CHAPTER 4

Vector Spaces

To criticize mathematics for its abstraction is to miss the point entirely. Abstraction is what makes mathematics work. — Ian Stewart

The main aim of this text is to study linear mathematics. In Chapter 2 we studied systems of linear equations, and the theory underlying the solution of a system of linear equations can be considered as a special case of a general mathematical framework for linear problems. To illustrate this framework, we discuss an example.

Consider the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, where

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ 3 & -3 & 6 \end{bmatrix}.$$

It is straightforward to show that this system has solution set

$$S = \{(r - 2s, r, s) : r, s \in \mathbb{R}\}.$$

Geometrically we can interpret each solution as defining the coordinates of a point in space or, equivalently, as the geometric vector with components

$$\mathbf{v} = (r - 2s, r, s).$$

Using the standard operations of vector addition and multiplication of a vector by a real number, it follows that $\mathbf{v}$ can be written in the form

$$\mathbf{v} = r(1, 1, 0) + s(-2, 0, 1).$$

We see that every solution to the given linear problem can be expressed as a linear combination of the two basic solutions (see Figure 4.0.1):

$$\mathbf{v}_1 = (1, 1, 0) \qquad \text{and} \qquad \mathbf{v}_2 = (-2, 0, 1).$$

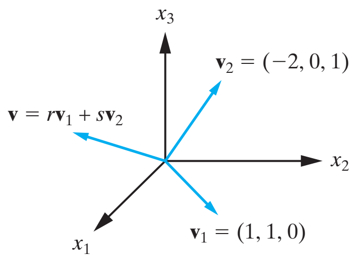

Figure 4.0.1: Two basic solutions to $A\mathbf{x} = \mathbf{0}$ and an example of an arbitrary solution to the system.

We will observe a similar phenomenon in Chapter 6, when we establish that every solution to the homogeneous second-order linear differential equation

$$y'' + a_1 y' + a_2 y = 0$$

can be written in the form

$$y(x) = c_1 y_1(x) + c_2 y_2(x),$$

where $y_1(x)$ and $y_2(x)$ are two nonproportional solutions to the differential equation on the interval of interest.

In each of these problems, we have a set of "vectors" V (in the first problem the vectors are ordered triples of numbers, whereas in the second, they are functions that are at least twice differentiable on an interval I) and a linear vector equation. Further, in both cases, all solutions to the given equation can be expressed as a linear combination of two particular solutions.

In the next two chapters we develop this way of formulating linear problems in terms of an abstract set of vectors, V, and a linear vector equation with solutions in V. We will find that many problems fit into this framework and that the solutions to these problems can be expressed as linear combinations of a certain number (not necessarily two) of basic solutions. The importance of this result cannot be overemphasized. It reduces the search for all solutions to a given problem to that of finding a finite number of solutions. As specific applications, we will derive the theory underlying linear differential equations and linear systems of differential equations as special cases of the general framework.

Before proceeding further, we give a word of encouragement to the more application-oriented reader. It will probably seem at times that the ideas we are introducing are rather esoteric and that the formalism is pure mathematical abstraction. However, in addition to its inherent mathematical beauty, the formalism incorporates ideas that pervade many areas of applied mathematics, particularly engineering mathematics and mathematical physics, where the problems under investigation are very often linear in nature. Indeed, the linear algebra introduced in the next two chapters should be considered an extremely important addition to one's mathematical repertoire, certainly on a par with the ideas of elementary calculus.

4.1 Vectors in $\mathbb{R}^n$

In this section, we use some familiar ideas about geometric vectors to motivate the more general and abstract idea of a vector space, which will be introduced in the next section. We begin by recalling that a geometric vector can be considered mathematically as a directed line segment (or arrow) that has both a magnitude (length) and a direction attached to it. In calculus courses, we define **vector addition** according to the parallelogram law (see Figure 4.1.1); namely, the sum of the vectors $\mathbf{x}$ and $\mathbf{y}$ is the diagonal of the parallelogram formed by $\mathbf{x}$ and $\mathbf{y}$. We denote the sum by $\mathbf{x}+\mathbf{y}$. It can then be shown geometrically that for all vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$,

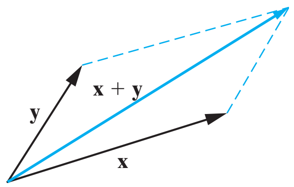

Figure 4.1.1: Parallelogram law of vector addition.

