major questions about the solutions to linear systems involve *consistency* and *uniqueness*, the next natural theorem to consider is related to this second concept.

Theorem 1.3.6 *Suppose that A is the augmented matrix corresponding to a consistent $m \times n$ linear system, and suppose that A is in reduced row-echelon form. Then the system has a unique solution if and only if there is a pivot in each of the first n columns of A.*

Proof. As with the proof of Theorem 1.3.5, we are not losing any generality by assuming that A is in RREF.
We first suppose that there is a pivot in each of the first n columns of A; this implies that $m \geq n$. We also recall that the linear system is assumed to be consistent, meaning that the first n rows of A have the following form:

$$\left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & b_1 \\ 0 & 1 & \cdots & 0 & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & b_n \end{array}\right].$$

(The matrix A may have rows of all zeros below the n rows here, but that will not affect our discussion.)
If the matrix A has the form we have just detailed, then the original linear system is equivalent to one with equations of the form $x_1 = b_1$, $x_2 = b_2$, …, $x_n = b_n$. That is, the system has a unique solution.
We will prove the contrapositive of the other implication. We suppose that there is at least one free variable (call it x_j) in the linear system. We recall (see Example 1.3.4) that when the matrix for a consistent system is in RREF, all solutions can be written by expressing basic variables in terms of any existing free variables. Therefore, this system has a solution in which all free variables are set equal to 0. Further, this system has a solution in which all free variables except x_j are set equal to 0 and $x_j = 1$. This may not change the value of any

of the basic variables, and there may not be any free variables aside from x_j, but these two solutions we have just described are not the same since x_j has a different value in each one. Therefore, this system has more than one solution, meaning that the system does not have a unique solution. ∎

Corollary 1.3.7 *If m and n are natural numbers and $m < n$, then an $m \times n$ linear system cannot have a unique solution.*

Proof. Suppose A is the augmented matrix corresponding to an $m \times n$ linear system with $m < n$. The RREF of A can have at most m pivots, so by Theorem 1.3.6 the system cannot have a unique solution. ∎

There are two natural definitions related to Corollary 1.3.7 which we now state.

Definition 1.3.8 An $m \times n$ linear system is called **underdetermined** if $m < n$. An $m \times n$ linear system is called **overdetermined** if $m > n$. ♢

The two previous theorems provide the last step to an algorithm for solving any linear system.

Algorithm 1.3.9 Algorithm for Solving Linear Systems. *Suppose that we have an $m \times n$ system of linear equations. Here are the steps to solve the system.*

1. *Form the augmented matrix of the linear system. We will call this matrix A.*

2. *Find the reduced row-echelon form of A. (If A is small, this can be done by hand; if A is not small, technology should be used to complete this step.)*

3. *Determine whether or not the system is consistent by observing the location of the pivots in the RREF of A. If there is a pivot in the rightmost column, the linear system is inconsistent and we need not proceed any further in the algorithm. If there is no pivot in the rightmost column, the system is consistent.*

4. *Determine whether or not the system has a unique solution. If there is a pivot in each of the first n columns, then the system has a unique solution which can be recorded. If there is not a pivot in each of the first n columns, then the system does not have a unique solution; in this case, a parametric description of the solution set can be recorded.*

The earlier examples in this section can be completed using this algorithm. We will include an additional example so the reader can practice using the algorithm once more.

Example 1.3.10 Consider the following linear system:

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

Does this system have a solution? If the system has a solution, write down the solution.

Solution. We follow Algorithm 1.3.9 and form the augmented matrix A for this system. When we row reduce this matrix, we find

$$A = \left[\begin{array}{ccc|c} -7 & -4 & 7 & -3 \\ -2 & 2 & -4 & 2 \\ 5 & 4 & 1 & 5 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1/9 \\ 0 & 1 & 0 & 4/3 \\ 0 & 0 & 1 & 2/9 \end{array}\right].$$

We will refer to the RREF of A as B. Since B has no pivot in the rightmost column, our linear system is consistent. Secondly, since B has a pivot in each of the first three columns, the solution to our linear system is unique. We record this solution as $x = -1/9$, $y = 4/3$, and $z = 2/9$. □

The rest of this section will be devoted to proving Theorem 1.3.14, which states that the RREF of a matrix is unique. (There are actually a couple of places so far in this section where we have been a bit sloppy and referred to *the* RREF of a matrix, but in each of these cases the uniqueness of the RREF was not essential to the argument.)

We will begin with a lemma. (Our approach to proving Theorem 1.3.14 has been heavily influenced by Kuttler's treatment[2].)

Lemma 1.3.11 *In a consistent $m \times n$ linear system, all solutions can be expressed by writing the basic variables as linear functions of the free variables (if they exist). Further, each basic variable x_i can be written as a linear function of only those free variables x_j with $j > i$.*

Proof. The first sentence in this lemma has essentially been proved in the discussion within Example 1.3.4. We will prove the second statement directly. We consider A as the augmented matrix of a consistent $m \times n$ linear system. Suppose A is row equivalent to B, where B is in RREF. Recall that part of the definition of RREF (Definition 1.2.9) is that pivots are the leftmost non-zero number in their row.
Consider the linear equation corresponding to row d of B; this equation will begin with a basic variable x_i and will possibly involve other variables x_j, with $j > i$, before the equals sign. However, all of these other variables x_j will be free variables, because any other basic variable x_k, with $k > i$, will correspond to a column in which that pivot is the only non-zero number. In other words, all entries b_{dk} along row d in B which correspond to pivot columns k, for $k > i$, will be zero. ■

The basic idea for the proof of Theorem 1.3.14 is to prove the result for homogeneous linear systems first and then to obtain the proof for general linear systems as an extension. We turn to our first result about homogeneous linear systems. (For a refresher on homogeneous systems, see Definition 1.1.5.)

Lemma 1.3.12 *If x_i is a basic variable of a homogeneous linear system, then any solution of the system with $x_j = 0$ for all free variables x_j with $j > i$, must have $x_i = 0$.*

Proof. We will prove this directly. Suppose that x_i is a basic variable of a homogeneous linear system and that in a solution of this system, $x_j = 0$ for all free variables x_j with $j > i$. From Lemma 1.3.11 we know that in the description of the solution to this system, x_j will be written as a linear function of the free variables with larger indices. But the nature of a homogeneous linear system demands that such a linear function will involve *only* free variables and no constants (the constants are all 0). Therefore, if $x_j = 0$ for all free variables x_j with $j > i$, we must have $x_i = 0$ as well. ■

We will now prove the uniqueness result for augmented matrices of homogeneous systems. (We should note here, perhaps, that while we introduced the notions of REF and RREF for augmented matrices, the row reduction algorithm can be applied to any matrix at all.)

Proposition 1.3.13 *Let A be the augmented matrix corresponding to a homogeneous linear system. Suppose that A is row equivalent to matrices B and C,*

[2] `math.libretexts.org/Bookshelves/Linear_Algebra/A_First_Course_in_Linear_Algebra_(Kuttler)/`

both of which are in reduced row-echelon form. Then $B = C$.

Proof. We proceed by contradiction and assume that $B \neq C$. Since B and C are row equivalent and both are in RREF, they must have the same pivot positions. (The reader is asked to prove this in Exercise 1.3.16.) Since $B \neq C$, these matrices must differ in some row, call it row k. Since B and C have the same pivot positions, we assume there is a pivot in column i of row k in both matrices. There must be some position j, with $j > i$, such that $b_{kj} \neq c_{kj}$. The variable x_j must not be a basic variable in the linear system, because if so, we would have $b_{kj} = c_{kj} = 0$. So x_j is a free variable.

Homogeneous linear systems are always consistent. (The reader is asked to prove this in Exercise 1.3.10.) There must exist a solution to the linear system where $x_j = 1$ and all other free variables take on the value of 0. In this solution, using the linear equations that correspond to the rows in B, we solve and obtain $x_i = -b_{kj}$. Using the linear equations that correspond to the rows in C, we find $x_i = -c_{kj}$. Since a solution is completely determined by the values of the free variables, this implies that $b_{kj} = c_{kj}$, which is a contradiction. ■

With this proposition in hand, we can state and prove our first large result.

Theorem 1.3.14 *Let A be an $m \times n$ matrix and let A be row equivalent to both B and C. If B and C are in reduced row-echelon form, then $B = C$.*

Proof. We first form the matrix A' by augmenting the matrix A with an additional column on the right consisting of all zeros. We similarly form the matrices B' and C' from B and C. We note that B' and C' are also in RREF and they are obtained from A' using the same row operations that reduced A to B and C.

We can consider A', B', and C' as augmented matrices corresponding to $m \times n$ homogeneous linear systems. By Proposition 1.3.13, since A' is row equivalent to both B' and C', where both B' and C' are in RREF, we must have $B' = C'$. By the construction of B' and C', this implies $B = C$. ■

Reading Questions

1. Consider the following linear system:

$$\begin{aligned} 4x_1 + 7x_2 + 17x_3 &= 23 \\ -3x_1 - 5x_2 - 12x_3 &= -17. \end{aligned}$$

Determine which of the variables are basic variables and which are free variables. Explain your answer.

2. Consider the following linear system:

$$\begin{aligned} x_1 - 2x_2 + 2x_3 &= -1 \\ -x_1 + 2x_2 - x_3 &= -1 \\ 3x_1 - 6x_2 + 7x_3 &= -5. \end{aligned}$$

Write a parametric description of the solution set of this system. (Follow Example 1.3.4.)

Exercises

1. In each of the following, suppose the augmented matrix for a linear system has been reduced to the following RREF. Write down the solution(s) to the system (if they exist).

(a) $\left[\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{array}\right]$

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

(c) $\left[\begin{array}{ccc|c} 1 & 0 & -3 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$

(d) $\left[\begin{array}{ccccc|c} 1 & 0 & 2 & -4 & 0 & 7 \\ 0 & 1 & 9 & -1 & 0 & -4 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$

2. Solve the following linear system.

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

3. Solve the following linear systems.

(a)

$$\begin{aligned} x - 2y + 3z &= 4 \\ 5x - 6y + 7z &= 5 \end{aligned}$$

(b)

$$\begin{aligned} x - 2y + 2z - w &= 1 \\ 3x - 7y - z + 5w &= 2 \end{aligned}$$

(c)

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

(d)

$$\begin{aligned} 4x - 2y &= -4 \\ 2x + 5y + 6z &= 2 \\ 3x - y + \tfrac{1}{2}z &= 0 \end{aligned}$$

4. Determine the values of a for which the following linear system has no solutions, exactly one solution, or infinitely many solutions. Explain your answers.

$$\begin{aligned} x + 2y + 2z &= -4 \\ 2x - 2y + 4z &= 7 \\ x + 2y - (a^2 - 5)z &= a + 1 \end{aligned}$$

5. Under what conditions is the following linear system consistent? Your answer should be an equation or equations that must be satisfied by the b_i. Explain your answer.

$$\begin{aligned} x + 2y &= b_1 \\ -2x - 5y + 3z &= b_2 \\ x + 4y - 6z &= b_3 \end{aligned}$$

6. Solve the following system for x, y, and z. (Hint: define new variables to produce a linear system.) Explain your solution.

$$\begin{aligned} -\frac{1}{x} + \frac{2}{y} - \frac{4}{z} &= 2 \\ -\frac{2}{x} + \frac{2}{y} - \frac{4}{z} &= 6 \\ -\frac{3}{x} + \frac{4}{y} + \frac{4}{z} &= -4 \end{aligned}$$

7. Answer the question posed in Exercise 1.1.1 by solving the linear system that was created in that exercise.

8. Answer the question posed in Exercise 1.1.2 by solving the linear system that was created in that exercise.

9. Suppose that the graph of the function $f(x) = ax^3 + bx^2 + cx + d$ passes through the points $(-1, 2)$, $(-2, -9)$, $(1, 4)$, and $(2, 15)$. Determine the values of a, b, c, and d.

Writing Exercises

10. Prove that every homogeneous linear system is consistent.

11. I do believe we need to replace or eliminate this exercise!

12.

(a) Prove that if $ad - bc \neq 0$, then the reduced row-echelon form of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

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

$$\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}$$

has exactly one solution.

13. Explain why every linear system over $\mathbb{R}$ has either no solutions, exactly one solution, or an infinite number of solutions.

14. Suppose that a 3×4 coefficient matrix for a linear system has three pivot columns. Is the system consistent? Explain.

15. What would you have to know about the pivot columns in an augmented matrix in order to know that the linear system has exactly one solution?

Explain.

16. Suppose that matrices A and B are row equivalent and both matrices are in REF. Prove that A and B have the same pivot positions.

1.4 Vectors

Up to this point in the chapter, we have been concerned with solving systems of linear equations, and we have used the tool of matrices to great effect. In this section we will connect linear systems to some basic geometric concepts, and this will result in alternate ways of writing linear systems which, at times, will be more helpful.

1.4.1 The Basics of Vectors

Although we will shortly connect "vectors" to geometric notions, at the beginning a vector will be strictly an algebraic object.

Definition 1.4.1 An **n-dimensional vector** over $\mathbb{R}$ is an ordered list of n real numbers. We will adopt the convention that, unless stated otherwise, vectors are column vectors written in this form:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

where $v_i \in \mathbb{R}$ for all i. (Column vectors are therefore matrices with a single column.) The set of all n-dimensional vectors over $\mathbb{R}$ is denoted $\mathbb{R}^n$. $\Diamond$

Note 1.4.2 We say that two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are equal if $u_i = v_i$ for each $i = 1, \ldots, n$. This means that an equation involving vectors in $\mathbb{R}^n$ captures the same information as n equations involving real numbers.

As we have said, vectors should be thought of first as algebraic objects, and there are several ways to combine these objects.

Definition 1.4.3 We can combine and modify vectors through addition and a form of multiplication. We will describe the multiplication first.

Let $\mathbf{v} \in \mathbb{R}^n$ have the form

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

and let $c \in \mathbb{R}$. Then the **scalar multiple** of $\mathbf{v}$ by c is the vector $c\mathbf{v}$ in $\mathbb{R}^n$ obtained by multiplying each entry of $\mathbf{v}$ by c; that is,

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

In this context, we will often refer to real numbers as **scalars**.

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$ with the form

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

then the **sum** of $\mathbf{u}$ and $\mathbf{v}$ is the vector $\mathbf{u} + \mathbf{v}$ in $\mathbb{R}^n$ obtained by adding the corresponding entries in $\mathbf{u}$ and $\mathbf{v}$. That is,

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

◊

Note 1.4.4 We will use the notation $\mathbf{0}$ to denote the **zero vector**—that is, the vector in $\mathbb{R}^n$ whose entries are all 0. We will also use the notation $-\mathbf{u}$ to indicate the scalar multiple $(-1)\mathbf{u}$.

We present some quick calculations in the following example.

Example 1.4.5 Let $\mathbf{u} \in \mathbb{R}^2$ and let $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ such that

$$\mathbf{u} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} -6 \\ -3 \\ 3 \end{bmatrix}.$$

Then we can calculate $2\mathbf{u}$ and $\mathbf{v} + \mathbf{w}$ using Definition 1.4.3:

$$2\mathbf{u} = \begin{bmatrix} -6 \\ 10 \end{bmatrix}, \qquad \mathbf{v} + \mathbf{w} = \begin{bmatrix} -4 \\ -2 \\ -1 \end{bmatrix}.$$

□

When combining vectors, we are limited to adding vectors of the same dimension—that is, vectors with the same number of entries. However, we are not limited to either addition *or* scalar multiplication; we can certainly do both at once. Nor are we limited to adding only two vectors at a time. The following definition provides the correct generalization.

Definition 1.4.6 Let $c_1, c_2, \ldots, c_m$ be real numbers and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. Then the **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ with **weights** $c_1, c_2, \ldots, c_m$ is

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m.$$

The **span** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ is the set of all possible linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ and is written $\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. In other words, the span is defined to be

$$\operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = \left\{ \sum_{i=1}^{m} c_i\mathbf{v}_i \mid c_1, \ldots, c_m \in \mathbb{R} \right\}.$$

◊

Example 1.4.7 Taking $\mathbf{v}$ and $\mathbf{w}$ from Example 1.4.5, we can calculate the linear combination of these vectors with weights 2 and -1:

$$2\mathbf{v} - \mathbf{w} = \begin{bmatrix} 4 \\ 2 \\ -8 \end{bmatrix} + \begin{bmatrix} 6 \\ 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \\ -11 \end{bmatrix}.$$

□

Note 1.4.8 We have defined addition and scalar multiplication for vectors here, but these concepts also make sense for matrices. We will set the stage briefly here for a return in Section 2.3.

If we think of an $m \times n$ matrix A in terms of its columns, then the n columns of A are all vectors in $\mathbb{R}^m$. For two $m \times n$ matrices A and B, we define the sum $A+B$ in this way: column j of $A+B$ is the sum of the two vectors in $\mathbb{R}^m$ which are the jth columns of A and B. Similarly, if $c \in \mathbb{R}$, then we can define the scalar multiple cA in terms of its columns: column j of cA is the scalar multiple of the jth column of A by c. In this way these algebraic notions for matrices are built upon the corresponding notions for vectors.

We include two initial calculations as examples. If A and B are defined as

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

then we have

$$A + B = \begin{bmatrix} -1 & 1 & 7 \\ 3 & 0 & -1 \end{bmatrix} \text{ and } -2A = \begin{bmatrix} -4 & 2 & 0 \\ -6 & -8 & 10 \end{bmatrix}.$$

We finally note that, as with vectors, the sum (and thus the linear combination) of two or more matrices only makes sense if all involved matrices are of the same size.

Before we explore the concept of span, we need to discuss the connection between systems of linear equations and vector equations. We saw a glimpse of this connection in Note 1.4.2, and we will develop it further here.

Let us consider the following simple linear system:

$$\begin{aligned} 3x_1 - 8x_2 &= 9 \\ -2x_1 + 5x_2 &= -1. \end{aligned} \tag{1.5}$$

Solving this linear system involves (as always) investigating whether there are values of x_1 and x_2 which satisfy both of these equations simultaneously. We will now arrive at this same question from a different angle.

Let's consider the following three vectors in $\mathbb{R}^2$:

$$\mathbf{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -8 \\ 5 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{w} = \begin{bmatrix} 9 \\ -1 \end{bmatrix}.$$

Since all of these vectors are in $\mathbb{R}^2$, we can ask this question: Can $\mathbf{w}$ be written as a linear combination of $\mathbf{u}$ and $\mathbf{v}$? (This question is equivalent to asking if $\mathbf{w}$ belongs to $\text{Span}\{\mathbf{u}, \mathbf{v}\}$.) In other words, do there exist scalars x_1 and x_2 such that $x_1\mathbf{u} + x_2\mathbf{v} = \mathbf{w}$?

The equation in this final question is equivalent to other vector equations:

$$\begin{gathered} x_1\mathbf{u} + x_2\mathbf{v} = \mathbf{w} \\ x_1 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + x_2 \begin{bmatrix} -8 \\ 5 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \end{bmatrix} \\ \begin{bmatrix} 3x_1 \\ -2x_1 \end{bmatrix} + \begin{bmatrix} -8x_2 \\ 5x_2 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \end{bmatrix} \end{gathered}$$

$$\begin{bmatrix} 3x_1 - 8x_2 \\ -2x_1 + 5x_2 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \end{bmatrix}.$$

Because of Note 1.4.2, this last vector equation is equivalent to the linear system in (1.5).

Example 1.4.9 Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix}, \qquad \text{and} \qquad \mathbf{v}_3 = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}.$$

Is $\mathbf{v}_3$ in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$?

We have written this question in terms of a vector equation, but by the previous discussion we should be able to translate this question to a more familiar one about solutions to linear systems and answer the question using Algorithm 1.3.9.

The question about vectors is the same as asking if this linear system is consistent:

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

The augmented matrix for this system is

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

and the RREF of this matrix is

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

Since there is a pivot in the final column of this matrix, we conclude (by invoking Theorem 1.3.5) that the linear system is inconsistent. This means that the answer to the original question is no, $\mathbf{v}_3$ is not in Span$\{\mathbf{v}_1, \mathbf{v}_2\}$. □

Example 1.4.9 illustrates a general workflow for questions like this. There is no need to explicitly write out the intermediate step involving linear systems; instead, we can form a matrix using the given vectors as the appropriate columns, row reduce the matrix, and answer the question.

Combining vectors in $\mathbb{R}^n$ is already (or soon will be) fairly natural for readers of this chapter. Under the operations of scalar multiplication and addition, vectors in $\mathbb{R}^n$ have some useful properties, which we record in the following fact. We will not spend time with these properties now, but we will look at them intently in Section 2.3. These properties can all be verified using the corresponding properties of addition and multiplication of real numbers. (And the fact that real numbers have these properties is *essential*!)

Fact 1.4.10 *For all vectors* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *in* $\mathbb{R}^n$, *and for all real numbers* c *and* d, *the following properties hold.*

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

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

3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$

4. $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$

5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$

6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$

7. $c(d\mathbf{u}) = (cd)\mathbf{u}$

8. $1\mathbf{u} = \mathbf{u}$

Thinking about linear systems through the lens of vectors also brings some structure to the solutions to consistent linear systems. When a linear system is consistent and the solution is unique, we have recorded this in terms of the variables involved. For example, we recorded the solution to the linear system in Example 1.3.10 as $x = -1/9$, $y = 4/3$, and $z = 2/9$. However, if we think of these variables as forming a vector in $\mathbb{R}^3$, then we can record the solution this way:

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1/9 \\ 4/3 \\ 2/9 \end{bmatrix}.$$

Similarly, when a consistent system does not have a unique solution, we can again use vector notation. (This would replace the parametric description of the solution set we presented in Example 1.3.4.) In Example 1.3.4 we considered such a linear system. The solution set can now be written this way:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 + 3x_3 \\ 7 - 2x_3 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$$

If we label the vectors $\mathbf{v}$ and $\mathbf{w}$ as

$$\mathbf{v} = \begin{bmatrix} 5 \\ 7 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix},$$

then all solutions to this system can be written as $\mathbf{v} + t\mathbf{w}$, where t can be any real number. As we learn more about the geometric interpretations of the solutions to linear systems, this phrasing in terms of vectors will be useful.

1.4.2 Vectors and Geometry

Before we leave this chapter, we need to introduce the connection between vectors and geometry. Readers who have taken multivariable calculus are likely aware of some of these concepts. For the sake of simplicity, we will restrict most of our discussion (and all of our illustrations) to $\mathbb{R}^2$ in this section.

We will visualize a vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ by identifying it with the point (a, b) in the plane. We will draw an arrow from the origin to (a, b) to aid this visualization. So, for example, the vectors $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ would be visualized in the following picture.

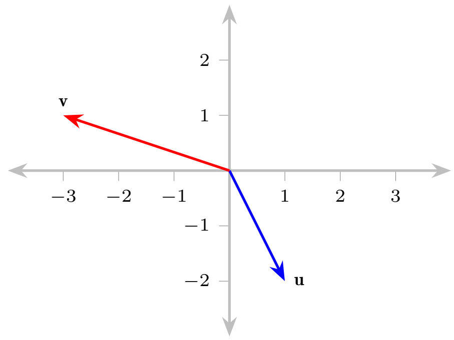

Figure 1.4.11 Illustration of two vectors in the plane

Scalar multiplication of a vector can be seen as stretching or contracting the vector, if the scalar is positive. If the scalar is negative, the vector will be pointing in the opposite direction, then stretched or contracted. Here is an example using the vector **v** previously defined.

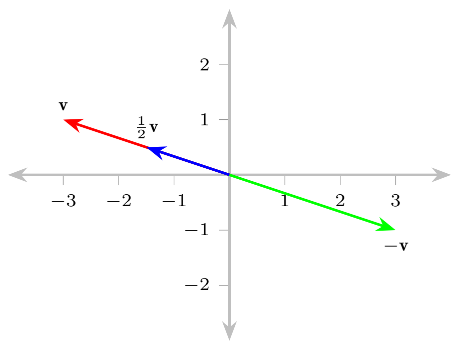

Figure 1.4.12 Illustration of scalar multiplication

Addition of vectors is also easy to visualize. To add two vectors, move the second vector from the origin so that its beginning coincides with the end of the first vector. The vector sum can be represented by the arrow from the origin to the end of this relocated second vector. (The dashed arrow in the following diagram is the relocated vector **v**.)

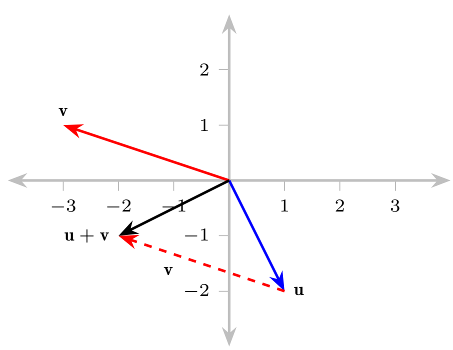

Figure 1.4.13 Illustration of vector addition

We can put these previous two ideas together to explain the way to visualize linear combinations (and spans). Let's consider the vectors **u** and **v** as previously defined, and we will also define **w** as $\mathbf{w} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$.

When we ask whether **w** is in Span$\{\mathbf{u}, \mathbf{v}\}$, this is the same as asking whether we can form **w** as a linear combination of **u** and **v**. We answer this question by forming a matrix with the appropriate columns and row reducing. We find the matrix

$$\left[\mathbf{u} \quad \mathbf{v} \mid \mathbf{w}\right] = \left[\begin{array}{cc|c} 1 & -3 & 2 \\ -2 & 1 & 2 \end{array}\right],$$

and the RREF of this matrix is

$$\left[\begin{array}{cc|c} 1 & 0 & -1.6 \\ 0 & 1 & -1.2 \end{array}\right].$$

This tells us that $\mathbf{w}$ *is* in the span of $\mathbf{u}$ and $\mathbf{v}$, and that

$$\mathbf{w} = -1.6\mathbf{u} - 1.2\mathbf{v}.$$

Here is the picture illustrating this.

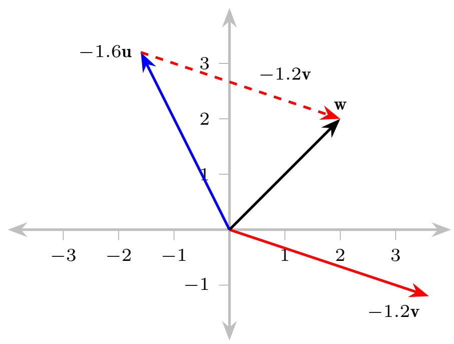

Figure 1.4.14 Illustration of a linear combination

The reader may be able to see that no matter what vector $\mathbf{w}$ in $\mathbb{R}^2$ was chosen in this last instance, that vector would be in $\operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$. (This means that $\operatorname{Span}\{\mathbf{u}, \mathbf{v}\} = \mathbb{R}^2$.) This has everything to do with the relationship between $\mathbf{u}$ and $\mathbf{v}$.

We will discuss this more carefully in Chapter 5, but this brief comment will suffice for now. As long as $\mathbf{u} \neq \mathbf{0}$ in $\mathbb{R}^2$, then $\operatorname{Span}\{\mathbf{u}\}$ is a line through the origin. (Remember that the span in this case consists of *all* multiples of $\mathbf{u}$.) Similarly, as long as neither $\mathbf{u}$ nor $\mathbf{v}$ is a multiple of the other vector, then $\operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$ is *all of* $\mathbb{R}^2$. (If we were talking about vectors in $\mathbb{R}^3$ instead of $\mathbb{R}^2$ this span would be a plane through the origin.)

1.4.3 Reading Questions

1. Let $\mathbf{u}$ and $\mathbf{v}$ be the following vectors:

$$\mathbf{u} = \begin{bmatrix} 2 \\ -7 \\ 4 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}.$$

Calculate each of the following vectors.

(a) $-2\mathbf{u}$

(b) $\mathbf{u} - \mathbf{v}$

(c) $3\mathbf{u} - 4\mathbf{v}$

2. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be the following vectors:

$$\mathbf{u} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}.$$

Is the vector

$$\begin{bmatrix} 7 \\ -1 \\ 8 \end{bmatrix}$$

in Span$\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$? Explain your answer using techniques from this section. (See especially Example 1.4.9.)

1.4.4 Exercises

1. Write a linear system that is equivalent to the following vector equation:

$$x \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix} + y \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ -1 \end{bmatrix}.$$

2. Write a vector equation that is equivalent to the following linear system:

$$\begin{aligned} 2x - 5z &= 9 \\ x + 3y + z &= -1 \\ -y - 7z &= 0. \end{aligned}$$

3. Let $\mathbf{u}$ and $\mathbf{v}$ be the following vectors:

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

List five vectors in Span$\{\mathbf{u}, \mathbf{v}\}$. For each vector, write down the specific linear combination of $\mathbf{u}$ and $\mathbf{v}$ used to generate the vector.

4. Describe all possible ways of writing $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as a linear combination of $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$, and $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$.

5. For each of the following, determine whether $\mathbf{b}$ is a linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$. Explain your answer.

(a)

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 5 \\ -4 \\ -2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}.$$

(b)

$$\mathbf{u} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 5 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 7 \\ 2 \\ -3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 10 \\ -1 \\ 4 \end{bmatrix}$$

6. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{b}$ be the following vectors:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ -6 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ -1 \\ h \end{bmatrix}.$$

For what value(s) of h is $\mathbf{b}$ in the plane spanned by $\mathbf{u}$ and $\mathbf{v}$? Explain your answer.

7. Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{b}$ be the following vectors:

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Under what conditions is $\mathbf{b}$ in $\mathrm{Span}\{\mathbf{u}, \mathbf{v}\}$? This will be an equation (or equations) satisfied by the b_i. Explain your answer.

Writing Exercises

8. Let $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$. Show that $\begin{bmatrix} h \\ k \end{bmatrix}$ is in $\mathrm{Span}\{\mathbf{u}, \mathbf{v}\}$ for all h and k. Explain your answer.

9. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are called *collinear* if $\mathbf{u} = k\mathbf{v}$ or $\mathbf{v} = k\mathbf{u}$ for some $k \in \mathbb{R}$. Show that the span of any two nonzero vectors in $\mathbb{R}^2$ which are not collinear is all of $\mathbb{R}^2$.

10. Give an example of three vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ in $\mathbb{R}^3$ such that no two of them are collinear but $\mathrm{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\} \neq \mathbb{R}^3$. Explain why your example works.

11. State criteria on vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in $\mathbb{R}^3$ such that $\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\} = \mathbb{R}^3$. Explain your answer.

12. Prove property 5 from Fact 1.4.10. Note which properties of the real numbers you use in this proof.

Chapter 2

Fields and Vector Spaces

2.1 Fields

In this section we will consider the real numbers and study their most important properties in a general setting. In the process, we will learn how to handle axioms and abstract algebraic concepts.

In Chapter 1, we used the real numbers as the familiar world within which all of our calculations took place. Before we offer any definitions or results in this chapter, we will ponder exactly what properties of the real numbers were essential to these calculations. (The reader is likely so familiar with the real numbers that its important properties have been internalized and are not in the conscious mind. In this section we will make those properties explicit.)

The most basic concept in this book is the linear equation. What properties of the real numbers are used to solve a simple linear equation? In what follows we will solve the equation $2x + 7 = 12$ and draw to the surface some of these important properties:

$$2x + 7 = 12 \tag{2.1}$$
$$(2x + 7) - 7 = 12 - 7 \tag{2.2}$$
$$2x + (7 - 7) = 5 \tag{2.3}$$
$$2x + 0 = 5 \tag{2.4}$$
$$2x = 5 \tag{2.5}$$
$$\frac{1}{2}(2x) = \frac{1}{2}5 \tag{2.6}$$
$$(\frac{1}{2}2)x = \frac{5}{2} \tag{2.7}$$
$$(1)x = \frac{5}{2} \tag{2.8}$$
$$x = \frac{5}{2}. \tag{2.9}$$

No student wants to write out all of these steps, and no instructor wants to grade such a solution! But the point is to notice the properties of the real numbers which are vital to solving an equation like this and which we usually ignore.

- In (2.3) we used the *associativity* of addition in $\mathbb{R}$. That is, we can move the parentheses around in addition and still have an equivalent expression. (The reader should see subtraction as a form of addition.)

- In (2.4) we used the fact that 7 has an *additive inverse*—a number we can add to 7 to get 0. (The additive inverse of 7 is -7.)
- In (2.5) we used the fact that 0 is an *additive identity* in $\mathbb{R}$—when we add 0 to any real number r, we get r again.
- In (2.7) we used the *associativity* of multiplication in $\mathbb{R}$. We can move the parentheses around in multiplication just like we can in addition.
- In (2.8) we used the fact that $\frac{1}{2}$ is the *multiplicative inverse* of 2 in $\mathbb{R}$. In other words, we can multiply 2 by $\frac{1}{2}$ to get 1.
- Finally, in (2.9) we used the fact that 1 is a *multiplicative identity* in $\mathbb{R}$—when we multiply any real number by 1, that real number is unchanged.

By identifying these properties, our goal is to envision other mathematical realms in which solving linear equations would work the same way it does within $\mathbb{R}$. Toward this end, we now define an algebraic object called a "field" which has all of the properties used above (plus a few we haven't yet mentioned).

Definition 2.1.1 A **field** is a set $\mathbb{F}$ with operations $+$ and $\cdot$ and distinct elements $0, 1 \in \mathbb{F}$ such that all of the following properties hold.

1. For all $a, b \in \mathbb{F}$, $a + b \in \mathbb{F}$. (We say that $\mathbb{F}$ is **closed under addition**.)
2. For all $a, b \in \mathbb{F}$, $a \cdot b \in \mathbb{F}$. (We say that $\mathbb{F}$ is **closed under multiplication**.)
3. For all $a, b \in \mathbb{F}$, $a + b = b + a$. (We say that addition in $\mathbb{F}$ is **commutative**.)
4. For all $a, b, c \in \mathbb{F}$, $a + (b + c) = (a + b) + c$. (We say that addition in $\mathbb{F}$ is **associative**.)
5. For each $a \in \mathbb{F}$, $a+0 = 0+a = a$. (We say that 0 is an **additive identity** in $\mathbb{F}$.)
6. For each $a \in \mathbb{F}$, there exists an element $b \in \mathbb{F}$, such that $a+b = b+a = 0$. (We say that each a has an **additive inverse** in $\mathbb{F}$.)
7. For each $a, b \in \mathbb{F}$, $a \cdot b = b \cdot a$. (We say that multiplication is **commutative** in $\mathbb{F}$.)
8. For each $a, b, c \in \mathbb{F}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (We say that multiplication is **associative** in $\mathbb{F}$.)
9. For each $a \in \mathbb{F}$, $a \cdot 1 = 1 \cdot a = a$. (We say that 1 is a **multiplicative identity** in $\mathbb{F}$.)
10. For each $a \in \mathbb{F}$ with $a \neq 0$, there exists an element $c \in \mathbb{F}$ such that $a \cdot c = c \cdot a = 1$. (We say that every nonzero element a in $\mathbb{F}$ has a **multiplicative inverse**.)
11. For each $a, b, c \in \mathbb{F}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$. (We say that addition and multiplication in $\mathbb{F}$ satisfy the **distributive law**.)

$\Diamond$

Note 2.1.2 Axioms 5 and 6 could perhaps be stated with a bit more precision. Some texts state axiom 5 like this: the set $\mathbb{F}$ must have an additive identity, that is, an element e such that $x + e = e + x = x$ for all $x \in \mathbb{F}$. Those same texts may also state axiom 6 without reference to the symbol 0: for each $a \in \mathbb{F}$, there exists an element $b \in \mathbb{F}$ such that $a + b = b + a = e$, where e is the additive identity of the field.

For some students, the use of the symbol 0 might be helpful, reminding them of the essential property of the additive identity through familiarity. But this may be confusing when handling a set which contains the symbol 0 when that symbol does *not* function as the relevant additive identity.

Our point here is that the use of 0 in the axioms for a field (and some of the subsequent theorems) is intended to stand in for a generic symbol for the additive identity. Similarly, the symbol 1 stands in as the generic symbol for the multiplicative identity. (A reader could imagine a restatement of axioms 9 and 10 above using a letter as the multiplicative identity in the way we have demonstrated axioms 5 and 6 can be restated.)

The reader should also note that in a generic field the term "nonzero" means "not the additive identity."

A note about axioms. What are presented in Definition 2.1.1 are known as the **axioms** of a field. This may be the reader's first exposure to axioms in mathematics, and this is worthy of a comment or two.

Much of theoretical mathematics is built upon axiomatic reasoning. The thinking goes like this: If we assume a limited number of properties are true, and we assume nothing beyond those properties, what else follows *necessarily?* So, we can ask what is true of a field, not just what is true of the real numbers. While the real numbers may have specific properties that a general field does not, anything that is true of a general field must be true of the real numbers.

Working through some examples (and non-examples) will help us make sense of this definition.

Example 2.1.3 The set of real numbers $\mathbb{R}$ is a field. (If the real numbers were not a field, then we wouldn't have done a very good job of abstracting the properties of the real numbers for this definition!) □

Example 2.1.4 The complex numbers are defined in this way:

$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\},$$

where $i^2 = -1$. Addition and multiplication are defined in this way:

$$\begin{aligned}(a + bi) + (c + di) &= (a + c) + (b + d)i \\ (a + bi) \cdot (c + di) &= (ac - bd) + (ad + bc)i.\end{aligned}$$

Notice that the addition and multiplication occuring within the parentheses on the right side of these equations are happening within $\mathbb{R}$. In this way, showing that some of the field axioms hold for $\mathbb{C}$ depends on $\mathbb{R}$ being a field.

