# CHAPTER 5 REAL AND COMPLEX NORMED, METRIC, AND INNER PRODUCT SPACES

So far, our studies have concentrated only on properties of vector spaces that follow from Definition 1.1. Spaces $\mathcal{G}^2$ and $\mathcal{G}^3$ of geometric vectors, however, have additional properties due to the fact that such vectors have lengths, dot products of vectors can be taken, and there is a distance function. In this chapter, we extend these ideas to other vector spaces. When a vector space is equipped with an inner product (the generalization of a dot product), it is called an *inner product* space; when it is equipped with a norm (the generalization of length), it is called a *normed* space; and when it is equipped with a metric (the generalization of distance), it is called a *metric* space. Norms and metrics are often defined in terms of inner products, but they can also be defined independently of inner products. To emphasize this, we discuss each type of space separately, and then show that an inner product induces a norm and a distance function so that every inner product space can be turned into a normed space and a metric space.

## §5.1 Normed Spaces

### Norms in Real Spaces

The length of a vector $\mathbf{v} = \langle v_x, v_y \rangle$ in $\mathcal{G}^2$ is

$$\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2}, \tag{5.1}$$

and the length of a vector $\mathbf{v} = \langle v_x, v_y, v_z \rangle$ in $\mathcal{G}^3$ is

$$\|\mathbf{v}\| = \sqrt{v_x^2 + v_y^2 + v_z^2}. \tag{5.2}$$

To introduce a norm into a real vector space (the generalization of length in $\mathcal{G}^2$ and $\mathcal{G}^3$), we ask what properties characterize lengths in equations 5.1 and 5.2. Convince yourself that these lengths satisfy the following properties:

$$\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|, \quad \text{where } c \text{ is a constant,} \tag{5.3a}$$

$$\|\mathbf{v}\| \ge 0, \quad \text{and } \|\mathbf{v}\| = 0 \text{ if, and only if, } \mathbf{v} = \mathbf{0}, \tag{5.3b}$$

$$\|\mathbf{u} + \mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|. \tag{5.3c}$$

The first of these requires the length of a multiple of a vector to be the absolute value of the constant times the length of the vector itself; the second requires lengths of nonzero vectors to be positive; the last is often called the **triangle inequality**. Geometrically, it requires the length of any side of a triangle to be less than the sum of the lengths of the other two sides. It is these properties that we use to define norms in other vector spaces.

**Definition 5.1** A **norm** on a real vector space $V$ is a real-valued function $\|\mathbf{v}\|$ of vectors $\mathbf{v}$ in the space that satisfies properties 5.3. A vector space equipped with a norm is called a **real normed space**.

The most common norm for the space $\mathcal{R}^n$ is the **Euclidean norm**. For a vector $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, it is defined as

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}. \tag{5.4}$$

It is straightforward to verify that the Euclidean norm satisfies properties 5.3. It is the generalization of length in $\mathcal{G}^2$ and $\mathcal{G}^3$. Equipped with the Euclidean norm, $\mathcal{R}^n$ is called **Euclidean $n$-space**, denoted by $\mathcal{E}^n$.

Many norms can often be associated with a vector space. Which is the most useful depends on the application in which the vector space arose. Here are some other norms for $\mathcal{R}^n$:

**1.** The **1-norm**,

$$\|\mathbf{v}\|_1 = |v_1| + |v_2| + \cdots + |v_n|. \tag{5.5}$$

**2.** For any integer $p \geq 1$, the $p$**-norm**

$$\|\mathbf{v}\|_p = [|v_1|^p + |v_2|^p + \cdots + |v_n|^p]^{1/p}. \tag{5.6}$$

**3.** The $\infty$**-norm** or **sup-norm**

$$\|\mathbf{v}\|_\infty = \underset{j=1,\ldots,n}{\text{maximum}} |v_j|. \tag{5.7}$$

The 1-norm and the Euclidean norm are the special cases of the $p$-norm when $p = 1$ and $p = 2$. We have used subscripts to denote these various norms in $\mathcal{R}^n$. When no subscript is present, it is assumed to be the Euclidean norm.

**Example 5.1** Verify that in the space $C^0[a, b]$ of continuous functions on the interval $a \leq x \leq b$, the maximum value of the absolute value of a function on the interval is a norm.

**Solution** It is not a trivial fact to prove, but you may recall that in your first calculus class, it was stated that a continuous function on a closed interval must attain a maximum value on the interval. This assures us that the norm is well-defined. If we denote the maximum value by

$$\|f(x)\| = \underset{a \leq x \leq b}{\text{maximum}} |f(x)|,$$

we must show that it satisfies properties 5.3. Since

$$\|cf(x)\| = \underset{a \leq x \leq b}{\text{maximum}} |cf(x)| = |c| \underset{a \leq x \leq b}{\text{maximum}} |f(x)| = |c|\, \|f(x)\|,$$

property 5.3a is satisfied. Clearly, $\|f(x)\| \geq 0$, and the only way for $\|f(x)\|$ to be equal to zero is for $f(x) \equiv 0$. In other words, we have property 5.3b. Finally,

$$\begin{aligned}\|f(x) + g(x)\| &= \underset{a \leq x \leq b}{\text{maximum}} |f(x) + g(x)| \\ &\leq \underset{a \leq x \leq b}{\text{maximum}} |f(x)| + \underset{a \leq x \leq b}{\text{maximum}} |g(x)| \\ &= \|f(x)\| + \|g(x)\|,\end{aligned}$$

the triangle inequality.•

Here is another norm for the space $C^0[a, b]$.

**Example 5.2** Verify that

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

is also a norm for $C^0[a, b]$.

**Solution** The fact that the definite integral of a continous functions exists once again assures us that the norm is well-defined. Properties 5.3 are straightforward to check:

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

$$\|f(x)\| = \int_a^b |f(x)|\, dx \geq 0, \quad \text{and} \quad \int_a^b |f(x)|\, dx = 0 \text{ if, and only if } f(x) = 0,$$

$$\begin{aligned}\|f(x) + g(x)\| &= \int_a^b |f(x) + g(x)|\, dx \leq \int_a^b [|f(x)| + |g(x)|]\, dx = \int_a^b |f(x)|\, dx + \int_a^b |g(x)|\, dx \\ &= \|f(x)\| + \|g(x)\|.\bullet\end{aligned}$$

In $\mathcal{G}^2$ and $\mathcal{G}^3$, the norm of a vector is geometric, the length of the vector as a line segment. In other vector spaces, the norm of a vector cannot be interpreted geometrically. For instance, in Example 5.2, there is no geometric interpretation for the norm of the function. It is not, for instance, the length of the graph of the function from $x = a$ to $x = b$.

When the norm of a vector is equal to 1, it is called a **unit** vector. We customarily indicate that a vector is a unit vector by placing a hat on it $\hat{\mathbf{v}}$ (just as we do for $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$). It is easy to find a unit vector in the same "direction" as any given vector $\mathbf{v}$, simply divide $\mathbf{v}$ by its norm,

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}. \tag{5.8}$$

When we have converted a vector into a unit vector in the same "direction", we say that we have **normalized** the vector. A unit vector in the direction "opposite" to $\mathbf{v}$ is

$$\hat{\mathbf{v}} = -\frac{\mathbf{v}}{\|\mathbf{v}\|}. \tag{5.9}$$

**Example 5.3** Normalize the vector (function) $f(x) = x^3 - 2x$ in $C^0[0,2]$ using the norm of Example 5.2.

**Solution** Since the norm of the function is

$$\begin{aligned}\|f(x)\| &= \int_0^2 |x^3 - 2x|\,dx = \int_0^{\sqrt{2}} (2x - x^3)\,dx + \int_{\sqrt{2}}^2 (x^3 - 2x)\,dx \\ &= \left\{x^2 - \frac{x^4}{4}\right\}_0^{\sqrt{2}} + \left\{\frac{x^4}{4} - x^2\right\}_{\sqrt{2}}^2 = 2,\end{aligned}$$

a normalized function is

$$\hat{f}(x) = \frac{1}{2}(x^3 - 2x).\bullet$$

**Norms in Complex Spaces**

Norms can also be introduced into complex vector spaces. They satisfy properties 5.3 where $|c|$ is interpreted as the modulus of $c$. For instance, the modulus of a complex number is a norm in the vector space $\mathcal{C}$ of complex numbers. The norm in $\mathcal{C}^n$ corresponding to the Euclidean norm in $\mathcal{R}^n$ is

$$\|(c_1, c_2, \ldots, c_n)\| = \sqrt{|c_1|^2 + |c_2|^2 + \cdots + |c_n|^2}, \tag{5.10}$$

where vertical bars denote moduli. The 1-norm, the $p$-norm, and the $\infty$-norm for $\mathcal{R}^n$ are also norms for $\mathcal{C}^n$, where once again absolute values must be interpreted as moduli.

**Example 5.4** Normalize the vector $\mathbf{v} = (3 + 4i, -2, 3i)$ in $\mathcal{C}^3$ using the norm in equation 5.10.

**Solution** Since the norm of the vector is

$$\|(3 + 4i, -2, 3i)\| = \sqrt{|3 + 4i|^2 + |-2|^2 + |3i|^2} = \sqrt{25 + 4 + 9} = \sqrt{38},$$

a normalized vector is

$$\hat{\mathbf{v}} = \frac{1}{\sqrt{38}}(3 + 4i, -2, 3i).\bullet$$

---

**EXERCISES 5.1**

**1.** Verify that the 1-norm and the infinity norm in $\mathcal{R}^n$ satisfy properties 5.3.

**2.** With the Euclidean norm in $\mathcal{G}^2$, tips of vectors described by the inequality $\|\mathbf{v}\| \le 1$ lie inside and on the unit circle centred at the origin. What does the equation describe with the 1-norm, the $p$-norm, and the infinity norm?

**3.** Repeat Exercise 2 with the Euclidean norm in $\mathcal{G}^3$.

**4.** (a) Show that in $P_2(x)$, the function

$$\|a + bx + cx^2\| = \sqrt{a^2 + b^2 + c^2},$$

is a norm.

(b) Find a unit vector (function) in the same "direction" as $p(x) = 1 - 2x + 3x^2$.

**5.** An important space in control theory is the space of complex, rational functions that do not have poles on the unit circle $|z| = 1$. Show that

$$\|f(z)\| = \left[\frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2 d\theta\right]^{1/2}$$

is well-defined and satisfies properties 5.3a,b.