$$\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x} \tag{4.1.1}$$

and

$$\mathbf{x}+(\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}. \tag{4.1.2}$$

These are the statements that the vector addition operation is commutative and associative. The *zero vector*, denoted $\mathbf{0}$, is defined as the vector satisfying

$$\mathbf{x}+\mathbf{0}=\mathbf{x}, \tag{4.1.3}$$

for all vectors $\mathbf{x}$. We consider the zero vector as having zero magnitude and arbitrary direction. Geometrically, we picture the zero vector as corresponding to a point in space. Let $-\mathbf{x}$ denote the vector that has the same magnitude as $\mathbf{x}$, but the opposite direction. Then according to the parallelogram law of addition,

$$\mathbf{x}+(-\mathbf{x})=\mathbf{0}. \tag{4.1.4}$$

The vector $-\mathbf{x}$ is called the *additive inverse* of $\mathbf{x}$. Properties (4.1.1)–(4.1.4) are the fundamental properties of vector addition.

The basic algebra of vectors is completed when we also define the operation of multiplication of a vector by a real number. Geometrically, if $\mathbf{x}$ is a vector and k is a real number, then $k\mathbf{x}$ is defined to be the vector whose magnitude is $|k|$ times the magnitude of $\mathbf{x}$ and whose direction is the same as $\mathbf{x}$ if $k > 0$, and opposite to $\mathbf{x}$ if $k < 0$. (See Figure 4.1.2.) If $k = 0$, then $k\mathbf{x} = \mathbf{0}$. This **scalar multiplication** operation has several important properties that we now list. Once more, each of these can be established geometrically using only the foregoing definitions of vector addition and scalar multiplication.

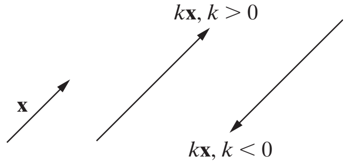

Figure 4.1.2: Scalar multiplication of $\mathbf{x}$ by k.

For all vectors $\mathbf{x}$ and $\mathbf{y}$, and all real numbers r, s and t,

$$1\mathbf{x} = \mathbf{x}, \tag{4.1.5}$$
$$(st)\mathbf{x} = s(t\mathbf{x}), \tag{4.1.6}$$
$$r(\mathbf{x}+\mathbf{y}) = r\mathbf{x}+r\mathbf{y}, \tag{4.1.7}$$
$$(s+t)\mathbf{x} = s\mathbf{x}+t\mathbf{x}. \tag{4.1.8}$$

It is important to realize that, in the foregoing development, we *have not* defined a "multiplication of vectors." In Chapter 3 we discussed the idea of a dot product and cross product of two vectors in space (see Equations (3.1.4) and (3.1.5)), but for the purposes of discussing abstract vector spaces we will essentially ignore the dot product and cross product. We will revisit the dot product in Section 4.11, when we develop inner product spaces.

We will see in the next section how the concept of a vector space arises as a direct generalization of the ideas associated with geometric vectors. Before performing this abstraction, we want to recall some further features of geometric vectors and give one specific and important extension.

We begin by considering vectors in the plane. Recall that $\mathbb{R}^2$ denotes the set of all ordered pairs of real numbers; thus,

$$\mathbb{R}^2 = \{(x, y) : x \in \mathbb{R}, y \in \mathbb{R}\}.$$

The elements of this set are called *vectors in* $\mathbb{R}^2$, and we use the usual vector notation to denote these elements. Geometrically we identify the vector $\mathbf{v} = (x, y)$ in $\mathbb{R}^2$ with

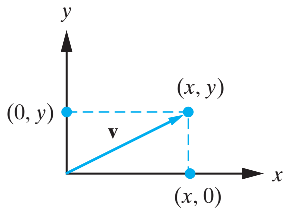

Figure 4.1.3: Identifying vectors in $\mathbb{R}^2$ with geometric vectors in the plane.

the geometric vector **v** directed from the origin of a Cartesian coordinate system to the point with coordinates (x, y). This identification is illustrated in Figure 4.1.3. The numbers x and y are called the **components** of the geometric vector **v**. The geometric vector addition and scalar multiplication operations are consistent with the addition and scalar multiplication operations defined in Chapter 2 via the correspondence with row (or column) vectors for $\mathbb{R}^2$:

If $\mathbf{v} = (x_1, y_1)$ and $\mathbf{w} = (x_2, y_2)$, and k is an arbitrary real number, then

$$\mathbf{v} + \mathbf{w} = (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \tag{4.1.9}$$

$$k\mathbf{v} = k(x_1, y_1) = (kx_1, ky_1). \tag{4.1.10}$$

These are the algebraic statements of the parallelogram law of vector addition and the scalar multiplication law, respectively. (See Figure 4.1.4.) Using the parallelogram law of vector addition and Equations (4.1.9) and (4.1.10), it follows that any vector $\mathbf{v} = (x, y)$ can be written as

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} = x(1, 0) + y(0, 1),$$

where $\mathbf{i} = (1, 0)$ and $\mathbf{j} = (0, 1)$ are the unit vectors pointing along the positive x- and y-coordinate axes, respectively.

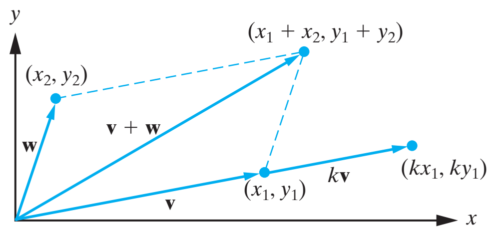

Figure 4.1.4: Vector addition and scalar multiplication in $\mathbb{R}^2$.

The properties (4.1.1)–(4.1.8) are now easily verified for vectors in $\mathbb{R}^2$. In particular, the zero vector in $\mathbb{R}^2$ is the vector

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

Furthermore, Equation (4.1.9) implies that

$$(x, y) + (-x, -y) = (0, 0) = \mathbf{0},$$

so that the additive inverse of the general vector $\mathbf{v} = (x, y)$ is $-\mathbf{v} = (-x, -y)$.

It is straightforward to extend these ideas to vectors in 3-space. We recall that

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

As illustrated in Figure 4.1.5, each vector $\mathbf{v} = (x, y, z)$ in $\mathbb{R}^3$ can be identified with the geometric vector **v** that joins the origin of a Cartesian coordinate system to the point with coordinates (x, y, z). We call x, y, and z the components of **v**.

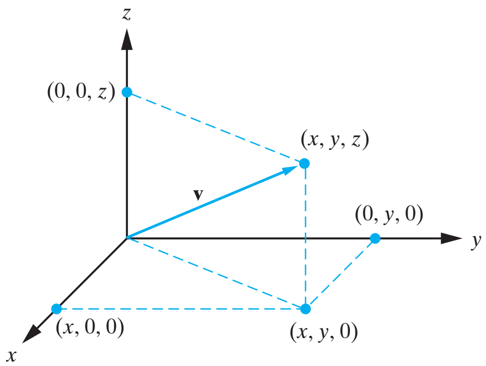

Figure 4.1.5: Identifying vectors in $\mathbb{R}^3$ with geometric vectors in space.

Recall that if $\mathbf{v} = (x_1, y_1, z_1)$, $\mathbf{w} = (x_2, y_2, z_2)$, and k is an arbitrary real number, then addition and scalar multiplication were given in Chapter 2 by

$$\mathbf{v} + \mathbf{w} = (x_1, y_1, z_1) + (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2), \tag{4.1.11}$$

$$k\mathbf{v} = k(x_1, y_1, z_1) = (kx_1, ky_1, kz_1). \tag{4.1.12}$$

Once more, these are, respectively, the component forms of the laws of vector addition and scalar multiplication for geometric vectors. It follows that an arbitrary vector $\mathbf{v} = (x, y, z)$ can be written as

$$\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x(1, 0, 0) + y(0, 1, 0) + z(0, 0, 1),$$

where $\mathbf{i} = (1, 0, 0)$, $\mathbf{j} = (0, 1, 0)$, and $\mathbf{k} = (0, 0, 1)$ denote the unit vectors which point along the positive x-, y-, and z-coordinate axes, respectively.

We leave it as an exercise to check that the properties (4.1.1)–(4.1.8) are satisfied by vectors in $\mathbb{R}^3$, where

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

and the additive inverse of $\mathbf{v} = (x, y, z)$ is $-\mathbf{v} = (-x, -y, -z)$.