The elements 0 and 1 in $\mathbb{C}$ are

$$\begin{aligned}0 &= 0 + 0i \\ 1 &= 1 + 0i,\end{aligned}$$

and these elements are not equal. (We recall that part of the definition of a field is that the elements 0 and 1 are distinct.)

With these definitions, one can check that $\mathbb{C}$ satisfies the properties of a field. We will prove a few of these properties, and we will assign a few of these proofs in the exercises.

The definitions of $+$ and $\cdot$ in $\mathbb{C}$ show that $\mathbb{C}$ is closed under addition and multiplication. (Notably, this relies on $\mathbb{R}$ being closed under addition and multiplication!) To prove that the addition in $\mathbb{C}$ is commutative, we consider two complex numbers $a+bi$ and $c+di$. Then

$$\begin{aligned}(a+bi)+(c+di) &= (a+c)+(b+d)i &(2.10)\\ &= (c+a)+(d+b)i &(2.11)\\ &= (c+di)+(a+bi). &(2.12)\end{aligned}$$

(We note that line (2.11) used the fact that addition of real numbers is commutative.) This proves that addition in $\mathbb{C}$ is commutative.

We will also prove that 0 is an additive identity in $\mathbb{C}$. If $a+bi \in \mathbb{C}$, then

$$(a+bi)+(0+0i) = (a+0)+(b+0)i = a+bi.$$

(This uses the fact that $0 \in \mathbb{R}$ is an additive identity.) We note that although the definition of a field requires us to examine addition by 0 "on both sides," since we just proved that addition is commutative, what we have already shown is sufficient. □

Example 2.1.5 The set of non-negative real numbers $\mathbb{R}^{\geq 0}$ is not a field. (When thinking about a subset of the real numbers, we will assume that the usual addition and multiplication are in view unless otherwise stated.) While this set is closed under addition and multiplication, it does not contain *additive inverses* for positive real numbers. For example, the number 9 has no additive inverse in this set. □

Example 2.1.6 The set of **rational numbers** $\mathbb{Q}$ is a field. The rational numbers are defined as the set of all quotients (hence the symbol $\mathbb{Q}$) of integers; more formally,

$$\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

We do not need to show that all of the field axioms hold in order to prove that $\mathbb{Q}$ is a field. Since $\mathbb{Q}$ is a subset of $\mathbb{R}$, and since it therefore inherits its operations from $\mathbb{R}$, some of the field axioms hold automatically. (These would include the associativity and commutativity of addition. What else holds "by inheritance?") Checking some of these details is left to the exercises. □

We refer the reader to Appendix **??** for the definitions of $\mathbb{Z}_n$ and $\mathbb{F}_p$, related notational conventions, and a refresher on modular arithmetic. All of this is necessary for the following example.

Example 2.1.7 When p is a prime integer, then $\mathbb{F}_p$ is a field.

Working within $\mathbb{F}_p$ can be a bit destabilizing at first, as the calculations take some practice. However, this effort pays off because the smallest finite fields offer some of the most tangible sandboxes in which to play. We will be using $\mathbb{F}_p$ for small p throughout this chapter to develop some interesting examples.

Part of the definition of a field is that one can divide by any nonzero element. But because we work within $\mathbb{R}$ so often, division within $\mathbb{F}_p$ is strange. In $\mathbb{F}_3$, for instance, the multiplicative inverse of 2 is 2. (Another way to say this is that, within $\mathbb{F}_3$, the element that acts the most like $\frac{1}{2}$ is 2.)

The field $\mathbb{F}_2$ contains only the elements 0 and 1, and it therefore is a model for anything that is *binary*. For this reason, working in $\mathbb{F}_2$ is often very useful in computer science. □

Having defined fields, we now turn to the consequences of this definition. In other words, if a set with two operations meets the definition of a field, what else must also be true? The following theorem presents some basic results that flow from the axioms of a field. Some of these will look familiar because they are true in $\mathbb{R}$ and that's where most people are comfortable working.

Theorem 2.1.8 *For any field $\mathbb{F}$, the following are true.*

1. *The additive identity in $\mathbb{F}$ is unique.*
2. *The multiplicative identity in $\mathbb{F}$ is unique.*
3. *Additive inverses in $\mathbb{F}$ are unique. (This means that for every element $x \in \mathbb{F}$, there exists a unique element $y \in \mathbb{F}$ which is the additive inverse of x. The truth of this statement justifies our use of the notation $-x$ for the additive inverse of an element $x \in \mathbb{F}$.)*
4. *Multiplicative inverses in $\mathbb{F}$ are unique. (The truth of this statement justifies our use of the notation x^{-1} for the multiplicative inverse of a nonzero element $x \in \mathbb{F}$.)*
5. *For every $x \in \mathbb{F}$, $-(-x) = x$.*
6. *For every nonzero $x \in \mathbb{F}$, $(x^{-1})^{-1} = x$.*
7. *For every $x \in \mathbb{F}$, $0 \cdot x = 0$.*
8. *For every $x \in \mathbb{F}$, $(-1) \cdot x = -x$.*
9. *If $x, y \in \mathbb{F}$ and $x \cdot y = 0$, then either $x = 0$ or $y = 0$.*

Proof. We will prove a few of these and leave the rest as exercises.
To prove (Item 1) that the additive identity is unique, we must prove that 0 is the only element in $\mathbb{F}$ which has the properties of an additive identity. We suppose that $a \in \mathbb{F}$ is such that $a + x = x + a = x$ for every element $x \in \mathbb{F}$. Since this must be true for every x, it must be true for $x = 0$. The previous equation then becomes

$$a + 0 = 0 + a = 0.$$

Since 0 is an additive identity, $a + 0 = a$, which combined with the fact that $a + 0 = 0$ means that $a = 0$. This proves that the additive identity, 0, is unique.
To prove (Item 3) that additive inverses are unique, we must prove that for any element $x \in \mathbb{F}$, there is a unique element that behaves like an additive inverse of x. We let $x \in \mathbb{F}$ and we suppose that $y, z \in \mathbb{F}$ are both additive inverses of x. We wish to show that $y = z$.
Our assumptions mean that

$$\begin{aligned} x + y &= y + x = 0 \\ x + z &= z + x = 0. \end{aligned}$$

We then use these assumptions, along with some of the properties of fields (the associativity of addition and the properties of the additive identity) in this calculation:

$$\begin{aligned} x + y &= 0 \\ z + (x + y) &= z + 0 \\ (z + x) + y &= z \\ 0 + y &= z \\ y &= z. \end{aligned}$$

This proves that additive inverses are unique. ∎

We offer one final note on notation. We will often use juxtaposition to denote multiplication within a field. That is, we may write xy instead of $x \cdot y$ to indicate the product of two elements in a field. We trust the reader will adjust quickly to this seismic shift.

Reading Questions

1. Carry out the following calculations within $\mathbb{F}_7$.

(a) $4 + 3$

(b) $5 \cdot 6$

(c) $2 \cdot (3+6)$ (Complete this calculation in the two different ways present in the distributive law and verify that they are equal.)

2. What is the additive inverse of 3 in $\mathbb{F}_5$? What is the multiplicative inverse of 3 in $\mathbb{F}_7$? Explain your answers.

3. Theorem **??** says that $\mathbb{Z}_9$ is not a field because 9 is not prime. Identify a nonzero element of $\mathbb{Z}_9$ that does not have a multiplicative inverse and explain why it does not have an inverse.

Exercises

1. Carry out the following calculations in $\mathbb{Z}_8$. (Remember that your answer for calculations in $\mathbb{Z}_n$ should be a number between 0 and $n-1$.)

(a) $6 + 7$

(b) $3 + 7 + 5$

(c) $7 \cdot 3$

(d) $6 \cdot 4 \cdot 3$

(e) $(-3) \cdot (5 + 3 + (-4))$

2. We know that since 12 is not a prime, $\mathbb{Z}_{12}$ is not a field. In particular, the axiom about multiplicative inverses does not hold. For each nonzero element of $\mathbb{Z}_{12}$, determine whether or not it has a multiplicative inverse. If the element has a multiplicative inverse, state that inverse.

3. For each nonzero element of $\mathbb{F}_{11}$, find the multiplicative inverse.

Writing Exercises

4. Finish Example 2.1.4. In other words, complete the proof begun in Example 2.1.4 that $\mathbb{C}$ is a field.

5. Consider Example 2.1.6. Which of the field axioms for $\mathbb{Q}$ hold "by inheritance" and for which of the axioms is there something that needs to be proved? Put each of the field axioms into one of these two categories. For each axiom that doesn't hold merely "by inheritance," prove that it holds for $\mathbb{Q}$.

6. For a set A, define the set of polynomials over A in the usual way:

$$A[x] = \{a_0 + a_1x + a_2x^2 + \cdots \mid a_i \in A\}.$$

(a) Is $\mathbb{Z}[x]$ a field? Justify your answer.

(b) Is $\mathbb{R}[x]$ a field? Justify your answer.

7. In this problem we consider "adding" an element to $\mathbb{F}_3$. If $\mathbb{F}_3[\alpha]$ is defined by
$$\mathbb{F}_3[\alpha] = \{a + b\alpha \mid a, b \in \mathbb{F}_3, \alpha^2 = -1\},$$
is $\mathbb{F}_3[\alpha]$ a field? Justify your answer.

8. In this problem we consider "adding" an element to $\mathbb{F}_5$.

(a) If $\mathbb{F}_5[\alpha]$ is defined by
$$\mathbb{F}_5[\alpha] = \{a + b\alpha \mid a, b \in \mathbb{F}_5, \alpha^2 = 4\},$$
is $\mathbb{F}_5[\alpha]$ a field? Justify your answer.

(b) If $\mathbb{F}_5[\alpha]$ is defined by
$$\mathbb{F}_5[\alpha] = \{a + b\alpha \mid a, b \in \mathbb{F}_5, \alpha^2 = 2\},$$
is $\mathbb{F}_5[\alpha]$ a field? Justify your answer.

9. Define $\mathbb{Q}[\sqrt{2}]$ by
$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$
Is $\mathbb{Q}[\sqrt{2}]$ a field? Justify your answer.

10. Consider the set $\mathbb{R}^2$ with operations defined as follows:
$$(a, b) \oplus (c, d) = (a + c, b + d)$$
$$(a, b) \otimes (c, d) = (ac - 2bd, ad + bc).$$
Is $\mathbb{R}^2$ with these operations a field? Justify your answer.

11. Prove Item 2 of Theorem 2.1.8.

12. Prove Item 4 of Theorem 2.1.8.

13. Prove Item 7 of Theorem 2.1.8.

14. Prove Item 8 of Theorem 2.1.8.

15. Prove Item 9 of Theorem 2.1.8.

2.2 Solving Linear Systems Over Fields

Having now defined a field, in this section we will show how the problems of chapter 1 can be solved in this general setting. We have laid the groundwork for reproducing the results of the first three sections of Chapter 1 for a field $\mathbb{F}$ instead of the real numbers.

We now return to where we began in Section 1.1: the humble linear equation. If fields are generalizations of the real numbers, and if we can solve linear equations when everything in sight comes from the real numbers, we should be able to solve linear equations when everything in sight comes from a general field.

Example 2.2.1 Consider the following equation where all variable values, constants, and coefficients are drawn from $\mathbb{F}_5$:
$$3x + 2 = 1.$$
Solving this equation in $\mathbb{R}$ would be easy; let's solve it in $\mathbb{F}_5$.

We first note that the additive inverse of 2 in $\mathbb{F}_5$ is 3, so our first step is to add 3 to both sides of the equation:
$$3x + 2 + 3 = 1 + 3$$

$$3x = 4.$$

We now need the multiplicative inverse of 3 in $\mathbb{F}_5$, which is 2. We multiply both sides by 2 to get our answer:

$$\begin{aligned} 2(3x) &= 2(4) \\ x &= 3. \end{aligned}$$

When we check that our solution works (we plug $x = 3$ back into the original equation and perform the computations in $\mathbb{F}_5$), we find that it does:

$$3x + 2 = 3(3) + 2 = 9 + 2 = 11 = 1.$$

□

The point of this section is that the algorithm for solving linear systems (Algorithm 1.3.9) which worked for $\mathbb{R}$ also works for a general field $\mathbb{F}$. In order to be comfortable with this notion, we need to talk quickly through the development of that algorithm in this general setting.

Because a general field contains both 0 and 1 (or rather, an additive and a multiplicative identity), and because within a field we can perform all of the operations needed to solve linear systems, everything we want to do is legitimate. The definitions of the coefficient and augmented matrices, the elementary row operations, the echelon forms, pivots, and the row reduction algorithm (all found in Section 1.2 and Section 1.3) are the same once we move away from $\mathbb{R}$. Similarly, the three important theorems we have encountered so far (which we will reproduce below) all hold over a general field $\mathbb{F}$. We will omit the proofs of these theorems because the earlier proofs, when translated from $\mathbb{R}$ to $\mathbb{F}$, are still valid.

Theorem 2.2.2 *Suppose A is the augmented matrix of a linear system over a field $\mathbb{F}$, and suppose that A is in reduced row-echelon form. Then the linear system is consistent if and only if there is no pivot in the final column of A.*

Theorem 2.2.3 *Suppose that A is the augmented matrix corresponding to a consistent $m \times n$ linear system over a field $\mathbb{F}$, and suppose that A is in reduced row-echelon form. Then the system has a unique solution if and only if there is a pivot in each of the first n columns of A.*

Theorem 2.2.4 *Let A be an $m \times n$ matrix with entries in a field $\mathbb{F}$ and let A be row equivalent to both B and C. If B and C are in reduced row-echelon form, then $B = C$.*

The algorithm for solving a linear system, supported by these theorems, remains the same as in Algorithm 1.3.9. We will finish this section with two examples where we go through this algorithm carefully.

Example 2.2.5 The following is a linear system over $\mathbb{F}_3$:

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

We will begin to solve this system by forming the augmented matrix:

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

Since working with fields other than $\mathbb{R}$ is still new, we will explain all of the steps needed to reduce this matrix to its RREF. We first add the first row to the second row to produce a 0 in the $(2,1)$ position. (Remember that 1 is the additive inverse of 2 in $\mathbb{F}_3$!) This matrix is the result:

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

We now add twice the first row to the third to produce a 0 in the $(3,1)$ position. Here is that matrix:

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

We now notice that the second and third rows are the same. This means the third row will end up being a row of zeros, and we can achieve this by adding twice the second row to the third (in $\mathbb{F}_3$, this is the same as subtracting the second row from the third):

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

The final step in reducing this matrix is to take care of the entry which is above the pivot in the $(2,2)$ position. We add the second row to the first, and this is the matrix which results:

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

Our matrix is now in RREF, and from Theorem 2.2.2 we can conclude that this linear system is consistent. Further, from Theorem 2.2.3 we can see that there is not a unique solution to this system. We can write the solutions to this system, however:

$$\begin{cases} x_1 = x_3 \\ x_2 = 2 + 2x_3 \\ x_3 \text{ is free.} \end{cases}$$

Since $\mathbb{F}_3$ has three elements, there are three possible values for x_3, meaning that there are three solutions to this linear system. □

Example 2.2.6 The following is a linear system over $\mathbb{C}$:

$$\begin{aligned} (2-i)x_1 + (2+4i)x_2 + (-7+6i)x_3 &= 2-6i \\ (1-2i)x_1 + (5+2i)x_2 + (-2+12i)x_3 &= -3-8i \\ -3x_1 + (1-5i)x_2 + 9x_3 &= -4+3i. \end{aligned}$$

This example may seem a bit intimidating at first, especially for readers who have not dealt much with $\mathbb{C}$. But when we follow our step-by-step approach, we should arrive at a solution with minimal problems.

First, we write down the augmented matrix of this system:

$$\left[\begin{array}{ccc|c} 2-i & 2+4i & -7+6i & 2-6i \\ 1-2i & 5+2i & -2+12i & -3-8i \\ -3 & 1-5i & 9 & -4+3i \end{array}\right].$$

To start row reducing this matrix, we need a 1 in the $(1,1)$ entry. Instead of following Algorithm 1.3.9 rigidly by exchanging rows and then dividing (or multiplying by an inverse), we will skip the first step and handle the rows as they are.

For a nonzero element $a + bi$ of $\mathbb{C}$, the multiplicative inverse is

$$\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i.$$

This means that the inverse of $2 - i$ is $\frac{2}{5} + \frac{1}{5}i$. So, in order to get a 1 in the $(1, 1)$ position, we multiply the first row of the matrix by $\frac{2}{5} + \frac{1}{5}i$. Here is the result:

$$\left[\begin{array}{ccc|c} 1 & 2i & -4+i & 2-2i \\ 1-2i & 5+2i & -2+12i & -3-8i \\ -3 & 1-5i & 9 & -4+3i \end{array}\right].$$

We now work to clear out the other entries in the first column. We add $(-1 + 2i)$ times the first row to the second and we add 3 times the first row to the third. (We are taking care of two steps at once here.) This is the result:

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

(For readers who are new to $\mathbb{C}$, verifying these calculations would be an excellent exercise!)

Since we already have a 1 in the $(2, 2)$ entry, we can use that to produce a zero below it in that column. We add $(-1 - i)$ times the second row to the third row, and we get this:

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

Even though this matrix is not yet in RREF, we do not need to continue with our algorithm. Theorem 2.2.2 tells us that the original system is inconsistent because of the pivot in the final column. Therefore, this system has no solution. □

Reading Questions

1. Solve the following linear equation over $\mathbb{C}$. List the steps you take in solving the equation in terms of the axioms of a field.

$$(1 - 3i)x + (4 + 2i) = -1 - 3i$$

2. The following is a matrix with entries from $\mathbb{F}_5$. Reduce this matrix to REF. (It is not necessary to reduce the matrix to RREF.) Describe each step you take.

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

Exercises

1. Solve the following linear system over $\mathbb{F}_2$:

$$\begin{aligned} x + z &= 1 \\ y + z &= 0 \end{aligned}$$

$$x+y+z=1.$$

2. Solve the following linear system over $\mathbb{F}_3$:

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

3. Solve the following linear system over $\mathbb{F}_7$:

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

4. The following matrix A is defined over $\mathbb{C}$:

$$A=\begin{bmatrix} i & 1+i & 3\\ 1-i & -3+2i & 6+10i \end{bmatrix}.$$

Put this matrix into RREF.

5. Solve the following linear system over $\mathbb{C}$:

$$\begin{aligned} (-i)x+(3-i)z&=-1\\ (3+2i)x+y+(-1+9i)z&=6-3i\\ (1+i)x+(2-i)y-(2i)z&=10-5i. \end{aligned}$$

6. Solve the following linear system over $\mathbb{F}_5$:

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

7. The following matrix A is defined over $\mathbb{Q}[\sqrt{2}]$ (see Exercise 2.1.9):

$$A=\begin{bmatrix} 1+\sqrt{2} & 2 & 3 & -1+2\sqrt{2}\\ 3\sqrt{2} & 1-\sqrt{2} & 0 & 2+\sqrt{2}\\ 4 & -5+2\sqrt{2} & 3+\sqrt{2} & -1 \end{bmatrix}.$$

Put this matrix into RREF.

8. The following matrix A is defined over $\mathbb{F}_3[\alpha]$ (see Exercise 2.1.7):

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

Put this matrix into RREF.

Writing Exercises

9.

(a) Suppose that the following is a linear system over $\mathbb{F}_3$:

$$\begin{aligned} ax+by&=e\\ cx+dy&=f. \end{aligned}$$

Show that if $ad + 2bc \neq 0$, then this linear system has a unique solution. (For reference, see Exercise 1.3.12.)

(b) Suppose that the system stated in part a. of this problem is a system over $\mathbb{F}_p$. What is the correct inequality the coefficients a, b, c, and d must satisfy in order for the system to have a unique solution? State your answer and prove it is correct.

10. Consider a linear system

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

with all coefficients and constants in the integers.

(a) Show that if the system has a solution $\mathbf{x} \in \mathbb{R}^n$ then it must have a solution $\mathbf{y} \in \mathbb{Q}^n$.

(b) Show that if the system has a *unique* solution $\mathbf{x} \in \mathbb{Q}^n$ then $\mathbf{x}$ is also the *unique* solution in $\mathbb{R}^n$.

11. Give descriptions of linear systems with each of the following properties, or state why such a system is impossible. Specific numbers and equations are not necessary, but your description should include what field is involved, the number of equations and variables, the number of free variables, etc. Explain your answers.

(a) A consistent system with exactly 8 solutions

(b) A consistent system with exactly 10 solutions

(c) A consistent system with exactly 9 solutions

(d) A consistent system with exactly 17 solutions

2.3 Vector Spaces

We now move to our next task of abstraction. We have generalized the real numbers and introduced the idea of a field, and we will now generalize the set and structure of the vectors $\mathbb{R}^n$ (see Section 1.4) to a vector space.

Definition 2.3.1 A **vector space** over a field $\mathbb{F}$ is a set V on which are defined the operations of *addition* and *scalar multiplication* such that all of the following properties hold.

1. For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} \in V$. (The sum of two vectors is a vector.)
2. For all $c \in \mathbb{F}$ and all $\mathbf{v} \in V$, $c\mathbf{v} \in V$. (The scalar multiple of a vector is a vector.)
3. For all $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. (Vector addition is commutative.)
4. For all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$. (Vector addition is associative.)
5. There is a vector $\mathbf{0} \in V$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for all $\mathbf{v} \in V$. (There is a vector which is the identity for vector addition. We call this the *zero vector*.)

6. For each $\mathbf{u} \in V$ there exists a vector $\mathbf{v} \in V$ such that $\mathbf{u} + \mathbf{v} = \mathbf{0}$. (Each vector has an additive inverse.)

7. For each $\mathbf{v} \in V$, $1\mathbf{v} = \mathbf{v}$. (Scalar multiplication by 1 is an identity.)

8. For each $\mathbf{v} \in V$ and all $c, d \in \mathbb{F}$, $c(d\mathbf{v}) = (cd)\mathbf{v}$. (Scalar multiplication of a vector is associative.)

9. For all $\mathbf{u}, \mathbf{v} \in V$ and each $c \in \mathbb{F}$, $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$. (Scalar multiplication distributes over the sum of vectors.)

10. For each $\mathbf{v} \in V$ and all $c, d \in \mathbb{F}$, $(c+d)\mathbf{v} = c\mathbf{v}+d\mathbf{v}$. (Scalar multiplication distributes over the sum of field elements.)

The elements of V are called *vectors* and the elements of $\mathbb{F}$ are called *scalars*. $\Diamond$

Note 2.3.2 The reader will want to note when the multiplication of field elements is in view and when scalar multiplication (of a scalar times a vector) is in view. The context should help, as should our practice of using bold notation for vectors. The product of vectors (in the way that we take the product of field elements) is not a defined construction in a general vector space.

Faced with this rather abstract definition, some examples are in order.

Example 2.3.3 The set $\mathbb{R}^n$, with the operations of scalar multiplication and vector addition defined in Definition 1.4.3, is a vector space over $\mathbb{R}$. As with our definition of a field, if $\mathbb{R}^n$ is not a vector space then we have carried out the enterprise of abstraction rather poorly. (See Fact 1.4.10 where we stated most of the properties of a vector space for $\mathbb{R}^n$, though we did not use that language.) $\square$

Example 2.3.4 Let $\mathbb{F}$ be a field and let $\mathbb{F}^n$ be defined in the following way:

$$\mathbb{F}^n = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \;\middle|\; a_i \in \mathbb{F} \right\}.$$

The operations of scalar multiplication and vector addition are defined for $\mathbb{F}^n$ over $\mathbb{F}$ in the same way that the operations of scalar multiplication and vector addition are defined for $\mathbb{R}^n$ over $\mathbb{R}$. Then $\mathbb{F}^n$ is a vector space.

This should be relatively easy for the reader to believe, but checking the details could be a helpful exercise in the definitions of both fields and vector spaces. $\square$

Example 2.3.5 Let A be a subset of the real numbers. Define the set V as the set of all functions $A \to \mathbb{R}$. Given appropriate operations, this is a vector space.

If f and g are elements of V, then we define the sum as

$$(f+g)(t) = f(t) + g(t)$$

for all $t \in A$. Additionally, if $c \in \mathbb{R}$ and $f \in V$, then we define scalar multiplication as

$$(cf)(t) = cf(t)$$

for all $t \in A$. We will now argue that V is a vector space.

Looking back at Definition 2.3.1, the first two axioms hold by the way we defined the operations. Axioms three and four hold because addition in $\mathbb{R}$ is both commutative and associative. (Since these vectors in V are functions which take values in $\mathbb{R}$, the properties of $\mathbb{R}$ are once again crucial.)

The zero vector in V is the function $f(t) = 0$, the function which is uniformly zero for each value $t \in A$. This function has the properties of the zero vector mandated by axiom five.

The additive inverse of a function $g \in V$ is the function $(-1)g = -g$ since

$$(g + (-g))(t) = g(t) + (-g)(t) = g(t) - g(t) = 0$$

for all $t \in A$. This shows that axiom six holds.

The rest of the axioms hold because of the definitions of vector addition and scalar multiplication and the properties of the real numbers. □

Note 2.3.6 The previous example can be difficult to digest, because we are considering functions to be vectors. It may take some adjustment to think of functions—rather than single real numbers (or even ordered lists of real numbers)—as the objects of study .

Example 2.3.7 Let P_n denote the set of all real-valued polynomials of degree less than or equal to n with real coefficients. (Recall that the *degree* of a polynomial is the largest exponent of the variable that has a nonzero coefficient.) This means that the polynomial $p(x) = 7x^{10} - \frac{1}{2}x^6 + 9$ has degree 10 and is an element of P_n for all $n \geq 10$.

We consider vector addition to be the usual algebraic sum of polynomials and scalar multiplication to be the usual product of a polynomial by a constant. Then P_n is a vector space over $\mathbb{R}$. □

Example 2.3.8 Consider the set $V = \mathbb{R}^2$ but with non-standard operations. (In this example and the next, we will use horizontal instead of vertical notation for $\mathbb{R}^2$. This is purely for ease of notation in these limited instances.) We define the sum of two vectors $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ to be

$$\mathbf{u} \oplus \mathbf{v} = (u_1 v_1, 1),$$

and we define scalar multiplication of a vector $\mathbf{u} = (u_1, u_2)$ by a real number c to be

$$c \odot \mathbf{u} = (cu_1, 1).$$

It is a good exercise to grapple with the axioms and determine whether or not V with these operations is a vector space.

(We are using the symbols $\oplus$ and $\odot$ because these operations we are defining are not familiar ones, and most readers likely have no previous associations with these symbols. For our purposes, these symbols have no overarching meaning; they will be redefined for the purposes of specific examples.)

With some work, the reader will find that this is *not* a vector space. All we need to do in order to show this is not a vector space is to find one axiom that does not hold. We will show that the fifth axiom (regarding the zero vector) does not hold.

We must be careful here, because the terminology can become confusing. By "zero vector" we do not always mean in $\mathbb{R}^2$ the ordered pair $(0, 0)$. Depending on the operations that are in view, the zero vector may exist in a form other than $(0, 0)$. To show this axiom does not hold, we must argue that *no* element of $\mathbb{R}^2$ can have the properties of a zero vector given these operations.

We will argue by contradiction. Suppose there is a vector $\mathbf{u} = (u_1, u_2)$ which has the properties of a zero vector. This means that, for any $\mathbf{v} \in V$, we have $\mathbf{u} \oplus \mathbf{v} = \mathbf{v} \oplus \mathbf{u} = \mathbf{v}$. We consider the vector $\mathbf{v} = (2, 2)$. If $\mathbf{u}$ is the zero vector, then we must have

$$(u_1, u_2) \oplus (2, 2) = (2, 2),$$

but the addition on this set means

$$(u_1, u_2) \oplus (2, 2) = (2u_1, 1).$$

Since it is impossible to have $(2, 2) = (2u_1, 1)$ as the second coordinates are not equal as elements of $\mathbb{R}$, we have a contradiction, and therefore there can be no zero vector in V with these operations. □

Example 2.3.9 We again consider the set $V = \mathbb{R}^2$. Vector addition will be the usual sum in $\mathbb{R}^2$, but scalar multiplication will be different. For any real number c and any $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$, we define $c \odot \mathbf{u}$ by

$$c \odot \mathbf{u} = (cu_1, \tfrac{1}{2}cu_2).$$

Once again we pose the question: Is this a vector space?

While it is good practice to check all of the axioms, again we only need to find one axiom which does not hold in order to show this is not a vector space. (If it were a vector space, we would need to prove that all of the axioms hold.) We will show that scalar multiplication by 1 is not an identity (axiom 7).

We will examine the scalar product $1 \odot (1, 1)$:

$$1 \odot (1, 1) = (1, \tfrac{1}{2}).$$

If axiom 7 were to hold, we would have $1 \odot (1, 1) = (1, 1)$. Since $(1, 1) \neq (1, \frac{1}{2})$, we have shown that axiom 7 does not hold. Therefore, this is not a vector space. □

Example 2.3.10 If we fix positive integers m and n and a field $\mathbb{F}$, then $M_{m,n}(\mathbb{F})$ is the set of all $m \times n$ matrices with entries in $\mathbb{F}$. When $m = n$, we use the notation $M_n(\mathbb{F})$.

With the standard addition and scalar multiplication of matrices (see Note 1.4.8 for a description), the set $M_{m,n}(\mathbb{F})$ is a vector space over $\mathbb{F}$ for a fixed m and n. The quick explanation here is that because of the way scalar multiplication and addition of matrices are defined, once a reader believes that $\mathbb{F}^n$ is a vector space (see Example 2.3.4), believing that $M_{m,n}(\mathbb{F})$ is a vector space just amounts to considering several columns at the same time. But since these columns have no interaction with each other as far as these operations are concerned, none of the vector space axioms will be violated. The zero vector for this vector space is the zero matrix. □

As evidence of the fact that $\mathbb{R}^n$ is the model for our general definition of a vector space, we now repeat the definitions of *linear combination* and *span*. These definitions in the general setting will be necessaary for the following section. (See the original definitions in Definition 1.4.6.)

Definition 2.3.11 Let $\mathbb{F}$ be a field and let $c_1, c_2, \ldots, c_m \in \mathbb{F}$. Let V be a vector space over $\mathbb{F}$ and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in V$. Then the **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ with **weights** $c_1, c_2, \ldots, c_m$ is

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m.$$

The **span** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ is the set of all possible linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ and is written $\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. In other words, the span is defined to be

$$\mathrm{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = \left\{ \sum_{i=1}^{m} c_i \mathbf{v}_i \mid c_1, \ldots, c_m \in \mathbb{F} \right\}.$$

◊

The final result in this section is similar to Theorem 2.1.8 in which we summarize some of the "obvious" facts which are true in any vector space. We will prove a few of these facts here and leave some others as exercises.

Theorem 2.3.12 *Let V be a vector space over a field $\mathbb{F}$. Then all of the following hold.*

1. *The zero vector in V is unique.*
2. *Additive inverses of vectors in V are unique.*
3. *For every $\mathbf{v} \in V$, $-(-\mathbf{v}) = \mathbf{v}$.*
4. *For every $\mathbf{v} \in V$, $0\mathbf{v} = \mathbf{0}$.*
5. *For every $c \in \mathbb{F}$, $c\mathbf{0} = \mathbf{0}$.*
6. *For every $\mathbf{v} \in V$, $(-1)\mathbf{v} = -\mathbf{v}$.*
7. *If $c \in \mathbb{F}$, $\mathbf{v} \in V$, and $c\mathbf{v} = \mathbf{0}$, then either $c = 0$ or $\mathbf{v} = \mathbf{0}$.*

Proof. We will first prove property 2. Suppose that $\mathbf{u} \in V$ and that both $\mathbf{v}, \mathbf{w} \in V$ have the properties of being additive inverses of $\mathbf{u}$. We will show that $\mathbf{v} = \mathbf{w}$.
Since $\mathbf{u} + \mathbf{v} = \mathbf{0}$ and $\mathbf{u} + \mathbf{w} = \mathbf{0}$, we have

$$\mathbf{u} + \mathbf{v} = \mathbf{u} + \mathbf{w}.$$

Adding $\mathbf{v}$ to both sides (and using both the associative law and the properties of the zero vector) we have

$$\begin{aligned} (\mathbf{v} + \mathbf{u}) + \mathbf{v} &= (\mathbf{v} + \mathbf{u}) + \mathbf{w} \\ \mathbf{0} + \mathbf{v} &= \mathbf{0} + \mathbf{w} \\ \mathbf{v} &= \mathbf{w}. \end{aligned}$$

We will now show that property 4 holds. Since $0 = 0 + 0$ in $\mathbb{F}$, we can use the distributive property (axiom 10 in Definition 2.3.1) to find the following:

$$0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}.$$

We have $0\mathbf{v} = 0\mathbf{v} + 0\mathbf{v}$, and subtracting $0\mathbf{v}$ from both sides gives us $\mathbf{0} = 0\mathbf{v}$. (The reader should recall that by "subtracting" we mean adding the additive inverse, which every vector in a vector space possesses.)
To prove property 6, we must show that $(-1)\mathbf{v}$ has the properties of the additive inverse of $\mathbf{v}$. (Then, by property 2, we can conclude what we want.) We will use the distributive law (axiom 10 again) as well as the just-proved property 4:

$$\begin{aligned} \mathbf{v} + (-1)\mathbf{v} &= 1\mathbf{v} + (-1)\mathbf{v} \\ &= (1 - 1)\mathbf{v} \\ &= 0\mathbf{v} = \mathbf{0}. \end{aligned}$$

■

Reading Questions

1. In this question we will explore $\mathbb{F}_5^3$. (See Example 2.3.4. The vector space $\mathbb{F}_5^3$ should be thought of as being the same as $\mathbb{R}^3$ except that the entries and the operations are coming from $\mathbb{F}_5$.)

(a) Write down two nonzero elements of $\mathbb{F}_5^3$ and compute their sum.

(b) Write down one nonzero element of $\mathbb{F}_5^3$ and compute the scalar multiple of this vector by the element 3 of $\mathbb{F}_5$.

2. Write down one element of P_4. (See Example 2.3.7.)

3. Consider the vector $\mathbf{u} = (1, 3)$ in $\mathbb{F}_5^2$. (Again, we are writing the vector horizontally for convenience.) How many elements does $\text{Span}\{\mathbf{u}\}$ contain? Write all of them down.

Exercises

1. For each of the following, determine whether the given subset $V \subseteq \mathbb{R}^2$ is a real vector space, using the usual operations of vector addition and scalar multiplication in $\mathbb{R}^2$.

(a) $V = \{(x, y) \mid x, y \geq 0\}$

(b) $V = \{(x, y) \mid xy \geq 0\}$

(c) $V = \{(x, y) \mid y = x\}$

2. Consider the set $V = \mathbb{F}_5^2$ with the usual scalar multiplication but with vector addition defined this way:

$$(u_1, v_1) \oplus (u_2, v_2) = (u_1 + u_2, v_1 + v_2 + 1).$$

This is not a vector space over $\mathbb{F}_5$. Determine which of the vector space axioms fail, and give an explanation for each such axiom.

3. Consider the set $V = \mathbb{Q}^2$ with the usual vector addition but with scalar multiplication defined this way:

$$c \odot (x, y) = (0, cy).$$

This is not a vector space over $\mathbb{Q}$. Determine which of the vector space axioms fail, and give an explanation for each such axiom.

4. Consider the following subset of $\mathbb{R}^2$:

$$V = \{(x, y) \mid y = 4x\}.$$

Prove that V, with the usual operations of vector addition and scalar multiplication in $\mathbb{R}^2$, is a vector space over $\mathbb{R}$. (Hint: Since V is a subset of $\mathbb{R}^2$ with the same operations as $\mathbb{R}^2$, some of the axioms may not need to be proved from scratch.)

5. Consider the following three elements of P_3:

$$\begin{aligned} p_1 &= 2t^2 - t + 6 \\ p_2 &= -t^3 - t^2 + 4t \\ p_3 &= 4t^3 + \tfrac{1}{2}t^2 + 2t + 1. \end{aligned}$$

Calculate the linear combination of p_1, p_2, and p_3 with weights 2, -3, and

-2, respectively.

6. Consider the following three elements of $\mathbb{F}_3^2$:

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Calculate the linear combination of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ with weights 1, 2, and 2, respectively.

7. Consider the vector space $\mathbb{F}_2^3$.

(a) How many distinct vectors are in $\mathbb{F}_2^3$?

(b) If $\mathbf{u}$ and $\mathbf{v}$ are the following vectors,

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix},$$

how many distinct vectors are in $\operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$?

8. Consider the vector space V of all functions $\mathbb{R} \to \mathbb{R}$. (This general setting was introduced in Example 2.3.5.) Let f and g be the following vectors in V:

$$f(t) = \sin^2(t), \quad g(t) = \cos^2(t).$$

Is the constant function $h(t) = 4$ an element of $\operatorname{Span}\{f, g\}$? Explain your answer.

9. Consider the following vectors in $\mathbb{F}_5^3$:

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

Is $\mathbf{w} \in \operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$? Explain your answer.

10. Consider the following vectors in $\mathbb{C}^3$:

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

Is $\mathbf{w} \in \operatorname{Span}\{\mathbf{u}, \mathbf{v}\}$? Explain your answer.