**Answers**

**2.** Inside and on the square $|v_x| + |v_y| = 1$; inside and on the curve $|v_x|^p + |v_y|^p = 1$; the square $-1 \le v_x, v_y \le 1$

**3.** Inside and on the cube $|v_x| + |v_y| + |v_z| = 1$; inside and on the surface $|v_x|^p + |v_y|^p + |v_z|^p = 1$; the cube $-1 \le v_x, v_y, v_z \le 1$

**4.**(b) $(1 - 2x + 3x^2)/\sqrt{14}$

## §5.2 Metric Spaces

When $\mathbf{u} = \langle u_x, u_y, u_z\rangle$ and $\mathbf{v} = \langle v_x, v_y, v_z\rangle$ are vectors in $\mathcal{G}^3$, the length of $\mathbf{u} - \mathbf{v}$ is

$$\|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2 + (u_z - v_z)^2}. \tag{5.11}$$

It can be interpreted as a distance in $\mathcal{G}^3$, the distance between the points $(u_x, u_y, u_z)$ and $(v_x, v_y, v_z)$ at the tips of the vectors. We also say that it is the distance between the vectors themselves. In other words, if we denote the distance between two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathcal{G}^3$ by $\mathrm{d}(\mathbf{u}, \mathbf{v})$, then

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) = \sqrt{(u_x - v_x)^2 + (u_y - v_y)^2 + (u_z - v_z)^2}. \tag{5.12}$$

Distances can be defined in many sets of entities that may, or may not, be vector spaces, and we will illustrate some examples shortly. But first we ask what are the properties that characterize distance. Convince yourself that the distance function $\mathrm{d}(\mathbf{u}, \mathbf{v})$ defined by equation 5.12 satisfies the following properties for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$:

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) \geq 0 \quad \text{and } \mathrm{d}(\mathbf{u}, \mathbf{v}) = 0 \text{ if, and only if, } \mathbf{u} = \mathbf{v}, \tag{5.13a}$$

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) = \mathrm{d}(\mathbf{v}, \mathbf{u}), \tag{5.13b}$$

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) \leq \mathrm{d}(\mathbf{u}, \mathbf{w}) + \mathrm{d}(\mathbf{w}, \mathbf{v}). \tag{5.13c}$$

The first of these requires the distance between unequal vectors to be positive; the second requires the distance function to be symmetric in its arguments; and the last is the triangle inequality that we saw for norms, but now in terms of distances. It is these properties that we use to define a *metric* in other contexts.

**Definition 5.2** A **distance function** or **metric**, $\mathrm{d}(\mathbf{u}, \mathbf{v})$ on elements $\mathbf{u}$ and $\mathbf{v}$ in a set $S$ is a real-valued function of pairs of elements that satisfies properties 5.13. The set is said to be equipped with a distance function or a metric.

Obviously, our intention is to introduce a distance function, or metric, into vector spaces, but it is important to realize that distances can be defined in sets other than vector spaces. In other words, although we have used vector notation to represent elements in Definition 5.2, these elements need not be vectors. Here are some examples, the first two of which are metrics in sets that are not vector spaces.

**Example 5.5** Let $S$ be the set of $n$-letter words in a $k$-character alphabet. Denote words as $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, where the $v_j$ are letters from the alphabet. Define the distance between two words $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ as the number of places in which letters are different; that is,

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n} x_j, \quad \text{where } x_j = \begin{cases} 0, & \text{if } u_j = v_j \\ 1, & \text{if } u_j \neq v_j. \end{cases}$$

Show that $\mathrm{d}(\mathbf{u}, \mathbf{v})$ satisfies properties 5.13.

**Solution** Properties 5.13a,b are clear. Suppose $\mathbf{u} = (u_1, u_2, \ldots, u_n)$, $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ are words in $S$, and consider the $j^{\text{rh}}$ entries. If $u_j = v_j$, then this entry contributes 0 to $\mathrm{d}(\mathbf{u}, \mathbf{v})$. The $j^{\text{th}}$ entries in $\mathrm{d}(\mathbf{u}, \mathbf{w}) + \mathrm{d}(\mathbf{w}, \mathbf{v})$ contribute either 0 or 2 depending on whether $w_j = u_j$ or $w_j \neq u_j$. If $u_j \neq v_j$, then this entry contributes 1 to $\mathrm{d}(\mathbf{u}, \mathbf{v})$. The $j^{\text{th}}$ entries in $\mathrm{d}(\mathbf{u}, \mathbf{w}) + \mathrm{d}(\mathbf{w}, \mathbf{v})$ contribute 1 when $w_j = u_j$ or $w_j = v_j$, or 2 when $w_j \neq u_j$ and $w_j \neq v_j$. Thus, the contribution of the $j^{\text{th}}$ entries of $\mathrm{d}(\mathbf{u}, \mathbf{w}) + \mathrm{d}(\mathbf{w}, \mathbf{v})$ is greater than or equal to that of $\mathrm{d}(\mathbf{u}, \mathbf{v})$. Since this is true for all $n$ entries, property 5.13c is verified.●

**Example 5.6** Let $V$ be a metric space, a vector space with a metric. Does the metric on $V$, endow any subset $S$ of $V$ with a metric? Would you call $S$ a metric space?

**Solution** The metric on $V$ certainly induces the same distances between vectors in $S$, so that $S$ is endowed with a metric. But $S$ is not defined to be a subspace of $V$, so that it would imprudent to call $S$ a metric space (although this is often done).•

**Example 5.7** Using the norm of Example 5.2, we can define a metric for $C^0[0,1]$ to be

$$\mathrm{d}\big(f(x), g(x)\big) = \|f(x) - g(x)\| = \int_0^1 |f(x) - g(x)|\, dx.$$

Find $\mathrm{d}(1 + x^2, 3x - x^3)$.

**Solution** Using the definition of the metric

$$\mathrm{d}(1 + x^2, 3x - x^3) = \int_0^1 |(1 + x^2) - (3x - x^3)|\, dx = \int_0^1 |-2x + x^2 - x^3|\, dx.$$

Since $-2x + x^2 - x^3$ is always negative between $x = 0$ and $x = 1$, we can write that

$$\mathrm{d}(1 + x^2, 3x - x^3) = \int_0^1 (2x - x^2 + x^3)\, dx = \left\{x^2 - \frac{x^3}{3} + \frac{x^4}{4}\right\}_0^1 = \frac{11}{12}.\bullet$$

The metric in $\mathcal{R}^n$ corresponding to metric 5.12 in $\mathcal{G}^3$ is

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) = \sqrt{(u_1 - v_1)^2 + (u_1 - v_2)^2 + \cdots + (u_n - v_n)^2}. \tag{5.14}$$

It is interpreted as the distance between the vectors in $\mathcal{R}^n$. A distance function for vectors $\mathbf{z}$ and $\mathbf{w}$ in $\mathcal{C}^n$ with complex components $(z_1, z_2, \ldots, z_n)$ and $(w_1, w_2, \ldots, w_n)$, respectively, is

$$\mathrm{d}(\mathbf{z}, \mathbf{w}) = \sqrt{|z_1 - w_1|^2 + |z_2 - w_2|^2 + \cdots + |z_n - w_n|^2}, \tag{5.15}$$

where vertical bars are moduli.

---

## EXERCISES 5.2

**1.** Let $S$ be the set of vectors in $\mathcal{G}^2$ which have their tips on the unit circle centred at the origin. Let $\mathrm{d}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$. Is this a metric on $S$? Would you call $S$ a metric subspace of $\mathcal{G}^2$?

**2.** Let $S$ be a finite collection of $n$ objects denoted by $x_j$, $j = 1, \ldots, n$. Show that the function

$$\mathrm{d}(x_j, x_k) = \begin{cases} 0, & \text{if } j = k \\ 1, & \text{if } j = k \end{cases}$$

defines a metric on $S$. Is it possible for $S$ to be a vector space?

**3.** When $n = 1$ in equation 5.15, does the distance function reduce to what you would expect?

### Answers

**1.** Yes, No **2.** No **3.** Yes

## §5.3 Inner Product Spaces

The dot or scalar product of two vectors $\mathbf{u} = u_x\hat{\mathbf{i}} + u_y\hat{\mathbf{j}} + u_z\hat{\mathbf{k}}$ and $\mathbf{v} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}$ in $\mathcal{G}^3$ is defined as

$$\mathbf{u} \cdot \mathbf{v} = u_x v_x + u_y v_y + u_z v_z. \tag{5.16}$$

This is not the only such product for $\mathcal{G}^3$, but it is the usual one. It is called the **standard dot product** or **standard scalar product**. In this section, we generalization the dot product to define what are called *inner* products in general vector spaces. Because they are defined differently in real, as opposed to complex, vector spaces, we consider these spaces separately.

### Real Inner Product Spaces

The notation that we use for the inner product of two vectors in more abstract vector spaces is $(\mathbf{u}, \mathbf{v})$. Although parentheses also denote scalar components of vectors, no confusion can arise since scalar components of vectors are scalars, and entries in the inner product are vectors. To see how to define inner products, we note that in the new notation, the dot product in $\mathcal{G}^2$ and $\mathcal{G}^3$ satisfies the following three properties. For any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$, and any real scalars $c_1$ and $c_2$:

$$(c_1\mathbf{u} + c_2\mathbf{v}, \mathbf{w}) = c_1(\mathbf{u}, \mathbf{w}) + c_2(\mathbf{v}, \mathbf{w}), \tag{5.17a}$$
$$(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u}), \tag{5.17b}$$
$$(\mathbf{u}, \mathbf{u}) \geq 0, \quad \text{and } (\mathbf{u}, \mathbf{u}) = 0 \text{ if, and only if, } \mathbf{u} = \mathbf{0}. \tag{5.17c}$$

The first property requires the inner product to be linear in its first argument. Due to the requirement of symmetry in property 5.17b, the inner product must also be linear in its second argument,

$$(\mathbf{u}, c_1\mathbf{v} + c_2\mathbf{w}) = c_1(\mathbf{u}, \mathbf{v}) + c_2(\mathbf{u}, \mathbf{w}).$$

It is these properties of the dot product that we use to define inner products in other vector spaces.

**Definition 5.3** If $V$ is a real vector space, then an **inner product** on $V$ is any real-valued function $(\mathbf{u}, \mathbf{v})$ of pairs of vectors in $V$ that satisfies conditions 5.17. A vector space endowed with an inner product is called an **inner product space**.

Obviously, $\mathcal{G}^2$ and $\mathcal{G}^3$ are inner product spaces, as is $\mathcal{R}^n$ with the inner product

$$(\mathbf{x}, \mathbf{y}) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \tag{5.18}$$

This is called the **standard inner product** for $\mathcal{R}^n$. Here are further examples of inner product spaces.

**Example 5.8** Show that when $u(x) = u_0 + u_1 x + u_2 x^2$ and $v(x) = v_0 + v_1 x + v_2 x^2$ are vectors in the space $P_2(x)$ of real polynomials of degree less than or equal to two, the function

$$(u, v) = u_0 v_0 + u_1 v_1 + u_2 v_2$$

is an inner product.

