CHAPTER 3

EUCLIDEAN VECTOR SPACES

Engineers and physicists distinguish between two types of physical quantities-**scalars**, which are quantities that can be described by a numerical value alone, and **vectors**, which are quantities that require both a number and a direction for their complete physical description. In this chapter our goal is to review some of the basic theory of vectors in two and three dimensions, and more generally in n dimensions.

3.1. Vectors In 2-Space, 3-Space, and n-Space

Engineers and physicists represent vectors in two dimensions (also called **2-space**) or in three dimensions (also called **3-space**) by arrows. The direction of the arrowhead specifies the **direction** of the vector and the **length** of the arrow specifies the magnitude. Mathematicians call these **geometric vectors**. The tail of the arrow is the **initial point** and the tip is the **terminal point** of the vector. When we want to indicate that a vector $\mathbf{v}$ has initial point A and terminal point B, then we write $\mathbf{v} = \overrightarrow{AB}$.

Vectors with the same length and direction are said to be **equivalent**. Equivalent vectors are also said to be **equal**, which we indicate by writing $\mathbf{u} = \mathbf{v}$.

The vector whose initial and terminal points coincide has length zero, so we call this the **zero vector** and denote it by $\mathbf{0}$. The zero vector has no natural direction.

There are a number of important algebraic operations on vectors.

- **Parallelogram rule for vector addition**

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* 2*-space or* 3*-space that are positioned so that their initial points coincide, then the two vectors form adjacent sides of a parallelogram, and the* ***sum*** $\mathbf{u}+\mathbf{v}$ *is the vector represented by the arrow from the common initial point of* $\mathbf{u}$ *and* $\mathbf{v}$ *to the opposite vertex of the parallelogram.*

Here is another way to form the sum of two vectors.

- **Triangular rule for vector addition**

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* 2*-space or* 3*-space that are positioned so that the initial point of* $\mathbf{v}$ *is the terminal point of* $\mathbf{u}$, *then the* ***sum*** $\mathbf{u}+\mathbf{v}$ *is the vector represented by the arrow from the initial point of* $\mathbf{u}$ *to the terminal point of* $\mathbf{v}$.

The sum obtained by the triangular rule is the same as the sum obtained by the parallelogram rule.

In ordinary arithmetic we can write $a - b = a + (-b)$ to express subtraction in terms of addition. There is an analogues idea in the vector arithmetic.

- **Vector subtraction**

  The ***negative*** *of a vector* $\mathbf{v}$, *denoted by* $-\mathbf{v}$, *is the vector that has the same length as* $\mathbf{v}$ *but is oppositely directed, and the* ***difference*** *of* $\mathbf{v}$ *from* $\mathbf{u}$, *denoted by* $\mathbf{u}-\mathbf{v}$, *is taken to be the sum* $\mathbf{u}+(-\mathbf{v})$.

Sometimes there is a need to change the length of a vector or change its length and reverse its direction. This is accomplished by a type of multiplication in which vectors are multiplied by scalars.

- **Scalar multiplication**

  If $\mathbf{v}$ *is a nonzero vector in* 2*-space or* 3*-space, and if* k *is a nonzero scalar, then we define the* ***scalar product of*** $\mathbf{v}$ ***by*** k *to be the vector whose length is* $|k|$ *times the length of* $\mathbf{v}$ *and whose direction is the same as that of* $\mathbf{v}$ *if* k *is positive and opposite to that of* $\mathbf{v}$ *if* k *is negative. If* $k = 0$ *or* $\mathbf{v} = \mathbf{0}$, *then we define* $k\mathbf{v}$ *to be* $\mathbf{0}$.

Suppose that $\mathbf{u}$ and $\mathbf{v}$ are vectors in 2-space or 3-space with a common initial point. If one of the vectors is a scalar multiple of the other, then the vectors lie on a common like, so that they are **collinear**.

Vector addition satisfies the **associative law for addition**, meaning that when we add three vectors, say $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, it does not matter which two we add first, that is,

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

If a vector $\mathbf{v}$ in 2-space or 3-space is positioned with its initial point at the origin of a rectangular coordinate system, then the vector is completely determined by the coordinates of its terminal point. We call these coordinates the **components** of $\mathbf{v}$ relative to the coordinate system. We write $\mathbf{v} = (v_1, v_2)$ or $\mathbf{v} = (v_1, v_2, v_3)$ in 2-space or 3-space.

Two vectors are equivalent if and only if their corresponding components are equal. Thus, for example, the vectors $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ in 3-space are equivalent if and only if $u_1 = v_1$, $u_2 = v_2$, and $u_3 = v_3$.

It is sometimes necessary to consider vectors whose initial points are not at the origin. If $\overrightarrow{P_1P_2}$ denotes the vector with initial point $P_1 = (x_1, y_1)$ and terminal point $P_2 = (x_2, y_2)$, then the components of this vector are given by the formula $\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1)$. Similarly, in 3-space, we have $\overrightarrow{P_1P_2} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$.

The notion of space can be extended to n dimensions. The set of all real numbers can be viewed geometrically as a line. It called the **real line** and denoted by $\mathbb{R}$ or $\mathbb{R}^1$. Intuitively, a line is one-dimensional. The set of all ordered pairs of real numbers (called **2-tuples**) and the set of all ordered triples of real numbers (called **3-tuples**) are denoted by $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Ordered pairs correspond to points in the plane (two-dimensional) and ordered triples to points in space (three-dimensional). The following definition extends this idea.

- **n-space**

  *If n is a positive integer, then an **ordered n-tuple** is a sequence of n real numbers $(v_1, v_2, \ldots, v_n)$. The set of all ordered n-tuples is called **n-space** and is denoted by $\mathbb{R}^n$.*

We denote a vector $\mathbf{v}$ in $\mathbb{R}^n$ using the notation $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and we call $\mathbf{0} = (0, 0, \ldots, 0)$ the **zero vector.**

- **Equivalent vectors**

  *Vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ are said to be **equivalent** (also called **equal**) if $u_1 = v_1, u_2 = v_2, \ldots, u_n = v_n$. We indicate this by writing $\mathbf{u} = \mathbf{v}$.*

- **Addition, subtraction, and scalar multiplication**

  If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in $\mathbb{R}^n$, and if k is any scalar, then we define

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n), \qquad -\mathbf{v} = (-v_1, -v_2, \ldots, -v_n),$$
$$k\mathbf{v} = (kv_1, kv_2, \ldots, kv_n), \qquad \mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = (u_1 - v_1, u_2 - v_2, \ldots, u_n - v_n).$$

The following result summarizes the most important properties of vector operations.