Writing Exercises

11. Is $\mathbb{R}$ a vector space over $\mathbb{Q}$? (Addition and scalar multiplication should be understood as the obvious operations in $\mathbb{R}$.) Explain your answer.

12. Is $\mathbb{Q}$ a vector space over $\mathbb{R}$? (Addition and scalar multiplication should be understood as the obvious operations in $\mathbb{Q}$.) Explain your answer.

13. Prove that $\mathbb{C}$ is a vector space over $\mathbb{R}$. (Addition and scalar multiplication should be understood as the obvious operations in $\mathbb{C}$.)

14. Prove that every vector space over $\mathbb{C}$ is a vector space over $\mathbb{R}$.

15. Prove Item 1 in Theorem 2.3.12.

16. Prove Item 5 in Theorem 2.3.12.

17. Prove Item 7 in Theorem 2.3.12.

2.4 Subspaces

Whenever we begin to deal with abstraction in mathematics, we also consider the relationship between these abstract concepts. We have already made it clear how the notions of fields and vector spaces are related—a vector space *requires* that we know what a field is! But mathematicians also consider how these abstract objects might fit inside or contain each other.

The most basic example of this is something with which all readers are likely familiar. When we learn about *sets*, we soon also learn about *subsets*. For two sets A and B, we say that A is a subset of B when all of the elements of A are contained in B. We denote this by $A \subseteq B$.

A vector space is a set with a lot more structure. So a "sub-vector space" must be a subset with the properties of a vector space. Here is the formal definition.

Definition 2.4.1 Suppose that V is a vector space over the field $\mathbb{F}$. Then a subset $W \subseteq V$ is a **subspace** of V if W is also a vector space over $\mathbb{F}$ with the same operations of vector addition and scalar multiplication that are used for V. ♢

To put this less formally: A subspace is a subset which is a vector space over the same field with the operations *inherited* from the larger space.

Since the operations of a subspace are inherited, we need not check the full list of properties (from Definition 2.3.1) to prove that a subset is a subspace. In fact, when we consult that definition, we see that axioms 3, 4, and 7–10 will automatically be satisfied—these are properties of the operations, not of the set on which the operation is taking place. This leads to the following fact.

Fact 2.4.2 *Let V be a vector space over the field $\mathbb{F}$ and let W be a subset of V. Then W is a subspace of V if the following properties hold.*

- *For all $\mathbf{u}, \mathbf{v} \in W$, $\mathbf{u} + \mathbf{v} \in W$. (The subset W is closed under addition.)*
- *For all $c \in \mathbb{F}$ and all $\mathbf{v} \in W$, $c\mathbf{v} \in W$. (The subset W is closed under scalar multiplication.)*
- *The zero vector $\mathbf{0}$ of V belongs to W.*
- *For each $\mathbf{u} \in W$ there exists a vector $\mathbf{v} \in W$ such that $\mathbf{u} + \mathbf{v} = \mathbf{0}$. (The subset W contains all additive inverses.)*

From this fact, there appear to be four conditions to check in order to prove that a subset is a subspace. However, we will refer the reader back to Theorem 2.3.12. By part 6 of that theorem, we know that if a subset of a vector space is closed under scalar multiplication it must also contain all additive inverses. This means that there are only three conditions to check to prove that a subset is a subspace. We summarize this as a theorem.

Theorem 2.4.3 *Let V be a vector space over the field $\mathbb{F}$ and let W be a subset of V. Then W is a subspace of V if the following properties hold.*

1. *For all $\mathbf{u}, \mathbf{v} \in W$, $\mathbf{u} + \mathbf{v} \in W$.*
2. *For all $c \in \mathbb{F}$ and all $\mathbf{v} \in W$, $c\mathbf{v} \in W$.*
3. *The zero vector $\mathbf{0}$ of V belongs to W.*

As always, some examples are in order!

Example 2.4.4 Every vector space V has a subspace consisting of a single vector—the **zero subspace**: $\{\mathbf{0}\}$. In a trivial way, this set meets all of the conditions listed in Theorem 2.4.3. $\square$

Example 2.4.5 Let V be the vector space of all functions $\mathbb{R} \to \mathbb{R}$ as defined in Example 2.3.5. Let W be the subset of all polynomials. (This means we include polynomials of any degree.) Then W is a subspace of V.

Since the sum of two polynomials is another polynomial and the scalar multiple of a polynomial is a polynomial, the first two conditions in Theorem 2.4.3 hold. Additionally, the zero function in V is the same as the zero polynomial in W. This proves that W is a subspace of V. $\square$

Example 2.4.6 The vector space $\mathbb{R}^2$ is *not* a subspace of the vector space $\mathbb{R}^3$. This may seem surprising, as the operations for these spaces are clearly compatible, and we often think of $\mathbb{R}^2$ as "living inside" of $\mathbb{R}^3$.

However, this common way of thinking is wrong because the space $\mathbb{R}^2$ is not even a subset of $\mathbb{R}^3$; thus, it is impossible for $\mathbb{R}^2$ to be a subspace of $\mathbb{R}^3$. (The set $\mathbb{R}^2$ consists of ordered pairs of real numbers, and $\mathbb{R}^3$ consists of ordered triples of real numbers.)

We can make a slight adjustment to match the way that many people think of $\mathbb{R}^2$ as "living inside" of $\mathbb{R}^3$. We define the set A in the following way:

$$A = \{(x, y, 0) \in \mathbb{R}^3\}.$$

Then A is what we usually call the "xy-plane" in $\mathbb{R}^3$. This subset A *is* a subspace of $\mathbb{R}^3$. $\square$

Example 2.4.7 Here is a collection of examples that generalize Example 2.4.6. Every line through the origin in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$, and every line or plane through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$. On the other hand, lines in $\mathbb{R}^2$ which do *not* pass through the origin are *not* subspaces of $\mathbb{R}^2$, and lines and planes which do *not* pass through the origin in $\mathbb{R}^3$ are *not* subspaces of $\mathbb{R}^3$. The details are left for the exercises. $\square$

The notion of the span of a set of vectors gives us another angle through which we can identify subspaces. We begin with a result.

Theorem 2.4.8 *Let V be a vector space over a field $\mathbb{F}$. If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are vectors in V, then $\text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a subspace of V.*

Proof. We let $W = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. By the definition of the span of a set of vectors, every element $\mathbf{w}$ of W can be written in the following form:

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n,$$

where $c_1, \ldots, c_n \in \mathbb{F}$. We first observe that $\mathbf{0} \in W$ by taking $c_1 = \cdots = c_n = 0$. Next, if $\mathbf{u}$ and $\mathbf{w}$ are elements of W, we can write these vectors as

$$\begin{aligned} \mathbf{u} &= c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \\ \mathbf{w} &= d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n, \end{aligned}$$

for some scalars c_i and d_i. The sum of these elements is

$$\begin{aligned} \mathbf{u} + \mathbf{w} &= (c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) + (d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n) \\ &= (c_1 + d_1)\mathbf{v}_1 + \cdots + (c_n + d_n)\mathbf{v}_n. \end{aligned}$$

(We are using some of the properties of a vector space from Definition 2.3.1 in order to carry out this algebraic manipulation.) Since $c_i + d_i \in \mathbb{F}$ for each i, this proves that $\mathbf{u} + \mathbf{w} \in W$.

Finally, we let $c \in \mathbb{F}$ and $\mathbf{w} \in W$. We want to show that $c\mathbf{w} \in W$. We can assume that $\mathbf{w}$ has the form

$$\mathbf{w} = d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n,$$

where $d_1, \ldots, d_n \in \mathbb{F}$. Then we have

$$\begin{aligned} c\mathbf{w} &= c(d_1\mathbf{v}_1 + \cdots + d_n\mathbf{v}_n) \\ &= c(d_1\mathbf{v}_1) + \cdots + (cd_n\mathbf{v}_n) \\ &= (cd_1)\mathbf{v}_1 + \cdots + (cd_n)\mathbf{v}_n. \end{aligned}$$

(Again, we are using properties of the vector space V here.) Since $cd_i \in \mathbb{F}$ for each i, this proves that W is closed under scalar multiplication.
Since W has all of the properties from Theorem 2.4.3, we have shown that W is a subspace of V. ■

In the final example of this section, we will use Theorem 2.4.8 to prove that a set is a subspace by realizing it as the span of a set of vectors.

Example 2.4.9 Consider the following subset of $\mathbb{R}^3$:

$$A = \left\{ \begin{bmatrix} 2a+b \\ 5b - \frac{1}{2}a \\ 6a \end{bmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}.$$

This notation means that every given pair of real numbers a and b specifies an element of A. For example, when $a = 1$ and $b = -1$, we have $(1, -\frac{11}{2}, 6) \in A$.

We will use Theorem 2.4.8 to prove that A is a subspace of $\mathbb{R}^3$. Any element $\mathbf{v} \in A$ can be written in the following way:

$$\mathbf{v} = \begin{bmatrix} 2a+b \\ 5b - \frac{1}{2}a \\ 6a \end{bmatrix} = a \begin{bmatrix} 2 \\ -\frac{1}{2} \\ 6 \end{bmatrix} + b \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}.$$

This proves that if

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ -\frac{1}{2} \\ 6 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix},$$

then $A = \operatorname{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. We conclude that A is a subspace of $\mathbb{R}^3$ by Theorem 2.4.8. □

Reading Questions

1. Let V be the vector space $\mathbb{R}^2$ with the usual addition and scalar multiplication. Let A be the following set:

$$A = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \text{ and } y \geq 0\}.$$

(We can identify A with the first quadrant in $\mathbb{R}^2$.) Determine whether or not A is a subspace of V. Explain your answer.

2. Consider the vector space P_4 and let A be the following set:

$$A = \{\text{polynomials in } P_4 \text{ with even degree}\}.$$

Determine whether or not A is a subspace of P_4. Explain your answer.

3. Let V be the vector space of all functions $\mathbb{R} \to \mathbb{R}$. Let f and g be the following functions:

$$f(t) = 2t^2 \qquad g(t) = \cos(3t).$$

Let $A = \mathrm{Span}\{f, g\}$. We know that A is a subspace of V by Theorem 2.4.8. Write down four distinct vectors in A.

Exercises

1. For each of the following, determine whether or not the subset W of $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$. Explain your answers. (We are writing vectors horizontally for convenience.)

(a) $W = \{(a, 0, 0) \mid a \in \mathbb{R}\}$

(b) $W = \{(a, b, c) \mid b = a + c\}$

(c) $W = \{(a, 1, c) \mid a, c \in \mathbb{R}\}$

(d) $W = \{(a, 0, c) \mid a, c \in \mathbb{R}\}$

2. For each of the following, determine whether or not the subset W of P_3 is a subspace of P_3. Explain your answers.

(a) $W = \{at^3 + at \mid a \in \mathbb{R}\}$

(b) $W = \{at^2 + bt + c \mid a, b, c \in \mathbb{R}\}$

(c) $W = \{p \in P_3 \mid p \text{ has rational coefficients}\}$

(d) $W = \{p \in P_3 \mid p(0) = 0\}$

(e) $W = \{a_3t^3 + a_2t^2 + a_1t + a_0 \mid a_3 + a_2 + a_1 + a_0 = 0\}$

3. Let $C[0, 10]$ be the vector space of continuous functions defined $[0, 10] \to \mathbb{R}$, and let $D[0, 10]$ be the set of differentiable functions defined $[0, 10] \to \mathbb{R}$.

For each of the following, determine whether or not the subset W of $C[0, 10]$ is a subspace of $C[0, 10]$. Explain your answers.

(a) $W = \{f \in C[0, 10] \mid f(1) = 0\}$

(b) $W = \{f \in C[0, 10] \mid f(1) = 1\}$

(c) $W = \{f \in D[0, 10] \mid f'(1) = 0\}$

(d) $W = \{f \in D[0, 10] \mid f'(1) = 1\}$

(e) $W = \{f \in D[0, 10] \mid f'(t) \text{ is constant}\}$

4. For each of the following, determine whether or not the subset W of $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$. Explain your answers. (We are writing vectors horizontally for convenience.)

(a) $W = \{(2a - b, a + 5b, -a) \mid a, b \in \mathbb{R}\}$

(b) $W = \{(a - b + 1, 2a + 3b, 0) \mid a, b \in \mathbb{R}\}$

(c) $W = \{(a + b, b - c, a + 2d) \mid a, b, c, d \in \mathbb{R}\}$

(d) $W = \{(2a - 4b, 5, 3c + b) \mid a, b, c \in \mathbb{R}\}$

Writing Exercises

5. Let V denote the vector space of all functions $\mathbb{R} \to \mathbb{R}$. (See Example 2.3.5 for the relevant definitions and operations.) Let a and b be real numbers with $a < b$, and let $C[a, b]$ denote the set of all continuous functions $[a, b] \to \mathbb{R}$. Prove that $C[a, b]$ is a subspace of V.

6. Let a and b be real numbers with $a < b$. Show that the set of all functions $f \in C[a, b]$ such that

$$\int_a^b f(x)\, dx = 0$$

is a subspace of $C[a, b]$. (See Exercise 2.4.5 for the definition of $C[a, b]$.)

7. Prove that a line in $\mathbb{R}^2$ is a subspace of $\mathbb{R}^2$ if and only if the line passes through the origin. (Note that all lines in $\mathbb{R}^2$ can be written in the form $ax + by + c = 0$.)

8. Prove that a plane in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$ if and only if the plane passes through the origin. (Note that all planes in $\mathbb{R}^3$ can be written in the form $ax + by + cz + d = 0$.)

9. Let V be a vector space and let W_1 and W_2 be subspaces of V. Must $W_1 \cup W_2$ be a subspace of V? Justify your answer.

10. Let V be a vector space and let W_1 and W_2 be subspaces of V. Must $W_1 \cap W_2$ be a subspace of V? Justify your answer.

11. Let V be a vector space and let W_1 and W_2 be subspaces of V. Define the *sum* of W_1 and W_2 like this:

$$W_1 + W_2 = \{\mathbf{v} \mid \mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 \text{ for some } \mathbf{w}_1 \in W_1 \text{ and some } \mathbf{w}_2 \in W_2\}.$$

Must $W_1 + W_2$ be a subspace of V? Justify your answer.

Chapter 3

Linear Transformations

3.1 Linear Transformations

Speaking broadly, mathematicians are often concerned about (mathematical) objects and the right sort of functions between those objects. The structure of specific objects can be illuminated by a look at the functions to and from those objects. In linear algebra, the objects in view are vector spaces (see Definition 2.3.1), and the functions between these objects are called *linear transformations*.

3.1.1 Introduction to Linear Transformations

Definition 3.1.1 If V and W are vector spaces over a field $\mathbb{F}$, then a function $T : V \to W$ is called a **linear transformation** if both of the following properties hold.

- For all $\mathbf{u}, \mathbf{v} \in V$, we have $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$.
- For all $\mathbf{v} \in V$ and all $c \in \mathbb{F}$, we have $T(c\mathbf{v}) = cT(\mathbf{v})$.

These functions are sometimes referred to as *linear maps* or *linear operators*.

If $T : V \to W$ is a linear transformation, then V is the **domain** of T and W is the **codomain** of T. ◊

Note 3.1.2 Many readers will be more familiar with the idea of the *range* of a function than the *codomain* of a function. The **range** of a linear transformation $T : V \to W$ is the set $\{T(\mathbf{v}) \in W \mid \mathbf{v} \in V\}$. In words, the range is the subset of the codomain consisting of the "outputs" of the function for all elements of the domain. We will often use the term **image** when discussing the range of a linear transformation.

Linear transformations are the "right" types of functions to study between vector spaces because they preserve the primary vector space operations. The first property of linear transformations means that such a function respects vector addition, and the second property means that such a function respects scalar multiplication.

Example 3.1.3 We consider the real vector spaces P_5 and P_4, along with the function $D : P_5 \to P_4$ which takes the derivative. That is, $D(p) = p'$ for all $p \in P_5$. So if $p = 3t^5 - 2t^3 + 10t$, then $D(p) = 15t^4 - 6t^2 + 10$. We note that $p \in P_5$ and $D(p) \in P_4$.

The fact that our function D is a linear transformation between these vector spaces is a consequence of calculus. For all differentiable functions f and g, and all real numbers c, it is true that

$$[f+g]' = f' + g'$$
$$[cf]' = cf'.$$

(If the reader doubts or has forgotten these facts, the closest textbook on single-variable calculus should be consulted *posthaste*.)

These calculus facts confirm that $D(p+q) = D(p)+D(q)$ and $D(cp) = cD(p)$ for all $p, q \in P_5$ and all $c \in \mathbb{R}$. This proves that $D : P_5 \to P_4$ is a linear transformation. □

Example 3.1.4 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the function which reflects a vector in the Cartesian plane across the x-axis. So $T(x, y) = (x, -y)$. Additionally, let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the function which rotates a vector counter-clockwise around the origin by $\frac{\pi}{2}$ radians. So $S(x, y) = (-y, x)$. Then both T and S are linear transformations.

We will supply two calculations here to give the sense of these functions. The reader should note that T takes the vector $(-3, 2)$ in the second quadrant and reflects it across the x-axis to the vector $(-3, -2)$ in the third quadrant. Also, S rotates the vector $(-3, 2)$ counter-clockwise around the origin by $\frac{\pi}{2}$ radians to the vector $(-2, -3)$. (It is fairly obvious that the length of the vectors $(-3, 2)$ and $(-2, -3)$ are the same. To check the claim about the angles, one would calculate the angles between the positive x-axis and both the vectors $(-3, 2)$ and $(-2, -3)$. The first angle is roughly 2.55 radians and the second is 4.12, giving a difference of 1.57 radians, or roughly $\frac{\pi}{2}$.)

We first check the additivity condition. Let $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$. Then we have

$$T((x_1, y_1) + (x_2, y_2)) = T(x_1 + x_2, y_1 + y_2) = (x_1 + x_2, -(y_1 + y_2))$$
$$T(x_1, y_1) + T(x_2, y_2) = (x_1, -y_1) + (x_2, -y_2) = (x_1 + x_2, -y_1 - y_2).$$

From the distributive property of the real numbers (in the second coordinate of these calculations), we can see that the additive property holds for T. (The calculation for S is similar.)

We now check the scalar multiplication property. (Again, the calculations for T and S are similar, so we will only show one of them.) Let $c \in \mathbb{R}$ and let $(x, y) \in \mathbb{R}^2$. Then we have

$$S(c(x, y)) = S(cx, cy) = (-cy, cx)$$
$$cS(x, y) = c(-y, x) = (-cy, cx).$$

Note that we used the commutativity of multiplication in $\mathbb{R}$ in this calculation.

These brief calculations show that both T and S are linear transformations. □

3.1.2 Linear Transformations and Matrices

While linear algebra is not *only* about matrices, matrices are valuable tools and provide a rich source of examples in this subject. In fact, matrices are so central to the notion of linear transformations that we will devote this subsection to their discussion.

Example 3.1.5 Let $\mathbb{F}$ be a field and let A be an $m \times n$ matrix with entries from $\mathbb{F}$. (We will refer to this in what follows as "a matrix over $\mathbb{F}$.") Then multiplication by A is a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$. (We will denote the function which is multiplication by A by $T_A : \mathbb{F}^n \to \mathbb{F}^m$.)

To justify this claim we must first explain what we mean by "multiplication by A." We will let $\mathbf{v} \in \mathbb{F}^n$ and denote entry (i, j) in A by a_{ij}. We will further denote the entries of $\mathbf{v}$ by

$$\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

Then the **matrix-vector product** $A\mathbf{v}$ is defined to be the following vector in $\mathbb{F}^m$:

$$A\mathbf{v} = \begin{bmatrix} a_{11}v_1 + \cdots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \cdots + a_{mn}v_n \end{bmatrix}. \tag{3.1}$$

One way to state this is that entry j in $A\mathbf{v}$ is the sum of the entry-wise product of row j in A with $\mathbf{v}$. Since $A\mathbf{v}$ is an element of $\mathbb{F}^m$, the domain and codomain of T_A are correct.

What we have defined is the product of a matrix and a vector. However, an alternate description of this product will be more useful in proving that T_A is a linear transformation.

If the columns of A are thought of as vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$, then the product $A\mathbf{v}$ is also

$$A\mathbf{v} = v_1\mathbf{a}_1 + \cdots + v_n\mathbf{a}_n = \sum_{i=1}^{n} v_i\mathbf{a}_i. \tag{3.2}$$

In words, $A\mathbf{v}$ is a linear combination of the columns of A with weights coming from the entries of $\mathbf{v}$. (We have reserved proving the equivalence of these two formulations to Exercise 3.1.7.16.)

With this equivalent definition, proving that T_A is a linear transformation is a snap. Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{F}^n$ and let $c \in \mathbb{F}$. We will further denote the entries of $\mathbf{u}$ and $\mathbf{v}$ by

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

Then we have the following:

$$\begin{aligned} T_A(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) &= \sum_{i=1}^{n} (u_i + v_i)\mathbf{a}_i \\ T_A(\mathbf{u}) + T_A(\mathbf{v}) = A\mathbf{u} + A\mathbf{v} &= \sum_{i=1}^{n} u_i\mathbf{a}_i + \sum_{i=1}^{n} v_i\mathbf{a}_i. \end{aligned}$$

These two expressions are equal due to the fact that $\mathbb{F}^m$ is a vector space.

We have one final calculation to prove that T_A is a linear transformation. Let $\mathbf{v}$ be a vector in $\mathbb{F}^n$ and let c be in $\mathbb{F}$. Then we have

$$\begin{aligned} T_A(c\mathbf{v}) = A(c\mathbf{v}) &= \sum_{i=1}^{n} (cv_i)\mathbf{a}_i \\ cT_A(\mathbf{v}) = cA(\mathbf{v}) &= c\sum_{i=1}^{n} v_i\mathbf{a}_i. \end{aligned}$$

Once again, thsse expressions are equal because $\mathbb{F}^m$ is a vector space.

These calculations prove that T_A is a linear transformation. $\square$

Note 3.1.6 To summarize, when $\mathbb{F}$ is a field, multiplication by an $m \times n$ matrix A is a linear transformation $T_A : \mathbb{F}^n \to \mathbb{F}^m$.

General matrices are rectangular, not necessarily square. When a matrix is square, however, we have additional properties to discuss.

Definition 3.1.7 Let A be an $n \times n$ matrix. (So A is *square*.) We say that A is a **diagonal matrix** if $a_{ij} = 0$ for all (i, j) such that $i \neq j$. If A is diagonal and $a_{ii} = 1$ for all $i = 1, \ldots, n$, then A is called an **identity matrix**. $\Diamond$

Note 3.1.8 We often use the notation $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to refer to the columns of the $n \times n$ identity matrix. In other words, $\mathbf{e}_j$ is the vector with a 1 in entry j and zeros elsewhere.

3.1.3 Properties of Linear Transformations

Recall that while linear transformations must have special properties, they are first of all *functions*. And, as functions, properties like injectivity and surjectivity can apply to linear transformations.

Definition 3.1.9 Let $T : V \to W$ be a linear transformation between vector spaces. We say that T is **injective** if $T(\mathbf{v}_1) = T(\mathbf{v}_2)$ implies $\mathbf{v}_1 = \mathbf{v}_2$ for all $\mathbf{v}_1, \mathbf{v}_2 \in V$. Injective linear transformations are also referred to as *one-to-one* since no two distinct elements of the domain may correspond to the same element of the range.

A linear transformation T is called **surjective** if for every $\mathbf{w} \in W$ there exists a vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. For surjective functions, the image/range is the same as the codomain. (The range is a subset of the codomain for every function, but these sets are equal if and only if the function is surjective.) Sometimes surjective functions are referred to as *onto* functions.

If a linear transformation is both injective and surjective, we say that it is **bijective**. $\Diamond$

Example 3.1.10 Let's reconsider the linear transformation $D : P_5 \to P_4$ which appeared in Example 3.1.3. We observe that D is surjective but not injective.

The transformation is surjective because we know about the antiderivative. Let $q \in P_4$ have the form

$$q(t) = a_4t^4 + a_3t^3 + a_2t^2 + a_1t + a_0.$$

This is a generic element of P_4, so we only need to supply an element $p \in P_5$ such that $D(p) = q$, and this will prove that D is surjective. Consider the element p defined as

$$p(t) = \tfrac{1}{5}a_4t^5 + \tfrac{1}{4}a_3t^4 + \tfrac{1}{3}a_2t^3 + \tfrac{1}{2}a_1t^2 + a_0t.$$

It is but the work of a Calculus I student to verify that $D(p) = q$, thus showing that D is surjective. (We note that we could have chosen p to have any constant term at all; we used the constant term of 0.)

Finally, we will show that D is not injective by looking at an example of two elements of P_5 which have the same image under D in P_4. Let $p_1 = t^2 + 10$ and $p_2 = t^2 + 20$. Then we see that even though $p_1 \neq p_2$, we have $D(p_1) = D(p_2)$, and this proves that D is not injective. $\square$

We will define one more property of linear transformations here that will resurface in Section 3.2.

Definition 3.1.11 Let $T : V \to W$ be a linear transformation between vector spaces. The **identity transformation** on V is the linear transformation $I_V : V \to V$ which is $I_V(\mathbf{v}) = \mathbf{v}$ for each $\mathbf{v} \in V$. (If the vector space we have in mind is clear, we will drop the subscript and use the notation I.)

We say that the linear transformation T is **invertible** if there exists a linear transformation $S : W \to V$ such that $S \circ T = I_V$. $\Diamond$

3.1.4 Isomorphisms

Bijective functions are important in almost all settings, and the linear algebra setting is no exception. We have a specific name for bijective linear transformations.

Definition 3.1.12 A bijective linear transformation T between vector spaces V and W is called an **isomorphism**. If there exists an isomorphism between vector spaces V and W, then these spaces are said to be **isomorphic**. $\Diamond$

The reader should think of isomorphic vector spaces as *essentially the same*. Such spaces will not be exactly the same, of course, in the same way that two finite sets of the same size are not necessarily identical. But the presence of an isomorphism means that the vector space operations are compatible in such a way that such spaces share many of the same properties.

Note 3.1.13 If V and W are vector spaces, then the set of all linear transformations from $V \to W$ is denoted $L(V, W)$. When $W = V$, we will write $L(V)$ instead of $L(V, V)$.

We can now prove that two concepts we have defined in this section are one and the same for linear transformations.

Proposition 3.1.14 *Let V and W be vector spaces over $\mathbb{F}$, and let $T \in L(V, W)$. Then T is an isomorphism if and only if T is invertible.*

Proof. This fact is true for functions without any of the linear transformation properties being involved. (A function is bijective if and only if it has an inverse.) $\blacksquare$

Proposition 3.1.15 *If $T \in L(V, W)$ is invertible, then $T^{-1} : W \to V$ is also a linear transformation.*

Proof. We will check the two properties of a linear transformation. (See Definition 3.1.1.) Suppose that $\mathbf{w}_1, \mathbf{w}_2 \in W$. Since $T \circ T^{-1} = I_W$, we have

$$\mathbf{w}_1 + \mathbf{w}_2 = T(T^{-1}(\mathbf{w}_1)) + T(T^{-1}(\mathbf{w}_2)) = T(T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2)).$$

When we apply T^{-1} to the beginning and end of this equality, using $T^{-1} \circ T = I_V$, we get

$$T^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = T^{-1}(\mathbf{w}_1) + T^{-1}(\mathbf{w}_2).$$

We will now check the scalar multiple property in a similar fashion. Let $\mathbf{w} \in W$ and let $c \in \mathbb{F}$. Then we have

$$c\mathbf{w} = cT(T^{-1}(\mathbf{w})) = T(cT^{-1}(\mathbf{w})).$$

Applying T^{-1} to both sides again we get

$$T^{-1}(c\mathbf{w}) = cT^{-1}(\mathbf{w}).$$

This proves that $T^{-1} \in L(W, V)$. $\blacksquare$

Before we leave this subsection, it is worth pointing out that when V and W are vector spaces, the set $L(V, W)$ itself has some important structure.

When V and W are vector spaces over $\mathbb{F}$, we can define the sum and scalar multiple of linear transformations since both of these operations happen on the level of elements. If $S, T \in L(V, W)$ and $c \in \mathbb{F}$, then we define $S + T$ and cT in the following way. For all $\mathbf{v} \in V$,

- $(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$, and
- $(cT)(\mathbf{v}) = cT(\mathbf{v})$.

These operations as defined make $L(V, W)$ into its own vector space. We will leave the proof of this theorem for the exercises.

Theorem 3.1.16 *Let V and W be vector spaces over a field $\mathbb{F}$, and let $S, T \in L(V, W)$ and $c \in \mathbb{F}$. Then $S + T$ and cT are both linear transformations from V to W, and $L(V, W)$ is a vector space with these operations.*

3.1.5 The Matrix-Vector Form of a Linear System

Having defined the product of a matrix and a vector in Example 3.1.5, we can reformulate one of the foundational (and introductory) matters of this book. We will now put the notion of a linear system—in particular, the solutions to linear systems—in a different context.

Let's consider the following system of linear equations over a field $\mathbb{F}$, as we saw in Section 2.2:

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

If we let A be the matrix $A = [a_{ij}]$, $\mathbf{x}$ be the vector of variables $\mathbf{x} = [x_j]$, and $\mathbf{b}$ be the vector of constants $\mathbf{b} = [b_j]$, then this linear system can be written efficiently as $A\mathbf{x} = \mathbf{b}$.

With this reformulation, the questions of the existence and uniqueness of solutions to a system of equations (see the end of Section 1.1) can now be stated in the language of the injectivity and surjectivity of linear transformations.

Example 3.1.17 Consider the linear transformation $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ which is multiplication by this matrix:

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

We will show that T_A is neither injective nor surjective.

Let $\mathbf{u}$ and $\mathbf{v}$ be the following vectors in $\mathbb{R}^3$:

$$\mathbf{u} = \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 7 \\ 0 \\ 2 \end{bmatrix}.$$

By forming and row-reducing the augmented matrices $[A \mid \mathbf{u}]$ and $[A \mid \mathbf{v}]$, we can determine how many solutions there are to the equations $T_A(\mathbf{x}) = \mathbf{u}$ and $T_A(\mathbf{x}) = \mathbf{v}$, respectively. Here are the calculations:

$$\begin{bmatrix} A \mid \mathbf{u} \end{bmatrix} \sim \left[\begin{array}{ccc|c} 1 & 0 & 4 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \tag{3.3}$$

$$[A \mid \mathbf{v}] \sim \left[\begin{array}{ccc|c} 1 & 0 & 4 & 2 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]. \tag{3.4}$$

From (3.3), since there is a pivot in the final column of the RREF of $[A \mid \mathbf{u}]$, we see that $\mathbf{u}$ is not in the image of T_A. This means that the matrix equation $A\mathbf{x} = \mathbf{u}$ has no solution, so T_A is not surjective; equivalently, the linear system which corresponds to the augmented matrix $[A \mid \mathbf{u}]$ is inconsistent.

From (3.4), we see that $\mathbf{v}$ is in the image of T_A. Since there is no pivot in the final column of the RREF of $[A \mid \mathbf{v}]$, and since there is a free variable in that same RREF, this means that the matrix equation $A\mathbf{x} = \mathbf{v}$ has multiple solutions, so T_A is not injective. Specifically, if

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix},$$

then we have both $T_A(\mathbf{x}_1) = \mathbf{v}$ and $T_A(\mathbf{x}_2) = \mathbf{v}$. (The vector $\mathbf{x}_1$ results from setting the free variable equal to 0, and we obtain $\mathbf{x}_2$ by setting the free variable equal to 1.) Finally, we note that the linear system which corresponds to the augmented matrix $[A \mid \mathbf{v}]$ is consistent with many solutions—that is, a solution is not unique. □

3.1.6 Reading Questions

1. For each of the following, determine the number of rows and columns that a matrix would have if multiplication by that matrix is a linear transformation with the given domain and codomain.

(a) domain: $\mathbb{R}^2$, codomain: $\mathbb{R}^3$

(b) domain: $\mathbb{Q}^4$, codomain: $\mathbb{Q}^2$

2. Let A, $\mathbf{u}$, $\mathbf{b}$, and $\mathbf{c}$ be defined as follows:

$$A = \begin{bmatrix} 2 & 0 & 6 \\ 1 & 3 & 6 \\ -1 & 5 & 2 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 3 \\ 6 \\ 6 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} -1 \\ -4 \\ 3 \end{bmatrix}.$$

Define a linear transformation $T_A : \mathbb{R}^3 \to \mathbb{R}^3$ to be multiplication by A.

(a) Find $T_A(\mathbf{u})$.

(b) Find an $\mathbf{x}$ in $\mathbb{R}^3$ such that $T_A(\mathbf{x}) = \mathbf{b}$.

(c) Is there more than one $\mathbf{x}$ whose image under T_A is $\mathbf{b}$? How do you know?

(d) Determine whether or not $\mathbf{c}$ is in the range of T_A.

3.1.7 Exercises

1. Consider the function $T : P_2 \to P_3$ defined by $T(p) = tp$. (So, for example, $T(2+t) = 2t + t^2$.) Is T a linear transformation? Justify your answer.

2. Consider the function $T : P_2 \to P_2$ defined by $T(p) = p(0) + p(1)t + p(2)t^2$. Is T a linear transformation? Justify your answer.

3. Consider the function $T : P_2 \to P_1$ defined by $T(p) = p(0) + p'(0)t$. Is T a linear transformation? Justify your answer.

4. Let $T : \mathbb{F}_5^3 \to \mathbb{F}_5^2$ be the function defined by $T(x, y, z) = (3x + y - 2z, -xy)$. Is T a linear transformation? Justify your answer.

5. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the function defined by $T(x, y) = (2x, x - 3y, 0)$. Is T a linear transformation? Justify your answer.

6. Consider the following matrix over $\mathbb{F}_3$:

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

For each of the following vectors $\mathbf{v}$, calculate the matrix-vector product $A\mathbf{v}$.

(a) $\mathbf{v} = (1, 1, 0)$

(b) $\mathbf{v} = (2, 1, 2)$

(c) $\mathbf{v} = (0, 2, 1)$

7. Consider the following matrix over $\mathbb{R}$:

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

(a) If T is the linear transformation which is multiplication by A, what are the domain and codomain of T?

(b) Calculate the image of the vector $\mathbf{v} = (-3, 1, 4)$ under the linear transformation T.

(c) Is the vector $\mathbf{w} = (-2, -1)$ in the image of T? Explain your answer.

8. Let A be the following matrix over $\mathbb{R}$:

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

Let T be the linear transformation which is multiplication by A.

(a) Is the vector $(1, 1, 1)$ in the image of T? Explain your answer.

(b) Is T surjective? Explain your answer.

9. Let A be the following matrix over $\mathbb{F}_7$:

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

Let T be the linear transformation which is multiplication by A.

(a) Is the vector $(3, 1, 1)$ in the image of T? Explain your answer.

(b) The vector $\mathbf{w} = (5, 4, 0)$ is in the image of T. Find one $\mathbf{x} \in \mathbb{F}_7^3$ such that $T(\mathbf{x}) = \mathbf{w}$.

(c) Is there more than one $\mathbf{x} \in \mathbb{F}_7^3$ such that $T(\mathbf{x}) = \mathbf{w}$? How do you know?

(d) Is T injective? Is T surjective? Explain your answers.

10. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation which is multiplication by the following matrix:

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

Give a description of all vectors $\mathbf{x} \in \mathbb{R}^3$ such that $T(\mathbf{x}) = \mathbf{0}$.

Writing Exercises

11. Define the function $T : C[0, \infty) \to C[0, \infty)$ to be the following:

$$(T(f))(x) = \int_0^x f(y)\, dy.$$

Prove that T is a linear transformation.

12. Let $T : U \to V$ and $S : V \to W$ be linear transformations between vector spaces over a field $\mathbb{F}$. Prove that $S \circ T$ is also a linear transformation.

13. Let $T : U \to V$ and $S : V \to W$ be linear transformations between vector spaces over a field $\mathbb{F}$.

(a) Prove that if $S \circ T$ is injective, then T must be injective.

(b) Prove that if $S \circ T$ is surjective, then S must be surjective.

14. Let $T : V \to W$ be a function between vector spaces over $\mathbb{F}$.

(a) If T is a linear transformation, must it be true that $T(\mathbf{0}_V) = \mathbf{0}_W$? Either prove this is true or produce a counterexample.

(b) If $T(\mathbf{0}_V) = \mathbf{0}_W$, must T be a linear transformation? Either prove this is true or produce a counterexample.

15. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be vectors which span a vector space V. If $T : V \to V$ is a linear transformation for which $T(\mathbf{v}_i) = \mathbf{0}$ for all $i = 1, \ldots, m$, prove that T is the zero transformation. (In other words, prove that $T(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in V$.)

16. Let A be an $m \times n$ matrix over a field $\mathbb{F}$, and let $\mathbf{v}$ be a vector in $\mathbb{F}^n$. Prove that the formulations of the matrix-vector product given in (3.1) and (3.2) are equivalent.

17. Prove Theorem 3.1.16.

3.2 The Matrix of a Linear Transformation

As we saw in the previous section, linear transformations can be defined using matrices and they can also be defined with no matrices in sight. In this section we will see that, for a certain class of linear transformations, there is *always a matrix in sight.*

3.2.1 Constructing the Matrix

Our claim might seem fanciful at first. *Can every linear transformation be realized using a matrix?* The surprising answer is *yes*, for a specific kind of linear transformation.