**Solution** If $w(x) = w_0 + w_1 x + w_2 x^2$ is a third vector, then

$$\begin{aligned}
(c_1 u + c_2 v, w) &= \big((c_1 u_0 + c_2 v_0) + (c_1 u_1 + c_2 v_1)x + (c_1 u_2 + c_2 v_2)x^2, w_0 + w_1 x + w_2 x^2\big) \\
&= (c_1 u_0 + c_2 v_0)w_0 + (c_1 u_1 + c_2 v_1)w_1 + (c_1 u_2 + c_2 v_2)w_2 \\
&= c_1(u_0 w_0 + u_1 w_1 + u_2 w_2) + c_2(v_0 w_0 + v_1 w_1 + v_2 w_2) \\
&= c_1(u, w) + c_2(v, w).
\end{aligned}$$

This verifies property 5.17a. The remaining properties 5.17b,c can be verified in a similar way.●

**Example 5.9** Show that when $u(x)$ and $v(x)$ are vectors in the space $C^0[a,b]$ of continuous functions on the interval $a \le x \le b$, the function

$$(u,v) = \int_a^b u(x)v(x)\,dx$$

is an inner product.

**Solution** If $w(x)$ is a third vector, then

$$(c_1u + c_2v, w) = \int_a^b (c_1u + c_2v)w\,dx = c_1\int_a^b uv\,dx + c_2\int_a^b vw\,dx = c_1(u,w) + c_2(v,w).$$

This confirms property 5.17a. The remaining properties 5.17b,c can be verified in a similar way.●

### Complex Inner Product Spaces

Inner products on complex vector spaces are defined as follows.

**Definition 5.4** An **inner product** on a complex vector space $V$ is a complex-valued function $(\mathbf{u},\mathbf{v})$ on pairs of vectors in $V$ such that for any three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in $V$, and any two complex scalars $c_1$ and $c_2$:

$$(c_1\mathbf{u} + c_2\mathbf{v}, \mathbf{w}) = \overline{c_1}(\mathbf{u},\mathbf{w}) + \overline{c_2}(\mathbf{v},\mathbf{w}), \tag{5.19a}$$

$$(\mathbf{w}, c_1\mathbf{u} + c_2\mathbf{v}) = c_1(\mathbf{w},\mathbf{u}) + c_2(\mathbf{w},\mathbf{v}), \tag{5.19b}$$

$$(\mathbf{u},\mathbf{v}) = \overline{(\mathbf{v},\mathbf{u})}, \tag{5.19c}$$

$$(\mathbf{u},\mathbf{u}) > 0, \quad \text{and } (\mathbf{u},\mathbf{u}) = 0 \text{ if, and only if, } \mathbf{u} = \mathbf{0}. \tag{5.19d}$$

A complex space equipped with an inner product is called a **complex inner product space**.

The overline on $\overline{c_1}$ denotes the complex conjugate of $c_1$. We say that the inner product is linear in its second argument, conjugate linear in its first argument, and conjugate symmetric. Although we have listed property 5.19b as a separate requirement for a complex inner product, in actual fact, it is a consequence of properties 5.19a,c (see Exercise 1).

Most mathematics texts replace properties 5.19a,b with

$$(c_1\mathbf{u} + c_2\mathbf{v}, \mathbf{w}) = c_1(\mathbf{u},\mathbf{w}) + c_2(\mathbf{v},\mathbf{w}), \tag{5.20a}$$

$$(\mathbf{w}, c_1\mathbf{u} + c_2\mathbf{v}) = \overline{c_1}(\mathbf{w},\mathbf{u}) + \overline{c_2}(\mathbf{w},\mathbf{v}). \tag{5.20b}$$

The inner product becomes linear in its first argument and conjugate linear in the second. Physicists use properties 5.19a,b. With either choice, the ensuing theory unfolds in much the same way. We will continue with the choice of physicists.

Corresponding to standard inner product 5.18 in $\mathcal{R}^n$, the **standard inner product** for two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathcal{C}^n$ is

$$(\mathbf{u},\mathbf{v}) = \overline{u_1}v_1 + \overline{u_2}v_2 + \cdots + \overline{u_n}v_n. \tag{5.21}$$

(For mathematicians requiring properties 5.20, definition 5.21 is replaced by

$$(\mathbf{u},\mathbf{v}) = u_1\overline{v_1} + u_2\overline{v_2} + \cdots + u_n\overline{v_n}.) \tag{5.22}$$

**Example 5.10** What is the standard inner product of the vectors $(1+2i, -3+i, 4-3i)$ and $(i, 2-3i, 4+5i)$ in $\mathcal{C}^3$?

**Solution** Using equation 5.21,

$$\big((1+2i,-3+i,4-3i),(i,2-3i,4+5i)\big) = (1-2i)(i)+(-3-i)(2-3i)+(4+3i)(4+5i) = \\ = -6+40i.\bullet$$

**Orthogonal Vectors**

In $\mathcal{G}^3$, the natural basis $\{\hat{\mathbf{i}},\hat{\mathbf{j}},\hat{\mathbf{k}}\}$ has a special property not shared by bases in spaces that do not have an inner product. The vectors are mutually perpendicular and all have length one. We know that the test for perpendicularity of vectors in this space is that their inner (dot) product is zero. In the following definition, we extend the idea of perpendicularity to all real and complex, inner product spaces, but give it a new name.

**Definition 5.5** Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are said to be **orthogonal** if their inner product is zero,

$$(\mathbf{u},\mathbf{v}) = 0. \tag{5.23}$$

A set of nonzero vectors $\{\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_m\}$ is said to be orthogonal if every vector is orthogonal to every other vector; that is, whenever $j \neq k$,

$$(\mathbf{v}_j,\mathbf{v}_k) = 0. \tag{5.24}$$

Orthogonal is the more general terminology for perpendicular in spaces where there is no geometric interpretation of what it means for vectors to be perpendicular.

**Example 5.11** Show that the vectors $p_1(x) = 1-x+3x^2$ and $p_2(x) = -1+2x+x^2$ are orthogonal in $P_2(x)$ with respect to the inner product in Example 5.8.

**Solution** Since the inner product of the vectors is

$$(1-x+3x^2,-1+2x+x^2) = (1)(-1)+(-1)(2)+(3)(1) = 0,$$

the vectors are orthogonal.●

**Example 5.12** The set of functions $\left\{\sin\dfrac{n\pi x}{L}\right\}$, where $n \geq 1$ is an integer, are used in finding the Fourier sine series for odd functions of period $2L$. The reason for this is that the functions are orthogonal on the interval $0 \leq x \leq L$ with respect to the inner product

$$\big(f(x),g(x)\big) = \int_0^L f(x)g(x)\,dx.$$

Verify this.

**Solution** The inner product of two of the functions is

$$\begin{aligned}\left(\sin\frac{n\pi x}{L},\sin\frac{m\pi x}{L}\right) &= \int_0^L \sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}dx \\ &= \int_0^L \left[-\cos\frac{(n+m)\pi x}{L}+\cos\frac{(n-m)\pi x}{L}\right]dx \\ &= \frac{1}{2}\left\{-\frac{L}{(n+m)\pi}\sin\frac{(n+m)\pi x}{L}+\frac{L}{(n-m)\pi}\sin\frac{(n-m)\pi x}{L}\right\}_0^L = 0.\bullet\end{aligned}$$

**Example 5.13** Are the vectors $(3+i,2-4i,5i)$ and $(2+2i,i,3)$ orthogonal with respect to the standard inner product 5.21 on $\mathcal{C}^3$?

**Solution** Since the inner product of the vectors is

$$\big((3+i,2-4i,5i),(2+2i,i,3)\big) = (3-i)(2+2i)+(2+4i)(i)+(-5i)(3) = 4-9i,$$

the vectors are not orthogonal.•

**Example 5.14** Show that the vectors $(1+i, 3-2i, 6)$ and $(2-i, 3+5i, -3i)$ are orthogonal with respect to the standard inner product 5.21 in $\mathcal{C}^3$.

**Solution** Since the inner product of the vectors is

$$\big((1+i, 3-2i, 6), (2-i, 3+5i, -3i)\big) = (1-i)(2-i) + (3+2i)(3+5i) + 6(-3i) = 0,$$

the vectors are orthogonal.•

According to the following theorem, orthogonal vectors are linearly independent.

**Theorem 5.1** If a set of vectors in an inner product space is orthogonal, then the set is linearly independent.

**Proof** Suppose the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ of vectors is orthogonal, and consider finding constants $c_1, \ldots, c_m$ so that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}.$$

If we take inner products of vectors on both sides of this equation with $\mathbf{v}_k$, we obtain

$$(\mathbf{v}_k, c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m) = 0.$$

Because the inner product is linear in its second argument, we can write that

$$c_1(\mathbf{v}_k, \mathbf{v}_1) + c_2(\mathbf{v}_k, \mathbf{v}_2) + \cdots + c_m(\mathbf{v}_k, \mathbf{v}_m) = 0.$$

Since the vectors are orthogonal, all inner products are zero, except the $k^{\text{th}}$ one,

$$c_k(\mathbf{v}_k, \mathbf{v}_k) = 0.$$

Property 5.17c (or 5.19d) implies that $(\mathbf{v}_k, \mathbf{v}_k) \neq 0$. Hence, $c_k = 0$, and this is true for $k = 1, \ldots, m$. The vectors are therefore linearly independent.∎

---

## EXERCISES 5.3

**1.** Verify that properties 5.19a,c imply property 5.19b.

**2.** Verify that inner product 5.21 satisfies properties 5.19.

**In Exercises 3–4 use the inner product of Example 5.9 to determine whether the functions are orthogonal on the interval** $0 \leq x \leq 1$**.**

**3.** $1-x$, $2+3x$

**4.** $3-5x$, $x^2-x^3$

**In Exercises 5–6 determine whether the pair of vectors is orthogonal with respect to the standard inner product in** $\mathcal{C}^2$**.**

**5.** $(2+4i, 3-2i)$, $(i, -1+2i)$

**6.** $(1-5i, 1+i)$, $(3+i, 9-7i)$

**In Exercises 7–8 determine whether the pair of vectors is orthogonal with respect to the standard inner product in** $\mathcal{C}^3$**.**

**7.** $(2+4i, 3-2i, 1+2i)$, $(i, -1+2i, 3)$

**8.** $(1-5i, 1+i, i)$, $(3+i, 9-7i, 1)$

**9.** Determine the value for the constant $a$ so that the functions $f(x) = 4x^2 - 1$ and $g(x) = 1 + ax$ are orthogonal on the interval $1 \leq x \leq 2$ with the inner product of Example 5.9.

**10.** Let $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_n)$ be vectors in $\mathcal{R}^n$. Verify that for any nonnegative constants $a_1, \ldots, a_n$, but not all zero,

$$(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n} a_j u_j v_j$$

is an inner product. Is there a reason for demanding that the constants not be negative?

**11.** Show that if $w(x) \ge 0$ is a continuous function on the interval $a \le x \le b$, and not identically equal to zero, then

$$\int_a^b w(x)f(x)g(x)\,dx$$

is an inner product on $C^0[a,b]$.

**12.** With the inner product of two functions $f(x)$ and $g(x)$

$$\int_0^{2L} f(x)g(x)\,dx,$$

show that the set of functions $\left\{1, \cos\frac{n\pi x}{L}, \sin\frac{n\pi x}{L}\right\}$ ($n \ge 1$ an integer) is orthogonal. These are the functions forming the basis for Fourier series of $2L$-periodic functions.