- **Properties of vector operations**

  If $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ are vectors in $\mathbb{R}^n$, and if k and m are scalars, then

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

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

(c) $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$

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

(e) $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$

(f) $(k+m)\mathbf{u}=k\mathbf{u}+m\mathbf{u}$

(g) $k(m\mathbf{u})=(km)\mathbf{u}$

(h) $1\mathbf{u}=\mathbf{u}$

Addition subtraction, and scalar multiplication are frequently used in combination to form new vectors.

- **Linear combination**

  If $\mathbf{u}$ *is a vector in* $\mathbb{R}^n$*,then* $\mathbf{u}$ *is said to be* ***linear combination*** *of the vectors* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *in* $\mathbb{R}^n$ *if it can be expressed in the form*

$$\mathbf{u}=k_1\mathbf{v}_1+k_2\mathbf{v}_2+\cdots+k_r\mathbf{v}_r,$$

  where $k_1, k_2, \ldots, k_r$ *are scalars.*

3.2. Norm, Dot Product, and Distance in $\mathbb{R}^n$

We denote the length of a vector $\mathbf{v}$ by the symbol $\|\mathbf{v}\|$, which is read as the **norm** (or the **length**, or the **magnitude**) of $\mathbf{v}$. It follows from the theorem of Pythagoras that the norm of a vector $\mathbf{v}=(v_1, v_2)$ in $\mathbb{R}^2$ is

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

Similarly, for a vector $\mathbf{v}=(v_1, v_2, v_3)$ in $\mathbb{R}^3$, it follows from two applications of the theorem of Pythagoras that

$$\|\mathbf{v}\|=\sqrt{v_1^2+v_2^2+v_3^2}.$$

In general, we make the following definition.

- **Norm of a vector**

  If $\mathbf{v}=(v_1, v_2, \ldots, v_n)$ *is a vector in* $\mathbb{R}^n$*,then the* ***norm*** *of* $\mathbf{v}$ *is denoted by* $\|\mathbf{v}\|$*, and defined by the formula*

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

Our first result in this section generalizes the following three familiar facts about vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ to $\mathbb{R}^n$:

·Distances are nonnegative.

·The zero vector is the only vector of length zero.

·Multiplying a vector by a scalar multiplies its length by the absolute value of that scalar.

- **Properties of norm**

  If $\mathbf{v}$ *is a vector in* $\mathbb{R}^n$ *and if* k *is any scalar, then*

  (a) $\|\mathbf{v}\| \geq 0.$

  (b) $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = 0.$

  (c) $\|k\mathbf{v}\| = |k| \, \|\mathbf{v}\|.$

A vector of norm 1 is called a **unit vector**. If $\mathbf{v}$ is any nonzero vector in $\mathbb{R}^n$, then

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$$

defines a unit vector that is in the same direction as $\mathbf{v}$. The process of multiplying a nonzero vector by the reciprocal of its norm to obtain a unit vector is called **normalizing**.

When a rectangular coordinate system is introduced in $\mathbb{R}^2$ or $\mathbb{R}^3$, the unit vectors in the positive directions of the coordinate axes are called **standard unit vectors**. In $\mathbb{R}^2$, these vectors are denoted by

$$\mathbf{i} = (1, 0) \quad \text{and} \quad \mathbf{j} = (0, 1),$$

and in $\mathbb{R}^3$ by

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \text{and} \quad \mathbf{k} = (0, 0, 1).$$

Every vector $\mathbf{v} = (v_1, v_2)$ in $\mathbb{R}^2$ and every vector $\mathbf{v} = (v_1, v_2, v_3)$ in $\mathbb{R}^3$ can be represented as a linear combination of standard unit vectors by writing

$$\begin{aligned}
\mathbf{v} &= (v_1, v_2) = v_1 (1, 0) + v_2 (0, 1) = v_1\mathbf{i} + v_2\mathbf{j}, \\
\mathbf{v} &= (v_1, v_2, v_3) = v_1 (1, 0, 0) + v_2 (0, 1, 0) + v_3 (0, 0, 1) = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}.
\end{aligned}$$

Moreover, we can generalize these formulas to $\mathbb{R}^n$ by defining the **standard unit vectors** in $\mathbb{R}^n$ to be

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0, 0), \ldots, \mathbf{e}_n = (0, 0, 0, \ldots, 0, 1),$$

in which case every vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ can be expressed as

$$\mathbf{v} = (v_1, v_2, \ldots, v_n) = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n.$$

If P_1 and P_2 are points in $\mathbb{R}^2$ or $\mathbb{R}^3$, then the length of the vector $\overrightarrow{P_1P_2}$ is equal to the **distance** d between the two points. More specifically, if $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, then

$$d = \left\|\overrightarrow{P_1P_2}\right\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

This is the familiar distance formula from analytic geometry. Similarly, the distance between the points $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ in 3-space is

$$d = \left\|\overrightarrow{P_1P_2}\right\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

Motivated by these formulas, we make the following definition.

- **Distance**

  If $\mathbf{u} = (u_1, u_2, \ldots, u_r)$ *and* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *are vectors in* $\mathbb{R}^n$*,then we denote the* ***distance*** *between* $\mathbf{u}$ *and* $\mathbf{v}$ *by* $d(\mathbf{u}, \mathbf{v})$ *and define it to be*

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}.$$

For example, if $\mathbf{u} = (1, 3, -2, 7)$ and $\mathbf{v} = (0, 7, 2, 2)$, then

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{(1-0)^2 + (3-7)^2 + (-2-2)^2 + (7-2)^2} = \sqrt{58}.$$

Our next objective is to define a useful multiplication operation on vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ and then extend that operation to $\mathbb{R}^n$.

- **Dot product in $\mathbb{R}^2$ and $\mathbb{R}^3$**

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are nonzero vectors in* $\mathbb{R}^2$ *or* $\mathbb{R}^3$*, and if* θ *is the angle between them, then the* ***dot product*** *(also called the* ***Euclidean inner product****) of* $\mathbf{u}$ *and* $\mathbf{v}$ *is a real number denoted by* $\mathbf{u} \cdot \mathbf{v}$ *and is defined as*

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta.$$

  If $\mathbf{u} = \mathbf{0}$ *or* $\mathbf{v} = \mathbf{0}$*, then we define* $\mathbf{u} \cdot \mathbf{v} = \mathbf{0}$.