We first make an observation related to the definition of the matrix-vector product in Example 3.1.5.

Note 3.2.1 If A is an $m \times n$ matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and if we recall the definition of $\mathbf{e}_j$ from Note 3.1.8, then

$$A(\mathbf{e}_j) = \mathbf{a}_j.$$

The truth of this equality comes by thinking of $A(\mathbf{e}_j)$ in the way expressed in (3.2), as a linear combination of the columns of A with weights from the entries in $\mathbf{e}_j$.

We now suppose that $\mathbb{F}$ is a field and that $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation. We claim that there is a unique $m \times n$ matrix A such that for every $\mathbf{v} \in \mathbb{F}^n$, $T(\mathbf{v}) = A\mathbf{v}$. In other words, we claim that the work of the linear transformation T can be carried out through multiplication by A.

We will define the matrix A which does the job. For each $j = 1, \ldots, n$, define the vector $\mathbf{a}_j$ by $\mathbf{a}_j = T(\mathbf{e}_j)$. We then define A as the matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_n$.

Since any vector $\mathbf{v} \in \mathbb{F}^n$, written as $\mathbf{v} = [v_i]$, has the property that

$$\mathbf{v} = \sum_{j=1}^{n} v_j \mathbf{e}_j,$$

we can verify that the action of T is the same as the action of multiplication by A:

$$T(\mathbf{v}) = T\left(\sum_{j=1}^{n} v_j \mathbf{e}_j\right) = \sum_{j=1}^{n} v_j T(\mathbf{e}_j) = \sum_{j=1}^{n} v_j \mathbf{a}_j = A\mathbf{v}.$$

Note that we used the fact that T is a linear transformation in this last string of equalities.

We have just proved the following theorem.

Theorem 3.2.2 *If $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation, then there exists a unique $m \times n$ matrix A over $\mathbb{F}$ such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} \in \mathbb{F}^n$.*

A scrupulous reader may protest our use of the word "unique" in the statement of this theorem. Here is the argument concerning uniqueness. If the theorem is true, then (for this theorem) there is only one way it could possibly work. If a matrix A exists, it must have the property that $A\mathbf{e}_j = T(\mathbf{e}_j)$ for all j. Since we have shown that such a construction does work, the matrix A we obtain must be unique.

This theorem is quite powerful. We will demonstrate that power through two examples that find their origin in Section 3.1.

Example 3.2.3 We take our notation from Example 3.1.4. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which reflects a vector in the Cartesian plane across the x-axis, and let $S : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which rotates a vector counter-clockwise around the origin by $\frac{\pi}{2}$ radians. In this example we will find the 2×2 matrices A and B such that $T(\mathbf{v}) = A\mathbf{v}$ and $S(\mathbf{v}) = B\mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^2$.

In the proof of Theorem 3.2.2, we saw that the way to form the matrix of a linear transformation is to calculate the image of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. In this context, we need to calculate the image of $\mathbf{e}_1$ and $\mathbf{e}_2$ under T and S.

The calculations we seek are below:

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad S(\mathbf{e}_1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad S(\mathbf{e}_2) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

This tells us that the matrices A and B are as follows:

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

Any curious reader can check that these matrices are correct by choosing a vector in $\mathbb{R}^2$ and multiplying by A and by B separately. The results should align with the actions of T and S, respectively. □

3.2.2 Composition and Matrix Multiplication

Since linear transformations are functions, we can *compose* them with other linear transformations. In order for this to make sense, we need to have the codomains and domains match up correctly.

If $T : U \to V$ and $S : V \to W$ are linear transformations between vector spaces, then we can define the function $S \circ T : U \to W$ by $(S \circ T)(\mathbf{u}) = S(T(\mathbf{u}))$ for each $\mathbf{u} \in U$. Note that such a definition is only possible when the *codomain of* T is the same as the *domain of* S. It is not difficult to show that $S \circ T$ is a linear transformation. (See Exercise 3.1.7.12.)

If $U = \mathbb{F}^n$, $V = \mathbb{F}^m$, and $W = \mathbb{F}^p$, then the linear transformation $S \circ T$ is defined $\mathbb{F}^n \to \mathbb{F}^p$, and Theorem 3.2.2 says that there is a unique matrix over $\mathbb{F}$ which carries out this linear transformation. What is that matrix?

Theorem 3.2.2 tells us that there are matrices A and B such that the transformations T and S are multiplication by B and A, respectively. The matrix B is $m \times n$ and A is $p \times m$. We will *define* the product of A and B so that the matrix of $S \circ T$ is the matrix product AB.

Definition 3.2.4 Let A be a $p \times m$ matrix over a field $\mathbb{F}$ and let B be an $m \times n$ matrix over $\mathbb{F}$. Then the **matrix product** AB is the unique $p \times n$ matrix over $\mathbb{F}$ such that for all $\mathbf{u} \in \mathbb{F}^n$,

$$A(B\mathbf{u}) = (AB)\mathbf{u}.$$

◊

Note 3.2.5 When we take the matrix product AB, the number of columns of A must match the number of rows of B. The matrix product makes no sense (and cannot be computed) otherwise. The matrix AB then has the same number of rows as A and the same number of columns as B.

Though we have defined matrix multiplication in terms of the composition of linear transformations, we can multiply matrices of the correct dimensions even when we have no specific linear transformations in mind. This is similar to our understanding of row-reducing a matrix—this arose in the context of solving linear systems, but the process can be carried out on any matrix.

We have defined matrix multiplication, but we have not specified how the entries in the matrix product are calculated. Fear not; the wait is over.

We will use the definition of matrix multiplication and the formula we have for the product of a matrix and a vector (see (3.1)). Since $(AB)\mathbf{u}$ is a vector, we will record a formula for entry i in this vector. In what follows, we assume $A = [a_{ij}]$ and the entries of $B = [b_{ij}]$; we also assume $\mathbf{u} = [u_i]$:

$$\begin{aligned} [(AB)\mathbf{u}]_i = [A(B\mathbf{u})]_i &= \sum_{k=1}^{m} a_{ik}[B\mathbf{u}]_k \\ &= \sum_{k=1}^{m} a_{ik} \sum_{j=1}^{n} b_{kj}u_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{m} a_{ik}b_{kj} \right) u_j. \end{aligned}$$

When we look again at the formula in (3.1) for the product of a matrix and a vector, we see that

$$[AB]_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj} \tag{3.5}$$

for all $1 \leq i \leq p$ and all $1 \leq j \leq n$. In words, this means that the (i, j)-entry of AB is the entry-wise product of row i in A with column j in B. (Later we will acknowledge this as the *dot product* of two vectors in $\mathbb{F}^m$.)

We will try to make this concrete with some examples.

Example 3.2.6 Let A and B be the following matrices over $\mathbb{R}$:

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

Note that the product AB makes sense since the number of columns of A is the same as the number of rows of B. Here is the matrix product:

$$AB = \begin{bmatrix} 2(-2) - 1(1) & 2(0) - 1(-3) \\ 3(-2) + 4(1) & 3(0) + 4(-3) \end{bmatrix} = \begin{bmatrix} -5 & 3 \\ -2 & -12 \end{bmatrix}.$$

Since the sizes of A and B allow it, we can also calculate BA in this example:

$$BA = \begin{bmatrix} -4 & 2 \\ -7 & -13 \end{bmatrix}.$$

Finally, we observe that $AB \neq BA$. □

Example 3.2.7 Let A and B be the following matrices over $\mathbb{F}_5$:

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

Since A is 3×2 and B is 2×2, we can calculate AB, which will be 3×2. (In this example we cannot calculate BA.) Here is the matrix product:

$$AB = \begin{bmatrix} 4(3) + 0(4) & 4(3) + 0(2) \\ 1(3) + 4(4) & 1(3) + 4(2) \\ 3(3) + 0(4) & 3(3) + 0(2) \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 4 & 1 \\ 4 & 4 \end{bmatrix}.$$

To obtain the last equality, we remember that we are working in $\mathbb{F}_5$. □

Since we defined matrix multiplication in the context of the composition of linear transformations, our next example picks up on this theme.

Example 3.2.8 We return to Example 3.2.3 and consider the linear transformations $S, T : \mathbb{R}^2 \to \mathbb{R}^2$, where T reflects a vector in the Cartesian plane across the x-axis and S rotates a vector counter-clockwise around the origin by $\frac{\pi}{2}$ radians. In the previous example, we calculated the matrices A and B for T and S, respectively. What is the matrix for $S \circ T$?

We have defined matrix multiplication to answer exactly this question. We only need to multiply the matrices in the proper order. The matrix for $S \circ T$ is

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

A related question in this context is whether or not linear transformations *commute*. In other words, is $S \circ T = T \circ S$? For this example, answering that question boils down to comparing the matrix product AB with the product

BA which we have just calculated:

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

From this we can see that $S \circ T$ and $T \circ S$ are two distinct linear transformations. □

As we start to deal more regularly with matrices in the context of linear transformations, we need to recall the notation $M_{m,n}(\mathbb{F})$ and $M_n(\mathbb{F})$ from Example 2.3.10.

The next theorem records some facts about matrix multiplication which will be useful later in the text. We will walk the reader through the proof of this theorem in the exercises at the end of this section.

Theorem 3.2.9 *Let $A, A_1, A_2 \in M_{m,n}(\mathbb{F})$, $B, B_1, B_2 \in M_{n,p}(\mathbb{F})$, and $C \in M_{p,q}(\mathbb{F})$. Then*

1. $A(BC) = (AB)C$,
2. $A(B_1 + B_2) = AB_1 + AB_2$, *and*
3. $(A_1 + A_2)B = A_1B + A_2B$.

This theorem says that, if all of the matrix products make sense, matrix multiplication is associative and obeys both of the distributive laws.

There is one other useful way to think about matrix multiplication—in terms of the columns of the matrix.

Proposition 3.2.10 *Let $A \in M_{m,n}(\mathbb{F})$, $B \in M_{n,p}(\mathbb{F})$, and let the columns of B be $\mathbf{b}_1, \dots, \mathbf{b}_p$. Then the columns of AB are $A\mathbf{b}_1, \dots, A\mathbf{b}_p$.*

Proof. By our definition of the matrix product, for each $j = 1, \dots, p$ we have

$$(AB)\mathbf{e}_j = A(B\mathbf{e}_j).$$

The observation in Note 3.2.1 means that $B\mathbf{e}_j = \mathbf{b}_j$, so we have

$$(AB)\mathbf{e}_j = A\mathbf{b}_j.$$

Since $(AB)\mathbf{e}_j$ is the jth column of AB, this proves the proposition. ■

From the understanding we developed in Example 3.1.5, this proposition means that every column of the matrix product AB is a linear combination of the columns of A, when the product AB is defined.

3.2.3 Invertible Matrices

With matrix multiplication defined in terms of the composition of linear transformations, we turn to a specific composition in this subsection. Specifically, we will think about matrices for linear transformations S and T when $S \circ T$ is the identity transformation.

In Definition 3.1.11, we called such linear transformations *invertible*. When such a linear transformation can be accomplished by matrix multiplication, we will refer to the connected matrix using this same term.

Definition 3.2.11 Let $A \in M_n(\mathbb{F})$. The matrix B is an **inverse** matrix for A if

$$AB = BA = I_n.$$

If a matrix A has an inverse, we say that A is **invertible** or **non-singular**. If A is a matrix for which no inverse matrix exists, we say that A is **singular** or not invertible. ◊

It may strike the reader as strange that only square matrices have a chance at being invertible—we have only defined invertibility for square matrices. There is a good reason for this, which we will explore in the exercises.

If we know an inverse B of a matrix A, then we can solve matrix-vector equations with ease:

$$\begin{aligned} A\mathbf{x} &= \mathbf{y} \\ B(A\mathbf{x}) &= B\mathbf{y} \\ (BA)\mathbf{x} &= B\mathbf{y} \\ I\mathbf{x} &= B\mathbf{y} \\ \mathbf{x} &= B\mathbf{y}. \end{aligned}$$

The next three propositions present some properties of matrix inverses.

Proposition 3.2.12 *If a matrix $A \in M_n(\mathbb{F})$ has an inverse, that inverse is unique.*

Proof. Let $A \in M_n(\mathbb{F})$ and suppose that both $B, C \in M_n(\mathbb{F})$ are inverses for A. Then we have

$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$

■

This proposition allows us to refer to *the inverse* of a matrix A and to use the notation A^{-1} for that matrix.

Proposition 3.2.13 *Suppose that $A, B \in M_n(\mathbb{F})$ are both invertible. Then AB is invertible as well and*

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

Proof. If A and B are both invertible, then both A^{-1} and B^{-1} exist. Since matrix multiplication is associative, the following calculations show that the matrix $B^{-1}A^{-1}$ satisfies the properties of the inverse of AB, thereby making AB invertible:

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = A(I_n)A^{-1} = AA^{-1} = I_n; \\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}(I_n)B = B^{-1}B = I_n. \end{aligned}$$

■

This final proposition states what may seem like an obvious fact, but which should still be justified. That justification is left to the exercises.

Proposition 3.2.14 *Let $A \in M_n(\mathbb{F})$ be invertible. Then A^{-1} is also invertible and $(A^{-1})^{-1} = A$.*

While we are not yet ready to calculate the inverse of a matrix (stay tuned!), we can provide examples of invertible matrices and their inverses.

Example 3.2.15 Consider the following matrix $A \in M_2(\mathbb{R})$:

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

We can verify that A^{-1} is

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

by computing the product in both orders and verifying that the result is I_2 in both cases.

Similarly, here is a 3×3 matrix over $\mathbb{F}_3$ which is invertible:

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

The reader is encouraged to verify that the following matrix satisfies the properties of the inverse of B:

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

□

3.2.4 The Transpose of a Matrix

The *transpose* of a matrix is useful notation for some formulas that will appear later.

Definition 3.2.16 If $A \in M_{m,n}(\mathbb{F})$, then the **transpose** of A, denoted A^T, is the element of $M_{n,m}(\mathbb{F})$ whose rows are the columns of A. In other words,

$$[A^T]_{ij} = [A]_{ji}$$

for all $1 \leq i \leq n$ and all $1 \leq j \leq m$. ♢

Note 3.2.17 The transpose is an easy way to turn a column vector into a row vector and vice versa.

Example 3.2.18 If A is the 2×3 matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -2 & 4 & 5 \end{bmatrix},$$

then A^T is the 3×2 matrix

$$A^T = \begin{bmatrix} 2 & -2 \\ -1 & 4 \\ 0 & 5 \end{bmatrix}.$$

□

Some matrices are unaffected by taking the transpose. These deserve a special designation!

Definition 3.2.19 A matrix which is equal to its own transpose is called a **symmetric matrix**. (All symmetric matrices must be square by necessity.) ♢

The following theorem collects some properties related to the transpose of a matrix.

Theorem 3.2.20 *Let $A, C \in M_{m,n}(\mathbb{F})$, let $B \in M_{n,p}(\mathbb{F})$, and let $k \in \mathbb{F}$. Then the following properties hold:*

1. $(A^T)^T = A$;
2. $(A + C)^T = A^T + C^T$;
3. $(kA)^T = kA^T$; *and*

4. $(AB)^T = B^T A^T$.

Proof. The first three parts of this theorem are immediate from the definitions and require no proof. To prove the fourth part, we will compare the (i,j)-entry of both $(AB)^T$ and $B^T A^T$. First, from the definition of the transpose and (3.5) we see that

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

To compare, entry (i,j) of $B^T A^T$ is

$$[B^T A^T]_{ij} = \sum_{k=1}^{n} [B^T]_{ik} [A^T]_{kj} = \sum_{k=1}^{n} b_{ki} a_{jk}.$$

Since multiplication is commutative in fields, these two expressions are equal. ■

Note 3.2.21 While it might be more aesthetically pleasing if we did not have to switch the order of the multiplication when taking the transpose of a product, this type of formula makes sense when considering the dimensions of the matrices involved. If A is $m \times n$ and B is $n \times p$, then the expression $A^T B^T$ wouldn't even make sense unless $m = p$. Further, reversing the order in a formula involving matrix multiplication is typical, as we have already seen in Proposition 3.2.13.

3.2.5 Reading Questions

1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is rotation *clockwise* around the origin by $\frac{\pi}{2}$ radians. Find the matrix for T. (Refer to Example 3.2.3.) Explain your process.

2. Consider the following two matrices A and B over $\mathbb{R}$:

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

Calculate both AB and BA.

3. Write down a 3×3 matrix over $\mathbb{F}_5$ which is symmetric. (See Definition 3.2.19.)

3.2.6 Exercises

1. Let A, B, and C be the following matrices over $\mathbb{R}$:

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

For each of the following, determine whether the given calculation makes sense. If it does, find the requested matrix. (Do this by hand, without technology.) If it doesn't make sense, explain why it doesn't.

(a) A^2

(b) AB

(c) AC

(d) BC

(e) BA

(f) B^2

2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which reflects a vector across the line $y = x$. Find the matrix for T.

3. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which projects a vector onto the line $y = x$. Find the matrix for T.

4. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which projects a vector onto the line $y = -x$. Find the matrix for T.

5. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which rotates a vector counter-clockwise around the origin by an angle of θ radians. Find the matrix for T. (Each entry in the matrix should be an expression involving θ.)

Writing Exercises

6. Let $A \in M_n(\mathbb{F})$ be invertible. Prove that A^T is invertible and that $(A^T)^{-1} = (A^{-1})^T$.

7. Let $A \in M_{m,n}(\mathbb{F})$.

(a) Suppose that A is **left-invertible**, meaning that there is an $n \times m$ matrix B such that $BA = I_n$. Prove that $m \geq n$.

(b) Suppose that A is **right-invertible**, meaning that there is an $n \times m$ matrix B such that $AB = I_m$. Prove that $m \leq n$.

(c) Prove that any A which is invertible must be a square matrix.

8. Prove Proposition 3.2.14.

9. In fields, we have the *cancellation law* for multiplication. If $ab = ac$ and $a \neq 0$, then $b = c$. Does matrix multiplication have this property?

Let A, B, and C be matrices over $\mathbb{F}$ such that AB and AC make sense and are the same size and A is not the zero matrix. If $AB = AC$, must it be true that $B = C$? Either prove this is true or provide a counter-example.

10. In fields, multiplication has the *no zero divisors* property. If $xy = 0$, then either $x = 0$ or $y = 0$. Does matrix multiplication have this property?

Let A and B be matrices over $\mathbb{F}$ such that AB makes sense. Let Z be the matrix of the same size as AB consisting of all zeros. If $AB = Z$, must it be true that either A or B is a matrix of all zeros? Either prove this is true or provide a counter-example.

11. Let $A \in M_2(\mathbb{F}_5)$ be of the form

$$A = \begin{bmatrix} a & 0 \\ b & c \end{bmatrix}.$$

(a) What conditions must a, b, and c satisfy so that $A^2 = I_2$?

(b) How many matrices in $M_2(\mathbb{F}_5)$ of this form have the property that $A^2 = I_2$?

3.3 Inverting a Matrix

In Section 3.2, we introduced the definition of an *invertible matrix* and discussed some properties of the *inverse* of a matrix. We will now introduce a method to determine whether or not a matrix is invertible. Additionally, when a matrix is invertible we will be able to calculate its inverse.

3.3.1 Elementary Matrices

The method we will present in this section begins with a simple definition.

Definition 3.3.1 An **elementary matrix** is one that is formed by performing a single elementary row operation on an identity matrix. ♢

Because elementary matrices are related to elementary row operations, there are three types of elementary matrices. The following example provides one elementary matrix of each type.

Example 3.3.2 Our first elementary matrix results from switching the second and third rows in I_3:

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

Next we will look at a matrix which comes about by adding 4 times the first row of I_2 to the second row:

$$\begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix}. \tag{3.6}$$

Finally, we have a matrix which is formed by multiplying the second row of I_4 by 7:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

□

Multiplying by an elementary matrix has the effect of carrying out an elementary row operation. In other words, if the $n \times n$ matrix E results from applying an elementary row operation to I_n, and if A is another $n \times n$ matrix, then EA is the matrix A after this same elementary row operation has been applied. We will demonstrate this in an example before stating the relevant theorem. (The proof of the theorem is saved for the exercises.)

Example 3.3.3 Let A be the following matrix over $\mathbb{R}$:

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

If we label as E the matrix in (3.6), then we can calculate EA:

$$EA = \begin{bmatrix} 1 & 0 \\ 4 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -5 & 6 \end{bmatrix}.$$

The reader can verify that EA is the result of adding four times the first row of A to the second row of A. □

Theorem 3.3.4 *If the elementary matrix E results from performing an elementary row operation on I_n, and if A is an $n \times n$ matrix, then EA is the* *matrix that results from applying this same elementary row operation to A.*

Each elementary row operation is "reversible" in the sense that there is another elementary row operation which reverses the work that was done by the first. (This appears as Exercise 1.2.9.) We can use this fact to establish the following useful proposition.

Proposition 3.3.5 *Every elementary matrix is invertible.*

Proof. Let E be an elementary matrix and let E' denote the elementary matrix which represents the reverse elementary row operation from E. By the definition of these matrices and Theorem 3.3.4, we see that

$$EE' = I_n, \text{ and } E'E = I_n.$$

This shows that E and E' are inverses of each other and, in particular, this proves that E is invertible. ∎

We will now connect elementary matrices to the RREF of a matrix in the following proposition. This is largely a restatement of Algorithm 1.2.12 using elementary matrices.

Proposition 3.3.6 *If $A \in M_{m,n}(\mathbb{F})$, then there exists $B \in M_{m,n}(\mathbb{F})$ in RREF and elementary matrices $E_1, \ldots, E_k \in M_m(\mathbb{F})$ such that*

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

Proof. Since each matrix can be reduced to a matrix in RREF, and since elementary row operations are accomplished by multiplying by elementary matrices, there exist elementary matrices $D_1, \ldots, D_k \in M_m(\mathbb{F})$ such that

$$B = D_k \cdots D_1 A.$$

Since elementary matrices are invertible, by repeated application of Proposition 3.2.13 we see that $D_k \cdots D_1$ is invertible and $(D_k \cdots D_1)^{-1} = D_1^{-1} \cdots D_k^{-1}$. Then we have

$$\begin{aligned} B &= (D_k \cdots D_1)A \\ (D_k \cdots D_1)^{-1}B &= (D_k \cdots D_1)^{-1}(D_k \cdots D_1)A \\ D_1^{-1} \cdots D_k^{-1}B &= A. \end{aligned}$$

We note that each D_i^{-1} is an elementary matrix, and if we define $E_i = D_i^{-1}$ for each $i = 1, \ldots, k$, we have our result. ∎

3.3.2 Finding the Inverse of a Matrix

We will now move on to develop an algorithm for finding the inverse of a matrix (when one exists). We need a lemma before stating our most important result of the section.

Lemma 3.3.7 *If $A \in M_n(\mathbb{F})$ is invertible, then for every $\mathbf{b} \in \mathbb{F}^n$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution.*

Proof. Let $\mathbf{b}$ be any vector in $\mathbb{F}^n$. Since A is invertible, A^{-1} exists, and we can show that $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution to the equation $A\mathbf{x} = \mathbf{b}$:

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}.$$

To show that this solution is unique, suppose $\mathbf{v}$ is another solution to this equation, so $A\mathbf{v} = \mathbf{b}$. Then we have

$$\begin{aligned} A\mathbf{v} &= \mathbf{b} \\ A^{-1}(A\mathbf{v}) &= A^{-1}\mathbf{b} \\ (A^{-1}A)\mathbf{v} &= A^{-1}\mathbf{b} \\ I_n\mathbf{v} &= A^{-1}\mathbf{b} \\ \mathbf{v} &= A^{-1}\mathbf{b}. \end{aligned}$$

■

Theorem 3.3.8 *A matrix $A \in M_n(\mathbb{F})$ is invertible if and only if A is row equivalent to I_n. When A is invertible, any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.*

Proof. If A is invertible, then by Lemma 3.3.7 the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbb{F}^n$. Now Theorem 1.3.6 means that the RREF of A has a pivot in each of its n columns. Since A is square, this means the RREF has a pivot in each row as well, meaning that the RREF of A must be I_n.

Conversely, suppose that A is row equivalent to I_n. By Proposition 3.3.6, there exist elementary matrices $E_1, \ldots, E_k$ such that

$$A = E_1 \cdots E_k I_n. \tag{3.7}$$

This means that $A = E_1 \cdots E_k$, and since the product of invertible matrices is invertible, this proves that A is invertible.

If we multiply both sides of (3.7) by $(E_1 \cdots E_k)^{-1}$, we get

$$E_k^{-1} \cdots E_1^{-1} A = I_n,$$

which shows the sequence of elementary row operations (through multiplication by elementary matrices) used to transform A into I_n. On the other hand, if we take the equation $A = E_1 \cdots E_k$ from the previous paragraph and invert both sides, we get

$$A^{-1} = E_k^{-1} \cdots E_1^{-1},$$

which we can easily adjust to

$$A^{-1} = E_k^{-1} \cdots E_1^{-1} I_n.$$

This establishes the final claim in the theorem. ■

This theorem provides an algorithm for us to determine when a matrix is invertible and, in the case it is invertible, to calculate its inverse.

Algorithm 3.3.9 *In order to determine whether or not a matrix $A \in M_n(\mathbb{F})$ is invertible, follow these steps.*

1. *Reduce the matrix $[A \mid I_n]$ to its RREF.*
2. *If the RREF has the form $[I_n \mid B]$, then A is invertible and $B = A^{-1}$.*
3. *If the RREF does not have I_n in its left n columns, then A is not invertible.*

We will end this section with several examples in which we work through this algorithm.

Example 3.3.10 Consider the following matrix $A \in M_3(\mathbb{R})$:

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

To determine whether or not this matrix is invertible, we row reduce the matrix $[A \mid I_3]$:

$$\left[\begin{array}{ccc|ccc} 2 & 0 & -3 & 1 & 0 & 0 \\ 1 & -1 & -2 & 0 & 1 & 0 \\ 0 & 3 & 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 9 & 3 \\ 0 & 1 & 0 & 2 & -4 & -1 \\ 0 & 0 & 1 & -3 & 6 & 2 \end{array}\right].$$

We see that A is invertible and that

$$A^{-1} = \begin{bmatrix} -4 & 9 & 3 \\ 2 & -4 & -1 \\ -3 & 6 & 2 \end{bmatrix}.$$

□

Example 3.3.11 Consider the following matrix $A \in M_4(\mathbb{Q})$:

$$A = \begin{bmatrix} 2 & 3 & 0 & 1 \\ 1 & 1 & -3 & 7 \\ 0 & -2 & 1 & 0 \\ -1 & -4 & 2 & -2 \end{bmatrix}.$$

We now row reduce $[A \mid I_4]$ to determine whether or not A is invertible. We find that $[A \mid I_4]$ is row equivalent to

$$\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 2 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & -1 & 0 & -1/5 & -1/5 & -1/5 \\ 0 & 0 & 1 & -2 & 0 & -2/5 & 3/5 & -2/5 \\ 0 & 0 & 0 & 0 & 1 & 3/5 & -17/5 & 13/5 \end{array}\right].$$

This calculation shows that A is not invertible. □

Example 3.3.12 Consider the following matrix $A \in M_2(\mathbb{F}_3)$:

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

We now row reduce $[A \mid I_2]$:

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cc|cc} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 2 \end{array}\right].$$

This proves that A is invertible and that

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

□

3.3.3 Reading Questions

1. Consider the following matrix defined over $\mathbb{R}$:

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

Write down elementary matrices $E_1, \ldots, E_n$ which reduce A to I_2. In other words, find elementary matrices $E_1, \ldots, E_n$ such that $I_2 = E_n \cdots E_1 A$.

2. Consider the following matrix defined over $\mathbb{F}_5$:

$$A = \begin{bmatrix} 4 & 0 & 4 \\ 3 & 3 & 1 \\ 1 & 0 & 3 \end{bmatrix}.$$

Determine whether or not A is invertible. Explain your answer. If A is invertible, find the inverse.

3.3.4 Exercises

1. For each of the following matrices A, find the RREF of A (call it B), and elementary matrices $E_1, \ldots, E_k$ such that $B = E_k \cdots E_1 A$.

(a) $A \in M_{2,3}(\mathbb{R})$:

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

(b) $A \in M_2(\mathbb{F}_5)$:

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

(c) $A \in M_2(\mathbb{C})$:

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

2. For each of the following matrices A, find the RREF of A (call it B), and elementary matrices $E_1, \ldots, E_k$ such that $B = E_k \cdots E_1 A$.

(a) $A \in M_3(\mathbb{R})$:

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

(b) $A \in M_3(\mathbb{F}_3)$:

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

3. For each of the following matrices A in $M_3(\mathbb{R})$, determine whether or not A is invertible using the algorithm from this section. In the cases where A is invertible, find the inverse.

(a) $A = \begin{bmatrix} -5 & -5 & 20 \\ 4 & -4 & 32 \\ 0 & -1 & 6 \end{bmatrix}$

(b) $A = \begin{bmatrix} -1 & -1 & -6 \\ -3 & 0 & -5 \\ 3 & 6 & 5 \end{bmatrix}$

4. For each of the following matrices A in $M_3(\mathbb{F}_5)$, determine whether or not A is invertible using the algorithm from this section. In the cases where A is invertible, find the inverse.

(a) $A = \begin{bmatrix} 4 & 3 & 0 \\ 2 & 3 & 1 \\ 2 & 4 & 3 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0 & 3 & 4 \\ 4 & 0 & 4 \\ 4 & 3 & 3 \end{bmatrix}$

5. Use an inverse matrix to solve the following linear system over $\mathbb{F}_3$:

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

6. Use an inverse matrix to solve the following linear system over $\mathbb{F}_5$:

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

7. Use an inverse matrix to solve the following linear system over $\mathbb{R}$:

$$\begin{aligned} -4x + 2y + 4z &= 1 \\ -2x + y - 6z &= -2 \\ -3x - y + 2z &= 3. \end{aligned}$$

Writing Exercises

8. Suppose $AB = AC$, where A is an $n \times n$ matrix and B and C are $n \times p$ matrices.

(a) Show that if A is invertible, then $B = C$.

(b) Provide an example where $AB = AC$ but A is not invertible and $B \neq C$.

9. Suppose that $(B - C)D = 0$, where B and C are $m \times n$ matrices, D is an $n \times n$ matrix, and 0 is the $m \times n$ zero matrix.

(a) Show that if D is invertible, then $B = C$.

(b) Provide an example where $(B - C)D = 0$ but D is not invertible and $B \neq C$.

10. Suppose that A and B are $n \times n$ matrices and that AB is invertible. Prove that A is invertible.

11. Suppose that $A \in M_n(\mathbb{F})$ is upper triangular and invertible. (A matrix is **upper triangular** when all entries below the main diagonal are 0.) Prove that A^{-1} is also upper triangular.

12. How many matrices in $M_2(\mathbb{F}_2)$ are invertible? What proportion of the matrices in $M_2(\mathbb{F}_2)$ are invertible?

13. How many matrices in $M_2(\mathbb{F}_3)$ are invertible? What proportion of the matrices in $M_2(\mathbb{F}_3)$ are invertible?

14. Prove Theorem 3.3.4. (You will need three cases.)

3.4 Subspaces and Linear Transformations

Every linear transformation between vector spaces brings with it some descriptions of related subspaces of the domain and codomain. We will explore some of these subspaces in this section.

3.4.1 The Kernel of a Linear Transformation

The *kernel* of a linear transformation T is the set of all vectors that T sends to the zero vector.

Definition 3.4.1 If $T : V \to W$ is a linear transformation between vector spaces, then the **kernel** of T is the set

$$\ker(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}.$$

$\Diamond$

While it may seem strange to single out the vectors that are sent to $\mathbf{0}$, this set reveals a lot about T.

Theorem 3.4.2 *Let $T : V \to W$ be a linear transformation. Then $\ker(T)$ is a subspace of V.*

Proof. We will prove this theorem using the criteria for a subspace spelled out in Theorem 2.4.3. Since $T(\mathbf{0}_V) = \mathbf{0}_W$, we have $\mathbf{0}_V \in \ker(T)$. (The fact that $T(\mathbf{0}_V) = \mathbf{0}_W$ is found in Exercise 3.1.7.14.)
Let $\mathbf{v}_1, \mathbf{v}_2 \in \ker(T)$. Using the additive property of T and the fact that these vectors are in the kernel, we have

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W.$$

Finally, we let $\mathbf{v} \in \ker(T)$ and $k \in \mathbb{F}$. Then we have

$$T(k\mathbf{v}) = kT(\mathbf{v}) = k\mathbf{0}_W = \mathbf{0}_W.$$

This calculation used the fact that $\mathbf{v}$ was assumed to be in the kernel and the scalar multiplication property of T. ■

Example 3.4.3 Let $V = D[a, b]$ be the set of all differentiable functions from $[a, b] \to \mathbb{R}$. Let $T : V \to V$ be the linear transformation which takes the derivative. (We proved a very similar function was a linear transformation in Example 3.1.3.) What is the kernel of T?

We recall from calculus that a function f has $f'(x) = 0$ for all x in an interval if and only if f is a constant function. This proves that $\ker(T)$ is the set of all constant functions, and it further establishes that set as a subspace of $D[a, b]$. □

While not all linear transformations are linked to matrices, some are. The kernel has an alternate name in those situations.

Definition 3.4.4 If $A \in M_{m,n}(\mathbb{F})$, then the **null space** of A is

$$\text{null}(A) = \{\mathbf{v} \in \mathbb{F}^n \mid A\mathbf{v} = \mathbf{0}\}.$$

$\Diamond$

Since we have shown that the kernel is a subspace, the word "space" in the term *null space* is justified.

We also recall the link between matrix-vector equations like $A\mathbf{x} = \mathbf{b}$ and linear systems. The definition of the null space shows that the set of solutions to a *homogeneous* linear system can be described as the null space of a matrix.

Example 3.4.5 Let A be the matrix

$$A = \begin{bmatrix} 3 & 1 & -3 \\ -1 & 0 & 2 \\ 2 & -2 & -10 \end{bmatrix}.$$

We can find null(A) by row reducing the matrix $[A \mid \mathbf{0}]$. Here is the RREF:

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

From this we see that $x_1 = 2x_3$, $x_2 = -3x_3$, and x_3 is free. In other words, any vector $\mathbf{x}$ in null(A) looks like

$$\mathbf{x} = \begin{bmatrix} 2x_3 \\ -3x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

So we have

$$\text{null}(A) = \text{Span}\{\mathbf{v}\}$$

where

$$\mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

□

Note 3.4.6 There is an important fact contained in this last example. When we have a homogeneous system, we can always pay attention to just the coefficient matrix instead of the augmented matrix. No elementary row operation can produce a non-zero entry in a column of zeros.

The following theorem is one of the reasons that the kernel is so useful.

Theorem 3.4.7 *Let $T : V \to W$ be a linear transformation. Then T is injective if and only if* $\ker(T) = \{\mathbf{0}\}$.

Proof. We first suppose that T is injective. (We recall the definition of injectivity from Definition 3.1.9.) Since $T(\mathbf{0}_V) = \mathbf{0}_W$, the injectivity of T implies that $\mathbf{v} = \mathbf{0}_V$ if $T(\mathbf{v}) = \mathbf{0}_W$ for any $\mathbf{v} \in V$. Therefore, $\ker(T) = \{\mathbf{0}\}$.
Next, we suppose that $\ker(T) = \{\mathbf{0}\}$. We want to prove that T is injective, so we let $\mathbf{v}_1, \mathbf{v}_2 \in V$ with $T(\mathbf{v}_1) = T(\mathbf{v}_2)$. We want to show that $\mathbf{v}_1 = \mathbf{v}_2$. By the linearity of T we have

$$\begin{aligned} T(\mathbf{v}_1) - T(\mathbf{v}_2) &= \mathbf{0} \\ T(\mathbf{v}_1 - \mathbf{v}_2) &= \mathbf{0}. \end{aligned}$$

Since $\ker(T) = \{\mathbf{0}\}$, we must have $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$, meaning that $\mathbf{v}_1 = \mathbf{v}_2$. This proves that T is injective. ■

Example 3.4.8 Consider the linear transformation $T : P_1 \to \mathbb{R}^2$ given by

$$T(p) = \begin{bmatrix} p(0) \\ p'(0) \end{bmatrix}.$$

To examine $\ker(T)$, we need to look at polynomials $p \in P_1$ such that $p(0) = 0$ and $p'(0) = 0$. If $p(t) = a + bt$, then $a = p(0)$ and $b = p'(0)$, so if $T(p) = \mathbf{0}$, we must have $a = 0$ and $b = 0$. This means that the only polynomial in $\ker(T)$ is the zero polynomial. Therefore, T is injective. □

We now present one final fact related to the kernel.

Corollary 3.4.9 *Suppose that the linear system represented by the equation* $A\mathbf{x} = \mathbf{b}$ *is consistent. Then this system has a unique solution if and only if* $\text{null}(A) = \{\mathbf{0}\}$.