We now come to our first major abstraction. Whereas the sets $\mathbb{R}^2$ and $\mathbb{R}^3$ and their associated algebraic operations arise naturally from our experience with Cartesian geometry, the motivation behind the algebraic operations in $\mathbb{R}^n$ for larger values of n does not come from geometry. Rather, we can view the addition and scalar multiplication operations in $\mathbb{R}^n$ for $n > 3$ as the natural extension of the component forms of addition and scalar multiplication in $\mathbb{R}^2$ and $\mathbb{R}^3$ in (4.1.9)–(4.1.12). Therefore, in $\mathbb{R}^n$ we have that if $\mathbf{v} = (x_1, x_2, \dots, x_n)$, $\mathbf{w} = (y_1, y_2, \dots, y_n)$, and k is an arbitrary real number, then

$$\mathbf{v} + \mathbf{w} = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n), \tag{4.1.13}$$

$$k\mathbf{v} = (kx_1, kx_2, \dots, kx_n). \tag{4.1.14}$$

Again, these definitions are direct generalizations of the algebraic operations defined in $\mathbb{R}^2$ and $\mathbb{R}^3$, but there is no geometric analogy when $n > 3$. It is easily established that these operations satisfy properties (4.1.1)–(4.1.8), where the **zero vector** in $\mathbb{R}^n$ is

$$\mathbf{0} = (0, 0, \dots, 0),$$

and the **additive inverse** of the vector $\mathbf{v} = (x_1, x_2, \dots, x_n)$ is

$$-\mathbf{v} = (-x_1, -x_2, \dots, -x_n).$$

The verification of this is left as an exercise.

Example 4.1.1 If $\mathbf{v} = (1.2, 3.5, 2, 0)$ and $\mathbf{w} = (12.23, 19.65, 23.22, 9.76)$, then

$$\mathbf{v} + \mathbf{w} = (1.2, 3.5, 2, 0) + (12.23, 19.65, 23.22, 9.76) = (13.43, 23.15, 25.22, 9.76)$$

and

$$2.35\mathbf{v} = (2.82, 8.225, 4.7, 0).$$ □

Exercises for 4.1

Key Terms

Vectors in $\mathbb{R}^n$, Vector addition, Scalar multiplication, Zero vector, Additive inverse, Components of a vector.

Skills

- Be able to perform vector addition and scalar multiplication for vectors in $\mathbb{R}^n$ given in component form.
- Understand the geometric perspective on vector addition and scalar multiplication in the cases of $\mathbb{R}^2$ and $\mathbb{R}^3$.
- Be able to formally verify the axioms (4.1.1)–(4.1.8) for vectors in $\mathbb{R}^n$.

True-False Review

For Questions 1–12, decide if the given statement is **true** or **false**, and give a brief justification for your answer. If true, you can quote a relevant definition or theorem from the text. If false, provide an example, illustration, or brief explanation of why the statement is false.

1. The vector (x, y) in $\mathbb{R}^2$ is the same as the vector $(x, y, 0)$ in $\mathbb{R}^3$.

2. Each vector (x, y, z) in $\mathbb{R}^3$ has exactly one additive inverse.

3. The solution set to a linear system of 4 equations and 6 unknowns consists of a collection of vectors in $\mathbb{R}^6$.

4. For every vector $(x_1, x_2, \dots, x_n)$ in $\mathbb{R}^n$, the vector $(-1) \cdot (x_1, x_2, \dots, x_n)$ is an additive inverse.

5. A vector whose components are all positive is called a "positive vector."

6. If s and t are scalars and $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^n$, then $(s+t)(\mathbf{x}+\mathbf{y}) = s\mathbf{x} + t\mathbf{y}$.

7. For every vector $\mathbf{x}$ in $\mathbb{R}^n$, the vector $0\mathbf{x}$ is the zero vector of $\mathbb{R}^n$.

8. The parallelogram whose sides are determined by vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^2$ have diagonals determined by the vectors $\mathbf{x}+\mathbf{y}$ and $\mathbf{x}-\mathbf{y}$.

9. If $\mathbf{x}$ is a vector in the first quadrant of $\mathbb{R}^2$, then any scalar multiple $k\mathbf{x}$ of $\mathbf{x}$ is still a vector in the first quadrant of $\mathbb{R}^2$.

10. The vector $5\mathbf{i} - 6\mathbf{j} + \sqrt{2}\mathbf{k}$ in $\mathbb{R}^3$ is the same as $(5, -6, \sqrt{2})$.

11. Three vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ in $\mathbb{R}^3$ always determine a 3-dimensional solid region in $\mathbb{R}^3$.

12. If $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^2$ whose components are even integers and k is a scalar, then $\mathbf{x}+\mathbf{y}$ and $k\mathbf{x}$ are also vectors in $\mathbb{R}^2$ whose components are even integers.