The sign of the dot product reveals information about the angle. Since

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

and $0 \le \theta \le \pi$, we have

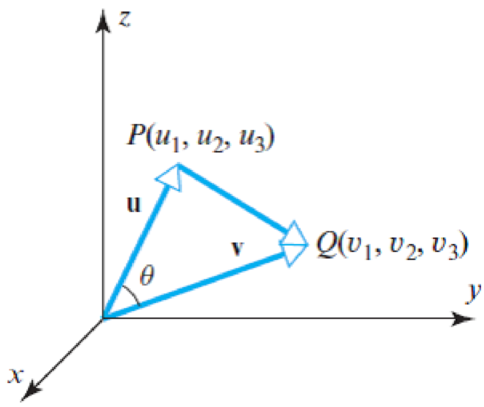
$\cdot\theta$ is acute if $\mathbf{u} \cdot \mathbf{v} > 0$.

$\cdot\theta$ is obtuse if $\mathbf{u}\cdot\mathbf{v} < 0$.

$\cdot\theta = \frac{\pi}{2}$ if $\mathbf{u}\cdot\mathbf{v} = 0$.

For computable purposes it is desirable to have a formula that expresses the dot product of two vectors in terms of components. We derive such a formula for vectors in 3-space. The derivation for vectors in 2-space is similar.

Let $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ be two nonzero vectors, and θ be the angle between them as shown in the following figure.



By the law of cosines, we have

$$\left\|\overrightarrow{PQ}\right\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta \Rightarrow \|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta,$$

that is,

$$\|\mathbf{u}\|\|\mathbf{v}\|\cos\theta = \frac{1}{2}\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right).$$

Substituting

$$\|\mathbf{u}\|^2 = u_1^2 + u_2^2 + u_3^2, \quad \|\mathbf{v}\|^2 = v_1^2 + v_2^2 + v_3^2, \quad \|\mathbf{u}-\mathbf{v}\|^2 = (v_1-u_1)^2 + (v_2-u_2)^2 + (v_3-u_3)^2$$

and simplifying, we obtain that

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3.$$

The corresponding formula for vectors in 2-space is $\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2$.

Motivated by the pattern in these formulas, we make the following definition.

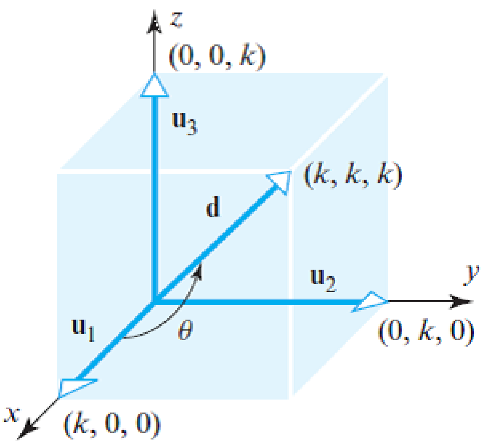
- **Dot product in $\mathbb{R}^n$**

  *If $\mathbf{u} = (u_1, u_2, \ldots, u_r)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are vectors in $\mathbb{R}^n$, then the **dot product** of $\mathbf{u}$ and $\mathbf{v}$ is defined by*

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

Example: Find the angle between a diagonal of a cube and one of its edges.

Solution: Consider the following figure.



Let k be the length of an edge. If we let $\mathbf{u}_1 = (k, 0, 0)$, $\mathbf{u}_2 = (0, k, 0)$, and $\mathbf{u}_3 = (0, 0, k)$, then the vector $\mathbf{d} = (k, k, k) = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$ is a diagonal of the cube. The angle between $\mathbf{d}$ and $\mathbf{u}_1$ satisfies

$$\cos\theta = \frac{\mathbf{u}_1 \cdot \mathbf{d}}{\|\mathbf{u}_1\|\,\|\mathbf{d}\|} = \frac{k^2}{k\sqrt{3k^2}} = \frac{1}{\sqrt{3}}.$$

Thus

$$\theta = \arccos\left(\frac{1}{\sqrt{3}}\right) \approx 54.74$$

degrees.■

In the special case where $\mathbf{u} = \mathbf{v}$ in the definition of dot product, we obtain the relationship

$$\mathbf{v} \cdot \mathbf{v} = v_1^2 + v_2^2 + \cdots + v_n^2,$$

so $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$.

Dot products have many of the same algebraic properties as products of real numbers.

- **Properties of dot product**

  If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^n$, *and if* k *is a scalar, then*

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

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

  (c) $k\,(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$

  (f) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$

Our next objective is to extend the notion of angle between two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ to $\mathbb{R}^n$. We do this by starting with the formula

$$\theta = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right).$$

Since dot products and norms have been defined for vectors in $\mathbb{R}^n$, it would seem that this formula has all the ingredients to serve as a definition of the angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. However, the problem being that the inverse cosine in the above formula is not defined unless its argument satisfies the inequalities

$$-1 \leq \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} \leq 1.$$

Fortunately, these inequalities do hold for all nonzero vectors in $\mathbb{R}^n$ as a result of the following fundamental result.

- **Cauchy-Schwarz inequality**

  If $\mathbf{u} = (u_1, u_2, \ldots, u_r)$ *and* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *are vectors in* $\mathbb{R}^n$*, then*

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|,$$

  or in terms of components

$$|u_1v_1 + u_2v_2 + \cdots + u_nv_n| \leq \left(u_1^2 + u_2^2 + \cdots + u_n^2\right)^{\frac{1}{2}} \left(v_1^2 + v_2^2 + \cdots + v_n^2\right)^{\frac{1}{2}}.$$

As we have already seen, some familiar results that can be visualized in $\mathbb{R}^2$ and $\mathbb{R}^3$ might be valid in $\mathbb{R}^n$ as well. For example, we have the following result.

- **Triangle inequalities**

  If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^n$*, then*

  (a) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *(triangle inequality for vectors)*

  (b) $d(\mathbf{u},\mathbf{v}) \leq d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$ *(triangle inequality for distances)*

Proof:

(a) We have

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = (\mathbf{u}\cdot\mathbf{u}) + 2(\mathbf{u}\cdot\mathbf{v}) + (\mathbf{v}\cdot\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2|\mathbf{u}\cdot\mathbf{v}| + \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2. \end{aligned}$$

This completes the proof since both sides of the inequality are nonnegative.