Proof. We first assume that the system has a unique solution. Since the linear system is consistent, then there exists a vector $\mathbf{v}$ such that $A\mathbf{v} = \mathbf{b}$. If $\mathbf{w} \in \text{null}(A)$, then $A(\mathbf{v}+\mathbf{w}) = A\mathbf{v}+A\mathbf{w} = \mathbf{b}+\mathbf{0} = \mathbf{b}$, so $\mathbf{v}+\mathbf{w}$ is also a solution. But since there is a unique solution, we must have $\mathbf{v} = \mathbf{v} + \mathbf{w}$, so $\mathbf{w} = \mathbf{0}$. This shows that $\text{null}(A) = \{\mathbf{0}\}$.
We now assume that $\text{null}(A) = \{\mathbf{0}\}$. From Theorem 3.4.7 we know that the associated linear transformation is injective. Since the system is consistent, there must be only one vector that the transformation sends to $\mathbf{b}$ so the system has a unique solution. ■

3.4.2 The Range as a Subspace

We have examined the kernel as a subspace of the domain of a linear transformation. We now turn our attention to a well-known subset of the codomain. The reader will be familiar with the range (or image) of a linear transformation. We can now prove that this is a subspace.

Theorem 3.4.10 *Let* $T : V \to W$ *be a linear transformation between vector spaces over* $\mathbb{F}$. *Then* $\text{range}(T)$ *is a subspace of* W.

Proof. Since we know that $T(\mathbf{0}_V) = \mathbf{0}_W$, it follows that $\mathbf{0}_W \in \text{range}(T)$. We now need to show the other properties demanded by Theorem 2.4.3.
If $\mathbf{w}_1, \mathbf{w}_2 \in \text{range}(T)$, then there exist vectors $\mathbf{v}_1, \mathbf{v}_2 \in V$ with $T(\mathbf{v}_1) = \mathbf{w}_1$ and $T(\mathbf{v}_2) = \mathbf{w}_2$. Then, using the linearity of T, we have

$$\mathbf{w}_1 + \mathbf{w}_2 = T(\mathbf{v}_1) + T(\mathbf{v}_2) = T(\mathbf{v}_1 + \mathbf{v}_2).$$

This proves that $\mathbf{w}_1 + \mathbf{w}_2 \in \text{range}(T)$, so $\text{range}(T)$ is closed under addition.
Finally, we let $k \in \mathbb{F}$ and $\mathbf{w} \in \text{range}(T)$, so there exists a vector $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Then, using the fact that T is a linear transformation, we have

$$k\mathbf{w} = kT(\mathbf{v}) = T(k\mathbf{v}),$$

which proves that $k\mathbf{w} \in \text{range}(T)$. Thus, $\text{range}(T)$ is closed under scalar multiplication. We conclude that $\text{range}(T)$ is a subspace of W. ■

As usual, when our linear transformation is linked to a matrix, we have more to say.

Definition 3.4.11 If $A \in M_{m,n}(\mathbb{F})$, then the **column space** of A, written $\text{col}(A)$, is the set of all linear combinations of the columns of A. If $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$, then

$$\text{col}(A) = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}.$$

◊

When we introduced the matrix-vector product in Example 3.1.5, we noted that $A\mathbf{v}$ is a linear combination of the columns of A with weights coming from the entries in $\mathbf{v}$. So, a vector in $\mathrm{col}(A)$ can be written as $A\mathbf{x}$ for some $\mathbf{x}$. Therefore, another way to write the column space is

$$\mathrm{col}(A) = \{\mathbf{w} \in \mathbb{F}^m \mid \mathbf{w} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{F}^n\}.$$

From this description, we can see that the *column space* of a matrix and the *range* of a linear transformation are the same.

Fact 3.4.12 *If $T : \mathbb{F}^n \to \mathbb{F}^m$ is a linear transformation given by multiplication by a matrix A, then*

$$\mathrm{range}(T) = \mathrm{col}(A).$$

We can also restate the consistency of linear systems using the language of the column space.

Fact 3.4.13 *If $A \in M_{m,n}(\mathbb{F})$, then the linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b} \in \mathrm{col}(A)$.*

Since the column space of a matrix is a subspace of the codomain of the associated linear transformation, there will be some occasions when that subspace is as large as it could be. The next theorem gives conditions for just that situation.

Theorem 3.4.14 *A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ spans $\mathbb{F}^m$ if and only if the RREF of the matrix $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ has a pivot in every row.*

Proof. The set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ spans $\mathbb{F}^m$ if and only if $\mathbf{b} \in \mathrm{col}(A)$ for every $\mathbf{b} \in \mathbb{F}^m$. This happens when the linear system with augmented matrix $[A \mid \mathbf{b}]$ is consistent for each $\mathbf{b}$.
We know from Theorem 2.2.2 that a linear system over $\mathbb{F}$ is consistent if and only if there is no pivot in the final column of the augmented matrix. If the RREF of A has a pivot in every row, then there cannot be a pivot in the final column of the RREF of $[A \mid \mathbf{b}]$ since each row already contains one pivot.
We will prove the contrapositive of the other implication. Suppose that the RREF of A does not have a pivot in every row. We will create a vector in $\mathbb{F}^m$ which is not in the span of this set of vectors. Since the RREF of A does not have a pivot in every row, let the smallest row number with no pivot be k. Form the augmented matrix with the RREF of A and the vector $\mathbf{e}_k$. Now reverse the elementary row operations that were taken to reduce A to its RREF. The result will be an augmented matrix $[A \mid \mathbf{b}]$ which is related to an inconsistent system. (There will be a pivot in the final column of the RREF of this matrix. We constructed it this way!) This proves that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ does not span $\mathbb{F}^m$. ■

Example 3.4.15 Define the vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $\mathbb{F}_5^3$ as

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

How large is $\mathrm{Span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$?

We form the matrix $[\mathbf{u}\ \mathbf{v}\ \mathbf{w}]$ and find the RREF:

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

Since there is a pivot in each row, the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ spans $\mathbb{F}_5^3$. □

Example 3.4.16 Consider the following vectors in $\mathbb{R}^3$:

$$\mathbf{u} = \begin{bmatrix} -2 \\ 3 \\ -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} -5 \\ 6 \\ 2 \end{bmatrix}.$$

When we row reduce the matrix $[\mathbf{u}\ \mathbf{v}\ \mathbf{w}]$, we get

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

Since there is not a pivot in every row, the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ does not span $\mathbb{R}^3$. □

3.4.3 Reading Questions

1. Let A be the following matrix:

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

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation which is multiplication by A.

(a) Calculate $\mathrm{null}(A)$.

(b) Is T injective? Explain.

2. Consider the following three vectors in $\mathbb{F}_3^2$:

$$\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

Does the set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ span $\mathbb{F}_3^2$? Explain.

3.4.4 Exercises

1. Consider the following matrix A in $M_{3,4}(\mathbb{F}_7)$:

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

For each of the following vectors $\mathbf{x} \in \mathbb{F}_7^4$, determine whether or not $\mathbf{x} \in \mathrm{null}(A)$.

(a) $\mathbf{x} = \begin{bmatrix} 6 \\ 1 \\ 4 \\ 2 \end{bmatrix}$

(b) $\mathbf{x} = \begin{bmatrix} 4 \\ 3 \\ 1 \\ 2 \end{bmatrix}$

2. For each of the following matrices A over $\mathbb{R}$, find null(A) by producing a set of vectors that spans null(A).

(a) $A = \begin{bmatrix} 2 & 3 & 4 & 2 \\ -5 & -1 & 5 & -3 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & -4 & -5 & -3 & 1 \\ 1 & 4 & -5 & 4 & 5 \\ 0 & -2 & 5 & 0 & 5 \end{bmatrix}$

3. Find a matrix A so that the following set is col(A):

$$\left\{ \begin{bmatrix} 2r+4s \\ 3s-7t \\ -r-s+5t \end{bmatrix} \mid r, s, t \in \mathbb{R} \right\}.$$

4. Let A and $\mathbf{x}$ be the following, with entries from $\mathbb{R}$:

$$A = \begin{bmatrix} 3 & -1 \\ -9 & 3 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

(a) Show that $\mathbf{x} \in \text{null}(A)$ and $\mathbf{x} \notin \text{col}(A)$.

(b) Explain why it is not possible, for this particular matrix A, to find a non-zero vector in null$(A) \cap$ col(A).

(c) Is it possible to find a matrix $A \in M_2(\mathbb{R})$ such that there exists a non-zero vector in null$(A) \cap$ col(A)? Justify your answer thoroughly.

5. Consider the following linear transformation $T : \mathbb{F}_7^4 \to \mathbb{F}_7^3$:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + 4x_2 + 6x_3 + x_4 \\ 4x_1 + 6x_2 + 5x_3 \\ 4x_1 + 2x_2 + 3x_3 \end{bmatrix}.$$

This T is not injective. Find distinct vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathbb{F}_7^4$ such that $T(\mathbf{x}_1) = T(\mathbf{x}_2)$.

6. Consider the following linear transformation $T : \mathbb{F}_7^3 \to \mathbb{F}_7^4$:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 3x_2 + 3x_3 \\ 2x_1 + 6x_2 + 3x_3 \\ x_2 + 3x_3 \\ 3x_1 + 5x_2 + 6x_3 \end{bmatrix}.$$

This T is not surjective. Find a vector $\mathbf{y} \in \mathbb{F}_7^4$ such that $\mathbf{y}$ is not in the image of T.

7. Consider the following linear transformation $T : \mathbb{R}^3 \to P_2$:

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = (a+b) + (b+c)t + (a+b+c)t^2.$$

(a) Either prove or disprove that T is injective.

(b) Either prove or disprove that T is surjective.

8. Consider the following linear transformation $T : \mathbb{R}^2 \to P_2$:

$$T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = a + bt + bt^2.$$

(a) Either prove or disprove that T is injective.

(b) Either prove or disprove that T is surjective.

9. Consider the following linear transformation $T : P_2 \to \mathbb{R}^2$:

$$T(a + bt + ct^2) = \begin{bmatrix} a - 2b \\ 3b + c \end{bmatrix}.$$

(a) Either prove or disprove that T is injective.

(b) Either prove or disprove that T is surjective.

10. For each of the following, consider a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ which is multiplication by the given $m \times n$ matrix A. In each case, determine whether or not T is injective and whether or not T is surjective. Explain your answers.

(a) $A = \begin{bmatrix} 2 & -4 & -14 \\ -4 & -1 & 10 \\ -2 & 2 & 10 \end{bmatrix}$

(b) $A = \begin{bmatrix} -3 & -5 & -2 & 3 \\ 1 & 1 & -5 & -5 \\ 1 & -3 & 1 & 2 \end{bmatrix}$

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

(d) $A = \begin{bmatrix} -2 & -3 & -1 \\ 5 & -5 & -2 \\ 0 & 4 & -4 \end{bmatrix}$

Writing Exercises

11. Suppose that $T : U \to V$ is a linear transformation between vector spaces over $\mathbb{F}$. If U' is a subspace of U, prove that $T(U')$ is a subspace of V, where

$$T(U') = \{T(\mathbf{u}) \in V \mid \mathbf{u} \in U'\}.$$

12. Suppose that $T : U \to V$ is a linear transformation between vector spaces over $\mathbb{F}$. If V' is a subspace of V, define the set U_1 by

$$U_1 = \{\mathbf{u} \in U \mid T(\mathbf{u}) \in V'\}.$$

Prove that U_1 is a subspace of U.

13. Let $A \in M_{m,n}(\mathbb{F})$ and let $B \in M_{n,p}(\mathbb{F})$. Prove that $\mathrm{col}(AB) \subseteq \mathrm{col}(A)$.

14. Prove that if $n < m$, then no set of n vectors can span $\mathbb{F}^m$.

Chapter 4

Determinants

4.1 Defining the Determinant

In the previous chapter, we learned an algorithm for finding the inverse of an invertible $n \times n$ matrix A. This algorithm also told us when a matrix is *not invertible*—that is, when it is singular. The centerpiece of the algorithm was the row reduction of an $n \times 2n$ matrix.

In this section, we will learn about the useful and powerful function called *the determinant*. The determinant by itself will not give us the inverse of a matrix, but it will reveal whether or not a matrix is invertible.

4.1.1 The Definition of the Determinant

In Exercise 1.3.12, we saw that when a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has the property that $ad - bc \neq 0$, A is row equivalent to I_2. (In that context we were working over $\mathbb{R}$, but this is true over any field.) The converse is also true, though we haven't yet established it.

If we connect this fact with Theorem 3.3.8, we see that when $ad - bc \neq 0$, the 2×2 matrix A is invertible. (We will see that the converse of this statement is also true.)

This quantity $ad - bc$ for A is special—it is called the *determinant* of A, and this relationship between the determinant and invertibility also holds for larger matrices. The goal of this section is to define a number for any $n \times n$ matrix which functions the same way that $ad - bc$ does for a 2×2 matrix.

Definition 4.1.1 For an $m \times n$ matrix A, we define A_{ij} to be the submatrix of A that results from deleting row i and column j from A. ◊

Example 4.1.2 If A is the following matrix,

$$A = \begin{bmatrix} 3 & -7 & -4 & -7 \\ 2 & 0 & -8 & -2 \\ 8 & 6 & 2 & 2 \end{bmatrix},$$

then the submatrices A_{13} and A_{22} are

$$A_{13} = \begin{bmatrix} 2 & 0 & -2 \\ 8 & 6 & 2 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 3 & -4 & -7 \\ 8 & 2 & 2 \end{bmatrix}.$$

□

Perhaps surprisingly, this bit of notation is all we need to define the determinant. We note that this is a *recursive* definition, which means that the calculation happens in stages by reducing the size of the matrix.

Definition 4.1.3 Suppose that $A \in M_n(\mathbb{F})$ with $n \geq 1$, and let the entries of A be denoted a_{ij}. Then the **determinant** of A, denoted $\det(A)$ or $|A|$, is defined as follows. If $n = 1$, then $\det(A) = a_{11}$. If $n \geq 2$, then

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

$\diamondsuit$

To calculate the determinant of an $n \times n$ matrix, we need to know the determinant of a lot of $(n-1)\times(n-1)$ submatrices—in fact, we need to know the determinant of n submatrices. The reader will appreciate that when we begin with a small matrix, this is a manageable calculation; when the matrix is larger, carrying out this calculation by hand would be a decidedly less pleasant task.

Note 4.1.4 When $n = 2$, the formula for the determinant of A reduces to the familiar expression with which we opened this section. Note that, in this case, $A_{11} = [a_{22}]$ and $A_{12} = [a_{21}]$ are both 1×1 matrices. We have

$$\begin{aligned}\det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12})\\ &= a_{11}a_{22} - a_{12}a_{21}.\end{aligned}$$

This last expression is precisely $ad - bc$ with different symbols.

We will put our new definition to use in the following example.

Example 4.1.5 Let A be the following matrix over $\mathbb{R}$:

$$A = \begin{bmatrix} 2 & -4 & -1\\ 6 & -6 & 0\\ -5 & -3 & -5\end{bmatrix}.$$

We will find $\det(A)$ using the formula:

$$\begin{aligned}\det(A) &= 2\begin{vmatrix} -6 & 0\\ -3 & -5\end{vmatrix} - (-4)\begin{vmatrix} 6 & 0\\ -5 & -5\end{vmatrix} + (-1)\begin{vmatrix} 6 & -6\\ -5 & -3\end{vmatrix}\\ &= 2(30-0) + 4(-30-0) - (-18-30)\\ &= 60 - 120 + 48 = -12.\end{aligned}$$

We see that $\det(A) = -12$. $\square$

Note 4.1.6 For the purposes of what follows, we will introduce another bit of notation. If $A = [a_{ij}]$, then the (i,j)**-cofactor** of A, by C_{ij}, is

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

Our first determinant formula could then be written

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

4.1.2 Additional Ways to Calculate the Determinant

The definition of the determinant uses the first row of the matrix as its "home base" for calculations. For this reason, the original formula is sometimes referred to as *cofactor expansion along the first row* of A. A perhaps surprising result, and one we will not prove here because of its complicated nature, asserts that the determinant can be calculated using expansion along any row, not just the first.

Theorem 4.1.7 *Suppose that $A = [a_{ij}]$ and that $A \in M_n(\mathbb{F})$ with $n \geq 2$. Then* $\det(A)$ *can be calculated using cofactor expansion across any row of A. In other words, for any $1 \leq i \leq n$, we have*

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

Because we can target any row of a matrix in our calculation of the determinant, rows which contain several zero entries are particularly attractive.

Example 4.1.8 We consider the following matrix A:

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

Since we can expand along any row to calculate the determinant, and since the second row contains two zeros, we will expand along the second row. The zeros make it unnecessary to calculate the cofactors C_{21} and C_{23}:

$$\begin{aligned} \det(A) &= 0 \cdot C_{21} + 4 \cdot C_{22} + 0 \cdot C_{23} \\ &= (-1)^{2+2}(4) \begin{vmatrix} 2 & -3 \\ -1 & -2 \end{vmatrix} \\ &= 4(-4-3) = -28. \end{aligned}$$

□

In addition to expanding along any row of a matrix to calculate the determinant, we can also use any column. Instead of proving this directly using the definition of the determinant, we will follow Beezer[3] and prove that a matrix and its transpose have the same determinant. The result about using columns to calculate the determinant will follow.

Theorem 4.1.9 *Let $A \in M_n(\mathbb{F})$. Then* $\det(A) = \det(A^T)$.

Proof. We will proceed using induction on the size of the matrix. When $n = 1$, a matrix A and its transpose are identical, so $\det(A) = \det(A^T)$ trivially.
We now suppose that for any square matrix of size $n-1$, the determinant of the matrix and its transpose are equal. In the calculation below, we will employ a trick at the beginning in order to have a second summation sign later on. The first equation here is true because taking the average of one number n times will give the same quantity again. We will use the notation $[A^T]_{ij}$ to indicate the (i,j)-entry of A^T. We have

$$\det(A^T) = \frac{1}{n} \sum_{i=1}^{n} \det(A^T)$$

[3] linear.ups.edu

$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{i+j}[A^T]_{ij}\det(A^T_{ij}).$$

By the definition of the transpose, we have $[A^T]_{ij} = a_{ji}$ and $A^T_{ij} = (A_{ji})^T$. So

$$\det(A^T) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{i+j}a_{ji}\det((A_{ji})^T).$$

We now invoke the induction hypothesis. Since A_{ji} is a square matrix of size $n-1$, we have $\det((A_{ji})^T) = \det(A_{ji})$. This, along with Theorem 4.1.7, allows us to finish the argument:

$$\begin{aligned}\det(A^T) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{i+j}a_{ji}\det(A_{ji})\\ &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(-1)^{j+i}a_{ji}\det(A_{ji})\\ &= \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n}(-1)^{j+i}a_{ji}\det(A_{ji})\\ &= \frac{1}{n}\sum_{j=1}^{n}\det(A) = \det(A).\end{aligned}$$

As a final note, we were able to switch the summation signs toward the end of this calculation due to the fact that addition is commutative in any field $\mathbb{F}$. ■

Corollary 4.1.10 *Suppose that $A = [a_{ij}]$ and that $A \in M_n(\mathbb{F})$ with $n \geq 2$. Then $\det(A)$ can be calculated using cofactor expansion along any column of A. In other words, for any $1 \leq i \leq n$, we have*

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

Proof. We note that column i of A is the same as row i of A^T. So, calculating $\det(A)$ using expansion along column i is the same as calculating $\det(A^T)$ using expansion along row i. But since $\det(A^T) = \det(A)$, this calculation will result in $\det(A)$. ■

The following example shows how we can use this flexibility in calculating the determinant.

Example 4.1.11 We will calculate the determinant of the following matrix:

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

Since the third column of A has three zeros, we will expand along that column:

$$\det(A) = \sum_{i=1}^{4} a_{i3}C_{i3} = 2\cdot C_{23}.$$

Just for this example, we will rename $A_{23} = B$, so $\det(A) = -2\det(B)$. Here is B:

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

We will take advantage of the zero in the $(3,3)$-entry of B by expanding along the third row to calculate $\det(B)$:

$$\begin{aligned} \det(B) &= (-1)^{3+1}(2)\begin{vmatrix} 1 & -1 \\ 3 & -3 \end{vmatrix} + (-1)^{3+2}(-2)\begin{vmatrix} 3 & -1 \\ 1 & -3 \end{vmatrix} \\ &= 2(0) + 2(-8) = -16. \end{aligned}$$

Since we determined that $\det(A) = -2\det(B)$, this means that $\det(A) = 32$. □

We will close this section with one final result, the proof of which will be Exercise 4.1.4.10. The truth of this proposition also relies on rows/columns with many zeros. By a "triangular" matrix we mean either an upper triangular or a lower triangular matrix.

Proposition 4.1.12 *Let $A \in M_n(\mathbb{F})$ be a triangular matrix. Then $\det(A)$ is the product of the entries along the main diagonal of A.*

4.1.3 Reading Questions

1. Compute the determinant of the matrix A by cofactor expansion across the first row. Write out all of your calculations.

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

2. Consider the following matrix B:

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

(a) Compute the determinant of B by cofactor expansion down the third column. Write out all of your calculations.

(b) There is a better row or column to use for this calculation. Which one is it? Choose that row or column and compute the determinant of B by cofactor expansion along that row or column. Write out all of your calculations. Why is this easier?

4.1.4 Exercises

1. For each of the following matrices over $\mathbb{R}$, find the determinant by hand using cofactor expansion along a row or column of your choice.

(a) $A = \begin{bmatrix} 3 & 3 & -0.5 \\ 2 & -3 & 0 \\ -1 & 2 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 2 & -4 & 4 & -1 \\ -1 & 0 & -4 & 3 \\ 2 & 2 & -4 & 0 \\ 4 & 0 & -2 & 1 \end{bmatrix}$

2. For each of the following matrices over $\mathbb{R}$, find the determinant by hand using cofactor expansion along a row or column of your choice.

(a) $A = \begin{bmatrix} -3 & 3 & 0 \\ 0 & 4 & -2 \\ -1 & -4 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} -2 & 4 & 0 & 0 \\ 0.5 & -3.5 & -2 & -4 \\ 0.5 & 0 & 0 & -2 \\ 1 & -2 & 3 & -3 \end{bmatrix}$

3. Consider the following matrix A with entries in $\mathbb{F}_5$:

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

Find $\det(A)$ by hand using cofactor expansion along a row or column of your choice. (The answer should be a number in $\mathbb{F}_5$.)

4. For each of the following, write down the elementary matrix that performs the given elementary row operation on a 3×3 matrix. Then, calculate the determinant of that elementary matrix.

(a) Multiply row 2 by -3

(b) Switch rows 2 and 3

(c) Replace the third row with the sum of the third row and seven times the first row

5. Let $A = \begin{bmatrix} 3 & 1 \\ 2 & -2 \end{bmatrix}$. Write down the matrix $5A$. How are $\det(A)$ and $\det(5A)$ related?

6. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix over $\mathbb{R}$ and let $k \in \mathbb{R}$. Find a formula that relates $\det(kA)$ to k and $\det(A)$.

7. Let $A = \begin{bmatrix} 1 & -2 \\ 2.5 & c \end{bmatrix}$ be a matrix over $\mathbb{R}$. Find a value of c so that $\det(A) = 0$ or explain why this is not possible.

8. Let A be the following matrix over $\mathbb{R}$:

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

Find a value of c so that $\det(A) = 0$ or explain why this is not possible.

9. Let A be the following matrix over $\mathbb{R}$:

$$A = \begin{bmatrix} x-5 & 1 \\ 4 & x-2 \end{bmatrix}.$$

Find all values of x such that $\det(A) = 0$.

Writing Exercises

10. Prove that the determinant of a triangular matrix is the product of the entries along the main diagonal.

11. Prove that if A is a 2×2 matrix and if B is the result of switching the rows in A, then $\det(B) = -\det(A)$.

12. Suppose that A is a 2×2 matrix and that B is the result of applying the *replace* row operation to A. Prove that $\det(B) = \det(A)$.

4.2 Properties of the Determinant

We have introduced the determinant, but we have not yet backed up our assertion that the determinant is *useful* or *powerful*. Our goal in this section is to establish just that. In particular, by the end of this section we will be able to conclude that the determinant gives a characterization of the invertibility of a square matrix.

4.2.1 The Determinant and Elementary Row Operations

In this subsection we will discover how elementary row operations affect the determinant of a matrix. These will be essential facts for proving the big theorems of this chapter. We begin with a result that is obvious in light of Theorem 4.1.7.

Proposition 4.2.1 *Let* $A \in M_n(\mathbb{F})$*. If* A *has a row of zeros, then* $\det(A) = 0$.

Proof. To calculate $\det(A)$, we use cofactor expansion along the row of zeros. This immediately shows that $\det(A) = 0$. ■

Note 4.2.2 We observe that Proposition 4.2.1 is also true if the word "row" is replaced by "column" since a matrix and its transpose have equal determinants. The reader should consider each result in this section and reflect on whether the statement would still hold after making the same word exchange.

Now, we examine the effect of the *switch* elementary row operation.

Theorem 4.2.3 *Let* $A \in M_n(\mathbb{F})$ *and let* B *be the result of switching two rows in* A*. Then* $\det(B) = -\det(A)$.

Proof. We will proceed by induction on n. This result only makes sense for $n \geq 2$, and the base case of $n = 2$ was covered in Exercise 4.1.4.11.
We let k be an integer such that $k \geq 2$ and we assume the result is true for all $k \times k$ matrices. Let A be a $(k+1) \times (k+1)$ matrix and let B be the result of switching two rows in A. We want to show that $\det(B) = -\det(A)$.
Since $k \geq 2$, we have $k + 1 \geq 3$, which means that we can calculate $\det(B)$ by expansion along a row that is not involved in the row exchange. Suppose that B was produced by switching rows p and q. We will calculate $\det(B)$ by

expanding along row i, where i is distinct from both p and q. We have

$$\det(B) = \sum_{j=1}^{k+1} (-1)^{i+j} [B]_{ij} \det(B_{ij}).$$

We note that since i will never be p or q, $[B]_{ij} = [A]_{ij}$ for all j. Additionally, for all j, B_{ij} can be obtained by performing a switch row operation on A_{ij}. This means that, by the inductive hypothesis, we have $\det(B_{ij}) = -\det(A_{ij})$ for all j since these matrices are $k \times k$. So, we have

$$\begin{aligned}\det(B) &= \sum_{j=1}^{k+1} (-1)^{i+j} [A]_{ij} (-1) \det(A_{ij}) \\ &= -\sum_{j=1}^{k+1} (-1)^{i+j} [A]_{ij} \det(A_{ij}) \\ &= -\det(A).\end{aligned}$$

This completes the inductive step.
We have shown that the result holds for all $n \geq 2$ by the Principle of Mathematical Induction. ■

The second elementary row operation we will consider is the *scale* operation. How is the determinant of a matrix affected if one row is multiplied by a non-zero element of the field?

Theorem 4.2.4 *Let $A \in M_n(\mathbb{F})$ and let B be the result of multiplying a row in A by a non-zero $c \in \mathbb{F}$. Then* $\det(B) = c\det(A)$*.*

Proof. We will not need induction for this argument. Suppose that B is formed by multiplying row i in A by $c \in \mathbb{F}$ where $c \neq 0$. We will calculate $\det(B)$ by expanding along row i. Note that since row i is the only row affected by this operation, $B_{ij} = A_{ij}$ for all $1 \leq j \leq n$. Additionally, we note that $[B]_{ij} = c[A]_{ij}$ for all $1 \leq j \leq n$. Now we have

$$\begin{aligned}\det(B) &= \sum_{j=1}^{n} (-1)^{i+j} [B]_{ij} \det(B_{ij}) \\ &= \sum_{j=1}^{n} (-1)^{i+j} c[A]_{ij} \det(A_{ij}) \\ &= c\sum_{j=1}^{n} (-1)^{i+j} [A]_{ij} \det(A_{ij}) \\ &= c\det(A).\end{aligned}$$

■

We now present the result related to the remaining elementary row operation, the *replace* operation.

Theorem 4.2.5 *Let $A \in M_n(\mathbb{F})$ and let B be the result of replacing a row in A with the sum of that row and c times another row in A. Then* $\det(B) = \det(A)$*.*

Proof. We proceed by induction on n. This result only makes sense when $n \geq 2$, and the base case of $n = 2$ is covered in Exercise 4.1.4.12.
We let k be an integer such that $k \geq 2$ and we assume the result is true for all $k \times k$ matrices. Let A be a $(k+1) \times (k+1)$ matrix and let B be the result of replacing row q in A with the sum of row q and c times row p in A. We want

to show that $\det(B) = \det(A)$.
We observe that k is large enough that we can calculate $\det(B)$ by expanding along a row which is not row q; we will call that row i. Since $i \neq q$, we have $[B]_{ij} = [A]_{ij}$ for all $1 \leq j \leq n$. Additionally, for each j, B_{ij} is a $k \times k$ matrix which has been obtained from A_{ij} by a *replace* row operation. The inductive hypothesis means that $\det(B_{ij}) = \det(A_{ij})$ for all $1 \leq j \leq n$. Therefore, we have the following:

$$\begin{aligned}\det(B) &= \sum_{j=1}^{n}(-1)^{i+j}[B]_{ij}\det(B_{ij})\\ &= \sum_{j=1}^{n}(-1)^{i+j}[A]_{ij}\det(A_{ij})\\ &= \det(A).\end{aligned}$$

This completes the inductive step.
We have shown that the result holds for all $n \geq 2$ by the Principle of Mathematical Induction. ■

The following example shows how these three theorems can be used to calculate the determinant of a matrix using row reduction.

Example 4.2.6 Let A be the following matrix:

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

We will find $\det(A)$ using row reduction. We first switch rows 1 and 2, which introduces a negative sign:

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

Once we reduce the matrix to a triangular form, we can use Proposition 4.1.12, so we do not need to reduce the matrix to RREF, only to REF. This means that the rest of the row reduction can be performed using only the *replace* operation, which does not change the determinant:

$$\det(A) = -\begin{vmatrix} 1 & -1 & 2 \\ 2 & 0 & -3 \\ 0 & 3 & -3 \end{vmatrix} = -\begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & -7 \\ 0 & 3 & -3 \end{vmatrix} = -\begin{vmatrix} 1 & -1 & 2 \\ 0 & 2 & -7 \\ 0 & 0 & \frac{15}{2} \end{vmatrix}.$$

We have reduced the matrix far enough so that we can calculate its determinant using the product of the entries along the main diagonal:

$$\det(A) = (-1)(1)(2)(\tfrac{15}{2}) = -15.$$

□

4.2.2 Invertibility and the Determinant

We will use the results that have accumulated thus far in this section to prove two major results. First, we need to record an easy fact.

Lemma 4.2.7 *For any $n \geq 1$, we have* $\det(I_n) = 1$.

Proof. Since the identity matrix is, among other things, a triangular matrix, Proposition 4.1.12 applies. The entries along the main diagonal are all 1, so $\det(I_n) = 1$. ■

We will now apply this lemma to record the determinant of any elementary matrix.

Proposition 4.2.8 *Let $E \in M_n(\mathbb{F})$ be an elementary matrix.*

1. *If E performs the* switch *elementary row operation, then* $\det(E) = -1$.
2. *If E performs the* scale *elementary row operation, and if the scaling is by $c \in \mathbb{F}$, then* $\det(E) = c$.
3. *If E performs the* replace *elementary row operation, then* $\det(E) = 1$.

Proof. Every elementary matrix in $M_n(\mathbb{F})$ is the result of performing a single elementary row operation on I_n. We have theorems in this section which tell us how these elementary row operations affect the determinant of a matrix, and since from Lemma 4.2.7 we know that $\det(I_n) = 1$, we will be able to arrive at our result.
If E performs a *switch* row operation, then by Theorem 4.2.3 we have $\det(E) = \det(EI_n) = -\det(I_n) = -1$.
If E scales one row of a matrix by a non-zero $c \in \mathbb{F}$, then by Theorem 4.2.4 we have $\det(E) = \det(EI_n) = c\det(I_n) = c$.
Finally, if E performs a *replace* row operation, then by Theorem 4.2.5 we have $\det(E) = \det(EI_n) = \det(I_n) = 1$, which completes the proof. ■

Example 4.2.9 Sometimes, the easiest way to find a determinant by hand is to use a combination of cofactor expansion and row reduction techniques. Let $A \in M_4(\mathbb{R})$ be the following matrix:

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

To find $\det(A)$, we first use the *replace* row operation, using the 1 in the $(2,1)$ position to put zeros in the column below it:

$$A \sim \begin{bmatrix} 0 & 1 & -1 & 2 \\ 1 & 3 & 0 & -2 \\ 0 & -2 & 1 & 3 \\ 0 & 6 & -1 & -7 \end{bmatrix} = B.$$

Since the *replace* row operation doesn't change the determinant, we have $\det(A) = \det(B)$. We now use cofactor expansion along the first column to calculate $\det(B)$. Since there is only one non-zero entry in that column, we have

$$\det(B) = -\begin{vmatrix} 1 & -1 & 2 \\ -2 & 1 & 3 \\ 6 & -1 & -7 \end{vmatrix}.$$

We can now use the *replace* row operation three more times, to produce zeros in the $(2,1)$, $(3,1)$, and $(3,2)$ positions of this 3×3 matrix:

$$\det(B) = -\begin{vmatrix} 1 & -1 & 2 \\ 0 & -1 & 7 \\ 0 & 5 & -19 \end{vmatrix} = -\begin{vmatrix} 1 & -1 & 2 \\ 0 & -1 & 7 \\ 0 & 0 & 16 \end{vmatrix}.$$

We now invoke Proposition 4.1.12 to see that $\det(B) = -(-1)(16) = 16$. Since $\det(B) = \det(A)$, we have $\det(A) = 16$. □

In this next result, we use Proposition 4.2.8 to show that the determinant respects matrix multiplication, at least when one of the factors is an elementary matrix.

Theorem 4.2.10 *Let $A, E \in M_n(\mathbb{F})$, and let E be an elementary matrix. Then*

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

Proof. This argument uses Proposition 4.2.8 and requires three cases. If E performs a *switch* row operation, then we know from Theorem 4.2.3 that $\det(EA) = -\det(A)$. Since we now know that $\det(E) = -1$, we have

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

If E performs a *scale* row operation, and if the scaling is by a non-zero $c \in \mathbb{F}$, then we know from Theorem 4.2.4 that $\det(EA) = c\det(A)$. Since $\det(E) = c$, we have

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

Finally, if E performs a *replace* row operation, then we know from Theorem 4.2.5 that $\det(EA) = \det(A)$. We know that $\det(E) = 1$, so

$$\det(EA) = \det(A) = 1 \cdot \det(A) = \det(E)\det(A).$$

■

Armed with this result, we can now prove one of the most useful facts about determinants.

Theorem 4.2.11 *For any $n \times n$ matrix A over $\mathbb{F}$, A is invertible if and only if $\det(A) \neq 0$.*

Proof. For $A \in M_n(\mathbb{F})$, let $B \in M_n(\mathbb{F})$ be the unique RREF of A. From Proposition 3.3.6, we know there exist elementary matrices $E_1, \ldots, E_k$ such that

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

We can apply Theorem 4.2.10 repeatedly to see that

$$\det(A) = \det(E_1 \cdots E_k B) = \det(E_1) \cdots \det(E_k)\det(B).$$

Since $\det(E_i) \neq 0$ for each i by Proposition 4.2.8, we conclude that $\det(A) \neq 0$ if and only if $\det(B) \neq 0$.

We now assume that A is invertible. Theorem 3.3.8 tells us that $B = I_n$, so $\det(B) \neq 0$. This proves one direction of the theorem.

We will prove the contrapositive of the other direction of the theorem. We now assume that A is not invertible, which (again by Theorem 3.3.8) means that $B \neq I_n$. Specifically, B must have fewer than n pivots, which means that B must have at least one row of zeros. By Proposition 4.2.1 we have $\det(B) = 0$. Therefore, we must also have $\det(A) = 0$. ■

If a casual math student spends some time away from linear algebra, this previous theorem might be the one and only fact they remember about the determinant. It is powerful and used *frequently*.

Example 4.2.12 Using this theorem, if $A \in M_3(\mathbb{R})$ is

$$A = \begin{bmatrix} 2 & -4 & 2 \\ 1 & 0 & 3 \\ 3.5 & 2 & 12.5 \end{bmatrix},$$

then we can say that A is not invertible since $\det(A) = 0$.

We can also analyze the invertibility of matrices over other fields. Consider the matrix $B \in M_3(\mathbb{F}_5)$ given by

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

We find that $\det(B) = 0$, so B is not invertible. (If B were a matrix over $\mathbb{R}$, we would have $\det(B) = 30$. But this means that, in $\mathbb{F}_5$, $\det(B) = 0$.)

Finally, we consider another matrix $C \in M_3(\mathbb{F}_5)$:

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

Since $\det(C) = 1$ in $\mathbb{F}_5$, C is invertible. □

We present one final, important result about determinants in the last theorem of this chapter.

Theorem 4.2.13 *Let* $A, B \in M_n(\mathbb{F})$. *Then* $\det(AB) = \det(A)\det(B)$.