Problems

1. If $\mathbf{x} = (3, 1)$, $\mathbf{y} = (-1, 2)$, determine the vectors $\mathbf{v}_1 = 2\mathbf{x}$, $\mathbf{v}_2 = 3\mathbf{y}$, $\mathbf{v}_3 = 2\mathbf{x} + 3\mathbf{y}$. Sketch the corresponding points in the xy-plane and the equivalent geometric vectors.

2. If $\mathbf{x} = (-1, -4)$ and $\mathbf{y} = (-5, 1)$, determine the vectors $\mathbf{v}_1 = 3\mathbf{x}$, $\mathbf{v}_2 = -4\mathbf{y}$, $\mathbf{v}_3 = 3\mathbf{x} + (-4)\mathbf{y}$. Sketch the corresponding points in the xy-plane and the equivalent geometric vectors.

3. If $\mathbf{x} = (3, -1, 2, 5)$, $\mathbf{y} = (-1, 2, 9, -2)$, determine $\mathbf{v} = 5\mathbf{x} + (-7)\mathbf{y}$ and its additive inverse.

4. If $\mathbf{x} = (1, 2, 3, 4, 5)$ and $\mathbf{z} = (-1, 0, -4, 1, 2)$, find $\mathbf{y}$ in $\mathbb{R}^5$ such that $2\mathbf{x} + (-3)\mathbf{y} = -\mathbf{z}$.

5. Verify the commutative law of addition for vectors in $\mathbb{R}^4$.

6. Verify the associative law of addition for vectors in $\mathbb{R}^4$.

7. Verify properties (4.1.5)–(4.1.8) for vectors in $\mathbb{R}^3$.

8. Show with examples that if $\mathbf{x}$ is a vector in the first quadrant of $\mathbb{R}^2$ (i.e., both coordinates of $\mathbf{x}$ are positive) and $\mathbf{y}$ is a vector in the third quadrant of $\mathbb{R}^2$ (i.e., both coordinates of $\mathbf{y}$ are negative), then the sum $\mathbf{x}+\mathbf{y}$ could occur in any of the four quadrants.

4.2 Definition of a Vector Space

In the previous section, we showed how the set $\mathbb{R}^n$ of all ordered n-tuples of real numbers, together with the addition and scalar multiplication operations defined on it, has the same algebraic properties as the familiar algebra of geometric vectors. We now push this abstraction one step further and introduce the idea of a vector space. Such an abstraction will enable us to develop a mathematical framework for studying a broad class of linear problems, such as systems of linear equations, linear differential equations, and systems of linear differential equations, which have far-reaching applications in all areas of applied mathematics, science, and engineering.

Let V be a nonempty set. For our purposes, it is useful to call the elements of V vectors and use the usual vector notation $\mathbf{u}, \mathbf{v}, \ldots,$ to denote these elements. For example, if V is the set of all 2×2 matrices, then the vectors in V are 2×2 matrices, whereas if V is the set of all positive integers, then the vectors in V are positive integers. We will be interested only in the case when the set V has an addition operation and a scalar multiplication operation defined on its elements in the following senses:

Vector Addition: A rule for combining any two vectors in V. We will use the usual $+$ sign to denote an addition operation, and the result of adding the vectors $\mathbf{u}$ and $\mathbf{v}$ will be denoted $\mathbf{u}+\mathbf{v}$.

Real (or Complex) Scalar Multiplication: A rule for combining each vector in V with any real (or complex) number. We will use the usual notation $k\mathbf{v}$ to denote the result of scalar multiplying the vector $\mathbf{v}$ by the real (or complex) number k.

To combine the two types of scalar multiplication, we let F denote the set of scalars for which the operation is defined. Thus, for us, F is either the set of all real numbers or the set of all complex numbers. For example, if V is the set of all 2×2 matrices with complex elements and F denotes the set of all complex numbers, then the usual operation of matrix addition is an addition operation on V, and the usual method of multiplying a matrix by a scalar is a scalar multiplication operation on V. Notice that the result of applying either of these operations is always another vector (2×2 matrix) in V.

As a further example, let V be the set of positive integers, and let F be the set of all real numbers. Then the usual operations of addition and multiplication within the real numbers define addition and scalar multiplication operations on V. Note in this case, however, that the scalar multiplication operation, in general, will not yield another vector in V, since when we multiply a positive integer by a real number, the result is not, in general, a positive integer.

We are now in a position to give a precise definition of a vector space.