(b) It follows from part (a) that

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \|\mathbf{u}-\mathbf{w}+\mathbf{w}-\mathbf{v}\| \leq \|\mathbf{u}-\mathbf{w}\| + \|\mathbf{w}-\mathbf{v}\| = d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$$

since $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$.■

It is proved in plane geometry that for any parallelogram the sum of the squares of diagonal is equal to the sum of the squares of the four sides. The following generalizes this result to $\mathbb{R}^n$.

- **Parallelogram equation for vectors**

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* $\mathbb{R}^n$, *then*

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right).$$

Proof: We have

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) + (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = 2(\mathbf{u}\cdot\mathbf{u}) + 2(\mathbf{v}\cdot\mathbf{v}) = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right)$$

as desired.■

The next result establishes a fundamental relationship between the dot product and norm in $\mathbb{R}^n$.

- **Polarization equality**

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* $\mathbb{R}^n$, *then*

$$\mathbf{u}\cdot\mathbf{v} = \frac{1}{4}\left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right).$$

Proof: We have

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2, \\
\|\mathbf{u}-\mathbf{v}\|^2 &= (\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v}) = \|\mathbf{u}\|^2 - 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2,
\end{aligned}$$

from which the result follows by simple algebra.■

There are various ways to express the dot product of vectors using matrix notation. The formulas depend on whether the vectors are expressed as row matrices or column matrices. More specifically, we have

· If $\mathbf{u}$ and $\mathbf{v}$ are column matrices, then $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$.

· If $\mathbf{u}$ is a row matrix and $\mathbf{v}$ is a column matrix, then $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\mathbf{v} = \mathbf{v}^T\mathbf{u}^T$.

· If $\mathbf{u}$ is a column matrix and $\mathbf{v}$ is a row matrix, then $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\mathbf{v} = \mathbf{u}^T\mathbf{v}^T$.

· If $\mathbf{u}$ and $\mathbf{v}$ are row matrices, then $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}\mathbf{v}^T = \mathbf{v}\mathbf{u}^T$.

If A is an $n\times n$ matrix and $\mathbf{u}$ and $\mathbf{v}$ are $n\times 1$ matrices, then it follows from above discussion that

$$A\mathbf{u}\cdot\mathbf{v} = \mathbf{v}^T(A\mathbf{u}) = \left(\mathbf{v}^T A\right)\mathbf{u} = \left(A^T\mathbf{v}\right)^T\mathbf{u} = \mathbf{u}\cdot A^T\mathbf{v}$$

and

$$\mathbf{u}\cdot A\mathbf{v} = (A\mathbf{v})^T\mathbf{u} = \left(\mathbf{v}^T A^T\right)\mathbf{u} = \mathbf{v}^T\left(A^T\mathbf{u}\right) = A^T\mathbf{u}\cdot\mathbf{v}.$$

These formulas provide an important link between multiplication by an $n \times n$ matrix A and multiplication by A^T.

Dot products provide another way of thinking about matrix multiplication. Recall that if $A = [a_{ij}]$ is an $m \times r$ matrix and $B = [b_{ij}]$ is an $r \times n$ matrix, then the (i, j)th entry of AB is

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj},$$

which is the dot product of the ith row vector of A

$$\begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{ir} \end{bmatrix}$$

and the jth column vector

$$\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{rj} \end{bmatrix}.$$

Thus, if the row vectors of A are $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ and the column vectors of B are $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$, then the matrix product AB can be expressed as

$$AB = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{c}_1 & \mathbf{r}_1 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{c}_n \\ \mathbf{r}_2 \cdot \mathbf{c}_1 & \mathbf{r}_2 \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{c}_n \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{r}_m \cdot \mathbf{c}_1 & \mathbf{r}_m \cdot \mathbf{c}_2 & \cdots & \mathbf{r}_m \cdot \mathbf{c}_n \end{bmatrix}$$

3.3. Orthogonality

In this section we focus on the notion of "perpendicularity". Perpendicular vectors in $\mathbb{R}^n$ play an important role in a wide variety of applications.

- **Orthogonal vectors**

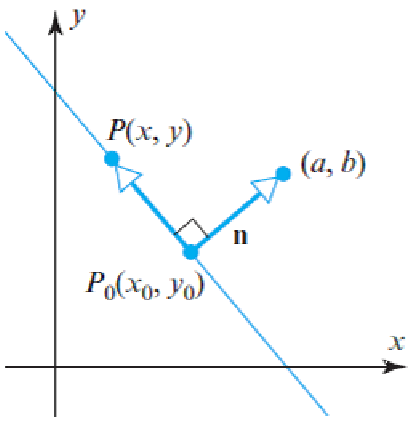
  Two nonzero vectors $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^n$ *are said to be* ***orthogonal*** *(or* ***perpendicular****) if* $\mathbf{u} \cdot \mathbf{v} = 0$*. The zero vector in* $\mathbb{R}^n$ *is orthogonal to every vector in* $\mathbb{R}^n$*.*

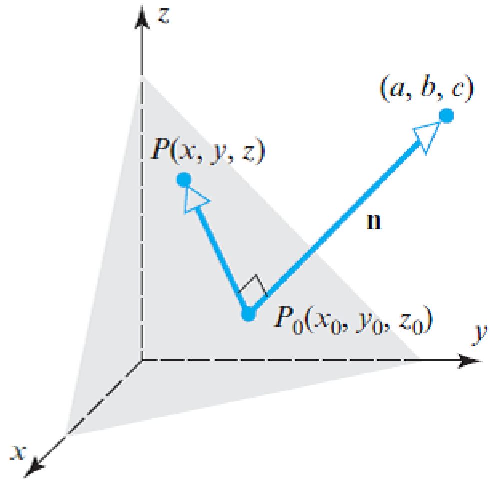
For example, if $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is the set of standard unit vectors in $\mathbb{R}^3$, then

$$\begin{aligned} \mathbf{i} \cdot \mathbf{j} &= (1,0,0) \cdot (0,1,0) = 0, \\ \mathbf{i} \cdot \mathbf{k} &= (1,0,0) \cdot (0,0,1) = 0, \\ \mathbf{j} \cdot \mathbf{k} &= (0,1,0) \cdot (0,0,1) = 0, \end{aligned}$$

so each pair of vectors in S is orthogonal.

A line in $\mathbb{R}^2$ is determined uniquely by its slope and one of its points, and a plane in $\mathbb{R}^3$ is determined uniquely by its inclination and one of its points. One way of specifying slope and inclination is to use a nonzero vector $\mathbf{n}$, called **normal**, that is orthogonal to the line or plane in the question. For example, consider the following figures.