**13.** With the inner product of Example 5.12, show that the set of functions $\left\{1, \cos\frac{n\pi x}{L}\right\}$ ($n \ge 1$ an integer) is orthogonal. These are the functions forming the basis for Fourier cosine series of even, $2L$-periodic functions.

**14.** (a) The $n^{\text{th}}$-degree Legendre polynomial is defined as

$$p_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k(2n-2k)!}{2^n k!(n-2k)!(n-k)!} x^{n-2k}, \quad n \ge 0,$$

where $\lfloor n \rfloor$ is the floor (or greatest integer) function. Find the first four polynomials.

(b) Show that if the inner product of two functions $f(x)$ and $g(x)$ is defined as

$$\int_{-1}^{1} f(x)g(x)\,dx,$$

the four polynomials in part (a) are orthogonal.

(c) All Legendre polynomials are in fact orthogonal on the interval $-1 \le x \le 1$. Find a set of functions that are orthogonal on the interval $a \le x \le b$ with respect to the same inner product, but with limits of integration being $a$ and $b$.

**15.** (a) The $n^{\text{th}}$-degree Hermite polynomial is defined as

$$h_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad n \ge 0.$$

Find the first four polynomials.

(b) Show that if the inner product of two functions $f(x)$ and $g(x)$ is defined as

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)g(x)\,dx,$$

the four polynomials in part (a) are orthogonal.

**16.** (a) The $n^{\text{th}}$-degree Chebyshev polynomial is defined by the recursive formula

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \qquad n \ge 1.$$

Find $T_2(x)$ and $T_3(x)$.

(b) Show that if the inner product of two functions $f(x)$ and $g(x)$ is defined as

$$\int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{1-x^2}}\,dx,$$

the four polynomials in part (a) are orthogonal.

**17.** The trace of a real, square matrix $A = (a_{ij})_{n\times n}$ is defined to be the sum of its diagonal entries,

$$\text{Trace}(A) = \sum_{i=1}^{n} a_{ii}.$$

(a) Show that the function

$$(A, B) = \text{Trace}(A^T B)$$

is an inner product on $M_{n,n}(\mathcal{R})$.

(b) Is the natural basis for $M_{2,2}(\mathcal{R})$ orthogonal with respect to this inner product?

(c) Find all matrices in $M_{2,2}(\mathcal{R})$ orthogonal to $\begin{pmatrix} -1 & -3 \\ 5 & 8 \end{pmatrix}$.

**18.** If $p(x)$ and $q(x)$ are polynomials in the space $P_n(x)$, is the function

$$\big(p(x), q(x)\big) = \sum_{j=0}^{n} p(j)q(j)$$

an inner product?

**19.** If $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$ are matrices in the space $M_{m,n}(\mathcal{R})$, is the function

$$(A, B) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}b_{ij}$$

an inner product?

**20.** In the space $P_2(z)$ of complex polynomials of degree less than or equal to two, is the function

$$(a_0 + a_1 z + a_2 z^2, b_0 + b_1 z + b_2 z^2) = \overline{a_0}b_0 + \overline{a_1}b_1 + \overline{a_2}b_2,$$

an inner product?

**21.** Show that if $\mathbf{u}$ and $\mathbf{w}$ are vectors in a finite-dimensional, inner product space and $(\mathbf{u}, \mathbf{v}) = (\mathbf{w}, \mathbf{v})$ for every vector $\mathbf{v}$ in the space, then $\mathbf{u} = \mathbf{w}$.

**Answers**

**3.** No **4.** Yes **5.** No **6.** Yes **7.** Yes **8.** No **9.** $-50/81$

**14.**(a) $p_0(x) = 1, \quad p_1(x) = x, \quad p_2(x) = \dfrac{3x^2-1}{2}, \quad p_3(x) = \dfrac{5x^3-3x}{2}$ (c) $p_n\left(\dfrac{2x}{b-a} - \dfrac{b+a}{b-a}\right)$

**15.**(a) $1,\ 2x,\ 4x^2-2,\ 8x^3-12x$ **16.**(a) $2x^2-1,\ 4x^3-3x$ **17.**(b) Yes (c) $\begin{pmatrix} 5c-3b+8d & b \\ c & d \end{pmatrix}$

**18.** Yes **19.** Yes **20.** Yes

## §5.4 Normed and Metric Spaces from Inner Product Spaces

We have seen that $\mathcal{G}^2$ and $\mathcal{G}^3$ are inner product spaces, normed spaces, and metric spaces. The reason for this is that in these spaces, the inner product is used to define a norm, and the norm is then used to define a metric. But this is not peculiar to these spaces. In any inner product space, the inner product induces a norm and a metric, so that any inner product space can be turned into a normed space and a metric space. To be precise, suppose that $V$ is an inner product space, and we define

$$\|\mathbf{v}\| = \sqrt{(\mathbf{v}, \mathbf{v})}. \tag{5.25}$$

Ideally, we should prove that this definition of a norm satisfies properties 5.3. Properties 5.3a,b are obvious, but verification of triangle property 5.3c requires the Cauchy-Schwarz inequality contained in the following theorem.

**Theorem 5.2 Cauchy-Schwarz Inequality** If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner product space $V$, then

$$|(\mathbf{u}, \mathbf{v})|^2 \le (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}), \tag{5.26a}$$

or, when a norm is defined by equation 5.25,

$$|(\mathbf{u}, \mathbf{v})| \le \|\mathbf{u}\| \|\mathbf{v}\|. \tag{5.26b}$$

Equality holds when $\mathbf{u} = \mathbf{0}$, or $\mathbf{v} = \mathbf{0}$, or $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent. Vertical bars on the left are absolute values when $V$ is a real space, and moduli when $V$ is complex.

**Proof** The result is obviously valid if either $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$ since both sides of the inequality are equal to 0. When $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent, then $\mathbf{v} = c\mathbf{u}$, where $c$ is a nonzero constant. In this case,

$$|(\mathbf{u}, \mathbf{v})|^2 = |(\mathbf{u}, c\mathbf{u})|^2 = |c(\mathbf{u}, \mathbf{u})|^2 = |c|^2 \, |(\mathbf{u}, \mathbf{u})|^2.$$

But,

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

and therefore

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

Consider now the case when neither $\mathbf{u}$ nor $\mathbf{v}$ is the zero vector, and they are not linearly dependent. We verify the inequality when $V$ is a real space. For every scalar $t$, the vector $\mathbf{u} + t\mathbf{v}$ is nonzero, and therefore

$$\begin{aligned}
0 &< (\mathbf{u} + t\mathbf{v}, \mathbf{u} + t\mathbf{v}) \\
&= (\mathbf{u}, \mathbf{u}) + (\mathbf{u}, t\mathbf{v}) + (t\mathbf{v}, \mathbf{u}) + (t\mathbf{v}, t\mathbf{v}) \\
&= (\mathbf{u}, \mathbf{u}) + t(\mathbf{u}, \mathbf{v}) + t(\mathbf{v}, \mathbf{u}) + t^2(\mathbf{v}, \mathbf{v}) \\
&= (\mathbf{u}, \mathbf{u}) + 2t(\mathbf{u}, \mathbf{v}) + t^2(\mathbf{v}, \mathbf{v}).
\end{aligned}$$

Because this quadratic expression in $t$ is always positive, it follows that the discriminant must be negative; that is,

$$0 > 4t^2(\mathbf{u}, \mathbf{v})^2 - 4t^2(\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}) = 4t^2[(\mathbf{u}, \mathbf{v})^2 - (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v})].$$

Hence,

$$(\mathbf{u}, \mathbf{v})^2 < (\mathbf{u}, \mathbf{u})(\mathbf{v}, \mathbf{v}). \blacksquare$$

We can now verify that norm 5.25 induced by the inner product does indeed satisfy the triangle inequality.

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v},\mathbf{u}+\mathbf{v}) && \text{(by equation 5.25)}\\
&= (\mathbf{u},\mathbf{u})+(\mathbf{u},\mathbf{v})+(\mathbf{v},\mathbf{u})+(\mathbf{v},\mathbf{v}) && \text{(by properties 5.17a,b)}\\
&= \|\mathbf{u}\|^2+2(\mathbf{u},\mathbf{v})+\|\mathbf{v}\|^2 && \text{(by property 5.17c)}\\
&\le \|\mathbf{u}\|^2+2\|\mathbf{u}\|\|\mathbf{v}\|+\|\mathbf{v}\|^2 && \text{(by the Cauchy-Schwarz inequality 5.26b)}\\
&= (\|\mathbf{u}\|+\|\mathbf{v}\|)^2.
\end{aligned}$$

The triangle inequality results when we take positive square roots of this equality. Thus, equation 5.25 does indeed turn an inner product space into a normed space. A metric on $V$ is

$$\mathrm{d}(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|. \tag{5.27}$$

This turns $V$ into a metric space. (You should prove that this definition of a metric satisfies properties 5.13.) In most instances, we continue to call $V$ an inner product space, rather than a normed space or a metric space. We do this because most of our discussions depend on a vector space having an inner product as well as norms (and perhaps metrics), and normed spaces (and metric spaces ) need not have inner products. Throughout the remainder of these notes, we assume that in an inner product space, norms and metrics are those induced by the inner product.

**Example 5.15** Prove that

$$\|f(x)\| = \sqrt{\int_a^b [f(x)]^2 dx}$$

defines a norm on the space of functions $C^0[a,b]$.

**Solution** Properties 5.3a,b are straightforward. But to verify the triangle inequality 5.3c directly is not easy. Instead, consider showing that

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

is an inner product on $C^0[a,b]$. It is not difficult to show that it satisfies properties 5.17, so that it is indeed an inner product. But the inner product induces the norm

$$\|f(x)\| = \sqrt{\big(f(x),f(x)\big)} = \sqrt{\int_a^b [f(x)]^2\,dx}$$

on the space, and this is what was to be proved.•

Here is a very simple example of using the Cauchy-Schwarz inequality in a problem that can be solved with multi-variable calculus, but with much more difficulty.

**Example 5.16** Find the largest and smallest values of the function $f(x,y,z)=4x-5y+2z$ on the sphere $x^2+y^2+z^2=1$.

**Solution** If we apply the Cauchy-Schwarz inequality to the vectors $\langle x,y,z\rangle$ and $\langle 4-5,2\rangle$ in $\mathcal{G}^3$, we obtain

$$|\big(\langle 4,-5,2\rangle,\langle x,y,z\rangle\big)| \le \|\langle 4,-5,2\rangle\|\,\|\langle x,y,z\rangle\|,$$

or,

$$|4x-5y+2z| \le \sqrt{16+25+4}\sqrt{x^2+y^2+z^2} = 3\sqrt{5}.$$

Hence $-3\sqrt{5} \le 4x - 5y + 2z \le 3\sqrt{5}$. Maximum and minimum values are attained when vectors $\langle x, y, z\rangle$ and $\langle 4, -5, 2\rangle$ are linearly dependent; that is when the vector from the origin to the point $(x, y, z)$ on the sphere is parallel to $\langle 4, -5, 2\rangle$. This occurs at the points $(\pm 4/(3\sqrt{5}), \mp 5/(3\sqrt{5}), \pm 2/(3\sqrt{5}))$.●

Equation 5.25 expresses the norm of a vector in terms of the inner product of the vector with itself. The following result expresses the inner product of two vectors in terms of the norms of the vectors.