Proof. We will prove this in two cases. First, if A is not invertible, then neither is AB, by Exercise 3.3.4.10. This means that $\det(AB) = \det(A)\det(B)$ since, by Theorem 4.2.11, both sides of the equation are zero.
If A is invertible, then A is row equivalent to I_n, and there exist elementary matrices $E_1, \ldots, E_k$ such that

$$A = E_1 \cdots E_k I_n = E_1 \cdots E_k.$$

In the calculation that follows, we use this factorization as well as repeated application of Theorem 4.2.10. We first use Theorem 4.2.10 to peel the determinant of elementary matrices away from $\det(B)$; we then use the same result to put those determinants back together to form $\det(A)$. Here is the argument:

$$\begin{aligned} \det(AB) &= \det(E_1 \cdots E_k B) \\ &= \det(E_1)\det(E_2 \cdots E_k B) \\ &= \det(E_1)\det(E_2)\det(E_3 \cdots E_k B) = \cdots \\ &= \det(E_1)\cdots\det(E_k)\det(B) \\ &= \det(E_1 E_2)\cdots\det(E_k)\det(B) = \cdots \\ &= \det(E_1 \cdots E_k)\det(B) = \det(A)\det(B). \end{aligned}$$

This completes the proof. ■

We take a step back for a moment to marvel at this theorem. We defined matrix multiplication in the context of the composition of linear transformations (see Subsection 3.2.2), and the calculations were quite involved. The definition of the determinant was also complicated, but in a different way, so the fact that these two notions fit together so nicely is worthy of our admiration.

Example 4.2.14 In this example, we will verify Theorem 4.2.13 for a specific example. Let A and B be the following matrices:

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

We calculate AB as

$$AB = \begin{bmatrix} -6 & -8 \\ -10 & -20 \end{bmatrix}.$$

We see that $\det(A) = -2$, $\det(B) = -20$, and $\det(AB) = 40$, so the relationship $\det(AB) = \det(A)\det(B)$ holds. □

4.2.3 Reading Questions

1. Consider the following three matrices:

$$A = \begin{bmatrix} 3 & -1 & -2 \\ 1 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 0 \\ 3 & -1 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & -1 & -2 \\ 1 & 2 & 0 \\ 0 & -1 & 2 \end{bmatrix}.$$

(a) Calculate $\det(A)$ using cofactor expansion along some row or column. Show your work.

(b) The matrix A_1 was obtained from A by a single elementary row operation. Which one?

(c) Knowing $\det(A)$ and given your answer to (b), what do you predict $\det(A_1)$ to be? (Consult Theorem 4.2.3.)

(d) Calculate $\det(A_1)$ using cofactor expansion along some row or column. Show your work.

(e) The matrix A_2 was obtained from A by a single elementary row operation. Which one?

(f) Knowing $\det(A)$ and given your answer to (e), what do you predict $\det(A_2)$ to be? (Consult Theorem 4.2.5.)

(g) Calculate $\det(A_2)$ using cofactor expansion along some row or column. Show your work.

2. Verify Theorem 4.2.13 for the following two matrices A and B:

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

(You should follow Example 4.2.14)

4.2.4 Exercises

1. Find the determinant of the matrix using row reduction.

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

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

2. Find the determinant of the matrix using row reduction.

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

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

3. Find the determinant using a combination of row reduction and cofactor expansion:

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

4. Find the determinant using a combination of row reduction and cofactor expansion:

$$A = \begin{bmatrix} -1 & 2 & 1 & 4 \\ 3 & -4 & 1 & -3 \\ 4 & -10 & -1 & 0 \\ -1 & 4 & 2 & 3 \end{bmatrix}.$$

5. Use the determinant to determine whether or not the matrix is invertible. (Note that not all fields are $\mathbb{R}$!)

(a) $A \in M_3(\mathbb{F}_3)$, $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 2 \\ 2 & 2 & 0 \end{bmatrix}$

(b) $A \in M_3(\mathbb{R})$, $A = \begin{bmatrix} -3 & -1 & -1 \\ 0 & -3 & -3 \\ 2 & -3 & 3 \end{bmatrix}$

(c) $A \in M_3(\mathbb{F}_5)$, $A = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 1 \\ 4 & 1 & 3 \end{bmatrix}$

(d) $A \in M_2(\mathbb{C})$, $A = \begin{bmatrix} 2+i & 2-3i \\ 4-i & -2+4i \end{bmatrix}$

(e) $A \in M_3(\mathbb{C})$, $A = \begin{bmatrix} 0 & 3-2i & -2-4i \\ -2 & 2+4i & 0 \\ 3+i & -1+i & 0 \end{bmatrix}$

6. Calculate $\det(A^3)$ if

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

7. Construct an invertible matrix $A \in M_3(\mathbb{R})$. For each entry of A, compute the corresponding cofactor. Create a new 3×3 matrix with these cofactors in the same position as the entry of A on which they were based; call this matrix C. Calculate AC^T. What do you observe?

Writing Exercises

8. Suppose that A is a square matrix with two identical columns. Prove that $\det(A) = 0$.

9. Suppose that $A \in M_n(\mathbb{F})$ is invertible. Prove that $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

10.

(a) Suppose that $A \in M_n(\mathbb{R})$ and that $\det(A^4) = 0$. Prove that A is not invertible.

(b) Does the result or your argument in part (a) change if $\mathbb{R}$ is replaced with $\mathbb{F}_7$? Explain.

11. Suppose that $A, B \in M_n(\mathbb{F})$. Show that $\det(AB) = \det(BA)$ regardless of whether or not $AB = BA$.

12. Let $A \in M_n(\mathbb{F})$ and let $k \in \mathbb{F}$. Find a formula for $\det(kA)$ and prove that your formula is correct.

13.

(a) Verify that $\det(A) = \det(B) + \det(C)$ where A, B, and C are

$$A = \begin{bmatrix} a+e & b+f \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad C = \begin{bmatrix} e & f \\ c & d \end{bmatrix}.$$

(b) Let A and B be

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Show that $\det(A + B) = \det(A) + \det(B)$ if and only if $a + d = 0$.

(c) Provide an example where $A, B \in M_3(\mathbb{R})$ to prove that $\det(A+B) = \det(A) + \det(B)$ is not always true.

14. Consider the following matrix (called a *Vandermond* matrix):

$$V = \begin{bmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{bmatrix}.$$

(a) Use row operations to explain why $\det(V) = (b - a)(c - a)(c - b)$.

(b) Explain why V is invertible if and only if a, b, and c are all distinct real numbers.

Chapter 5

The Dimension of a Vector Space

Thus far in this text, the only way we have related vector spaces to each other is through linear transformations between those spaces. But we have not had any *intrinsic* quality of a vector space that enables comparison between spaces.

The notion of the *dimension* of a vector space allows just such a comparison. In this chapter we will develop the necessary machinery for defining dimension, and we will relate this concept to matrices, linear transformations, and more.

5.1 Linear Independence

Linear independence—or, rather, its opposite—is related to the idea of *redundancy*. If there is a *linear dependence* among a set of vectors, then we don't need all of those vectors to produce the same span.

Definition 5.1.1 Consider a set of vectors $V' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ in a vector space V. When $n \geq 2$, we say that V' is **linearly dependent** if, for some i, $1 \leq i \leq n$, $\mathbf{v}_i$ is a linear combination of the other vectors in the set. When $n = 1$, the set V' is **linearly dependent** if and only if $\mathbf{v}_1 = \mathbf{0}$. ♢

Example 5.1.2 Consider the following three vectors in $\mathbb{F}_3^3$:

$$\mathbf{u} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}.$$

The set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent since $\mathbf{v} = 2\mathbf{u} + \mathbf{w}$. □

The definition of *linear dependence* we have given is the intuitive one, but it is not the one most widely used. The following result provides an equivalent definition of linear dependence which is much easier to deploy.

Proposition 5.1.3 *A set of vectors $V' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ in a vector space V is linearly dependent if and only if there exist $c_1, \dots, c_n \in \mathbb{F}$, not all of which are 0, such that*

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Proof. We will first dispatch with the case where $n = 1$. If $n = 1$ and V' is linearly dependent, then $\mathbf{v}_1 = \mathbf{0}$. Then the equation $1\mathbf{v}_1 = \mathbf{0}$ is satisfied. Conversely, if $c_1\mathbf{v}_1 = \mathbf{0}$ for some $c_1 \neq 0$, then by Theorem 2.3.12, Item 7, we

must have $\mathbf{v}_1 = \mathbf{0}$, meaning V' is linearly dependent.
We now consider the case where $n \geq 2$. If V' is linearly dependent, then some vector in V', call it $\mathbf{v}_j$, is a linear combination of the other vectors in V'. This means we have

$$\mathbf{v}_j = c_1\mathbf{v}_1 + \cdots + c_{j-1}\mathbf{v}_{j-1} + c_{j+1}\mathbf{v}_{j+1} + \cdots + c_n\mathbf{v}_n.$$

If we subtract $\mathbf{v}_j$ from both sides, we have

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_{j-1}\mathbf{v}_{j-1} - \mathbf{v}_j + c_{j+1}\mathbf{v}_{j+1} + \cdots + c_n\mathbf{v}_n.$$

Since we now have written $\mathbf{0}$ as a non-trivial linear combination of the vectors in V'—that is, the coefficients in the linear combination are not all zero—we have completed half of the proof.
We now suppose that there is a linear combination of the vectors in V',

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n,$$

where not all of the coefficients are zero. If $c_j \neq 0$, then we can write

$$\begin{aligned} -c_j\mathbf{v}_j &= c_1\mathbf{v}_1 + \cdots + c_{j-1}\mathbf{v}_{j-1} + c_{j+1}\mathbf{v}_{j+1} + \cdots + c_n\mathbf{v}_n \\ \mathbf{v}_j &= \left(-\frac{c_1}{c_j}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{j-1}}{c_j}\right)\mathbf{v}_{j-1} \\ &\quad + \left(-\frac{c_{j+1}}{c_j}\right)\mathbf{v}_{j+1} + \cdots + \left(-\frac{c_n}{c_j}\right)\mathbf{v}_n. \end{aligned}$$

Thus we have written $\mathbf{v}_j$ as a linear combination of the other vectors in V', so V' is linearly dependent. ■

We will often use this statement in Proposition 5.1.3 as our definition of linear dependence.

Definition 5.1.4 A set of vectors $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in a vector space V is **linearly independent** if it is not linearly dependent. ♢

Note 5.1.5 In practice, we will think about linear independence in the following way. A set $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent if the vector equation

$$x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0}$$

has only the trivial solution.

Further, when a set V' is linearly dependent, then we will call a non-trivial linear combination

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

a **linear dependence relation** for the vectors in V'.

We will try to make the notions of linear dependence and linear independence more concrete with some examples.

Example 5.1.6 Consider the set $V' = \{p_1, p_2\}$ in P_2, where

$$p_1 = 1 + t \text{ and } p_2 = 3t^2.$$

We can see that the set V' is linearly independent, because the only way to produce the zero polynomial from the combination $c_1p_1 + c_2p_2$ is if $c_1 = c_2 = 0$. This is relatively easy to see in this example, since the degrees of t are not at all shared between p_1 and p_2. If the coefficient of t^2 must be zero in the sum $c_1p_1 + c_2p_2$, then we must have $c_2 = 0$. And if the coefficient of t must be zero in the sum $c_1p_1 + c_2p_2$, then we must have $c_1 = 0$. □

Example 5.1.7 Consider the following vectors in $\mathbb{R}^2$:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ -2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ -5 \end{bmatrix}.$$

We can show that $\mathbf{v}_3 \in \mathrm{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$ by a familiar matrix reduction:

$$\begin{bmatrix} 2 & -4 & 5 \\ 4 & -2 & -5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2.5 \\ 0 & 1 & -2.5 \end{bmatrix}.$$

This shows us that $\mathbf{v}_3 = -2.5\mathbf{v}_1 - 2.5\mathbf{v}_2$, which proves that the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent. Further, we can conclude that the following is a linear dependence relation for the vectors in the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$:

$$\mathbf{0} = 2.5\mathbf{v}_1 + 2.5\mathbf{v}_2 + \mathbf{v}_3.$$

□

We will see now that sets of *two vectors* are particularly nice when it comes to determining linear independence. (This means that there was an easier way to complete Example 5.1.6.)

Consider a set $V' = \{\mathbf{v}, \mathbf{w}\}$ in a vector space V. If V' is linearly dependent, then we can write $\mathbf{v} = c\mathbf{w}$ or $\mathbf{w} = d\mathbf{v}$ for some $c, d \in \mathbb{F}$. Conversely, if V' is linearly independent, then we cannot have either $\mathbf{v} = c\mathbf{w}$ or $\mathbf{w} = d\mathbf{v}$. This means that we have a handy characterization of linear dependence for sets of two vectors.

Fact 5.1.8 *A set of two vectors $\{\mathbf{v}, \mathbf{w}\}$ is linearly dependent if and only if at least one of the vectors is a multiple of the other. The set is linearly independent if and only if neither vector is a multiple of the other.*

Example 5.1.9 If $\mathbf{v}$ and $\mathbf{w}$ are the following vectors in $\mathbb{R}^3$,

$$\mathbf{v} = \begin{bmatrix} 4 \\ 3 \\ -3 \end{bmatrix} \quad \text{and} \quad \mathbf{w} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix},$$

then the set $\{\mathbf{v}, \mathbf{w}\}$ is linearly independent since neither $\mathbf{v}$ nor $\mathbf{w}$ is a multiple of the other. □

There is one other notable fact that will allow us to determine whether particular sets of vectors are linearly dependent.

Fact 5.1.10 *Any set of vectors that contains the zero vector is linearly dependent. This is true because a linear dependence relation is easy to construct. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of vectors in a vector space V, and if $\mathbf{v}_i = \mathbf{0}$, then*

$$\mathbf{0} = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_n$$

is a non-trivial linear combination of the vectors in the set which produces the zero vector.

A reader may guess that we will occasionally need to figure out whether or not a given set of vectors is linearly independent. As with questions about span, this turns out to be easier when the vector space is $\mathbb{F}^n$ for some field $\mathbb{F}$. For other sorts of vector spaces, we will need different methods.

Proposition 5.1.11 *Let $A \in M_{m,n}(\mathbb{F})$. The columns of A are linearly independent if and only if $\mathbf{x} = \mathbf{0}$ is the only solution to the linear system represented by the equation $A\mathbf{x} = \mathbf{0}$. That is, the columns of A are linearly independent if and only if* $\mathrm{null}(A) = \{\mathbf{0}\}$.

Proof. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{F}^m$ be the columns of A. Then the vector form of the equation $A\mathbf{x} = \mathbf{0}$ is

$$x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0}.$$

If the columns of A are linearly independent, then the only solution to this equation is $\mathbf{x} = \mathbf{0}$, which means that $\mathrm{null}(A) = \{\mathbf{0}\}$. Alternatively, if the columns of A are linearly dependent, then $\mathrm{null}(A)$ contains a non-zero vector—namely, the vector of scalars which provides a linear dependence relation for these vectors. ■

Note 5.1.12 The reader may note the slight abuse of terminology in the previous proof. We referred to the columns of a matrix being linearly dependent or independent instead of the set containing the columns of the matrix. We trust that the reader will forgive and overlook this misstep since the meaning is clear and the verbal gymnastics needed to be precise at all times can prove tiresome.

The following proposition provides another test of the linear dependence of a set of vectors.

Corollary 5.1.13 *If $n > m$, then every set of n vectors in $\mathbb{F}^m$ is linearly dependent.*

Proof. Let $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in $\mathbb{F}^m$ and let A be the matrix with the elements of V' as its columns. By Corollary 1.3.7 (or, technically, by the version of this result generalized to any field $\mathbb{F}$), we know that the $m \times n$ linear system represented by $A\mathbf{x} = \mathbf{0}$ cannot have a unique solution. Since $\mathbf{x} = \mathbf{0}$ is a known solution, the presence of another solution means that the columns of A must be linearly dependent by Proposition 5.1.11. ■

Proposition 5.1.11 provides us with a convenient algorithm to determine whether or not a set of vectors in $\mathbb{F}^m$ is linearly independent.

Algorithm 5.1.14 *Suppose $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a set of vectors in a vector space $\mathbb{F}^m$. In order to determine whether or not V' is linearly independent, we follow these steps.*

1. *Form the matrix $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ and row reduce it to REF. Call this matrix B.*

2. *The set V' is linearly independent if and only if B has a pivot in every column.*

Example 5.1.15 Consider the following vectors in $\mathbb{F}_5^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}.$$

The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent because the matrix $A = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3]$ has I_3 as its RREF. □

Example 5.1.16 We consider the following vectors in $\mathbb{R}^4$:

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 4.5 \\ 3.5 \\ -4.5 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3.5 \\ 5 \\ 4 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -5 \\ 28 \\ 22 \\ -14 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 5 \\ 0 \\ 0.5 \\ -4 \end{bmatrix}.$$

The set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is linearly dependent, because the matrix A which has the vectors $\mathbf{v}_i$ as its columns has the following RREF:

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

□

Algorithm 5.1.14 only covers the situations when our vectors come from some $\mathbb{F}^m$. In the case of other vector spaces, we will need to do more work.

Example 5.1.17 Consider the following three elements of P_2:

$$p_1 = 1 + t, \quad p_2 = t - t^2, \quad p_3 = 2 + 2t + t^2.$$

To determine whether or not the set $\{p_1, p_2, p_3\}$ is linearly dependent, we need to return to the definition. Suppose that we have

$$c_1 p_1 + c_2 p_2 + c_3 p_3 = 0$$

for some $c_1, c_2, c_3 \in \mathbb{R}$. In other words, this linear combination is the zero polynomial, so we have

$$c_1 p_1 + c_2 p_2 + c_3 p_3 = 0 + 0t + 0t^2.$$

For these specific polynomials, this means we have

$$\begin{aligned} c_1(1+t) + c_2(t - t^2) + c_3(2 + 2t + t^2) &= 0 + 0t + 0t^2 \\ (c_1 + 2c_3)1 + (c_1 + c_2 + 2c_3)t + (-c_2 + c_3)t^2 &= 0 + 0t + 0t^2. \end{aligned}$$

Since the coefficients of the corresponding powers of t must be equal on both sides of this equation, we have a linear system to solve:

$$\begin{aligned} c_1 + 2c_3 &= 0 \\ c_1 + c_2 + 2c_3 &= 0 \\ -c_2 + c_3 &= 0. \end{aligned}$$

Our variables in this system are c_1, c_2, and c_3, and we solve the system using the techniques from Section 1.3. We find that

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

This shows that the only solution to this linear system is the trivial one: $c_1 = c_1 = c_3 = 0$. That means that the set $\{p_1, p_2, p_3\}$ is linearly independent. □

We end this section with two additional results.

Corollary 5.1.18 *A set of n vectors in $\mathbb{F}^n$ is linearly independent if and only if that set spans $\mathbb{F}^n$.*

Proof. Let $A = [\mathbf{v}_1 \cdots \mathbf{v}_n] \in M_n(\mathbb{F})$. By Theorem 3.4.14, we know that the set $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ spans $\mathbb{F}^n$ if and only if the RREF of A has a pivot in every row. On the other hand, Algorithm 5.1.14 says that V' is linearly independent if and only if the RREF of A has a pivot in every column. Since A is a square matrix, each of these happen exactly when the RREF of A is I_n. ■

The following result appears to be little more than a slight restatement of the definition of linear dependence. However, the precise wording used in this theorem turns out to be quite useful in proving some results later in the text.

Theorem 5.1.19 The Linear Dependence Lemma. *Suppose* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is a linearly dependent set in a vector space* V *and that* $\mathbf{v}_1 \neq \mathbf{0}$. *Then there exists* $j \in \{2, \ldots, n\}$ *such that* $\mathbf{v}_j \in \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}\}$.

Proof. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a linearly dependent set of vectors in a vector space V, and suppose that $\mathbf{v}_1 \neq \mathbf{0}$. Then there exist scalars $c_1, \ldots, c_n$, not all of which are zero, such that

$$\mathbf{0} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Let k be the largest index such that $c_k \neq 0$. It must be that $k \geq 2$ since we assumed $\mathbf{v}_1 \neq \mathbf{0}$. Then

$$c_k\mathbf{v}_k = -c_1\mathbf{v}_1 - \cdots - c_{k-1}\mathbf{v}_{k-1}.$$

Since $c_k \neq 0$, we have

$$\mathbf{v}_k = \left(-\frac{c_1}{c_k}\right)\mathbf{v}_1 + \cdots + \left(-\frac{c_{k-1}}{c_k}\right)\mathbf{v}_{k-1}.$$

This shows that $\mathbf{v}_k \in \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\}$ and completes the proof. ■

Note 5.1.20 It is important to record that Theorem 5.1.19 doesn't say that in linearly dependent sets *every* vector is a combination of the vectors that preceed it. We merely have the *existence* of a vector with that property.

Reading Questions

1. For each of the following, determine whether the given set of vectors in $\mathbb{R}^3$ is linearly dependent or linearly independent. (You should NOT need to do any matrix row reduction to figure this out.) Refer to a fact or theorem from the section when you are giving your answer.

(a) $\{\mathbf{v}_1, \mathbf{v}_2\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 4 \\ -10 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -6 \\ 15 \end{bmatrix}$$

(b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -8 \\ -9 \\ 3 \end{bmatrix}$$

(c) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 5 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -8 \\ -9 \\ 3 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 10 \\ 7 \\ -2 \end{bmatrix}$$

(d) $\{\mathbf{v}_1, \mathbf{v}_2\}$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -7 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

2. Determine whether the following sets in P_3 are linearly independent. Explain your answers. You should not need to do any calculations.

(a) $\{1+t^2, t^3\}$

(b) $\{1, 2t^2, -7+6t^2\}$

(c) $\{2-5t^2, -4+10t^2\}$

Exercises

1. For each of the following, determine by inspection (without doing any calculation) whether the given set is linearly dependent or linearly independent. Explain your answers.

(a) $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{R}^3$ where

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 8 \\ -4 \end{bmatrix}$$

(b) $\{p_1, p_2\}$ in P_2 where

$$p_1 = 2t - 4t^2, \quad p_2 = -t + 2t^2$$

(c) $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{F}_5^2$ where

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

(d) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{F}_7^2$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

(e) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -3 \\ 9 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ -2 \\ -4 \end{bmatrix}$$

2. Determine the value(s) of c, if any, that will make the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ linearly independent in $\mathbb{R}^3$.

(a) $\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ -4 \\ -3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 9 \\ -11 \\ c \end{bmatrix}$

(b) $\mathbf{v}_1 = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ -6 \\ -4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -2 \\ c \end{bmatrix}$

3. Determine the value(s) of c, if any, that will make the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ linearly dependent in $\mathbb{R}^3$, if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ -4 \\ -4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -2 \\ 0 \\ c \end{bmatrix}.$$

4. How many pivot columns must a 6×4 matrix have if its columns are linearly independent? Explain.

5. Determine whether the following statements are true or false. Justify your answer either way.

(a) If $V' = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a subset of a vector space V and $\mathbf{v}_2$ is not a scalar multiple of $\mathbf{v}_1$, then V' is linearly independent.

(b) If $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ is a subset of a vector space V and $\mathbf{v}_3$ is not a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_4$, then V' is linearly independent.

(c) If $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ is a subset of a vector space V and both $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\{\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ are linearly independent, then V' is linearly independent.

(d) If $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_4\}$ is a subset of a vector space V and V' is linearly independent, then $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is also linearly independent.

6. Determine whether or not the following set of vectors is linearly independent in the given vector space.

(a) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^4$ if

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ -2.5 \\ -5 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 4 \\ 1 \\ -1 \\ -4.5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 4 \\ -1 \\ 3 \\ 2.5 \end{bmatrix}$$

(b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{F}_5^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

(c) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{F}_3^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

7. For each of the following subsets $\{p_1, p_2, p_3\}$ of P_2, determine whether the set is linearly dependent or linearly independent. Explain your answers.

(a) $p_1 = 3 + 5t^2$, $p_2 = -5 - 3t + 2t^2$, $p_3 = -4 - 5t - 2t^2$

(b) $p_1 = 2 - t + t^2$, $p_2 = -3 + 5t - 12t^2$, $p_3 = -2 - 2t + 8t^2$

Writing Exercises

8. Let $T : V \to W$ be a linear transformation between vector spaces.

(a) Prove that if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly dependent set in V, then $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is a linearly dependent set in W.

(b) Prove that if T is injective and if $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is a linearly dependent set in W, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a linearly dependent set in V.

9. Suppose that V_1 and V_2 are subsets of a vector space V. Prove that if $V_1 \subseteq V_2$ and V_1 is linearly dependent, then V_2 is linearly dependent.

10. Let $A \in M_n(\mathbb{F})$ and let $T : \mathbb{F}^n \to \mathbb{F}^n$ be a linear transformation.

(a) Prove that $\text{null}(A) = \{\mathbf{0}\}$ if and only if $\text{col}(A) = \mathbb{F}^n$.

(b) Prove that T is injective if and only if it is surjective.

11. Let V be a vector space over $\mathbb{Q}$ and let $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a subset of V. Prove that V' is linearly dependent if and only if there exist *integers* $c_1, \ldots, c_n$, not all of which are zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}.$$

12. Let $A \in M_{m,n}(\mathbb{F})$ and suppose that $\text{null}(A) \neq \{\mathbf{0}\}$. Prove that the set of vectors that spans $\text{null}(A)$ is linearly independent.

5.2 Basis of a Vector Space

We have previously examined when a set of vectors spans a vector space. In this section, we will learn how to work with the most efficient spanning set possible.

5.2.1 The Definition of a Basis

We begin with the notion of finite- and infinite-dimensional vector spaces.

Definition 5.2.1 A vector space V is **finite-dimensional** if there is a finite set of vectors which spans V. A vector space is **infinite-dimensional** if it is not finite-dimensional. ♢

We recall that *linear independence* in Section 5.1 was introduced as a way to eliminate redundancy. We pick up on this idea in the next definition.

Definition 5.2.2 Let V be a finite-dimensional vector space. Then a set $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a **basis** for V if B is a linearly independent set and if $V = \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. ♢

Note 5.2.3 The notion of a basis exists for infinite-dimensional vector spaces, but since the overwhelming majority of our work will be with finite-dimensional spaces, we have only given the definition in that setting.

Example 5.2.4 We recall that $\mathbf{e}_i$ is the vector in $\mathbb{F}^n$ with a 1 in the ith coordinate and zeros elsewhere. Then the set $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{F}^n$. If we form the $n \times n$ matrix with these vectors as columns, we see that it is the $n \times n$ identity matrix. Since there is a pivot in every column, E is linearly independent according to Algorithm 5.1.14. Then Corollary 5.1.18 tells us that E also spans $\mathbb{F}^n$. This proves that E is a basis for $\mathbb{F}^n$.

We call this basis the **standard basis** for $\mathbb{F}^n$. □

Example 5.2.5 We now consider the set $B = \{1, t, t^2\}$ within the vector space P_2. Since any vector in P_2 can be written as $a(1) + b(t) + c(t^2)$, it is clear that B spans P_2. It is also true that B is linearly independent: the set $\{1, t\}$ is linearly independent since neither vector is a scalar multiple of the other. And then since t^2 is not a linear combination of 1 and t, we conclude that B is linearly independent by (the contrapositive of) the Linear Dependence Lemma (Theorem 5.1.19). This proves that B is a basis for P_2.

The analogous basis for P_n, $\{1, t, \ldots, t^n\}$, is often called the **standard basis** for P_n. □

Example 5.2.6 Consider the following matrix $A \in M_{3,5}(\mathbb{R})$:

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

We will find a basis for null(A).

Following the procedure we first encountered in Example 3.4.5, we start by finding the RREF of A:

$$A \sim \begin{bmatrix} 1 & 0 & 0 & -5/2 & -2 \\ 0 & 1 & 0 & -7/4 & -5/2 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

We see that x_4 and x_5 are free variables, and that any vector $\mathbf{x}$ in null(A) can be written as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} (5/2)x_4 + 2x_5 \\ (7/4)x_4 + (5/2)x_5 \\ x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_4 \begin{bmatrix} 5/2 \\ 7/4 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 5/2 \\ 1 \\ 0 \\ 1 \end{bmatrix}.$$

If we label the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 5/2 \\ 7/4 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 5/2 \\ 1 \\ 0 \\ 1 \end{bmatrix},$$

then we can see that null$(A) = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Further, we see that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent (neither vector is a scalar multiple of the other), so $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for null(A). □

Note 5.2.7 What we observed in Example 5.2.6 is true more generally. Since the method we use to find a spanning set for null(A) always produces a linearly independent set (see Exercise 5.1.12), this method will always produce a basis for null(A).

Here is an example where we are looking at whether a set of two vectors is a basis.

Example 5.2.8 It turns out that it is fairly easy to tell whether a set of two vectors in $\mathbb{R}^2$ forms a basis for $\mathbb{R}^2$. Since linear independence is easy to check with two vectors—is either vector a scalar multiple of the other?—we can focus on this characteristic. This means that the set $\{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}$$

is a basis for $\mathbb{R}^2$. Neither vector is a scalar multiple of the other, so the set is linearly independent. And then Corollary 5.1.18 tells us that this set must also span $\mathbb{R}^2$. (We could also easily see this by row reducing the matrix $[\mathbf{v}_1\ \mathbf{v}_2]$.)

On the other hand, the set $W' = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{w}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_2 = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

is not a basis for $\mathbb{R}^2$. Since $\mathbf{w}_2 = -2\mathbf{w}_1$, W' is not linearly independent, so it cannot be a basis. □

Putting some facts together, there is a fairly straightforward condition for when a list of vectors in $\mathbb{F}^n$ is a basis for that space.

Proposition 5.2.9 *The set $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a basis for $\mathbb{F}^n$ if and only if the RREF of the matrix $[\mathbf{v}_1 \cdots \mathbf{v}_m]$ is I_n.*

Proof. From Theorem 3.4.14 we know that the set $V' = \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ spans $\mathbb{F}^m$ if and only if the RREF of $A = [\mathbf{v}_1 \cdots \mathbf{v}_m]$ has a pivot in every row. Additionally, Algorithm 5.1.14 tells us that V' is linearly independent if and only if the RREF of A has a pivot in each column. The only way a matrix in RREF can have a pivot in every row and every column is if that RREF is the identity matrix. ■

We put this proposition into action in the following example.

Example 5.2.10 Let $A \in M_3(\mathbb{F}_5)$ be the following matrix:

$$A = \begin{bmatrix} 3 & 4 & 4 \\ 3 & 0 & 1 \\ 4 & 3 & 4 \end{bmatrix}.$$

We will label column i in A as the vector $\mathbf{v}_i \in \mathbb{F}_5^3$.

Since the RREF of A is

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

the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is not a basis for $\mathbb{F}_5^3$.

On the other hand, if $B \in M_3(\mathbb{F}_5)$ is the matrix

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

then the columns of B form a basis for $\mathbb{F}_5^3$ since the RREF of B is I_3. □

5.2.2 The Properties of a Basis

Having a basis is a powerful tool. In particular, it guarantees a uniqueness that is quite useful.

Theorem 5.2.11 The Unique Representation Theorem. *A set of vectors $V' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ in a vector space V is a basis for V if and only if each element of V can be uniquely represented as a linear combination of the vectors in V'.*

Proof. We will prove the forward direction of this biconditional statement directly. Suppose that $V' = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis of V. Since $V = \mathrm{Span}(V')$, every vector in V can be written as a linear combination of the vectors in V'. Let $\mathbf{v}$ be a vector in V, and suppose that $\mathbf{v}$ can be written as a linear combination of the vectors in V' in two ways:

$$\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{v}_i \quad \text{and} \quad \mathbf{v} = \sum_{i=1}^{n} b_i \mathbf{v}_i.$$

We want to show that $a_i = b_i$ for each i, $1 \le i \le n$. Since both of these representations are equal to $\mathbf{v}$, they are equal to each other, so we have

$$\mathbf{0} = \sum_{i=1}^{n} a_i \mathbf{v}_i - \sum_{i=1}^{n} b_i \mathbf{v}_i = \sum_{i=1}^{n} (a_i - b_i)\mathbf{v}_i.$$

Since V' is a linearly independent set (since we are assuming it is a basis), it must be that $a_i - b_i = 0$ for each i. Therefore, $a_i = b_i$ and the representation of $\mathbf{v}$ is unique.
For the other direction, we suppose that every element of V can be uniquely represented as a linear combination of the vectors in V'. Since every element of V can be represented as a linear combination of the vectors in V', we see that V' spans V. Since every element in V can be represented *uniquely* as a linear combination of the vectors in V', and since $\mathbf{0} \in V$ can be represented as the trivial linear combination of the vectors in V', this means that V' is linearly independent. (The trivial linear combination of vectors in V' is the *only* way to obtain $\mathbf{0}$ as a linear combination of the vectors in V'.) Since V' is linearly independent and spans V, this proves that V' is a basis for V. ■

This next result shows us how to trim a spanning set down until we reach a basis.

Theorem 5.2.12 The Spanning Set Theorem. *Suppose that V is a nonzero vector space and that $V = \mathrm{Span}(B)$ for some set of vectors $B \subseteq V$.*

1. *If B is a linearly dependent set and a vector $\mathbf{w} \in B$ can be written as a linear combination of the rest of the vectors in B, then $\mathrm{Span}(B) = \mathrm{Span}(B - \{\mathbf{w}\})$.*

2. *A subset of B is a basis for V.*

Proof. We suppose that $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. If B is linearly dependent, then by the Linear Dependence Lemma (Theorem 5.1.19), there exists a vector $\mathbf{v}_k \in B$ such that $\mathbf{v}_k$ can be written as a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}$. We suppose this combination is

$$\mathbf{v}_k = a_1\mathbf{v}_1 + \cdots + a_{k-1}\mathbf{v}_{k-1}. \tag{5.1}$$

Now suppose $\mathbf{v}$ is a vector in V. We have

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_{k-1}\mathbf{v}_{k-1} + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n. \tag{5.2}$$

Using (5.1), we can substitute this expression in for $\mathbf{v}_k$ in (5.2) and, once the algebraic dust settles, we will have $\mathbf{v}$ written as a combination of the vectors in $B - \{\mathbf{v}_k\}$. This shows that $\mathrm{Span}(B) = \mathrm{Span}(B - \{\mathbf{v}_k\})$. (Since $B - \{\mathbf{v}_k\} \subseteq B$, it is true that $\mathrm{Span}(B - \{\mathbf{v}_k\}) \subseteq \mathrm{Span}(B)$. The argument thus far in this proof has established the other subset containment.)
If B is linearly independent, then it is already a basis for V. If it is linearly dependent, then we can remove a vector according to the above procedure to obtain a set $B_1 = B - \{\mathbf{w}\}$ which still spans V. As long as there are two or more vectors in the spanning set, we can repeat this process until we are left with a linearly independent set and thus a basis. If the spanning set is eventually reduced to a single vector, that vector will be nonzero since V is nonzero, and therefore that set will be linearly independent and therefore a basis. ■

Corollary 5.2.13 *Every finite-dimensional vector space has a basis.*

Proof. Since a finite-dimensional vector space by definition has a finite spanning set B, Theorem 5.2.12 tells us that a subset of B will be a basis for the vector space. ■

While the proof of Theorem 5.2.12 provides a way to trim a spanning set down to a basis, it does not offer a practical method for this process. The following algorithm provides such a method for certain vector spaces.

Algorithm 5.2.14 *Let $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of vectors in $\mathbb{F}^m$. The following steps result in a basis B for* $\operatorname{Span}(V')$*.*

1. *Put the matrix $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$ in RREF.*
2. *If column i in the RREF contains a pivot, then include $\mathbf{v}_i$ in B.*

Proof. We form the matrix $A = [\mathbf{v}_1 \cdots \mathbf{v}_n]$. If A is already in RREF, then the non-pivot columns are linear combinations of the pivot columns that preceed them. So, those can be discarded and the pivot columns will be a basis, according to Theorem 5.2.12.

We will complete the proof with a reminder about the effect of elementary row operations on the columns of a matrix. If a column $\mathbf{v}_k$ of A is a linear combination of the columns that preceed it, then

$$\mathbf{v}_k = \sum_{i=1}^{k-1} c_i \mathbf{v}_i$$

for some scalars c_i. This means that the column vector $[c_i]$ is a solution to the linear system represented by the augmented matrix $[\mathbf{v}_1 \cdots \mathbf{v}_{k-1} \mid \mathbf{v}_k]$. One of the earliest facts we learned about elementary row operations is that they preserve the solution sets of linear systems, so the same vector $[c_i]$ will be a solution to the linear system represented by the RREF of $[\mathbf{v}_1 \cdots \mathbf{v}_{k-1} \mid \mathbf{v}_k]$. This proves that the relationships between the columns of a matrix are the same as the relationships between the columns of the RREF of that matrix.

So, if A is not in RREF, we can find the RREF of A, call it C. The non-pivot columns of C indicate that the corresponding columns of A should not be included in the basis. In other words the pivot columns of C indicate that the corresponding columns of A are the ones that should remain to form the basis. ■

Note 5.2.15 We emphasize here that the pivot columns in the reduced matrix do not provide the vectors for the basis! The pivot columns merely provide the *instructions* for which of the original vectors should be kept to form the basis.

Example 5.2.16 Consider the following matrix $A \in M_{4,5}(\mathbb{F}_5)$:

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

We will find a basis for $\operatorname{col}(A)$ using Algorithm 5.2.14. When we put A into RREF, we find

$$A \sim \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The pivots are in columns 1, 3, and 5, so a basis for col(A) is $\{\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_5\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_5 = \begin{bmatrix} 3 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \quad .$$

□

We arrive at the end of this section with two helpful perspectives on a basis. A basis can be formed by trimming a spanning set down until it is linearly independent. Thus, a basis is a spanning set that is as small as possible. On the other hand, a linearly independent set can always be enlarged until it spans. Therefore, a basis is also a linearly independent set that is as large as possible.

5.2.3 Reading Questions

1. Consider the set $V' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{R}^3$ where

$$\mathbf{v}_1 = \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 6 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \\ -1 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 2 \\ -2 \\ 7 \end{bmatrix}.$$

Find a basis for Span(V'). Follow Example 5.2.16 and explain your answer.