In the first figure the line through the point $P_0 = (x_0, y_0)$ has normal $\mathbf{n} = (a, b)$ and in the second figure the plane through the point $P_0 = (x_0, y_0, z_0)$ has normal $\mathbf{n} = (a, b, c)$. Both the line and the plane are represented by the vector equation

$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0,$$

where P is either an arbitrary point (x, y) on the line or an arbitrary point (x, y, z) in the plane. The vector $\overrightarrow{P_0P}$ can be expressed in terms of components as

$$a\,(x - x_0) + b\,(y - y_0) = 0$$

and

$$a\,(x - x_0) + b\,(y - y_0) + c\,(z - z_0) = 0.$$

These are called the **point-normal** equations of the line and the plane. Conversely, if a and b are constants that are not both zero, then an equation of the form $ax + by + c = 0$ represents a line in $\mathbb{R}^2$ with normal $\mathbf{n} = (a, b)$, and if a, b, and c are constants that are not all zero, then an equation of the form $ax + by + cz + d = 0$ represents a plane in $\mathbb{R}^3$ with normal $\mathbf{n} = (a, b, c)$.

Recall that

$$ax + by = 0 \quad \text{and} \quad ax + by + cz = 0$$

are called homogeneous equations. Since they can be written as

$$(a, b) \cdot (x, y) = 0 \quad \text{and} \quad (a, b, c) \cdot (x, y, z) = 0,$$

or, alternatively, as

$$\mathbf{n}_1 \cdot (x, y) = 0 \quad \text{and} \quad \mathbf{n}_2 \cdot (x, y, z) = 0,$$

these homogeneous equations can be written as

$$\mathbf{n} \cdot \mathbf{x} = 0,$$

where $\mathbf{n}$ is the vector of coefficients and $\mathbf{x}$ is the vector of unknowns. In $\mathbb{R}^2$, this is called the **vector form of a line through the origin** and in $\mathbb{R}^3$, it is called the **vector form of a plane through the origin**.

In many applications it is necessary to decompose a vector $\mathbf{u}$ into a sum of two terms, one term being a scalar multiple of a special nonzero vector $\mathbf{a}$ and the other term being orthogonal to $\mathbf{a}$. For example, if $\mathbf{u}$ and $\mathbf{a}$ are vectors in $\mathbb{R}^2$ that are positioned so their initial points coincide at a point Q, then we can create such a decomposition as follows:

· Drop a perpendicular from the tip of $\mathbf{u}$ to the line through $\mathbf{a}$.

· Construct the vector $\mathbf{w}_1$ from Q to the foot of the perpendicular.

· Construct the vector $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1$.

Since

$$\mathbf{w}_2 + \mathbf{w}_1 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u},$$

we have decomposed $\mathbf{u}$ into a sum of two orthogonal vectors, the first term being a scalar multiple of $\mathbf{a}$ and the second being orthogonal to $\mathbf{a}$.

The following theorem shows that the foregoing results, which we explained using vectors in $\mathbb{R}^2$, apply as well in $\mathbb{R}^n$.

- **Projection theorem**

  If $\mathbf{u}$ and $\mathbf{a}$ are vectors in $\mathbb{R}^n$, and if $\mathbf{a} \neq \mathbf{0}$, then $\mathbf{u}$ can be expressed in exactly one way in the form $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1$ is a scalar multiple of $\mathbf{a}$ and $\mathbf{w}_2$ is orthogonal to $\mathbf{a}$.

Proof: Since the vector $\mathbf{w}_1$ is to be a scalar multiple of $\mathbf{a}$, it must have the form $\mathbf{w}_1 = k\mathbf{a}$. Our goal is to find a value of the scalar k and a vector $\mathbf{w}_2$ that is orthogonal to $\mathbf{a}$ such that $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$. We can determine k by

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2 = k\mathbf{a} + \mathbf{w}_2.$$

We obtain that

$$\mathbf{u}\cdot\mathbf{a} = (k\mathbf{a}+\mathbf{w}_2)\cdot\mathbf{a} = k\left\|\mathbf{a}\right\|^2 + (\mathbf{w}_2\cdot\mathbf{a}).$$

Since $\mathbf{w}_2$ is to be orthogonal to $\mathbf{a}$, we have $\mathbf{w}_2\cdot\mathbf{a}=0$, so k must satisfy the equation

$$\mathbf{u}\cdot\mathbf{a} = k\left\|\mathbf{a}\right\|^2 \Rightarrow k = \frac{\mathbf{u}\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}.$$

Thus

$$\mathbf{w}_2 = \mathbf{u}-\mathbf{w}_1 = \mathbf{u}-k\mathbf{a} = \mathbf{u}-\frac{\mathbf{u}\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}\mathbf{a},$$

and

$$\mathbf{w}_2\cdot\mathbf{a} = \mathbf{u}\cdot\mathbf{a}-\frac{\mathbf{u}\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}(\mathbf{a}\cdot\mathbf{a}) = \mathbf{u}\cdot\mathbf{a}-\mathbf{u}\cdot\mathbf{a}=0,$$

that is, $\mathbf{w}_2$ is orthogonal to $\mathbf{a}$.■

The vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in the projection theorem have associative names. The vector $\mathbf{w}_1$ is called the **orthogonal projection of u on a**, and the vector $\mathbf{w}_2$ is called the **vector component of u orthogonal to a**. The vector $\mathbf{w}_1$ is commonly denoted by the symbol $\text{proj}_{\mathbf{a}}\mathbf{u}$, in which case $\mathbf{w}_2=\mathbf{u}-\text{proj}_{\mathbf{a}}\mathbf{u}$. In summary,

$$\text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}\mathbf{a} \quad \text{and} \quad \mathbf{u}-\text{proj}_{\mathbf{a}}\mathbf{u} = \mathbf{u}-\frac{\mathbf{u}\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}\mathbf{a}.$$

Example: Find the orthogonal projections of the vectors $\mathbf{e}_1=(1,0)$ and $\mathbf{e}_2=(0,1)$ on the line L that makes an angle θ with the positive x-axis in $\mathbb{R}^2$.