**Theorem 5.3** If $\mathbf{v}$ and $\mathbf{w}$ are vectors in an inner product space, then

$$(\mathbf{v}, \mathbf{w}) = \frac{1}{2}\left[\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}\|^2 - \|\mathbf{w}\|^2\right]. \tag{5.28}$$

**Proof**

$$\begin{aligned}\|\mathbf{v}+\mathbf{w}\|^2 - \|\mathbf{v}\|^2 - \|\mathbf{w}\|^2 &= (\mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}) - (\mathbf{v}, \mathbf{v}) - (\mathbf{w}, \mathbf{w}) \\ &= (\mathbf{v}, \mathbf{v}) + (\mathbf{v}, \mathbf{w}) + (\mathbf{w}, \mathbf{v}) + (\mathbf{w}, \mathbf{w}) - (\mathbf{v}, \mathbf{v}) - (\mathbf{w}, \mathbf{w}) \\ &= 2(\mathbf{v}, \mathbf{w}).\bullet\end{aligned}$$

### Orthonormal Vectors

In $\mathcal{G}^3$, the natural basis $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ has a special property not shared by bases in spaces that do not have an inner product. The vectors are mutually perpendicular and all have length one. In Section 5.3, we generalized the concept of perpendicularity to orthogonality. Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are orthogonal if their inner product vanishes, $(\mathbf{u}, \mathbf{v}) = 0$. In the following definition, we introduce the adjective to describe orthogonal vectors that also have length one.

**Definition 5.6** A set of nonzero vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ in an inner product space is said to be **orthonormal** if every vector is orthogonal to every other vector, and each vector has unit norm. This is represented algebraically by

$$(\mathbf{v}_j, \mathbf{v}_k) = \delta_{jk}. \tag{5.29}$$

The symbol $\delta_{jk}$ is called the **Kronecker delta**; it has value zero when $j \ne k$, and value 1 when $j = k$.

Given an orthogonal set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$, it is easy to construct an orthonormal set, simply divide each vector by its norm; that is, replace $\mathbf{v}_j$ with

$$\hat{\mathbf{v}}_j = \frac{\mathbf{v}_j}{\|\mathbf{v}_j\|}. \tag{5.30}$$

We say that we have **normalized** the vectors.

**Example 5.17** In Example 5.11 of Section 5.3, we showed that the vectors $p_1(x) = 1 - x + 3x^2$ and $p_2(x) = -1 + 2x + x^2$ are orthogonal in $P_2(x)$ with respect to the inner product in Example 5.8. Normalize the vectors.

**Solution** Since squares of the norms of the polynomials are

$$\|1 - x + 3x^2\|^2 = (1)^2 + (-1)^2 + (3)^2 = 11 \quad \text{and} \quad \|-1 + 2x + x^2\|^2 = (-1)^2 + (2)^2 + (1)^2 = 6,$$

normalized vectors are

$$\frac{1}{\sqrt{11}}(1 - x + 3^2) \quad \text{and} \quad \frac{1}{\sqrt{6}}(-1 + 2x + x^2).\bullet$$

**Example 5.18** In Example 5.12 of Section 5.3, we showed that the set of functions $\left\{\sin\dfrac{n\pi x}{L}\right\}$, where $n \ge 1$ is an integer, are orthogonal on the interval $0 \le x \le L$ with respect to the inner product

$$\big(f(x), g(x)\big) = \int_0^L f(x)g(x)\,dx.$$

Normalize the functions.

**Solution** The square of the norm of the $n^{\text{th}}$ function is

$$\left\|\sin\frac{n\pi x}{L}\right\|^2 = \int_0^L \sin^2\frac{n\pi x}{L}dx = \int_0^L \frac{1}{2}\left(1-\cos\frac{2n\pi x}{L}\right)dx$$
$$= \frac{1}{2}\left\{x - \frac{L}{2n\pi}\sin\frac{2n\pi x}{L}\right\}_0^L = L.$$

Consequently, an orthonormal set is $\left\{\frac{1}{\sqrt{L}}\sin\frac{n\pi x}{L}\right\}$.●

---

**EXERCISES 5.4**

**1.** With the inner product of Example 5.18, show that the set of functions $\left\{1, \cos\frac{n\pi x}{L}\right\}$ ($n \geq 1$ an integer) is orthogonal. These are the functions forming the basis for Fourier cosine series of even, $2L$-periodic functions. Normalize the functions.

**2.** In Exercise 14 in Section 5.3, the first four Legendre polynomials were shown to be orthogonal. Normalize the functions.

**3.** In Exercise 15 in Section 5.3, the first four Hermite polynomials were shown to be orthogonal. Normalize the functions. You will need the fact that

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

**4.** In Exercise 16 in Section 5.3, the first four Chebyshev polynomials were shown to be orthogonal. Normalize the functions.

**5.** Find largest and smallest values for the function $f(x,y,z) = 3x+10y-2z$ on the sphere $x^2+y^2+z^2 = 5$.

**6.** Find largest and smallest values for the function $f(x,y,z) = 3x+2y-5z$ on the surface $x^2-2x+y^2+4y+z^2-10z+24 = 0$.

**7.** Find largest and smallest values for the function $f(x,y,z) = 2x-y+5z$ on the ellipsoid $3x^2+4y^2+7z^2 = 4$.

**8.** In $\mathcal{E}^3$, the parallelogram law states that

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 \leq 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2.$$

Verify that it is valid for norms in any inner product space (although it loses its geometric interpretation).

**9.** (a) Verify that when $\mathbf{v}$ and $\mathbf{w}$ are vectors in a real, inner product space, then

$$\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2(\mathbf{v},\mathbf{w}) + \|\mathbf{w}\|^2.$$

(b) Is this is valid in a complex, inner product space? If not, what is its replacement?

**10.** Verify that when $\mathbf{v}$ and $\mathbf{w}$ are orthogonal vectors in an inner product space, then

$$\|\mathbf{v}+\mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2.$$

**11.** In $\mathcal{E}^2$ and $\mathcal{E}^3$, the angle $\theta$ between two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is defined by

$$\cos\theta = \frac{\mathbf{v}_1\cdot\mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}.$$

We can do the same thing in any real inner product space, although there is no geometric visualization of the angle. We simply replace the dot product with the inner product

$$\cos\theta = \frac{(\mathbf{v}_1, \mathbf{v}_2)}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|}.$$

This only makes sense if the right side is between $\pm 1$. How can we be sure that this is true?

**12.** Use the Cauchy-Schwarz inequality to prove that for any $n$ real numbers $r_1, \ldots, r_n$,

$$(r_1 + r_2 + \cdots + r_n)^2 \le n(r_1^2 + \cdots + r_n^2).$$

**13.** (a) Show that for any three real numbers $r_1$, $r_2$, and $r_3$,

$$r_1 r_2 + r_2 r_3 + r_3 r_1 \le r_1^2 + r_2^2 + r_3^2.$$

(b) Show that the result in part (a) is not always true for more than three numbers.

**14.** A metric on a vector space is said to be **translation invariant** if it satisfies

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

Verify that when a metric is induced by a norm, then this property is always satisfied.

**15.** Show that when $V$ is a normed linear space, then

$$\mathrm{d}(\mathbf{u}, \mathbf{v}) = \frac{\|\mathbf{u} - \mathbf{v}\|}{1 + \|\mathbf{u} - \mathbf{v}\|}$$

is a metric. With this metric, all distances are less than one.

**Answers**

**1.** $\left\{\frac{1}{\sqrt{2L}}, \frac{1}{\sqrt{L}}\cos\frac{n\pi x}{L}, \frac{1}{\sqrt{L}}\sin\frac{n\pi x}{L}\right\}$ **2.** $\frac{1}{\sqrt{2}}$, $\frac{\sqrt{3}x}{\sqrt{2}}$, $\frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1)$, $\frac{\sqrt{7}}{2\sqrt{2}}(5x^3 - 3x)$

**3.** $\frac{1}{\pi^{1/4}}$, $\frac{\sqrt{2}x}{\pi^{1/4}}$, $\frac{2x^2 - 1}{\sqrt{2}\pi^{1/4}}$, $\frac{2x^3 - 3x}{\sqrt{3}\pi^{1/4}}$ **4.** $\frac{1}{\sqrt{\pi}}$, $\sqrt{\frac{2}{\pi}}x$, $\sqrt{\frac{2}{\pi}}(2x^2 - 1)$, $\sqrt{\frac{2}{\pi}}(4x^3 - 3x)$,

**5.** $\pm\sqrt{565}$ **6.** $-26 \pm 2\sqrt{57}$ **7.** $\pm\sqrt{433/21}$ **9.**(b) $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + (\mathbf{v}, \mathbf{w}) + (\mathbf{w}, \mathbf{v}) + \|\mathbf{w}\|^2$

**11.** The Cauchy-Schwarz inequality guarantees it.

## §5.5 Orthogonal Complements and Orthogonal Components of Vectors

In Section 1.7, we learned how to take vector components along subspaces of a vector space when the space is the direct sum of two subspaces. In this section, we specialize this operation to the situation where every vector in one subspace is orthogonal to every vector in the other subspace.

### Orthogonal Complements

The following definition describes what it means for one subspace of a vector space to be the *orthogonal complement* of another subspace.

**Definition 5.7** Let $W$ be a subspace in an inner product space $V$. The **orthogonal complement** of $W$, denoted by $W^\perp$, is the set of vectors in $V$ that are orthogonal to every vector in $W$.

According to the following theorem, $W^\perp$ is a subspace of $V$.

**Theorem 5.4** The orthogonal complement $W^\perp$ of a subspace $W$ in an inner product space $V$ is a subspace of $V$.

**Proof** Suppose that $\mathbf{u}$ and $\mathbf{v}$ are any two vectors in $W^\perp$, and $a$ is a scalar. If $\mathbf{w}$ is any vector in $W$, then

$$(\mathbf{w}, \mathbf{u} + \mathbf{v}) = (\mathbf{w}, \mathbf{u}) + (\mathbf{w}, \mathbf{v}) = 0, \qquad (\mathbf{w}, a\mathbf{u}) = a(\mathbf{w}, \mathbf{u}) = 0.$$

This shows that $\mathbf{u} + \mathbf{v}$ and $a\mathbf{u}$ are both in $W^\perp$; that is, $W^\perp$ is closed under vector addition and scalar multiplication, and must be a subspace of $V$.∎

In $\mathcal{G}^3$, the orthogonal complement of the subspace of vectors along the $x$-axis is all vectors in the $yz$-plane. The orthogonal complement of vectors $(x_1, -x_1, x_3)$ in $\mathcal{E}^3$ are vectors of the form $(w_1, w_1, 0)$. This is obvious once you see it; try deriving it.