2. Determine whether or not the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ forms a basis for $\mathbb{F}_7^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 5 \\ 6 \\ 2 \end{bmatrix}.$$

Explain your answer.

5.2.4 Exercises

1. For each of the following, determine whether the given set of vectors forms a basis for the indicated vector space.

(a) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{R}^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 0.5 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -4.5 \\ -2.5 \\ 4.5 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 4.5 \\ 1.5 \\ -4.5 \end{bmatrix}$$

(b) $\{\mathbf{v}_1, \mathbf{v}_2\}$ in $\mathbb{F}_5^2$ if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

(c) $\{p_1, p_2, p_3\}$ in P_2 if

$$p_1 = 4 + 2t + 4t^2, \quad p_2 = -3 + 4t, \quad p_3 = 2 - 2t - 4t^2$$

2. For each of the following, determine whether the given set of vectors forms a basis for the indicated vector space.

(a) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{R}^4$ if

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ -2 \\ 1 \\ -3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -4 \\ -2 \\ 5 \\ -13 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

(b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbb{F}_3^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

(c) $\{p_1, p_2, p_3\}$ in P_2 if

$$p_1 = 3 + 3t - 4t^2, \quad p_2 = 3 - 3t, \quad p_3 = 12 + 6t - 12t^2$$

3. For each matrix A, find a basis for null(A) and col(A).

(a) $A \in M_{3,4}(\mathbb{R})$,

$$A = \begin{bmatrix} -3 & 2 & -2 & 0.5 \\ 2.5 & 5 & 15 & 5 \\ -5 & -1.5 & -13 & 0.5 \end{bmatrix}$$

(b) $A \in M_{4,6}(\mathbb{F}_5)$,

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

4. Produce a matrix $A \in M_{3,4}(\mathbb{F}_5)$ which has two vectors in a basis for null(A) and two vectors in a basis for col(A).

5. Find a basis for Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\} \subseteq \mathbb{R}^4$, if

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ -3 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 0 \\ -1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 9 \\ 10 \\ 1 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 3 \\ 2 \\ -4 \\ 9 \end{bmatrix}, \quad \mathbf{v}_5 = \begin{bmatrix} 12 \\ -16 \\ 5 \\ 1 \end{bmatrix}.$$

6. Find a basis for Span$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\} \subseteq \mathbb{F}_5^4$, if

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 3 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 2 \end{bmatrix}, \quad \mathbf{v}_5 = \begin{bmatrix} 1 \\ 4 \\ 3 \\ 3 \end{bmatrix}.$$

7. Let V be the vector space of all functions $\mathbb{R} \to \mathbb{R}$. Find a basis for the subspace H, if

$$H = \text{Span}\{\sin(t), \sin(2t), \sin(t)\cos(t)\}.$$

8. Find a matrix $A \in M_2(\mathbb{R})$ such that

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

Is A unique? Explain.

9. Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation which satisfies the following:

$$T\left(\begin{bmatrix}1\\1\\3\end{bmatrix}\right) = \begin{bmatrix}2\\-1\end{bmatrix},\quad T\left(\begin{bmatrix}-1\\0\\2\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix},\quad T\left(\begin{bmatrix}3\\-2\\0\end{bmatrix}\right) = \begin{bmatrix}3\\-2\end{bmatrix}.$$

Calculate

$$T\left(\begin{bmatrix}-1\\4\\6\end{bmatrix}\right).$$

10. Find a subset of the following set which is a basis for P_2:

$$\{t-1, t^2-2t, t^2-2, t^2+1\}.$$

Writing Exercises

11. Let $T : V \to W$ be a linear transformation between vector spaces, and let B be a basis for V.

(a) Produce an example to show that $T(B)$ does not need to be a basis of W.

(b) Suppose that T is injective. Must $T(B)$ be a basis for W? If so, prove it. If not, produce a counter-example.

12. Suppose that $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for a vector space V. Prove that

$$\{\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \dots, \mathbf{v}_{n-1} + \mathbf{v}_n, \mathbf{v}_n\}$$

is also a basis for V.

13. Prove or disprove: Every basis of P_2 must contain a polynomial of degree 2, a polynomial of degree 1, and a constant polynomial.

14. Write down a basis for $M_2(\mathbb{R})$. Prove that your set is a basis. (There is no need to prove that $M_2(\mathbb{R})$ is a vector space as this was covered in Example 2.3.10.)

5.3 Dimension

In this section we will define the *dimension* of a vector space, finally delivering on the promise made in the introduction to this chapter to describe an intrinsic quality of vector spaces that allows a comparison between spaces.

5.3.1 The Dimension of a Vector Space

We are on the threshold of the definition of dimension. We will first present a result that connects (for a finite-dimensional space) the number of vectors needed for a spanning set to the concept of linear independence. We will omit the proof.

Lemma 5.3.1 *Suppose that V is a vector space and that $V \subseteq \operatorname{Span}\{\mathbf{v}_1, \dots \mathbf{v}_m\}$. If $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ is a linearly independent subset of V, then $m \geq n$.*

We will now use this lemma to prove a result related to dimension.

Theorem 5.3.2 *Suppose that, for every $n \geq 1$, a vector space V contains a linearly independent subset of size n. Then V is infinite-dimensional.*

Proof. We will prove the contrapositive. If V is finite-dimensional, then there exists a set $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ such that $V = \text{Span}(V')$. By Lemma 5.3.1, for any $n > m$, V cannot contain a linearly independent set of n vectors. This completes the proof of the contrapositive. ■

This theorem gives us an introduction to our first infinite-dimensional example.

Example 5.3.3 Let P be the vector space of all polynomials with real coefficients. (We do not restrict the degree of polynomials in P.) For any n, P contains the linearly independent set

$$\{1, t, \ldots, t^n\}.$$

Therefore, by Theorem 5.3.2, P is infinite-dimensional. □

We now come to the bedrock result of this section, the result that makes the definition of *dimension* possible.

Theorem 5.3.4 *Suppose that $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ are both bases for a vector space V. Then $m = n$.*

Proof. Since $V \subseteq \text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ and $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a linearly independent set, Lemma 5.3.1 implies that $m \geq n$. However, since $V \subseteq \text{Span}\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a linearly independent set, Lemma 5.3.1 also implies that $n \geq m$. Therefore, $m = n$. ■

Even though a vector space may have a huge number of bases, all of those bases have the same size. This is a number intrinsic to the vector space, not to any specific basis of that vector space. This is what we mean by the *dimension* of a vector space.

Definition 5.3.5 Let V be a finite-dimensional vector space. If $V \neq \{\mathbf{0}\}$, then the **dimension** of V, written $\dim(V)$, is the size of any basis of V. If $\dim(V) = n$, we say that V is n**-dimensional**.

If $V = \{\mathbf{0}\}$, then we define the dimension of V to be 0. ◊

Two of the families of vector spaces we frequently discuss have easy-to-determine dimensions, as the next two examples illustrate.

Example 5.3.6 Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis for $\mathbb{F}^n$, then $\dim(\mathbb{F}^n) = n$. □

Example 5.3.7 Since $\{1, t, \ldots, t^n\}$ is a basis for P_n, then $\dim(P_n) = n + 1$. □

The proofs of the next two results are a consequence of Lemma 5.3.1 and will appear in the exercises.

Proposition 5.3.8 *The dimension of any vector space is less than or equal to the size of any spanning set.*

Proposition 5.3.9 *If a vector space V is finite-dimensional and $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ is a linearly independent set in V, then $\dim(V) \geq n$.*

We will now begin to discuss dimension as a tool to compare vector spaces. Linear transformations are the main way we relate vector spaces to each other, so these next results will rely on that machinery.

Theorem 5.3.10 *Suppose that $V' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V. Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be a subset of a vector space W. Then there is a unique linear transformation $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i, $1 \leq i \leq n$.*

Proof. Given $\mathbf{v} \in V$, there exists a unique linear combination

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i$$

by Theorem 5.2.11. We define the function T by

$$T(\mathbf{v}) = \sum_{i=1}^{n} c_i \mathbf{w}_i.$$

In words, we send a vector $\mathbf{v}$ to the linear combination of the $\mathbf{w}_i$ vectors using the same weights as those needed to form $\mathbf{v}$ from the basis V'. This gives $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i, so we only need to show that T is a linear transformation. Suppose that $\mathbf{u}, \mathbf{v} \in V$ with

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i \quad \text{and} \quad \mathbf{u} = \sum_{i=1}^{n} d_i \mathbf{v}_i.$$

Then we have

$$\begin{aligned} T(\mathbf{u} + \mathbf{v}) &= T\left(\sum_{i=1}^{n} (c_i + d_i)\mathbf{v}_i\right) \\ &= \sum_{i=1}^{n} (c_i + d_i)\mathbf{w}_i \\ &= \sum_{i=1}^{n} c_i \mathbf{w}_i + \sum_{i=1}^{n} d_i \mathbf{w}_i \\ &= T(\mathbf{u}) + T(\mathbf{v}). \end{aligned}$$

Now we let $\mathbf{v} \in V$ and $c \in \mathbb{F}$. Then, if

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i,$$

we have

$$\begin{aligned} T(c\mathbf{v}) &= T\left(\sum_{i=1}^{n} (cc_i)\mathbf{v}_i\right) \\ &= \sum_{i=1}^{n} (cc_i)\mathbf{w}_i \\ &= c\sum_{i=1}^{n} c_i \mathbf{w}_i \\ &= cT(\mathbf{v}). \end{aligned}$$

We will complete the proof by justifying the claim that T is unique. Suppose that $T' \in L(V, W)$ with $T'(\mathbf{v}_i) = \mathbf{w}_i$ for each i. Then, if $\mathbf{v} \in V$ with

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i,$$

we have

$$T'(\mathbf{v}) = T'\left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i T'(\mathbf{v}_i) = \sum_{i=1}^{n} c_i \mathbf{w}_i.$$

This shows that $T'(\mathbf{v}) = T(\mathbf{v})$ for every $\mathbf{v} \in V$, so $T' = T$ and T is unique. ■

The notion of *alikeness* that we use in linear algebra is when two vector spaces are isomorphic. The reader may wish to consult Definition 3.1.12 for a refresher.

Theorem 5.3.11 *Let $T \in L(V, W)$ and let $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V. Then T is an isomorphism if and only if $T(B) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is a basis for W.*

Proof. We first suppose that T is an isomorphism. We want to show that $T(B)$ is a basis for W, so we begin with linear independence. Suppose that $c_1, \ldots, c_n \in \mathbb{F}$ such that

$$\mathbf{0} = \sum_{i=1}^{n} c_i T(\mathbf{v}_i).$$

Then we have

$$\mathbf{0} = \sum_{i=1}^{n} T(c_i \mathbf{v}_i) = T\left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right).$$

Since T is injective, by Theorem 3.4.7 we must have

$$\mathbf{0} = \sum_{i=1}^{n} c_i \mathbf{v}_i.$$

But since B is a linearly independent set, we have $c_i = 0$ for all i. This proves that $T(B)$ is linearly independent.
We now prove that $T(B)$ spans W. Let $\mathbf{w} \in W$. Since T is surjective, there exists $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Since B is a basis for V, we have

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i.$$

Then

$$\mathbf{w} = T(\mathbf{v}) = T\left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i T(\mathbf{v}_i).$$

This proves that $W = \operatorname{Span}(T(B))$, so $T(B)$ is a basis for W.
We now need to prove the other implication, and we assume that $T(B)$ is a basis for W. We need to show that T is an isomorphism. To show that T is injective, suppose that $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{0}$. We have

$$\mathbf{v} = \sum_{i=1}^{n} c_i \mathbf{v}_i,$$

so

$$\mathbf{0} = T(\mathbf{v}) = T\left(\sum_{i=1}^{n} c_i \mathbf{v}_i\right) = \sum_{i=1}^{n} c_i T(\mathbf{v}_i).$$

But since $T(B)$ is a linearly independent set by assumption, this implies that $c_i = 0$ for all i. This means that $\mathbf{v} = \mathbf{0}$, so T is injective.
To prove that T is surjective, we assume that $\mathbf{w} \in W$. Since $T(B)$ spans W, we have

$$\mathbf{w} = \sum_{i=1}^{n} d_i T(\mathbf{v}_i)$$

for some $d_i \in \mathbb{F}$. We claim that if

$$\mathbf{v} = \sum_{i=1}^{n} d_i \mathbf{v}_i,$$

then $T(\mathbf{v}) = \mathbf{w}$. Here is the justification:

$$T(\mathbf{v}) = T\left(\sum_{i=1}^{n} d_i \mathbf{v}_i\right) = \sum_{i=1}^{n} d_i T(\mathbf{v}_i) = \mathbf{w}.$$

This proves that T is surjective and is thus an isomorphism. ■

When we view dimension as an intrinsic quality of a vector space that allows comparison between spaces, we find something surprising about vector spaces with the same dimension. They are essentially the same!

Theorem 5.3.12 *Let V and W be finite-dimensional vector spaces over $\mathbb{F}$. Then $\dim(V) = \dim(W)$ if and only if V and W are isomorphic.*

Proof. Suppose that $\dim(V) = \dim(W) = n$. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V and let $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ be a basis for w. By Theorem 5.3.10, we can find $T \in L(V, W)$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i, $1 \leq i \leq n$. Then by Theorem 5.3.11, T is an isomorphism.
To prove the claim in the other direction, suppose that $T \in L(V, W)$ is an isomorphism. If $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V, then $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ is a basis for W by Theorem 5.3.11. Thus $\dim(V) = \dim(W)$. ■

Here is an immediate consequence of this result.

Corollary 5.3.13 *Every finite-dimensional vector space over $\mathbb{F}$ of dimension n is isomorphic to $\mathbb{F}^n$.*

Example 5.3.14 Since P_2 is a three-dimensional vector space over $\mathbb{R}$, $\mathbb{R}^3$ and P_2 are isomorphic. □

5.3.2 Dimension and Subspaces

If we know the dimension of a vector space, then we sometimes have a quicker path to finding a basis for that space. This next result says that if we have a spanning set of the same size as a basis, then it must be a basis.

Theorem 5.3.15 *Suppose that V is a vector space with $\dim(V) = n > 0$. If $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is such that $V = \operatorname{Span}(B)$, then B is a basis for V.*

Proof. By The Spanning Set Theorem (Theorem 5.2.12), we know that a subset B' of B will be a basis for V. But since $\dim(V) = n$, the size of B' must be n. Therefore, $B' = B$ and B is a basis for V. ■

What is true in Theorem 5.3.15 for a spanning set is also true for a linearly independent set. To prove that, however, we first need the analog to The Spanning Set Theorem for linearly independent sets.

Theorem 5.3.16 *Suppose that V is a finite-dimensional vector space and that V' is a linearly independent set of vectors in V. Then there is a basis of V which contains V'.*

Proof. Let $V' = \{\mathbf{v}_1, \ldots \mathbf{v}_n\}$ be a linearly independent set of vectors in V. If $V = \operatorname{Span}(V')$, then V' is a basis and we are done. If $V \neq \operatorname{Span}(V')$, then there exists some vector $\mathbf{v}_{n+1} \in V - \operatorname{Span}(V')$. By the Linear Dependence Lemma (Theorem 5.1.19), the set $V_1 = V' \cup \{\mathbf{v}_{n+1}\}$ is linearly independent.
We can repeat this process. If $V = \operatorname{Span}(V_1)$, we are done; otherwise, we create $V_2 = V_1 \cup \{\mathbf{v}_{n+2}\}$ in the same fashion that we created V_1. We can continue doing this, adding one vector at a time to this set and maintaining linear independence. Eventually we must reach the point where $V = \operatorname{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}, \ldots, \mathbf{v}_{n+k}\}$, since otherwise Lemma 5.3.1 would imply that V is infinite-dimensional. ■

We now have the machinery necessary to state the following theorem. The proof will appear in the exercises.

Theorem 5.3.17 *Suppose that V is a vector space with $\dim(V) = n$ and that B is a linearly independent subset of V of size n. Then B is a basis for V.*

The final result of this section collects some facts about dimension and subspaces which we will use in some of the sections that follow.

Theorem 5.3.18 *Suppose that V is a finite-dimensional vector space and that U is a subspace of V. Then the following hold.*

1. *The subspace U is finite-dimensional.*
2. *We have $\dim(U) \leq \dim(V)$.*
3. *We have $\dim(U) = \dim(V)$ if and only if $U = V$.*

Proof. We will prove these facts in order. If the subspace U is $\{\mathbf{0}\}$, then we have nothing to prove. If not, then there is some non-zero vector $\mathbf{u}_1 \in U$. If $U = \mathrm{Span}\{\mathbf{u}_1\}$, we are done; if not, then there exists $\mathbf{u}_2 \in U - \mathrm{Span}\{\mathbf{u}_1\}$. By Theorem 5.1.19, the set $\{\mathbf{u}_1, \mathbf{u}_2\}$ is linearly independent. We can continue to repeat this process. At each stage we have a linearly independent set $U_k = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$, and this cannot continue indefinitely since U is a subspace of V, which is finite-dimensional. Thus this process must eventually stop when $U = \mathrm{Span}(U_j)$ for some j, and that proves that U is finite dimensional.
The space U is finite dimensional, so it has a basis B. This is a linearly independent set of vectors in V, so Theorem 5.3.16 says that B can be extended to a basis B' of V. This means that B' will have at least as many vectors in it as B, so $\dim(V) \geq \dim(U)$.
If $U = V$ it is obvious that $\dim(U) = \dim(V)$, so we only need to prove the claim in the other direction. We will prove the contrapositive, so we assume $U \neq V$. Let $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be a basis for U. Since $U \neq V$, there exists a vector $\mathbf{v}_1 \in V - \mathrm{Span}(B)$. By Theorem 5.1.19 the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_n, \mathbf{v}_1\}$ is linearly independent in V, implying that $\dim(V) \geq n+1$. Therefore $\dim(U) \neq \dim(V)$. ∎

Example 5.3.19 We can apply this latest result to the vector space $\mathbb{R}^3$. The familiar subspaces of $\mathbb{R}^3$ are *all* of the subspaces of $\mathbb{R}^3$.

1. The only subspace of dimension 0 in $\mathbb{R}^3$ is the zero subspace $\{\mathbf{0}\}$.
2. One-dimensional subspaces of $\mathbb{R}^3$ are lines through the origin. These can all be written as the span of a single (non-zero) vector.
3. Two-dimensional subspaces of $\mathbb{R}^3$ are planes through the origin. These are all spanned by sets of two linearly independent vectors.
4. The only three-dimensional subspace of $\mathbb{R}^3$ is $\mathbb{R}^3$ itself.

□

5.3.3 Reading Questions

1. Consider the following vectors in $\mathbb{R}^2$:

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

By inspection, why is the set $\{\mathbf{v}_1, \mathbf{v}_2\}$ a basis for $\mathbb{R}^2$? Explain your answer.

2. Let $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$, and $\mathbf{v}_4$ be vectors in $\mathbb{R}^3$.

(a) The set $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is not a basis for $\mathbb{R}^3$, and there's a very short argument why. What is that argument?

(b) Must there be a subset of S which is a basis of $\mathbb{R}^3$? Why or why not?

5.3.4 Exercises

1. Find the dimension of the subspace of $\mathbb{R}^3$ consisting of all vectors whose first and third coordinates are equal.

2. For each of the following sets of vectors in the given vector space, find the dimension of the subspace spanned by that set of vectors.

(a) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{R}^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} -8 \\ 6 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 3 \\ -3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 44 \\ -6 \\ -9 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} -31 \\ 3 \\ 7 \end{bmatrix}$$

(b) $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ in $\mathbb{F}_5^3$ if

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_4 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

3. For each of the following matrices A, determine the dimensions of $\operatorname{null}(A)$ and $\operatorname{col}(A)$.

(a) $A \in M_{4,5}(\mathbb{R})$,

$$A = \begin{bmatrix} 7 & 5 & 73 & 0 & 11 \\ 2 & -3 & 12 & -7 & -26 \\ -3 & 6 & -15 & 3 & 39 \\ 1 & 6 & 21 & -3 & 25 \end{bmatrix}$$

(b) $A \in M_{2,4}(\mathbb{F}_5)$,

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

4. Determine whether the following statements are true or false. Justify your answer either way.

(a) If a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ spans a finite-dimensional space V, and if V' is a set of more than m vectors in V, then V' is linearly dependent.

(b) The vector space $\mathbb{R}^2$ is a subspace of $\mathbb{R}^3$.

(c) A vector space is infinite-dimensional if it is spanned by an infinite set.

5. Determine whether the following statements are true or false. Justify your answer either way.

(a) If $\dim(V) = m$, then there exists a spanning set of $m + 1$ vectors in V.

(b) If every set of m vectors in V fails to span V, then $\dim(V) > m$.

(c) If $m \geq 2$ and $\dim(V) = m$, then every set of $m-1$ non-zero vectors in V is linearly independent.

6. The first four Hermite polynomials are 1, $2t$, $-2+4t^2$, and $-12t+8t^3$. Show that the set of these polynomials is a basis for P_3.

7. The first four Laguerre polynomials are 1, $1-t$, $2-4t+t^2$, and $6-18t+9t^2-t^3$. Show that the set of these polynomials is a basis for P_3.

Writing Exercises

8. Let A be a matrix.

(a) Prove that $\dim(\mathrm{null}(A))$ is the number of non-pivot columns in A.

(b) Prove that $\dim(\mathrm{col}(A))$ is the number of pivot columns of A.

9. Let V be the set of all functions $\mathbb{R} \to \mathbb{R}$. Prove that V is infinite-dimensional.

10. Suppose that $T : V \to W$ is a linear transformation between vector spaces and that V is finite-dimensional. Prove that $\dim(\mathrm{range}(T)) \leq \dim(V)$.

11. Prove that $\mathbb{C}^2$ is two-dimensional as a vector space over $\mathbb{C}$ but four-dimensional as a vector space over $\mathbb{R}$.

12. Prove Proposition 5.3.8.

13. Prove Proposition 5.3.9.

14. Prove Theorem 5.3.17.

5.4 Rank and Nullity

In this section we will connect dimension with the subspaces associated with linear transformations (see Section 3.4).

5.4.1 Defining Rank and Nullity

We begin by defining the dimension of the range of a linear transformation.

Definition 5.4.1 Let T be a linear transformation. Then the **rank** of T, denoted $\mathrm{rank}(T)$, is the dimension of the range of T:

$$\mathrm{rank}(T) = \dim(\mathrm{range}(T)).$$

The **rank** of a matrix A is the dimension of the column space of A:

$$\mathrm{rank}(A) = \dim(\mathrm{col}(A)).$$

$\Diamond$

It may seem strange to define the same word in two ways. However, since the range of T is exactly the column space of A when T is multiplication by A, these two definitions coincide.

When A is an $m \times n$ matrix over $\mathbb{F}$, its rows are vectors in $\mathbb{F}^n$ and its columns are vectors in $\mathbb{F}^m$. This is why the column space of A is a subspace of $\mathbb{F}^m$. We can also examine the analogous space for the rows.

Definition 5.4.2 The set of all linear combinations of the rows of a matrix A is called the **row space** of A. We denote this by $\mathrm{row}(A)$. $\Diamond$

Note 5.4.3 Since the rows of A are the columns of A^T, it is immediate that $\text{row}(A) = \text{col}(A^T)$.

With the definition of the row space it is natural to wonder how the sizes of the row and column spaces compare to each other. The following results will help us settle this matter.

Proposition 5.4.4 *If A and B are row equivalent matrices, then we have* $\text{row}(A) = \text{row}(B)$. *Further, if B is in REF, then the non-zero rows of B form a basis for* $\text{row}(B)$ *(and* $\text{row}(A)$*).*

Proof. We will first show that $\text{row}(A) = \text{row}(B)$ as sets. If A is reduced to B, then the rows of B are linear combinations of the rows of A. (The elementary row operations produce linear combinations of the original rows.) Therefore, any linear combination of the rows of B can be written as a linear combination of the rows of A. This proves that $\text{row}(B) \subseteq \text{row}(A)$. Since all row operations are reversible, we can use row operations to produce A from B, and this same argument shows that $\text{row}(A) \subseteq \text{row}(B)$. This proves that $\text{row}(A) = \text{row}(B)$.
If the matrix B is in REF, the nonzero rows are linearly independent because no nonzero row is a linear combination of the rows below it. Here we are applying the Linear Dependence Lemma (Theorem 5.1.19) to the nonzero rows from bottom to top. Since the rows of B span $\text{row}(B)$ by definition, the fact that they are linearly independent means that they form a basis for $\text{row}(B)$. ■

Theorem 5.4.5 *Let $A \in M_{m,n}(\mathbb{F})$. Then* $\text{rank}(A) = \dim(\text{row}(A))$.

Proof. If we put A into REF, then Proposition 5.4.4 tells us that the number of pivots is $\dim(\text{row}(A))$ since that is the number of vectors in a basis of $\text{row}(A)$. However, the number of pivots in a REF (or the RREF) of A is also $\text{rank}(A)$. (See Exercise 5.3.4.8.) This proves that $\text{rank}(A) = \dim(\text{row}(A))$. ■

Note 5.4.6 This theorem says that $\text{rank}(A) = \text{rank}(A^T)$. This theorem also answers the question about the relative sizes of $\text{col}(A)$ and $\text{row}(A)$—they are the same!

Example 5.4.7 Consider the following matrix $A \in M_{4,5}(\mathbb{F}_5)$:

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

We will find a basis for $\text{row}(A)$. Here is the RREF of A:

$$A \sim \begin{bmatrix} 1 & 0 & 0 & 4 & 4 \\ 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We have used the RREF of a matrix in the past to find bases for the null space and column space of a matrix. Now, we will use it to find a basis for the row space. Proposition 5.4.4 tells us that the nonzero rows of this RREF are the basis we seek, therefore a basis for $\text{row}(A)$ is $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\begin{aligned} \mathbf{v}_1 &= [1\ 0\ 0\ 4\ 4] \\ \mathbf{v}_2 &= [0\ 1\ 0\ 2\ 2] \\ \mathbf{v}_3 &= [0\ 0\ 1\ 1\ 4]. \end{aligned}$$

□

We have defined the dimension of the range of a linear transformation T, so we now turn to the kernel.

Definition 5.4.8 If T is a linear transformation, then the **nullity** of T is the dimension of the kernel of T. If A is a matrix, then the **nullity** of A is the dimension of the null space of A. ♢

Note 5.4.9 As we saw with rank, these two uses of "nullity" coincide for the situation when T is multiplication by A.

Some other texts use the notation $\mathrm{null}(A)$ to indicate nullity instead of null space, as we have. We will not introduce any additional notation for the nullity, but we will use $\dim(\ker(T))$ or $\dim(\mathrm{null}(A))$ as appropriate.

5.4.2 The Rank-Nullity Theorem

The following theorem brings together the rank and nullity of a matrix/linear transformation.

Theorem 5.4.10 The Rank-Nullity Theorem. *If $A \in M_{m,n}(\mathbb{F})$, then*

$$\mathrm{rank}(A) + \dim(\mathrm{null}(A)) = n.$$

If $T \in L(V, W)$, then

$$\mathrm{rank}(T) + \dim(\ker(T)) = \dim(V).$$

Proof. We will prove the result for matrices. The proof for linear transformations is a bit more technical. (The reader should note that the result for linear transformations implies the result for matrices!)
If $A \in M_{m,n}(\mathbb{F})$, let B be the RREF of A. Then $\mathrm{rank}(A)$ is the number of pivot columns in B. Further, $\dim(\mathrm{null}(A))$ is the number of non-pivot columns in B. (See Exercise 5.3.4.8.) Since each of the n columns of B must be either a pivot or a non-pivot column, and since A and B have the same number of columns, this proves the theorem. ■

Example 5.4.11 If A is a 5×6 matrix with a three-dimensional null space, this theorem tells us that the rank of A is $6 - 3 = 3$.

Let us consider an additional scenario: Could a 6×8 matrix A have a one-dimensional null space? If such a matrix existed, it would have a rank of $8 - 1 = 7$, according to Theorem 5.4.10. But the largest rank that a 6×8 matrix can have is 6, since there cannot be more pivots than there are rows. So the answer is no, a 6×8 matrix cannot have a one-dimensional null space. □

When the dimensions of the domain and codomain of a linear transformation are equal, some properties of such a transformation coincide.

Corollary 5.4.12 *If $T \in L(V, W)$ and $\dim(V) = \dim(W)$, then the following are equivalent.*

1. *The transformation T is injective.*
2. *The transformation T is surjective.*
3. *The transformation T is an isomorphism.*

Proof. By Theorem 3.4.7, T is injective if and only if $\ker(T) = \{\mathbf{0}\}$. In other words, T is injective if and only if $\dim(\ker(T)) = 0$. By Theorem 5.4.10, this happens if and only if $\mathrm{rank}(T) = \dim(V)$, and if $\dim(V) = \dim(W)$, $\mathrm{rank}(T) = \dim(W)$ if and only if T is injective. This proves that T is injective if and only if T is surjective. Since a bijective linear transformation is an isomorphism, our proof is complete. ■

To close out this section, we present a long theorem with many equivalent statements. We will omit a proof, because the equivalence of most of these statements has been already established at various places in this text. (In other books, this theorem forms the central focus of the text. It is certainly important, but we have chosen a different emphasis.)

Theorem 5.4.13 The Invertible Matrix Theorem. *Let $A \in M_n(\mathbb{F})$. Then the following statements are equivalent.*

1. *The matrix A is invertible.*
2. *The matrix A is row equivalent to I_n.*
3. *The matrix A has n pivots.*
4. *The equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
5. *The columns of A form a linearly independent set.*
6. *If $T_A : \mathbb{F}^n \to \mathbb{F}^n$ is multiplication by A, then T is injective.*
7. *The equation $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b} \in \mathbb{F}^n$.*
8. *The columns of A span $\mathbb{R}^n$.*
9. *The linear transformation T_A is surjective.*
10. *There is an $n \times n$ matrix B such that $AB = BA = I_n$.*
11. *The matrix A^T is invertible.*
12. *We have $\det(A) \neq 0$.*
13. *The columns of A form a basis for $\mathbb{F}^n$.*
14. *We have $\mathrm{col}(A) = \mathbb{F}^n$.*
15. *We have $\mathrm{rank}(A) = n$.*
16. *We have $\mathrm{null}(A) = \{\mathbf{0}\}$.*
17. *We have $\dim(\mathrm{null}(A)) = 0$.*

This theorem ties together threads from almost every section we've covered, which is quite an achievement! The reader should note that this result only applies to *square* matrices.

5.4.3 Reading Questions

1. Consider the following matrix A:

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

Find a basis for $\mathrm{row}(A)$. Explain your answer.

2. Suppose that $T : V \to W$ is a linear transformation and that $\dim(V) = 4$ and $\dim(W) = 5$. What are the possible values for $\dim(\ker(T))$? Explain.

5.4.4 Exercises

1. Find the rank and nullity of each of the following matrices.

(a) $A \in M_{3,4}(\mathbb{R})$,

$$A = \begin{bmatrix} 2 & -3 & 4 & 1 \\ -1 & 2 & -3 & 1 \\ 4 & -3 & 2 & 11 \end{bmatrix}$$

(b) $A \in M_{4,5}(\mathbb{F}_3)$,

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

2. Let $D : P_n \to P_{n-1}$ be the differentiation linear transformation. Calculate $\text{rank}(D)$ and $\dim(\ker(D))$.

3.

(a) If a 5×3 matrix A has rank 3, find $\dim(\text{null}(A))$, $\dim(\text{row}(A))$, and $\text{rank}(A^T)$.

(b) If the null space of a 7×6 matrix A is 5-dimensional, what is the dimension of the column space of A?

(c) If A is a 7×9 matrix, what is the smallest possible dimension of $\text{null}(A)$?

4. Suppose a nonhomogeneous linear system of nine equations and ten variables has a solution for all possible constants on the right side of the equations. Is it possible to find two nonzero solutions of the associated homogeneous system that are not multiples of each other? Explain.

5. Suppose $A \in M_{m,n}(\mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^m$. What has to be true about the two numbers $\text{rank}([A\,\mathbf{b}])$ and $\text{rank}(A)$ in order for the equation $A\mathbf{x} = \mathbf{b}$ to be consistent? Explain.

Writing Exercises

6. If A and B are matrices, prove that $\text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}$.

7. Suppose that $T \in L(V, W)$ and that V and W are both finite-dimensional.

(a) Prove that T is surjective if and only if $\text{rank}(T) = \dim(W)$.

(b) Prove that T is injective if and only if $\text{rank}(T) = \dim(V)$.

8. Suppose that $A\mathbf{x} = \mathbf{b}$ is a 6×6 linear system which is consistent but which does *not* have a unique solution. Prove that there must be a vector $\mathbf{c} \in \mathbb{F}^6$ such that the system $A\mathbf{x} = \mathbf{c}$ is inconsistent.

9. Prove that if $T \in L(V, W)$, then $\text{rank}(T) \leq \min\{\dim(V), \dim(W)\}$.

10. Prove that if $A \in M_{m,n}(\mathbb{F})$, then $\text{rank}(A) \leq \min\{m, n\}$.

11. Let $T \in L(V, W)$.

(a) Prove that if $\dim(V) > \dim(W)$, then T cannot be injective.

(b) Prove that if $\dim(W) > \dim(V)$, then T cannot be surjective.

5.5 Coordinates

We have recently shown (Corollary 5.3.13) that all n-dimensional vector spaces over $\mathbb{F}$ are isomorphic to $\mathbb{F}^n$. In this section we explore the vast implications of this isomorphism.

5.5.1 Coordinates of Vectors

If V is a finite-dimensional vector space over $\mathbb{F}$, then it has a basis $\mathcal{B}$. We have seen (Theorem 5.2.11) that each vector in V then has a *unique* representation as a linear combination of these basis vectors. In the definition that follows, we focus on the coefficients in these linear combinations.

Definition 5.5.1 The **coordinates** of a vector $\mathbf{v} \in V$ with respect to a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are the unique scalars $c_1, \ldots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

The **coordinate vector of v with respect to** $\mathcal{B}$ is the vector $[\mathbf{v}]_{\mathcal{B}} \in \mathbb{F}^n$,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

$\diamondsuit$

Note 5.5.2 When the basis $\mathcal{B}$ we are using is unambiguous, we may drop a bit of the cumbersome terminology contained in the phrase "coordinate vector of $\mathbf{v}$ with respect to $\mathcal{B}$" and simply refer to the "coordinate vector of $\mathbf{v}$."

This process of assigning to a vector $\mathbf{v} \in V$ a vector $[\mathbf{v}]_{\mathcal{B}} \in \mathbb{F}^n$ is sometimes called a **coordinate mapping**, and it defines a function $V \to \mathbb{F}^n$. This function is actually an isomorphism of vector spaces.

Theorem 5.5.3 *Let V be an n-dimensional vector space over $\mathbb{F}$, and let $\mathcal{B}$ be a basis for V. Consider the coordinate map $C_{\mathcal{B}} : V \to \mathbb{F}^n$ given by*

$$C_{\mathcal{B}}(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}.$$

Then $C_{\mathcal{B}}$ is an isomorphism.

Proof. The function $C_{\mathcal{B}}$ is a linear transformation. (We ask the reader to verify this in the exercises.) We note that $C_{\mathcal{B}}$ maps the basis vectors in $\mathcal{B}$ to the standard basis in $\mathbb{F}^n$. So, by Theorem 5.3.11, $C_{\mathcal{B}}$ is an isomorphism. ∎

The existence of coordinate vectors means that just about everything for finite-dimensional vector spaces can be accomplished with vectors and matrices over $\mathbb{F}$. We explore this in the following examples.

Example 5.5.4 Let $\mathcal{B} = \{1, t, t^2\}$ be the standard basis of the vector space P_2. If p_1 and p_2 are

$$p_1 = 2 - t + 4t^2 \quad \text{and} \quad p_2 = -3t^2 + 10,$$

then the coordinate vectors of p_1 and p_2 are

$$[p_1]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix} \quad \text{and} \quad [p_2]_{\mathcal{B}} = \begin{bmatrix} 10 \\ 0 \\ -3 \end{bmatrix}.$$

Note that the order of the coordinates really matters, so in this case the terms in p_2 had to be reordered (in increasing powers of t) before the coefficients were entered as the coordinate vector. □

Example 5.5.5 Within $\mathbb{F}_5^3$, consider $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}.$$

Since neither of these vectors is a scalar multiple of the other, $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a linearly independent set and therefore a basis for W. If we let $\mathbf{v}_3$ be

$$\mathbf{v}_3 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix},$$

we can verify that $\mathbf{v}_3 \in W$ by row-reducing the appropriate matrix:

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

Since there is no pivot in the final column, we see that $\mathbf{v}_3 \in W$. Further, we can write down the coordinate vector of $\mathbf{v}_3$ with respect to $\mathcal{B}$ by studying this row-reduced matrix. We see that

$$[\mathbf{v}_3]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

It may seem strange for a vector in the three-dimensional space $\mathbb{F}_5^3$ to have a coordinate vector with only two entries, but this is due to the fact that W is two-dimensional. (It has a basis of only two vectors!) The coordinate mapping in this case says that W is isomorphic to $\mathbb{F}_5^2$, and this is why the coordinate vector for any vector in W has only two entries. □

There are some consequences of Theorem 5.5.3 that we want to spell out explicitly because of their usefulness. The proof of the following proposition can be found as part of the proof of Theorem 5.3.11.