Solution: The vector $\mathbf{a}=(\cos\theta,\sin\theta)$ is a unit vector along the line L, so our first problem is to find the orthogonal porjection of $\mathbf{e}_1$ along $\mathbf{a}$. Since

$$\left\|\mathbf{a}\right\| = \sqrt{\cos^2\theta+\sin^2\theta} = 1 \quad \text{and} \quad \mathbf{e}_1\cdot\mathbf{a} = (1,0)\cdot(\cos\theta,\sin\theta) = \cos\theta,$$

it follows that this projection is

$$\text{proj}_{\mathbf{a}}\mathbf{e}_1 = \frac{\mathbf{e}_1\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}\mathbf{a} = \cos\theta(\cos\theta,\sin\theta) = \left(\cos^2\theta,\cos\theta\sin\theta\right).$$

Similarly, since $\mathbf{e}_2\cdot\mathbf{a}=(0,1)\cdot(\cos\theta,\sin\theta)=\sin\theta$, it follows that

$$\text{proj}_{\mathbf{a}}\mathbf{e}_2 = \frac{\mathbf{e}_2\cdot\mathbf{a}}{\left\|\mathbf{a}\right\|^2}\mathbf{a} = \sin\theta(\cos\theta,\sin\theta) = \left(\cos\theta\sin\theta,\sin^2\theta\right),$$

the desired projection.■

Example: Let $\mathbf{u}=(2,-1,3)$ and $\mathbf{a}=(4,-1,2)$. Find the vector component of $\mathbf{u}$ along $\mathbf{a}$ and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.

Solution: We have

$$\|\mathbf{a}\|^2 = 4^2 + (-1)^2 + 2^2 = 21 \quad \text{and} \quad \mathbf{u}\cdot\mathbf{a} = (2)(4) + (-1)(-1) + (3)(2) = 15.$$

Thus,

$$\text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \frac{15}{21}(4,-1,2) = \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right)$$

is the vector component of $\mathbf{u}$ along $\mathbf{a}$ and

$$\mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = (2,-1,3) - \left(\frac{20}{7}, -\frac{5}{7}, \frac{10}{7}\right) = \left(-\frac{6}{7}, -\frac{2}{7}, \frac{11}{7}\right)$$

is the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.■

Sometimes it is more interesting to study the norm of the vector component of $\mathbf{u}$ along $\mathbf{a}$ than the norm of the vector component itself. A formula for this norm can be derived as follows:

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \left\|\frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}\right\| = \left|\frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\right| \|\mathbf{a}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|^2}\|\mathbf{a}\| \Rightarrow \|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|}.$$

If θ denotes the angle between $\mathbf{u}$ and $\mathbf{a}$, then $\mathbf{u}\cdot\mathbf{a} = \|\mathbf{u}\|\,\|\mathbf{a}\|\cos\theta$, so

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \|\mathbf{u}\|\,|\cos\theta|\,.$$

Many theorems about vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ also hold in $\mathbb{R}^n$. Another example of this is the following generalization of the theorem of Pythagoras.

- **Theorem of Pythagoras in** $\mathbb{R}^n$

  If $\mathbf{u}$ *and* $\mathbf{v}$ *are orthogonal vectors in* $\mathbb{R}^n$*, then* $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

Proof: Since $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, we have $\mathbf{u}\cdot\mathbf{v} = 0$, and then

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2,$$

as desired.■

We now show how orthogonal projections can be used to solve the following three distance problems:

· Find the distance between a point and a line in $\mathbb{R}^2$.

· Find the distance between a point and a plane in $\mathbb{R}^3$.

· Find the distance between two parallel planes in $\mathbb{R}^3$.

- **Distance formulas**

· *In* $\mathbb{R}^2$*, the distance* D *between the point* $P_0 = (x_0, y_0)$ *and the line* $ax + by + c = 0$ *is*

$$D = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}.$$

· *In* $\mathbb{R}^3$*, the distance* D *between the point* $P_0 = (x_0, y_0, z_0)$ *and the plane* $ax + by + cz + d = 0$ *is*

$$D = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Proof: We prove the first formula. The second one can be proved similarly. Let $Q = (x_1, y_1)$ be any point on the line, and let $\mathbf{n} = (a, b)$ be a normal vector to the line that is positioned with its initial point at Q. It is now evident that the distance D between P_0 and the line is simply the length (or norm) of the orthogonal projection of the vector $\overrightarrow{QP_0}$ on $\mathbf{n}$, which is

$$D = \left\| \text{proj}_{\mathbf{n}} \overrightarrow{QP_0} \right\| = \frac{\left| \overrightarrow{QP_0} \cdot \mathbf{n} \right|}{\|\mathbf{n}\|}.$$

Since $\overrightarrow{QP_0} = (x_0 - x_1, y_0 - y_1)$, $\overrightarrow{QP_0} \cdot \mathbf{n} = a\,(x_0 - x_1) + b\,(y_0 - y_1)$, and $\|\mathbf{n}\| = \sqrt{a^2 + b^2}$, we find that

$$D = \frac{|ax_0 + by_0 - (ax_1 + by_1)|}{\sqrt{a^2 + b^2}}.$$

But $ax_1 + by_1 + c = 0$ implies that $c = -\,(ax_1 + by_1)$, and we are done.■

The distance posed in the third problem above can be obtained by determining a point on one plane and find its distance to the second plane. For example, the planes

$$x + 2y - 2z = 3 \quad \text{and} \quad 2x + 4y - 4z = 7$$

are parallel since their normals, $(1, 2, -2)$ and $(2, 4, -4)$, are parallel vectors. To find the distance D between the planes, we set $y = z = 0$ in the equation $x + 2y - 2z = 3$, and obtain the point $P_0 = (3, 0, 0)$ in this plane. Thus, the distance between P_0 and the plane $2x + 4y - 4z = 7$ is

$$D = \frac{|(2)\,(3) + (4)\,(0) + (-4)\,(0) - 7|}{\sqrt{2^2 + 4^2 + (-4)^2}} = \frac{1}{6}$$

units.

3.4. The Geometry of Linear Systems

In this section we use parametric and vector methods to study general systems of linear equations. This work enables us to interpret solution sets of linear systems with n unknowns as geometric objects in $\mathbb{R}^n$ and we interpret solution sets of linear systems with two and three unknowns as points, lines, and planes in $\mathbb{R}^2$ and $\mathbb{R}^3$.

In the last section we have obtained equations of lines and planes that are determined by a point and a normal vector. However, there are other useful ways of specifying lines and planes. For example, a unique line in $\mathbb{R}^2$ or $\mathbb{R}^3$ is determined by a point $\mathbf{x}_0$ on the line and a nonzero vector $\mathbf{v}$ parallel to the line, and a unique plane in $\mathbb{R}^3$ is determined by a point $\mathbf{x}_0$ in the plane and two noncollinear vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ parallel to the plane.