**Example 5.19** Let $W$ be the subspace of a 4-dimensional inner product space spanned by the vectors $\mathbf{u} = (1, 1, 1, 1)$ and $\mathbf{v} = (1, 2, 3, 4)$. Find a basis for $W^\perp$.

**Solution** We need two vectors orthogonal to $\mathbf{u}$ and $\mathbf{v}$, because, if a vector is orthogonal to these vectors, it is orthogonal to every linear combination of them. If $(a, b, c, d)$ is orthogonal to $\mathbf{u}$ and $\mathbf{v}$, then

$$0 = \big((a, b, c, d), (1, 1, 1, 1)\big) = a + b + c + d, \quad 0 = \big((a, b, c, d), (1, 2, 3, 4)\big) = a + 2b + 3c + 4d.$$

These can be solved for $a$ and $b$ in terms of $c$ and $d$, $a = c + 2d$, $b = -2c - 3d$. If we set $c = 0$ and $d = 1$, we obtain the vector $(2, -3, 0, 1)$, and if we set $c = 1$ and $d = 0$, we get $(1, -2, 1, 0)$. These two vectors are a basis for $W^\perp$.•

The following result provides a superior way to find orthogonal complements.

**Theorem 5.5** The orthogonal complement of the row space of a matrix is the null space of the matrix.

**Proof** Let $A$ be an $m \times n$ matrix. A vector $\mathbf{v}$ is in the orthogonal complement of the row space of the matrix if the (standard) inner product of $\mathbf{v}$ with every row of $A$ is zero. But the product $A\mathbf{v}$ of matrices yields the $m$ inner products of $\mathbf{v}$ with the rows of $A$. Hence, $\mathbf{v}$ is orthogonal to every row of $A$ if $A\mathbf{v} = \mathbf{0}$; that is, $\mathbf{v}$ is in the null space of the matrix.∎

What this means is that to find the orthogonal complement of a subspace $W$ of a vector space, find the null space of a matrix whose rows are a basis for $W$.

**Example 5.20** Find a basis for the orthogonal complement of the subspace $W$ in Example 5.19.

**Solution** The reduced row echelon form for the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{pmatrix} \quad \text{is} \quad \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

Components $(x_1, x_2, x_3, x_4)$ of vectors in the null space satisfy

$$x_1 = x_3 + 2x_4, \qquad x_2 = -2x_3 - 3x_4.$$

Vectors in $W^\perp$ are therefore of the form

$$\begin{pmatrix} x_3 + 2x_4 \\ -2x_3 - 3x_4 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}.$$

In other words, $W^\perp$ is spanned by the vectors $(1, -2, 1, 0)$ and $(2, -3, 0, 1)$.•

Since a subspace $W$ and its orthogonal complement $W^\perp$ have only the zero vector in common, it seems reasonable to expect that a vector space can be thought of as the direct sum $W \oplus W^\perp$. According to the corollary to Theorem 1.18 in Section 1.7, this will be the case if we can show that every vector in the space can be expressed as the sum of a vector from $W$ and a vector from $W^\perp$. The following discussion provides us with the tool to do this.

**Orthogonal Components of Vectors**

In Section 1.7, we defined subspace components of vectors. A special case of this is when the subspaces are orthogonal complements. The present discussion shows how to take what are called *orthogonal components* of a vector. We begin with the familiar space $\mathcal{G}^2$.

Suppose that $\mathbf{w}$ is a fixed vector in $\mathcal{G}^2$, and $\mathbf{v}$ is any other vector in the space (Figure 5.1). Geometrically, we can decompose $\mathbf{v}$ into a vector along $\mathbf{w}$ and a vector $\mathbf{u}$ perpendicular to $\mathbf{w}$ by dropping a perpendicular from the tip of $\mathbf{v}$ to $\mathbf{w}$, extended if necessary. Vector $\mathbf{v}$ is then the sum of $\mathbf{u}$ and a scalar multiple of $\lambda\mathbf{w}$ of $\mathbf{w}$.

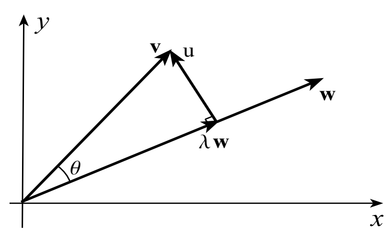


**Figure 5.1**

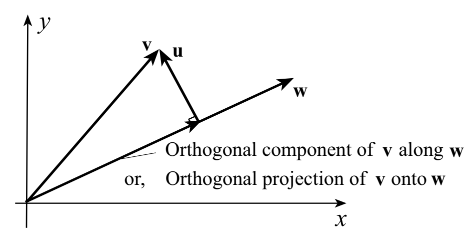


**Figure 5.2**

If $\hat{\mathbf{w}}$ denotes a unit vector in the direction of $\mathbf{w}$, then the length of this scalar multiple is

$$\|\lambda\mathbf{w}\| = \|\mathbf{v}\| \cos\theta = \|\mathbf{v}\|\|\hat{\mathbf{w}}\| \cos\theta = \hat{\mathbf{w}} \cdot \mathbf{v} = \frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{w}\|}.$$

To be specific, then,

$$\mathbf{v} = \mathbf{u} + \frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{w}\|}\hat{\mathbf{w}} = \mathbf{u} + \frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{w}\|^2}\mathbf{w}. \tag{5.31}$$

What we have done is express every vector $\mathbf{v}$ in $\mathcal{G}^2$ as the sum of two vectors

$$\mathbf{v} = \mathbf{u} + \mathbf{r}, \qquad \text{where} \quad \mathbf{r} = \frac{\mathbf{w} \cdot \mathbf{v}}{\|\mathbf{w}\|^2}\mathbf{w}. \tag{5.32}$$

Vector $\mathbf{r}$ is in the subspace $W$ of vectors along $\mathbf{w}$, and $\mathbf{u}$ is in the subspace $W^\perp$ of vectors perpendicular to $W$. This shows that $\mathcal{G}^2$ is the direct sum $W \oplus W^\perp$, and in the terminology of Section 1.7, vector $\mathbf{r}$ is the vector component of $\mathbf{v}$ along the subspace $W$ as determined by the subspace $W^\perp$. We shorten this description and call $\dfrac{\mathbf{w}\cdot\mathbf{v}}{\|\mathbf{w}\|^2}\mathbf{w}$ the **orthogonal component of v along w** (or the subspace spanned by $\mathbf{w}$). It is also called the **orthogonal projection of v onto w** (Figure 5.2), denoted by

$$\text{Oproj}_{\mathbf{w}}\mathbf{v} = \frac{\mathbf{w}\cdot\mathbf{v}}{\|\mathbf{w}\|^2}\mathbf{w}. \tag{5.33}$$

The following theorem indicates that we can do the same thing in any inner product space, real or complex.

**Theorem 5.6** If $\mathbf{w}$ is a fixed vector in an inner product space $V$, and $\mathbf{v}$ is any other vector in the space, then $\mathbf{v}$ can expressed in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{r}, \tag{5.34a}$$

where $\mathbf{u}$ is orthogonal to $\mathbf{w}$, and $\mathbf{r}$ is the orthogonal component of $\mathbf{v}$ along $\mathbf{w}$,

$$\mathbf{r} = \frac{(\mathbf{w},\mathbf{v})}{\|\mathbf{w}\|^2}\mathbf{w}. \tag{5.34b}$$

**Proof** Since $\mathbf{r}$ is uniquely defined by equation 5.34b, all that we need show is that the vector $\mathbf{u}$ defined by equation 5.34a is orthogonal to $\mathbf{w}$. If we solve for $\mathbf{u}$ and take inner products with $\mathbf{w}$, we get

$$(\mathbf{w},\mathbf{u}) = \left(\mathbf{w}, \mathbf{v} - \frac{(\mathbf{w},\mathbf{v})}{\|\mathbf{w}\|^2}\mathbf{w}\right) = (\mathbf{w},\mathbf{v}) - \frac{(\mathbf{w},\mathbf{v})(\mathbf{w},\mathbf{w})}{\|\mathbf{w}\|^2} = (\mathbf{w},\mathbf{v}) - (\mathbf{w},\mathbf{v}) = 0.$$

This implies that $\mathbf{u}$ is orthogonal to $\mathbf{w}$.∎

It is worthwhile noticing that had we taken the vectors $\mathbf{w}$ and $\mathbf{u}$ in the reverse order in the proof, the calculation would have gone as follows:

$$(\mathbf{u},\mathbf{w}) = \left(\mathbf{v} - \frac{(\mathbf{w},\mathbf{v})}{\|\mathbf{w}\|^2}\mathbf{w}, \mathbf{w}\right) = (\mathbf{v},\mathbf{w}) - \frac{\overline{(\mathbf{w},\mathbf{v})}(\mathbf{w},\mathbf{w})}{\|\mathbf{w}\|^2} = (\mathbf{v},\mathbf{w}) - (\mathbf{v},\mathbf{w}) = 0.$$

Two points need to be stressed about this result:

1. Space $V$ need not be 2-dimensional. When $V$ has dimension $n > 2$, vector $\mathbf{u}$ is in the $(n-1)$-dimensional subspace $W^\perp$, where $W$ is the subspace of multiples of $\mathbf{w}$.
2. When the vector space is real, the order of $\mathbf{v}$ and $\mathbf{w}$ in the inner product defining the orthogonal component $\mathbf{r}$ is immaterial. When the space is complex, however, it is crucial that the order specified be maintained. Suppose for instance that we wrote the inner product in the reverse order, and attempt to show that $\mathbf{u}$ as defined by equation 5.34 is orthogonal to $\mathbf{w}$ as in Theorem 5.6,

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

In a complex space, the inner product is not symmetric, so that this does not vanish, and the vectors are not orthogonal.

Theorem 5.6 shows that a vector $\mathbf{v}$ can always be expressed as the sum of a vector along a given vector $\mathbf{w}$ and a vector orthogonal to $\mathbf{w}$. We now generalize this result to show that vector $\mathbf{w}$ can be replaced by a set of orthogonal vectors.