Proposition 5.5.6 *Let V be a vector space of dimension m over $\mathbb{F}$. Since the coordinate mapping $C_{\mathcal{B}} : V \to \mathbb{F}^m$ is an isomorphism, then the following statements are true.*

1. *A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in V is linearly independent if and only if the set of coordinate vectors $\{C_{\mathcal{B}}(\mathbf{v}_1), \ldots, C_{\mathcal{B}}(\mathbf{v}_n)\}$ is linearly independent in $\mathbb{F}^m$.*

2. *A set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ spans V if and only if the set of coordinate vectors $\{C_{\mathcal{B}}(\mathbf{v}_1), \ldots, C_{\mathcal{B}}(\mathbf{v}_n)\}$ spans $\mathbb{F}^m$.*

Hopefully the reader can now see exactly how helpful the coordinate mapping isomorphism is. The following example should help to connect the dots.

Example 5.5.7 Consider the set of vectors $Y = \{p_1, p_2, p_3\}$ in P_3, where

$$\begin{aligned} p_1 &= 1 - t - 3t^2 + 2t^3 \\ p_2 &= -5 + 4t + 2t^2 - t^3 \\ p_3 &= 1 + 3t + 4t^2 - 3t^3. \end{aligned}$$

With respect to the standard basis $\mathcal{B}$ of P_3, these are the coordinate vectors:

$$[p_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -1 \\ -3 \\ 2 \end{bmatrix}, \quad [p_2]_{\mathcal{B}} = \begin{bmatrix} -5 \\ 4 \\ 2 \\ -1 \end{bmatrix}, \quad [p_3]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 4 \\ -3 \end{bmatrix}.$$

By row-reducing the matrix which has these coordinate vectors as its columns, we can see that the set of coordinate vectors $\{[p_1]_{\mathcal{B}}, [p_2]_{\mathcal{B}}, [p_3]_{\mathcal{B}}\}$ is linearly independent in $\mathbb{R}^4$:

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

This shows that the set Y is linearly independent in P_3.

For dimension reasons, we already knew that the set Y cannot span P_3, however this row-reduced matrix confirms it. Since there is not a pivot in each row, the set of coordinate vectors does not span $\mathbb{R}^4$, and this means that Y does not span P_3. $\square$

5.5.2 Coordinates and Linear Transformations

Back in Section 3.2, we showed how every linear transformation $\mathbb{F}^n \to \mathbb{F}^m$ could be realized as multiplication by a matrix over $\mathbb{F}$. We now bring that understanding into contact with coordinate vectors. While not every linear transformation between vector spaces is multiplication by a matrix, every linear transformation between finite-dimensional vector spaces can be represented as multiplication by a matrix when considering the relevant coordinate vectors.

Definition 5.5.8 Let V and W be n- and m-dimensional vector spaces over $\mathbb{F}$, respectively, and let $T : V \to W$ be a linear transformation. Further, suppose that $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V and $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ is a basis for W. If, for each j, $1 \le j \le n$, we have $a_{1j}, \ldots, a_{mj}$ as the coordinates of $T(\mathbf{v}_j)$ with respect to $\mathcal{C}$, then the **matrix of T with respect to $\mathcal{B}$ and $\mathcal{C}$** is the matrix $A = [a_{ij}]$. (In other words, column j of this matrix is the coordinate vector $[T(\mathbf{v}_j)]_{\mathcal{C}}$.) We denote this matrix as $[T]_{\mathcal{B},\mathcal{C}}$.

When $V = W$ and $\mathcal{B} = \mathcal{C}$, then we use the notation $[T]_{\mathcal{B}}$ and refer to the **matrix of T with respect to $\mathcal{B}$**.

Finally, when the basis/bases we are using are unambiguous, we may refer to $[T]_{\mathcal{B}}$ or $[T]_{\mathcal{B},\mathcal{C}}$ as the **coordinate matrix** of T. $\Diamond$

The point of this rather long (and cumbersome!) definition is that we can represent a linear transformation T as multiplication by a matrix. That's what the following proposition shows.

Proposition 5.5.9 *Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces. Suppose that V and W have bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Let $A = [T]_{\mathcal{B},\mathcal{C}}$. Then, for any $\mathbf{v} \in V$,*

$$[T(\mathbf{v})]_{\mathcal{C}} = A[\mathbf{v}]_{\mathcal{B}}.$$

Proof. Let the bases $\mathcal{B}$ and $\mathcal{C}$ be $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ and $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$. For $\mathbf{v} \in V$, suppose that

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n,$$

or, in other words,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

We also assume that, for each j, $1 \le j \le n$, the coordinates of $T(\mathbf{v}_j)$ with respect to $\mathcal{C}$ are $a_{1j}, \ldots, a_{mj}$.
Then, using the linearity of T, we have

$$T(\mathbf{v}) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j) = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij} \mathbf{w}_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} c_j \right) \mathbf{w}_i.$$

This says that the ith coordinate of $T(\mathbf{v})$ with respect to $\mathcal{C}$ is $\sum a_{ij} c_j$, which is the same as the ith entry of $A[\mathbf{v}]_{\mathcal{B}}$. ■

Note 5.5.10 According to this proposition, here is the way to realize a linear transformation as a matrix. Form $[T]_{\mathcal{B},\mathcal{C}}$ by calculating the coordinate vector $[T(\mathbf{v}_j)]_{\mathcal{C}}$ for every vector $\mathbf{v}_j \in \mathcal{B}$. Then, to use this matrix to determine what happens to a vector $\mathbf{v} \in V$, find the coordinate vector $[\mathbf{v}]_{\mathcal{B}}$. After multiplying this vector by $[T]_{\mathcal{B},\mathcal{C}}$, the result will be the coordinate vector $[T(\mathbf{v})]_{\mathcal{C}}$. In order to recover the value of $T(\mathbf{v})$, use the basis vectors in $\mathcal{C}$ and this coordinate vector to find the correct linear combination.

Example 5.5.11 Let $D : P_3 \to P_2$ be the differentiation function. (We proved that a very similar function was a linear transformation in Example 3.1.3.) Let $\mathcal{B}$ be the standard basis for P_3, and let $\mathcal{C}$ be the standard basis for P_2. Here we calculate the coordinate vectors for the derivative of each of the polynomials in $\mathcal{B}$:

$$[D(1)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [D(t)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [D(t^2)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \quad [D(t^3)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}.$$

These coordinate vectors form the columns of the matrix $[D]_{\mathcal{B},\mathcal{C}}$.

We will now use this matrix to carry out the action of D. Let's take the derivative of $p = -2-4t-t^2-t^3$. Since the coordinate vector of p with respect to $\mathcal{B}$ is

$$[p]_{\mathcal{B}} = \begin{bmatrix} -2 \\ -4 \\ -1 \\ -1 \end{bmatrix},$$

we can multiply this vector by $[D]_{\mathcal{B},\mathcal{C}}$ to get $[D(p)]_{\mathcal{C}}$:

$$[D(p)]_{\mathcal{C}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ -4 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \\ -3 \end{bmatrix}.$$

This tells us that the coordinates for $D(p)$ with respect to $\mathcal{C}$ are -4, -2, and -3. In other words,

$$D(p) = -4(1) - 2(t) - 3(t^2),$$

and this matches what we know to be the derivative of p. □

Example 5.5.12 We consider a linear transformation $T : \mathbb{R}^3 \to P_2$ defined by

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = (a+2b) + (-3a+4b-c)t + (2a-4c)t^2.$$

We let $\mathcal{B}$ be the standard basis for $\mathbb{R}^3$ and $\mathcal{C}$ be the standard basis for P_2. We now write $[T(\mathbf{e}_i)]_{\mathcal{C}}$ for each $\mathbf{e}_i \in \mathcal{B}$:

$$[T(\mathbf{e}_1)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \quad [T(\mathbf{e}_2)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix}, \quad [T(\mathbf{e}_3)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \\ -4 \end{bmatrix}.$$

These coordinate vectors make up the columns of the matrix $[T]_{\mathcal{B},\mathcal{C}}$. If we wanted to calculate $T(\mathbf{v})$, where

$$\mathbf{v} = \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix},$$

we could do so using coordinate vectors and the matrix $[T]_{\mathcal{B},\mathcal{C}}$. Since the coordinate vector of $\mathbf{v}$ with respect to $\mathcal{B}$ is fairly obvious—it is $\mathbf{v}$ itself—we can proceed with this calculation:

$$[T(\mathbf{v})]_{\mathcal{C}} = [T]_{\mathcal{B},\mathcal{C}}[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 1 & 2 & 0 \\ -3 & 4 & -1 \\ 2 & 0 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} -5 \\ -19 \\ -14 \end{bmatrix}.$$

This tells us that $T(\mathbf{v}) = -5 - 19t - 14t^2$. □

We will end this section with two results related to coordinate matrices. This first result says that the composition of linear transformation really does match up with the multiplication of matrices.

Theorem 5.5.13 *Let U, V, and W be finite-dimensional vector spaces with bases $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$, respectively. Suppose that $T \in L(U,V)$ and that $S \in L(V,W)$. Then*

$$[ST]_{\mathcal{B},\mathcal{D}} = [S]_{\mathcal{C},\mathcal{D}}[T]_{\mathcal{B},\mathcal{C}}.$$

This final result states that the invertibility of a linear transformation and the invertibility of its coordinate matrix are tied together in the predictable way.

Corollary 5.5.14 *Let V and W be finite-dimensional vector spaces with bases $\mathcal{B}$ and $\mathcal{C}$, respectively. Suppose that $T \in L(V,W)$ and that $A = [T]_{\mathcal{B},\mathcal{C}}$. Then A is invertible if and only if T is invertible, and in that case, $[T^{-1}]_{\mathcal{C},\mathcal{B}} = A^{-1}$.*

5.5.3 Reading Questions

1. Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

(a) The set $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbb{R}^2$. Without doing any calculations, explain why this is so. (I'm not looking for the definition of a basis, I want an explanation as to why this set satisfies that definition.)

(b) Let $\mathbf{w} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$. What is the coordinate vector of $\mathbf{w}$ with respect to $\mathcal{B}$?

2. Let $T : P_2 \to P_2$ be the following function:

$$T(p) = p(-1) + p(0)t + p(1)t^2.$$

Let $\mathcal{B}$ be the standard basis for P_2.

(a) Find the coordinate matrix $[T]_{\mathcal{B}}$ for T.

(b) Use this coordinate matrix to calculate $T(q)$, if

$$q = -3 - 5t + 3t^2.$$

5.5.4 Exercises

1. For the given basis $\mathcal{B}$ of $\mathbb{R}^2$ and the given coordinate vector $[\mathbf{v}]_{\mathcal{B}}$, find $\mathbf{v}$.

(a) $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$, $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix} \right\}$

(b) $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$, $\mathcal{B} = \left\{ \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \end{bmatrix} \right\}$

2. For the basis $\mathcal{B} = \{p_1, p_2, p_3\}$ of P_2 and the coordinate vector $[p]_{\mathcal{B}}$, find p if

$$p_1 = 2 - 4t^2, \quad p_2 = -1 - t, \quad p_3 = 3t + 2t^2$$

and

$$[p]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix}.$$

3. Find the coordinate vectors $[\mathbf{v}]_{\mathcal{B}}$ for each of the following vectors $\mathbf{v}$ with respect to the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of $\mathbb{R}^3$, if

$$\mathbf{v}_1 = \begin{bmatrix} -2 \\ 5 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 7 \\ 4 \\ -4 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2 \\ -7 \\ 2 \end{bmatrix}.$$

(a) $\mathbf{v} = \begin{bmatrix} 16 \\ 5 \\ -16 \end{bmatrix}$

(b) $\mathbf{v} = \begin{bmatrix} 6 \\ -23 \\ 11 \end{bmatrix}$

4. Find the coordinate vectors $[p]_{\mathcal{B}}$ for each of the following polynomials p with respect to the basis $\mathcal{B} = \{p_1, p_2, p_3\}$ of P_2, if

$$p_1 = 8 + 4t - 4t^2, \quad p_2 = 5 + 8t + 3t^2, \quad p_3 = -6 - 2t - 5t^2.$$

(a) $p = -2 + 2t - 23t^2$

(b) $p = 23 + 28t + 5t^2$

5. Use coordinate vectors to test the linear independence of the following sets of polynomials in P_3.

(a) $\{p_1, p_2, p_3\}$ if

$$\begin{aligned} p_1 &= -6 + 7t + 6t^2 + 3t^3 \\ p_2 &= 2t - 4t^2 + 7t^3 \\ p_3 &= 2 + 6t - t^2 - 5t^3 \end{aligned}$$

(b) $\{p_1, p_2, p_3\}$ if

$$\begin{aligned} p_1 &= 6 + 7t - t^2 - 2t^3 \\ p_2 &= -5 - 7t - 6t^2 + 8t^3 \\ p_3 &= 7 + 7t - 8t^2 + 4t^3 \end{aligned}$$

6. Use coordinate vectors to test whether the following sets of vectors span P_2.

(a) $\{p_1, p_2, p_3, p_4\}$ if

$$\begin{aligned} p_1 &= -4 + t + t^2 \\ p_2 &= 3 + 5t + t^2 \\ p_3 &= -2 - 4t + 2t^2 \\ p_4 &= 2 - 4t - t^2 \end{aligned}$$

(b) $\{p_1, p_2, p_3, p_4\}$ if

$$\begin{aligned} p_1 &= 4 + 6t + 5t^2 \\ p_2 &= -3t^2 \\ p_3 &= 4 + 6t - 4t^2 \\ p_4 &= 8 + 12t + t^2 \end{aligned}$$

7. Let $T : P_2 \to \mathbb{R}^2$ be the linear transformation

$$T(p) = \begin{bmatrix} p(0) + p(1) \\ p(1) - p(2) \end{bmatrix}.$$

Let $\mathcal{B}$ be the standard basis for P_2 and let $\mathcal{E}$ be the standard basis for $\mathbb{R}^2$.

(a) Find the coordinate matrix $[T]_{\mathcal{B},\mathcal{E}}$.

(b) Use this coordinate matrix to calculate $T(-10 + 3t^2)$.

8. Let $T : \mathbb{R}^3 \to P_2$ be the linear transformation

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = (2a - b) + (b - 3c)t + (a - b + c)t^2.$$

Let $\mathcal{E}$ be the standard basis for $\mathbb{R}^3$ and let $\mathcal{B}$ be the standard basis for P_2.

(a) Find the coordinate matrix $[T]_{\mathcal{E},\mathcal{B}}$.

(b) Use this coordinate matrix to calculate $T(\mathbf{v})$ for

$$\mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}.$$

9. Let $T : P_2 \to P_2$ be the linear transformation

$$T(p) = p' + p(1)t^2.$$

Let $\mathcal{B}$ be the standard basis for P_2.

(a) Choose a basis $\mathcal{C}$ for P_2 which is *not* the standard basis. Prove that your set of polynomials is a basis.

(b) Find the coordinate matrix $[T]_{\mathcal{C},\mathcal{B}}$.

(c) Use this coordinate matrix to calculate $T(2 + t - 4t^2)$.

10. Let $D : P_3 \to P_2$ be the derivative and let $T : P_2 \to P_3$ be the linear transformation which is multiplication by t. Let $\mathcal{B}$ be the standard basis for P_2 and let $\mathcal{C}$ be the standard basis for P_3.

(a) Find the coordinate matrix $[T]_{\mathcal{B},\mathcal{C}}$.

(b) Find the coordinate matrix $[DT]_{\mathcal{B}}$.

(c) Find the coordinate matrix $[TD]_{\mathcal{C}}$.

11. Consider the plane P in $\mathbb{R}^3$ defined by $x - 2y + 3z = 0$.

(a) Find a basis for P.

(b) Determine whether each of the following vectors is in P, and for each one that is, find its coordinate vector in terms of the basis you gave in part a.

i. $\mathbf{v}_1 = (1, -1, -1)$
ii. $\mathbf{v}_2 = (2, 3, 1)$
iii. $\mathbf{v}_3 = (5, -2, -3)$

Writing Exercises

12. Prove that the coordinate mapping in Theorem 5.5.3 is a linear transformation.

13. Without using Theorem 5.3.11, prove that the coordinate mapping in Theorem 5.5.3 is injective.

14. Without using Theorem 5.3.11, prove that the coordinate mapping in Theorem 5.5.3 is surjective.

15. Let $T : V \to W$ be a linear transformation between finite-dimensional vector spaces, and let $\mathcal{B}$ and $\mathcal{C}$ be bases for V and W, respectively. Prove that $\operatorname{rank}(T)$ (the rank of T as a linear transformation) is the same as $\operatorname{rank}([T]_{\mathcal{B},\mathcal{C}})$ (the rank of the coordinate matrix of T).

16. Prove Corollary 5.5.14.

5.6 Change of Basis

Every basis for a vector space gives a different angle on that space—we get a different coordinate system for each basis. Since any finite-dimensional vector space has many bases, in this section we explain how to move between bases.

5.6.1 The Change-of-Basis Matrix

We will first describe a situation in which this technique will be useful. Consider the following two bases for $\mathbb{R}^2$: $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \mathbf{w}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

We can verify that $\mathcal{B}$ and $\mathcal{C}$ are bases for $\mathbb{R}^2$ since they are both linearly independent sets of two vectors in a two-dimensional space.

If we have a vector $\mathbf{v} \in \mathbb{R}^2$, it is straightforward to calculate both $[\mathbf{v}]_{\mathcal{B}}$ and $[\mathbf{v}]_{\mathcal{C}}$. The question for us is: how do these two coordinate vectors relate to each other? Specifically, how might we calculate one coordinate vector from the other one?

It turns out that we already have the necessary machinery for this calculation. We summarize the process in the following proposition.

Proposition 5.6.1 *Let $\mathcal{B}$ and $\mathcal{C}$ be two bases for a finite-dimensional vector space V. Then, for any $\mathbf{v} \in V$, we have*

$$[\mathbf{v}]_{\mathcal{C}} = [I]_{\mathcal{B},\mathcal{C}}[\mathbf{v}]_{\mathcal{B}},$$

where $I : V \to V$ is the identity transformation.

Proof. This is a simple application of Proposition 5.5.9 to the identity transformation I:

$$[\mathbf{v}]_{\mathcal{C}} = [I(\mathbf{v})]_{\mathcal{C}} = [I]_{\mathcal{B},\mathcal{C}}[\mathbf{v}]_{\mathcal{B}}.$$

■

Definition 5.6.2 If $\mathcal{B}$ and $\mathcal{C}$ are two bases for a finite-dimensional vector space V, then the matrix $P_{\mathcal{B},\mathcal{C}} = [I]_{\mathcal{B},\mathcal{C}}$ is called the **change-of-basis matrix** from $\mathcal{B}$ to $\mathcal{C}$. ♢

Example 5.6.3 We will continue the example begun earlier in this section. If $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2\}$, then we can calculate $P_{\mathcal{B},\mathcal{C}}$ by determining the coordinate vectors $[\mathbf{v}_1]_{\mathcal{C}}$ and $[\mathbf{v}_2]_{\mathcal{C}}$. We need only to row-reduce two matrices:

$$\begin{bmatrix} -2 & -1 & 1 \\ 1 & 3 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2/5 \\ 0 & 1 & -1/5 \end{bmatrix}, \quad \begin{bmatrix} -2 & -1 & 0 \\ 1 & 3 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2/5 \\ 0 & 1 & 4/5 \end{bmatrix}.$$

From these calculations, we can see how to write the $\mathcal{B}$-basis vectors in terms of the vectors in $\mathcal{C}$, and these form the columns of our change-of-basis matrix:

$$P_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} -2/5 & -2/5 \\ -1/5 & 4/5 \end{bmatrix}.$$

We now consider a vector $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ in $\mathbb{R}^2$. We can calculate $[\mathbf{v}]_{\mathcal{B}}$ in this way:

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

Now that we have $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ 5/2 \end{bmatrix}$, we can use the change-of-basis matrix to find $[\mathbf{v}]_{\mathcal{C}}$:

$$[\mathbf{v}]_{\mathcal{C}} = P_{\mathcal{B},\mathcal{C}}[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} -2/5 & -2/5 \\ -1/5 & 4/5 \end{bmatrix} \begin{bmatrix} 2 \\ 5/2 \end{bmatrix} = \begin{bmatrix} -9/5 \\ 8/5 \end{bmatrix}.$$

We can verify that this is the correct coordinate vector for $\mathbf{v}$ by calculating it directly:

$$\begin{bmatrix} -2 & -1 & 2 \\ 1 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -9/5 \\ 0 & 1 & 8/5 \end{bmatrix}.$$

□

Example 5.6.4 We consider the vector space P_2. Let $\mathcal{C}$ be the standard basis for P_2 and let $\mathcal{B}$ be the set $\{p_1, p_2, p_3\}$, where

$$p_1 = 2 + t, \quad p_2 = -1 - t + 2t^2, \quad p_3 = -2t + 3t^2.$$

In order to find the change-of-basis matrix, we need to write the coordinate vectors of the basis vectors of $\mathcal{B}$ with respect to $\mathcal{C}$. But since $\mathcal{C}$ is the standard basis of P_2, this is an easy task to complete. Here is the change-of-basis matrix:

$$P_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 2 & -1 & 0 \\ 1 & -1 & -2 \\ 0 & 2 & 3 \end{bmatrix}.$$

□

What we saw in Example 5.6.4 is an indication that some change-of-basis matrices are easier to calculate than others. In particular, when the standard basis is the *target* (not the *source*) basis, the matrix is almost immediate.

Lemma 5.6.5 *Let $\mathcal{E}$ be the standard basis of $\mathbb{F}^n$, and let $\mathcal{B}$ be any other basis of $\mathbb{F}^n$. Then the columns of $P_{\mathcal{B},\mathcal{E}}$ are the vectors of $\mathcal{B}$, in order.*

Proof. If $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, then column j of $P_{\mathcal{B},\mathcal{E}}$ is $[\mathbf{v}_j]_{\mathcal{E}}$. But since $\mathcal{E}$ is the standard basis, then $[\mathbf{v}_j]_{\mathcal{E}} = \mathbf{v}_j$. ■

The next lemma also shows that the change-of-basis matrices from one basis to another and back again have the inverse relationship we might expect.

Lemma 5.6.6 *Let $\mathcal{B}$ and $\mathcal{C}$ be two bases for a finite-dimensional vector space V. Then the relationship between the two change-of-basis matrices is*

$$P_{\mathcal{C},\mathcal{B}} = \left(P_{\mathcal{B},\mathcal{C}}\right)^{-1}.$$

Proof. Since $P_{\mathcal{B},\mathcal{C}} = [I]_{\mathcal{B},\mathcal{C}}$, by Theorem 5.5.13 we have

$$P_{\mathcal{B},\mathcal{C}}[I]_{\mathcal{C},\mathcal{B}} = [I]_{\mathcal{B},\mathcal{C}}[I]_{\mathcal{C},\mathcal{B}} = [I]_{\mathcal{C},\mathcal{C}} = I_n.$$

Since $P_{\mathcal{B},\mathcal{C}}$ is square, this proves that $[I]_{\mathcal{C},\mathcal{B}} = (P_{\mathcal{B},\mathcal{C}})^{-1}$. ■

Example 5.6.7 We consider two bases for $\mathbb{F}_5^3$: the standard basis $\mathcal{E}$ and $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}.$$

(The reader should verify that $\mathcal{B}$ is a basis for $\mathbb{F}_5^3$.)

Lemma 5.6.5 tells us that the change-of-basis matrix $P_{\mathcal{B},\mathcal{E}}$ is easy to write down:

$$P_{\mathcal{B},\mathcal{E}} = \begin{bmatrix} 4 & 3 & 0 \\ 2 & 1 & 4 \\ 2 & 0 & 1 \end{bmatrix}.$$

Then Lemma 5.6.6 says that $P_{\mathcal{E},\mathcal{B}} = (P_{\mathcal{B},\mathcal{E}})^{-1}$, so we can find that matrix without too much difficulty as well:

$$P_{\mathcal{E},\mathcal{B}} = \begin{bmatrix} 3 & 1 & 1 \\ 3 & 2 & 2 \\ 4 & 3 & 4 \end{bmatrix}.$$

□

The final results of this section deal with linear transformations. This theorem relates the coordinate matrix for a linear transformation to the situation in which we want to change bases in the domain and codomain.

Theorem 5.6.8 *Let V and W be finite-dimensional vector spaces, and let $T \in L(V, W)$. Additionally, let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for V, and let $\mathcal{C}_1$ and $\mathcal{C}_2$ be bases for W. Then*

$$[T]_{\mathcal{B}_2,\mathcal{C}_2} = [I]_{\mathcal{C}_1,\mathcal{C}_2}[T]_{\mathcal{B}_1,\mathcal{C}_1}[I]_{\mathcal{B}_2,\mathcal{B}_1}.$$

Proof. We will use Theorem 5.5.13:

$$[T]_{\mathcal{B}_2,\mathcal{C}_2} = [ITI]_{\mathcal{B}_2,\mathcal{C}_2} = [I]_{\mathcal{C}_1,\mathcal{C}_2}[T]_{\mathcal{B}_1,\mathcal{C}_1}[I]_{\mathcal{B}_2,\mathcal{B}_1}.$$

■

The most important (and most common) use of this theorem happens when $V = W$, $\mathcal{B}_1 = \mathcal{C}_1$, and $\mathcal{B}_2 = \mathcal{C}_2$.

Corollary 5.6.9 *Let $\mathcal{B}$ and $\mathcal{C}$ be bases for a finite-dimensional vector space V, and let $T \in L(V)$. Then*

$$[T]_{\mathcal{C}} = (P_{\mathcal{C},\mathcal{B}})^{-1}[T]_{\mathcal{B}}P_{\mathcal{C},\mathcal{B}}.$$

Proof. This result is due to Theorem 5.6.8 and Lemma 5.6.6. ■

We will end this section with an example which takes advantage of Corollary 5.6.9.

Example 5.6.10 We consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which is reflection across the line $y = \frac{1}{2}x$. While the action of T is not impossible to write down in the usual coordinate system, it is even easier using the alternate basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

To see why this linear transformation is easier to describe in the $\mathcal{B}$-coordinates, we recall how easy reflection across the y-axis is to describe relative to the standard basis—simply negate the first coordinate! The $\mathcal{B}$-basis vectors in this case lie on the axis of reflection and along the line perpendicular to that axis.

We note that $T(\mathbf{v}_1) = \mathbf{v}_1$ and that $T(\mathbf{v}_2) = -\mathbf{v}_2$. This shows that the coordinate matrix of T with respect to $\mathcal{B}$ is

$$[T]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

(Writing the action of T this way makes it especially easy to see that performing this transformation twice puts us back where we started.) We will use

Corollary 5.6.9 to calculate the matrix for T relative to the standard basis. That is, we wish to calculate $[T]_{\mathcal{E}}$.

We first note that the matrix $P_{\mathcal{B},\mathcal{E}}$ is, according to Lemma 5.6.5,

$$P_{\mathcal{B},\mathcal{E}} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}.$$

Then, with the help of Lemma 5.6.6, we have

$$P_{\mathcal{E},\mathcal{B}} = (P_{\mathcal{B},\mathcal{E}})^{-1} = \begin{bmatrix} 2/5 & 1/5 \\ -1/5 & 2/5 \end{bmatrix}.$$

We can put these together to find $[T]_{\mathcal{E}}$:

$$[T]_{\mathcal{E}} = P_{\mathcal{B},\mathcal{E}}[T]_{\mathcal{B}}P_{\mathcal{E},\mathcal{B}} = \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 2/5 & 1/5 \\ -1/5 & 2/5 \end{bmatrix} = \begin{bmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{bmatrix}.$$

The action of the transformation, as written in the final line here, is perhaps better understood in words rather than symbols. To reflect across the line $y = \frac{1}{2}x$, first shift from the standard coordinates to the alternate $\mathcal{B}$-coordinates. (This is accomplished by $P_{\mathcal{E},\mathcal{B}}$.) In this new coordinate system, the action of T is easily described. (Thus, $[T]_{\mathcal{B}}$.) After that action is carried out, then we shift back to the standard coordinate system. (That is the work of $P_{\mathcal{B},\mathcal{E}}$.) From start to finish, this gives us a matrix which carries out the action of T relative to $\mathcal{E}$. □

5.6.2 Reading Questions

1. Let $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\mathbf{w}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$, and $\mathbf{w}_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$. Consider the bases $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2\}$ of $\mathbb{R}^2$.

Find the change-of-basis matrix $P_{\mathcal{B},\mathcal{C}}$.

2. Using the definitions of the vectors and bases from the previous reading question, find $P_{\mathcal{C},\mathcal{B}}$.

5.6.3 Exercises

1. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2\}$ be bases of $\mathbb{R}^2$, where

$$\mathbf{v}_1 = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \quad \mathbf{w}_1 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -5 \\ -3 \end{bmatrix}.$$

Find the change-of-basis matrices $P_{\mathcal{B},\mathcal{C}}$ and $P_{\mathcal{C},\mathcal{B}}$.

2. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ and $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ be bases of $\mathbb{F}_5^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix}$$
$$\mathbf{w}_1 = \begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}.$$

Find the change-of-basis matrices $P_{\mathcal{B},\mathcal{C}}$ and $P_{\mathcal{C},\mathcal{B}}$.

3. Consider the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbb{F}_3^2$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

(a) If $\mathcal{E}$ is the standard basis for $\mathbb{F}_3^2$, find $P_{\mathcal{B},\mathcal{E}}$ and $P_{\mathcal{E},\mathcal{B}}$.

(b) Use your work in part (a) to find $[\mathbf{v}]_{\mathcal{B}}$ if $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

4. Let $\mathcal{B} = \{p_1, p_2, p_3\}$ be a basis for P_2, where

$$p_1 = 3t - 5t^2, \quad p_2 = 3 - 3t + 5t^2, \quad p_3 = -1 - t - t^2.$$

(a) If $\mathcal{E}$ is the standard basis for P_2, find $P_{\mathcal{B},\mathcal{E}}$ and $P_{\mathcal{E},\mathcal{B}}$.

(b) Use your work in part (a) to find $[p]_{\mathcal{B}}$ if $p = -3 + \frac{1}{2}t + \frac{3}{2}t^2$.

5. Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which is projection onto the line $y = -x$.

(a) Propose a basis $\mathcal{B}$ for $\mathbb{R}^2$ where $[T]_{\mathcal{B}}$ will be easy to determine.

(b) Find $[T]_{\mathcal{B}}$.

(c) If $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$, find $P_{\mathcal{B},\mathcal{E}}$ and $P_{\mathcal{E},\mathcal{B}}$.

(d) Using your work in previous parts of this problem, find $[T]_{\mathcal{E}}$.

6. Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ which is reflection across the line $y = -\frac{1}{4}x$.

(a) Propose a basis $\mathcal{B}$ for $\mathbb{R}^2$ where $[T]_{\mathcal{B}}$ will be easy to determine.

(b) Find $[T]_{\mathcal{B}}$.

(c) If $\mathcal{E}$ is the standard basis for $\mathbb{R}^2$, find $P_{\mathcal{B},\mathcal{E}}$ and $P_{\mathcal{E},\mathcal{B}}$.

(d) Using your work in previous parts of this problem, find $[T]_{\mathcal{E}}$.

Writing Exercises

7. Let V be an n-dimensional vector space over $\mathbb{F}$, and let $\mathcal{B}$ and $\mathcal{C}$ be two bases for V. Prove that the columns of the matrix $P_{\mathcal{B},\mathcal{C}}$ are linearly independent.

8. Let V be an n-dimensional vector space over $\mathbb{F}$, and let $\mathcal{B}$ and $\mathcal{C}$ be two bases for V. Prove that the columns of the matrix $P_{\mathcal{B},\mathcal{C}}$ span $\mathbb{F}^n$.

9. Let $A \in M_n(\mathbb{F})$. Prove that there exist bases $\mathcal{B}$ and $\mathcal{C}$ for $\mathbb{F}^n$ such that $A = P_{\mathcal{B},\mathcal{C}}$.

Chapter 6

Eigenvalues and Eigenvectors

6.1 Eigenvalues and Eigenvectors

For linear transformations $T : V \to W$, there isn't often a connection between $\mathbf{v}$ and $T(\mathbf{v})$ that is easy to describe. These vectors, after all, live in different vector spaces, so they need not have any obvious relationship to each other. When $W = V$, we sometimes have a different story to tell (for some vectors).

For a transformation $T \in L(V)$, $\mathbf{v}$ and $T(\mathbf{v})$ live in the same vector space, so there is occasionally an easy-to-define relationship between these two vectors. Sometimes, the action of T on a vector turns out to be rather simple.

6.1.1 Defining Eigenvalues and Eigenvectors

We first define the sorts of vectors we alluded to in the previous paragraphs.

Definition 6.1.1 Let V be a vector space over $\mathbb{F}$, and let $T \in L(V)$. A nonzero vector $\mathbf{v} \in V$ is an **eigenvector** for T if $T(\mathbf{v}) = \lambda\mathbf{v}$ for some $\lambda \in \mathbb{F}$. A scalar λ is called an **eigenvalue** of T if there is a nontrivial solution to the equation $T(\mathbf{x}) = \lambda\mathbf{x}$. Such a solution is called an **eigenvector for T corresponding to λ**.

If $A \in M_n(\mathbb{F})$, the eigenvectors and eigenvalues of A are the eigenvectors and eigenvalues of the transformation $T \in L(\mathbb{F}^n)$ defined by $T(\mathbf{x}) = A\mathbf{x}$. $\Diamond$

Informally, eigenvectors for T are nonzero vectors on which T acts by scalar multiplication. The next example shows that for a T that has eigenvectors, it is not (always) *every* vector in V that has this special property.

Example 6.1.2 When we are given a matrix A and a vector $\mathbf{v}$, it is easy to determine whether or not $\mathbf{v}$ is an eigenvector for A. Consider the following:

$$A = \begin{bmatrix} 3 & 0 \\ 7 & -1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

We take the product $A\mathbf{u}$,

$$\begin{bmatrix} 3 & 0 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 12 \\ 21 \end{bmatrix}.$$

Since $A\mathbf{u} = 3\mathbf{u}$, $\mathbf{u}$ is an eigenvector for A with eigenvalue 3. Also, since

$$A\mathbf{v} = \begin{bmatrix} 3 & 0 \\ 7 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -6 \\ -15 \end{bmatrix},$$

we can see that $\mathbf{v}$ is not an eigenvector for A, because $A\mathbf{v}$ is not a scalar multiple of $\mathbf{v}$. $\square$

When $\mathbf{v}$ is an eigenvector of T, then applying T may change the *length* of $\mathbf{v}$ but it will not change the *direction* of $\mathbf{v}$. (To say this we must include "pointing in the exact opposite direction" as being in the same direction.) This is a simplification, because not every vector space has a neat, geometric interpretation.

Example 6.1.3 Let $T : P_2 \to P_2$ be the following linear transformation:

$$T(a + bt + ct^2) = (4a - b + 6c) + (2a + b + 6c)t + (2a - b + 8c)t^2.$$

If $p = 1 + t + t^2$, it is not difficult to check that

$$T(p) = 9 + 9t + 9t^2 = 9p.$$

Therefore, p is an eigenvector for T with eigenvalue 9. □

Example 6.1.4 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is counterclockwise rotation about the origin by an angle of θ. We can see that T will have an eigenvector if and only if θ is an integer multiple of π radians. If θ is an even integer multiple of π, then every vector in $\mathbb{R}^2$ is an eigenvector for T with eigenvalue 1, and if θ is an odd integer multiple of π, then every vector in $\mathbb{R}^2$ is an eigenvector for T with eigenvalue -1. □

We take a slightly different approach in our next example. Instead of verifying that a vector is an eigenvector, we provide the eigenvalue and then search for the eigenvector(s).

Example 6.1.5 We consider the matrix A from Example 6.1.2. Let's show that -1 is an eigenvalue of A and find the corresponding eigenvectors.

We know that -1 is an eigenvalue of A if the equation $A\mathbf{x} = -\mathbf{x}$ has a nontrivial solution for some $\mathbf{x} \in \mathbb{R}^2$. This is equivalent to saying that the equation $A\mathbf{x} + \mathbf{x} = \mathbf{0}$ has a nontrivial solution. We can also view $\mathbf{x}$ as $I\mathbf{x}$, so if -1 is an eigenvalue of A, there is a nonzero vector $\mathbf{x}$ which satisfies

$$(A + I)\mathbf{x} = \mathbf{0}. \tag{6.1}$$

Viewed from the correct angle, we have reduced this problem to finding the null space of a matrix.

We will calculate $A + I$:

$$A + I = \begin{bmatrix} 3 & 0 \\ 7 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 7 & 0 \end{bmatrix}.$$

We can see that the columns of $(A + I)$ are linearly dependent, so we know that (6.1) has nontrivial solutions. This proves that -1 is an eigenvalue of A.

In order to find the eigenvectors of A that correspond to $\lambda = -1$, we describe the null space of the appropriate matrix. We row-reduce $(A + I)$:

$$\begin{bmatrix} 4 & 0 \\ 7 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

This shows that every eigenvector of A corresponding to $\lambda = -1$ has the form $x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, as long as $x_2 \neq 0$. The interested/vigilant reader can check that, for example, $A \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$. □

The process we undertook in the previous example showed that there are almost always multiple eigenvectors for a linear transformation which corre-