We start by deriving an equation for the line L that contains the point $\mathbf{x}_0$ and is parallel to $\mathbf{v}$. If $\mathbf{x}$ is a general point on such a line, then the vector $\mathbf{x} - \mathbf{x}_0$ is one scalar multiple of $\mathbf{v}$, say

$$\mathbf{x} - \mathbf{x}_0 = t\mathbf{v} \Rightarrow \mathbf{x} = \mathbf{x}_0 + t\mathbf{v}.$$

As the variable t, called **parameter**, varies from $-\infty$ to ∞, the point $\mathbf{x}$ traces out the line L.

- **Line equation**

  Let L be a line in $\mathbb{R}^2$ or $\mathbb{R}^3$ that contains the point $\mathbf{x}_0$ and is parallel to the nonzero vector $\mathbf{v}$. Then the equation of L is

  $$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}.$$

  If $\mathbf{x}_0 = \mathbf{0}$, then the line passes through the origin and the equation has the form $\mathbf{x} = t\mathbf{v}$.

Similarly, we have the following results for planes in $\mathbb{R}^3$.

- **Plane equation**

  Let W be a plane in $\mathbb{R}^3$ that contains the point $\mathbf{x}_0$ and is parallel to the noncollinear vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. Then an equation of W is

  $$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2.$$

  If $\mathbf{x}_0 = \mathbf{0}$, then the plane passes through the origin and the equation has the form $\mathbf{x} = t_1\mathbf{v}_1 + t_2\mathbf{v}_2$.

Example:

(a) Let L be the line in $\mathbb{R}^2$ passes through the origin and is parallel to $\mathbf{v} = (-2, 3)$. Then

$$(x, y) = \mathbf{x} = t\mathbf{v} = t\,(-2, 3) = (-2t, 3t)\,,$$

so $x = -2t$ and $y = 3t$.

(b) Let L be the line in $\mathbb{R}^3$ passes through the point $P_0 = (1, 2, -3)$ and is parallel to $\mathbf{v} = (4, -5, 1)$. Then

$$(x, y, z) = \mathbf{x} = \mathbf{x}_0 + t\mathbf{v} = (1, 2, -3) + (4t, -5t, t) = (4t + 1, -5t + 3, t - 3)\,,$$

so $x = 4t + 1$, $y = -5t + 2$, $z = t - 3$.

(c) Let W be the line in $\mathbb{R}^4$ passes through the point $\mathbf{x}_0 = (2, -1, 0, 3)$ and is parallel to both $\mathbf{v}_1 = (1, 5, 2, -4)$ and $\mathbf{v}_2 = (0, 7, -8, 6)$. Then

$$(x_1, x_2, x_3, x_4) = \mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2 = (2, -1, 0, 3) + (t_1, 5t_1, 2t_1, -4t_1) + (0, 7t_2, -8t_2, 6t_2),$$

so $x_1 = t_1 + 2$, $x_2 = 5t_1 + 7t_2 - 1$, $x_3 = 2t_1 - 8t_2$, $x_4 = -4t_1 + 6t_2 + 3$.■

If $\mathbf{x}_0$ and $\mathbf{x}_1$ are distinct points in $\mathbb{R}^n$, then the line determined by these points is parallel to the vector $\mathbf{v} = \mathbf{x}_1 - \mathbf{x}_0$. Thus,

$$\mathbf{x} = \mathbf{x}_0 + t(\mathbf{x}_1 - \mathbf{x}_0) \Rightarrow \mathbf{x} = (1 - t)\,\mathbf{x}_0 + t\mathbf{x}_1$$

are **two-point vector equations** of a line in $\mathbb{R}^n$.

Example: Consider the line $\mathbb{R}^2$ passes through the points $P_1 = (0, 7)$ and $P_2 = (5, 0)$. Then

$$\mathbf{v} = \mathbf{x}_1 - \mathbf{x}_0 = (5, 0) - (0, 7) = (5, -7)$$

and

$$\mathbf{x} = (0, 7) + t(5, -7) = (5t, 7 - 7t),$$

so $x = 5t$ and $y = 7 - 7t$, and we have the equation of the line as $7x + 5y = 35$.■

Our next objective is to show how to express linear equations and linear systems in dot product notation.

Recall that a linear equation in the variables $x_1, x_2, \ldots, x_n$ has the form

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

and the corresponding homogeneous equation is

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0,$$

where $a_1, a_2, \ldots, a_n$ are not all zero. If we write $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, then the linear equation and homogeneous equation can be written as $\mathbf{a} \cdot \mathbf{x} = b$ and $\mathbf{a} \cdot \mathbf{x} = \mathbf{0}$, respectively. In general, if

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

then

$$\begin{aligned} \mathbf{r}_1 \cdot \mathbf{x} &= \mathbf{0} \\ \mathbf{r}_2 \cdot \mathbf{x} &= \mathbf{0} \\ &\vdots \\ \mathbf{r}_m \cdot \mathbf{x} &= \mathbf{0} \end{aligned}$$

where $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ are the successive row vectors of the coefficient matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

As a result we note that if A is an $m \times n$ matrix, then the solution set of the homogeneous linear system $Ax = \mathbf{0}$ consists of all vectors in $\mathbb{R}^n$ that are orthogonal to every row vector of A.

3.5. Cross Product

In this section we define a type of vector multiplication that produces another vector but is applicable only to vectors in $\mathbb{R}^3$.

- **Cross product**

  *If $\mathbf{u} = (u_1, u_2, u_3)$ and $\mathbf{v} = (v_1, v_2, v_3)$ are vectors in $\mathbb{R}^3$, then the **cross product** $\mathbf{u} \times \mathbf{v}$ of $\mathbf{u}$ and $\mathbf{v}$ is defined by*

$$\mathbf{u} \times \mathbf{v} = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1),$$

  or, in determinant notation,

$$\mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}, - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \right).$$

For example, if $\mathbf{u} = (1, 2, -2)$ and $\mathbf{v} = (3, 0, 1)$, then

$$\mathbf{u} \times \mathbf{v} = \left(\begin{vmatrix} 2 & -2 \\ 0 & 1 \end{vmatrix}, - \begin{vmatrix} 1 & -2 \\ 3 & 1 \end{vmatrix}, \begin{vmatrix} 1 & 2 \\ 3 & 0 \end{vmatrix} \right) = (2, -7, -6).$$

The following result gives some important relationships between the dot and the cross product and also shows that $\mathbf{u} \times \mathbf{v}$ is orthogonal to both $\mathbf{u}$ and $\mathbf{v}$.