**Theorem 5.7** If $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ is a set of orthogonal vectors in an inner product space $V$, and $\mathbf{v}$ is any vector in the space, then $\mathbf{v}$ can be expressed in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{r}, \tag{5.35a}$$

where $\mathbf{u}$ is orthogonal to every vector in the subspace spanned by the $\mathbf{w}_j$, and $\mathbf{r}$ is the sum of the orthogonal components of $\mathbf{v}$ along the $\mathbf{w}_j$; that is,

$$\mathbf{r} = \sum_{j=1}^{m} \frac{(\mathbf{w}_j, \mathbf{v})}{\|\mathbf{w}_j\|^2}\mathbf{w}_j. \tag{5.35b}$$

**Proof** Since $\mathbf{r}$ is uniquely defined by equation 5.35b, all that we need show is that the vector $\mathbf{u}$ defined by equation 5.35a is orthogonal to each of the $\mathbf{w}_j$. If we solve for $\mathbf{u}$ and take inner products with $\mathbf{w}_k$, we get

$$(\mathbf{w}_k, \mathbf{u}) = \left(\mathbf{w}_k, \mathbf{v} - \sum_{j=1}^{m} \frac{(\mathbf{w}_j, \mathbf{v})}{\|\mathbf{w}_j\|^2}\mathbf{w_j}\right) = (\mathbf{w}_k, \mathbf{v}) - \sum_{j=1}^{m} \frac{(\mathbf{w}_j, \mathbf{v})(\mathbf{w}_k, \mathbf{w}_j)}{\|\mathbf{w}_j\|^2}.$$

Since the $\mathbf{w}_j$ are orthogonal, $(\mathbf{w}_k, \mathbf{w}_j)$ equals zero whenever $j \neq k$. Hence,

$$(\mathbf{w}_k, \mathbf{u}) = (\mathbf{w}_k, \mathbf{v}) - \frac{(\mathbf{w}_k, \mathbf{v})(\mathbf{w}_k, \mathbf{w}_k)}{\|\mathbf{w}_k\|^2} = (\mathbf{w}_k, \mathbf{v}) - (\mathbf{w}_k, \mathbf{v}) = 0.$$

Thus, $\mathbf{u}$ is orthogonal to each of the $\mathbf{w}_j$.∎

Since $\mathbf{r}$ is a linear combination of the vectors in the set $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$, it is therefore a vector in the subspace $W$ spanned by the $\mathbf{w}_j$. Vector $\mathbf{u}$ being orthogonal to each of the $\mathbf{w}_j$ is in $W^\perp$. This verifies that $V = W \oplus W^\perp$. Vector $\mathbf{r}$ is the vector component of $\mathbf{v}$ along $W$ as determined by $W^\perp$. We call it the **orthogonal component** of $\mathbf{v}$ along the subspace $W$ spanned by the $\mathbf{w}_j$, or, the **orthogonal projection** of $\mathbf{v}$ onto subspace $W$,

$$\text{OProj}_W \mathbf{v} = \sum_{j=1}^{m} \frac{(\mathbf{w}_j, \mathbf{v})}{\|\mathbf{w}_j\|^2}\mathbf{w}_j. \tag{5.36}$$

**Example 5.21** In $\mathcal{E}^2$, find the orthogonal component of $(4, -1)$ along $(2, 3)$.

**Solution** The orthogonal component of $(4, -1)$ along $(2, 3)$ is

$$\frac{\big((2,3),(4,-1)\big)}{13}(2,3) = \frac{5}{13}(2,3).\bullet$$

**Example 5.22** If all three vectors $\mathbf{u} = (1, 2, 5)$, $\mathbf{v} = (-2, 1, 4)$, and $\mathbf{w} = (1, -2, 1)$ are in $\mathcal{E}^3$, find the orthogonal component of $\mathbf{u}$ along the subspace spanned by $\mathbf{v}$ and $\mathbf{w}$.

**Solution** Since $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, the orthogonal component is

$$\begin{aligned}\frac{\big((-2,1,4),(1,2,5)\big)}{21}(-2,1,4) + \frac{\big((1,-2,1),(1,2,5)\big)}{6}(1,-2,1) &= \frac{20}{21}(-2,1,4) + \frac{1}{3}(1,-2,1)\\ &= \frac{1}{21}(-33,16,87).\bullet\end{aligned}$$

### The Need for Orthogonal Bases

In Chapter 4, we demonstrated the value of eigenvalue bases in a number of applications. There are advantages to using orthogonal bases for inner product spaces, or even better, orthonormal bases. We demonstrate with three situations.

### Scalar Components of Vectors in Inner Product Spaces

If $\mathbf{v} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{j}}$ is a vector in $\mathcal{E}^3$, its (scalar) components can be obtained from inner products of the vector with the basis vectors,

$$v_x = \mathbf{v} \cdot \hat{\mathbf{i}} = (\mathbf{v}, \hat{\mathbf{i}}), \qquad v_y = \mathbf{v} \cdot \hat{\mathbf{j}} = (\mathbf{v}, \hat{\mathbf{j}}), \qquad v_z = \mathbf{v} \cdot \hat{\mathbf{k}} = (\mathbf{v}, \hat{\mathbf{k}}). \tag{5.37}$$

The vectors $\mathbf{b}_1 = 3\hat{\mathbf{i}}$, $\mathbf{b}_2 = -2\hat{\mathbf{j}}$, and $\mathbf{b}_3 = 4\hat{\mathbf{k}}$ also form a basis for $\mathcal{E}^3$; it is orthogonal, but not orthonormal. To obtain formulas for the components of a vector $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + v_3\mathbf{b}_3$, we can once again take dot products of $\mathbf{v}$ with the basis vectors. When we do so with $\mathbf{b}_1$, we obtain

$$\mathbf{v}\cdot\mathbf{b}_1 = v_1(\mathbf{b}_1\cdot\mathbf{b}_1) = v_1\|\mathbf{b}_1\|^2 \qquad \Longrightarrow \qquad v_1 = \frac{\mathbf{v}\cdot\mathbf{b}_1}{\|\mathbf{b}_1\|^2}.$$

When we write this with inner product notation, and include results for $v_2$ and $v_3$, we obtain

$$v_1 = \frac{(\mathbf{v},\mathbf{b}_1)}{\|\mathbf{b}_1\|^2}, \qquad v_2 = \frac{(\mathbf{v},\mathbf{b}_2)}{\|\mathbf{b}_2\|^2}, \qquad \frac{(\mathbf{v},\mathbf{b}_3)}{\|\mathbf{b}_3\|^2}. \tag{5.38}$$

These equations are more complicated than 5.37 due to the fact that the basis $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ is orthogonal, but not orthonormal.

Now consider the basis consisting of the vectors $\mathbf{d}_1 = 3\hat{\mathbf{i}} - 2\hat{\mathbf{j}}$, $\mathbf{d}_2 = \hat{\mathbf{i}} - 2\hat{\mathbf{k}}$, and $\mathbf{d}_3 = \hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$. These vectors are not orthogonal and they do not have length one. If we take dot products of a vector $\mathbf{v} = v_1\mathbf{d}_1 + v_2\mathbf{d}_2 + v_2\mathbf{d}_3$ with the basis vectors, we obtain

$$\begin{aligned}
\mathbf{v}\cdot\mathbf{d}_1 &= v_1(\mathbf{d}_1\cdot\mathbf{d}_1) + v_2(\mathbf{d}_1\cdot\mathbf{d}_2) + v_3(\mathbf{d}_1\cdot\mathbf{d}_3),\\
\mathbf{v}\cdot\mathbf{d}_2 &= v_1(\mathbf{d}_2\cdot\mathbf{d}_1) + v_2(\mathbf{d}_2\cdot\mathbf{d}_2) + v_3(\mathbf{d}_2\cdot\mathbf{d}_3),\\
\mathbf{v}\cdot\mathbf{d}_3 &= v_1(\mathbf{d}_3\cdot\mathbf{d}_1) + v_2(\mathbf{d}_3\cdot\mathbf{d}_2) + v_3(\mathbf{d}_3\cdot\mathbf{d}_3).
\end{aligned} \tag{5.39}$$

We have a system of three equations to solve for the components of $\mathbf{v}$. These are definitely more complicated than equations 5.37 and 5.38.

What we have just seen in $\mathcal{E}^3$ occurs in every finite-dimensional inner product space $V$. If $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthonormal basis for the space, and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are the components of any vector, with respect to this basis, then

$$v_i = (\mathbf{v}, \mathbf{b}_i). \tag{5.40}$$

If $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_n\}$ is only an orthogonal basis, then

$$(\mathbf{v},\mathbf{d}_i) = v_i(\mathbf{d}_i,\mathbf{d}_i) \qquad \Longrightarrow \qquad v_i = \frac{(\mathbf{v},\mathbf{d}_i)}{(\mathbf{d}_i,\mathbf{d}_i)} = \frac{(\mathbf{v},\mathbf{d}_i)}{\|\mathbf{d}_i\|^2}. \tag{5.41}$$

If the basis is not orthogonal, then taking inner products of the vector with the basis vectors yields a system of $n$ equations in the $n$ components of the vector.

**Linear Operators on Inner Product Spaces**

When $L$ is a linear operator on an $n$-dimensional space $V$, its associated matrix, relative to some basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ has columns that are images of the $\mathbf{b}_i$ under $L$. When $V$ is an inner product space, and the basis is orthonormal, we can give a formula for the entries of the matrix. We can demonstrate this with an operator on $\mathcal{G}^3$ such as

$$L: \quad \begin{aligned}
v_1' &= 3v_1 - 2v_2 + 4v_3,\\
v_2' &= -v_1 + v_2,\\
v_3' &= 2v_1 - 2v_2 + 3v_3.
\end{aligned}$$

Because no mention of basis has been made, we assume that these are natural components of $\mathbf{v}$ and $\mathbf{v}'$. The matrix associated with the operator is

$$A = \begin{pmatrix} 3 & -2 & 4\\ -1 & 1 & 0\\ 2 & -2 & 3 \end{pmatrix}.$$

The first column is the image of $\hat{\mathbf{i}}$ under $L$; that is

$$\begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} = L(\hat{\mathbf{i}}).$$

Inner products of this vector with the basis vectors are the components of the vector,

$$\big(\hat{\mathbf{i}}, L(\hat{\mathbf{i}})\big) = 3, \qquad \big(\hat{\mathbf{j}}, L(\hat{\mathbf{i}})\big) = -2, \qquad \big(\hat{\mathbf{k}}, L(\hat{\mathbf{i}})\big) = 4.$$

This is not a peculiarity of the first column; it is also valid for the second and third columns. In other words, the $(i,j)^{\text{th}}$ entry of $A$ is the inner product of the $i^{\text{th}}$ basis vector with the image of the $j^{\text{th}}$ basis vector. That this result is valid for any linear operator on any finite-dimensional inner product vector space (provided an orthonormal basis of the space is used to find the matrix associated with the operator) is proved in the next theorem.