- **Cross product and dot product**

  If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^3$, *then*

  (a) $\mathbf{u}\cdot(\mathbf{u}\times\mathbf{v})=0$

  (b) $\mathbf{v}\cdot(\mathbf{u}\times\mathbf{v})=0$

  (c) $\|\mathbf{u}\times\mathbf{v}\|^2=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-(\mathbf{u}\cdot\mathbf{v})^2$

  (d) $\mathbf{u}\times(\mathbf{v}\times\mathbf{w})=(\mathbf{u}\cdot\mathbf{w})\cdot\mathbf{v}-(\mathbf{u}\cdot\mathbf{v})\cdot\mathbf{w}$

  (e) $(\mathbf{u}\times\mathbf{v})\times\mathbf{w}=(\mathbf{u}\cdot\mathbf{w})\cdot\mathbf{v}-(\mathbf{v}\cdot\mathbf{w})\cdot\mathbf{u}$

The main arithmetic properties of the cross product are listed in the next result.

- **Arithmetic properties of cross product**

  If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^3$, *and if* k *is a scalar, then*

  (a) $\mathbf{u}\times\mathbf{v}=-(\mathbf{v}\times\mathbf{u})$

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

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

  (d) $k(\mathbf{u}\times\mathbf{v})=(k\mathbf{u})\times\mathbf{v}=\mathbf{u}\times(k\mathbf{v})$

  (e) $\mathbf{u}\times\mathbf{0}=\mathbf{0}\times\mathbf{u}=\mathbf{0}$

  (f) $\mathbf{u}\times\mathbf{u}=\mathbf{0}$

Example: We have

$$\mathbf{i}\times\mathbf{j}=(1,0,0)\times(0,1,0)=\left(\begin{vmatrix}0&0\\1&0\end{vmatrix},-\begin{vmatrix}1&0\\0&0\end{vmatrix},\begin{vmatrix}1&0\\0&1\end{vmatrix}\right)=(0,0,1)=\mathbf{k},$$

and we have

$$\begin{array}{ccc}\mathbf{i}\times\mathbf{i}=\mathbf{0} & \mathbf{j}\times\mathbf{j}=\mathbf{0} & \mathbf{k}\times\mathbf{k}=\mathbf{0}\\ \mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{j}\times\mathbf{k}=\mathbf{i} & \mathbf{k}\times\mathbf{i}=\mathbf{j}\\ \mathbf{j}\times\mathbf{i}=-\mathbf{k} & \mathbf{k}\times\mathbf{j}=-\mathbf{i} & \mathbf{i}\times\mathbf{k}=-\mathbf{j}\end{array}$$

by the similar argument.■

It is also worth noting that a cross product can be represented in the form

$$\mathbf{u}\times\mathbf{v}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\u_1&u_2&u_3\\v_1&v_2&v_3\end{vmatrix}=\begin{vmatrix}u_2&u_3\\v_2&v_3\end{vmatrix}\mathbf{i}-\begin{vmatrix}u_1&u_3\\v_1&v_3\end{vmatrix}\mathbf{j}+\begin{vmatrix}u_1&u_2\\v_1&v_2\end{vmatrix}\mathbf{k}.$$

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in $\mathbb{R}^3$, then the norm of $\mathbf{u}\times\mathbf{v}$ has a useful geometric interpretation. If θ denotes the angle between $\mathbf{u}$ and $\mathbf{v}$, then $\mathbf{u}\cdot\mathbf{v}=\|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta$, so the identity $\|\mathbf{u}\times\mathbf{v}\|^2=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-(\mathbf{u}\cdot\mathbf{v})^2$ can be written as

$$\|\mathbf{u}\times\mathbf{v}\|^2=\|\mathbf{u}\|^2\|\mathbf{v}\|^2-\|\mathbf{u}\|^2\|\mathbf{v}\|^2\cos^2\theta=\|\mathbf{u}\|^2\|\mathbf{v}\|^2\left(1-\cos^2\theta\right)=\|\mathbf{u}\|^2\|\mathbf{v}\|^2\sin^2\theta.$$

Since $0\le\theta\le\pi$, it follows that $\sin\theta\ge 0$, so we have

$$\|\mathbf{u}\times\mathbf{v}\|=\|\mathbf{u}\|\,\|\mathbf{v}\|\sin\theta.$$

But $\|\mathbf{v}\| \sin\theta$ is the altitude of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$. Thus, the area A of this parallelogram is

$$A = (\text{base})\,(\text{altitude}) = \|\mathbf{u}\|\,\|\mathbf{v}\| \sin\theta = \|\mathbf{u}\times\mathbf{v}\|\,.$$

Example: Find the area of the triangle determined by the points $P_1 = (2,2,0)$, $P_2 = (-1,0,2)$, and $P_3 = (0,4,3)$.

Solution: We have

$$\overrightarrow{P_1P_2} = (-3,-2,2) \quad \text{and} \quad \overrightarrow{P_1P_3} = (-2,2,3)\,.$$

It follows that

$$\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3} = (-10,5,-10)\,,$$

and consequently

$$A = \frac{1}{2}\left\|\overrightarrow{P_1P_2} \times \overrightarrow{P_1P_3}\right\| = \frac{15}{2}$$

is the desired area.■

- **Scalar triple product**

  If $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^3$, $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w})$ *is called the* ***scalar triple product*** *of* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$.

The scalar triple product of $\mathbf{u} = (u_1,u_2,u_3)$, $\mathbf{v} = (v_1,v_2,v_3)$, and $\mathbf{w} = (w_1,w_2,w_3)$ can be calculated from the formula

$$\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}.$$

We conclude this section by noting two geometric interpretations determinants.

- **Determinants**

  · *The absolute value of the determinant*

$$\begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}$$

  is equal to the area of the parallelogram in $\mathbb{R}^2$ *determined by the vectors* $\mathbf{u} = (u_1,u_2)$ *and* $\mathbf{v} = (v_1,v_2)$.

· *The absolute value of the determinant*

$$\begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

is equal to the volume of the parallelepiped in $\mathbb{R}^3$ *determined by the vectors* $\mathbf{u} = (u_1, u_2, u_3)$, $\mathbf{v} = (v_1, v_2, v_3)$, *and* $\mathbf{w} = (w_1, w_2, w_3)$.