**Theorem 5.8** The $(i,j)^{\text{th}}$ entry of the matrix $A$ associated with a linear operator $L$ on an $n$-dimensional, inner product space $V$ with orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is

$$a_{ij} = \big(\mathbf{b}_i, L(\mathbf{b}_j)\big). \tag{5.42}$$

**Proof** The $j^{\text{th}}$ column of $A$ is the image $L(\mathbf{b}_j)$ of the $j^{\text{th}}$ basis vector $\mathbf{b}_j$. But according to equation 5.40, the $i^{\text{th}}$ component of this vector is $\big(\mathbf{b}_i, L(\mathbf{b}_j)\big)$.∎

**The Matrix Associated With Inner Products**

The action of a linear transformation between finite-dimensional vector spaces can be accomplished with the matrix associated with the transformation. The same can be done with inner products; that is, we can associate matrices with inner products on finite-dimensional vector spaces so that taking the inner product reduces to matrix multiplication. This is very simple when the basis is orthonormal, not quite so simple when the basis is orthogonal, and not at all simple when the basis is not orthogonal. We demonstrate with some simple examples. The inner product of two vectors $\mathbf{u} = u_x\hat{\mathbf{i}} + u_y\hat{\mathbf{j}} + u_z\hat{\mathbf{k}}$ and $\mathbf{v} = v_x\hat{\mathbf{i}} + v_y\hat{\mathbf{j}} + v_z\hat{\mathbf{k}}$ in $\mathcal{E}^3$ is

$$(\mathbf{u}, \mathbf{v}) = u_x v_x + u_y v_y + u_z v_z.$$

We can use matrices to write this in the form

$$(\mathbf{u}, \mathbf{v}) = (u_x, u_y, u_z)\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = (u_x, u_y, u_z)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}. \tag{5.43}$$

The $3 \times 3$ identity matrix $I_3$ is called the **matrix of the inner product** associated with the natural basis. In $\mathcal{E}^n$, the matrix associated with the natural basis is the $n \times n$ identity $I_n$. This is not a peculiarity of natural bases. The following theorem shows that $I_n$ is the matrix associated with any orthonormal basis in a real vector space.

**Theorem 5.9** Let $V$ be a real, $n$-dimensional inner product space with orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$. If the components of two vectors with respect to this basis are $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$, then their inner product is

$$(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} u_i v_i. \tag{5.44}$$

**Proof** Using the fact that an inner product is linear in its arguments, we can write that

$$(\mathbf{u}, \mathbf{v}) = \left(\sum_{i=1}^{n} u_i \mathbf{b}_i, \sum_{j=1}^{n} v_i \mathbf{b}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} u_i v_j (\mathbf{b}_i, \mathbf{b}_j).$$

Since the basis is orthonormal, $(\mathbf{b}_i, \mathbf{b}_j) = \delta_{ij}$, and therefore

$$(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^{n} u_i v_i. \blacksquare$$

Because we can write equation 5.9 in the form

$$(\mathbf{u}, \mathbf{v}) = (u_1, u_2, \ldots, u_n) I_n \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \tag{5.45}$$

the matrix associated with the orthonormal basis is $I_n$.

Analogous to Theorem 5.9 for real inner product spaces, we have the following theorem for complex inner product spaces.

**Theorem 5.10** If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ are components of vectors with respect to an orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ in a complex, $n$-dimensional inner product space, their inner product is

$$(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n} \overline{u_j} v_j. \tag{5.46}$$

**Proof** Using properties 5.19 for the inner product, we can write that

$$(\mathbf{u}, \mathbf{v}) = \left( \sum_{j=1}^{n} u_j \mathbf{b}_j, \sum_{k=1}^{n} v_k \mathbf{b}_k \right) = \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{u_j} v_k (\mathbf{b}_j, \mathbf{b}_k).$$

Since the basis is orthonormal, $(\mathbf{b}_j, \mathbf{b}_k) = \delta_{jk}$, and therefore

$$(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{n} \overline{u_j} v_j. \blacksquare$$

This shows that the matrix associated with an orthonormal basis is once again the identity,

$$(\mathbf{u}, \mathbf{v}) = (\overline{u_1}, \overline{u_2}, \ldots, \overline{u_n}) I_n \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}. \tag{5.47}$$

The vectors $\mathbf{b}_1 = 3\hat{\mathbf{i}}$, $\mathbf{b}_2 = -2\hat{\mathbf{j}}$, and $\mathbf{b}_3 = 4\hat{\mathbf{k}}$ also form a basis for $\mathcal{E}^3$; the vectors are orthogonal, but not orthonormal. The inner product of two vectors $\mathbf{u} = u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + u_3\mathbf{b}_3$ and $\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + v_3\mathbf{b}_3$ is

$$\begin{aligned}
(\mathbf{u}, \mathbf{v}) &= (u_1\mathbf{b}_1 + u_2\mathbf{b}_2 + u_3\mathbf{b}_3, v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + v_3\mathbf{b}_3) \\
&= u_1v_1(\mathbf{b}_1, \mathbf{b}_1) + u_1v_2(\mathbf{b}_1, \mathbf{b}_2) + u_1v_3(\mathbf{b}_1, \mathbf{b}_3) \\
&\qquad + u_2v_1(\mathbf{b}_2, \mathbf{b}_1) + u_2v_2(\mathbf{b}_2, \mathbf{b}_2) + u_2v_3(\mathbf{b}_2, \mathbf{b}_3) \\
&\qquad + u_3v_1(\mathbf{b}_3, \mathbf{b}_1) + u_3v_2(\mathbf{b}_3, \mathbf{b}_2) + u_3v_3(\mathbf{b}_3, \mathbf{b}_3) \\
&= 9u_1v_1 + 4u_2v_2 + 16u_3v_3.
\end{aligned}$$

We can write this in matrix form as

$$(\mathbf{u}, \mathbf{v}) = (u_1, u_2, u_3) \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 16 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

The $3 \times 3$ diagonal matrix is the matrix associated with the orthogonal basis.

Now consider the basis consisting of the vectors $\mathbf{d}_1 = 3\hat{\mathbf{i}} - 2\hat{\mathbf{j}}$, $\mathbf{d}_2 = \hat{\mathbf{i}} - 2\hat{\mathbf{k}}$, and $\mathbf{d}_3 = \hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$. These vectors are not orthogonal and they do not have length one. The inner product of two vectors $\mathbf{u} = u_1\mathbf{d}_1 + u_2\mathbf{d}_2 + u_3\mathbf{d}_3$ and $\mathbf{v} = v_1\mathbf{d}_1 + v_2\mathbf{d}_2 + v_3\mathbf{d}_3$ is

$$\begin{aligned}
(\mathbf{u}, \mathbf{v}) &= (u_1\mathbf{d}_1 + u_2\mathbf{d}_2 + u_3\mathbf{d}_3, v_1\mathbf{d}_1 + v_2\mathbf{d}_2 + v_3\mathbf{d}_3) \\
&= u_1v_1(\mathbf{d}_1, \mathbf{d}_1) + u_1v_2(\mathbf{d}_1, \mathbf{d}_2) + u_1v_3(\mathbf{d}_1, \mathbf{d}_3) \\
&\qquad + u_2v_1(\mathbf{d}_2, \mathbf{d}_1) + u_2v_2(\mathbf{d}_2, \mathbf{d}_2) + u_2v_3(\mathbf{d}_2, \mathbf{d}_3) \\
&\qquad + u_3v_1(\mathbf{d}_3, \mathbf{d}_1) + u_3v_2(\mathbf{d}_3, \mathbf{d}_2) + u_3v_3(\mathbf{d}_3, \mathbf{d}_3) \\
&= u_1(13v_1 + 3v_2 + 7v_3) + u_2(3v_1 + 5v_2 - v_3) + u_3(7v_1 - v_2 + 6v_3).
\end{aligned}$$

This could be simplified further, but by leaving it in this form, we can write it in matrix form

$$(\mathbf{u}, \mathbf{v}) = (u_1, u_2, u_3)\begin{pmatrix} 13 & 3 & 7 \\ 3 & 5 & -1 \\ 7 & -1 & 6 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

This time the matrix is not diagonal, but in all cases, the matrix is symmetric. When the basis is orthogonal, the matrix is diagonal; and when the basis is orthonormal, the matrix is the identity. If we denote the matrix by $G$, then its $(i,j)^{\text{th}}$ entry is the inner product of the $i^{\text{th}}$ and $j^{\text{th}}$ basis vectors. If they are denoted by $\mathbf{b}_i$ and $\mathbf{b}_j$, then

$$G = \big((\mathbf{b}_i, \mathbf{b}_j)\big). \tag{5.48}$$

These discussions have shown three advantages in using an orthogonal basis for an inner product space, or even better, an orthonormal one. In Section 5.6, we give a systematic procedure on how to develop an orthonormal basis from any given basis. The procedure is not always necessary however, and as we shall see, it does not always lead to the simplest orthonormal basis.

**Example 5.23** Find an orthonormal basis for $\mathcal{G}^3$ if it must contain the vector $(1, -2, 3)/\sqrt{14}$.

**Solution** We need two orthogonal vectors that are also orthogonal to $(1, -2, 3)$; they can always be normalized later. Let two such vectors be $(a, b, c)$ and $(d, e, f)$. Orthogonality requires

$$a - 2b + 3c = 0, \qquad d - 2e + 3f = 0, \qquad ad + be + cf = 0.$$

We have three equations, two linear and one nonlinear, in six unknowns. This would seem to give plenty of freedom to make other demands. Suppose for instance that we demand that $b = 1$, $c = 0$ and $e = 1$. The equations then reduce to

$$a - 2 = 0, \quad d - 2 + 3f = 0, \quad ad + 1 = 0 \qquad \Longrightarrow \qquad a = 2, \quad d = -\frac{1}{2}, \quad f = \frac{5}{6}.$$

Thus, the vectors $(2, 1, 0)$ and $(-1/2, 1, 5/6)$ are orthogonal to each other, and both are orthogonal to $(1, -2, 3)$. An orthonormal basis for $\mathcal{G}^3$ that contains $(1, -2, 3)/\sqrt{14}$ is

$$\frac{(1, -2, 3)}{\sqrt{14}}, \quad \frac{(2, 1, 0)}{\sqrt{5}}, \quad \frac{(-3, 6, 5)}{\sqrt{70}}. \bullet$$

---

**EXERCISES 5.5**

**1.** Find scalar components of the polynomial $x^3 + 2x + 1$, as a vector in $P_3(x)$, with respect to the basis of polynomials in part (a) of Exercise 14 in Section 5.3.

**2.** Find scalar components of the polynomial $x^3 + 2x^2 + 1$, as a vector in $P_3(x)$, with respect to the basis of polynomials in part (a) of Exercise 16 in Section 5.3.

**3.** Find scalar components of the polynomial $x^3 + 2x^2 - x$, as a vector in $P_3(x)$, with respect to the basis of polynomials in part (a) of Exercise 15 in Section 5.3.

**4.** Find an orthonormal basis for $\mathcal{G}^3$ that contains vectors of unit length in the directions of the vectors $(3, 5, 1)$ and $(2, -2, 4)$.

**5.** If $W$ is the subspace of $\mathcal{E}^5$ spanned by the vectors $(1, -3, 5, 0, 5)$, $(-1, 1, 2, -2, 3)$, and $(0, -1, 4, -1, 5)$, find a basis for $W^{\perp}$.

**In Exercises 6–8 find the orthogonal component of vector v along the subspace spanned by w. All vectors are in $\mathcal{E}^n$ for an appropriate value of $n$.**

**6.** $\mathbf{v} = (2, 1)$; $\mathbf{w} = (4, -3)$

**7.** $\mathbf{v} = (-1, 2, 5)$; $\mathbf{w} = (3, 2, -1)$

**8.** $\mathbf{v} = (3, -2, 1, 4)$; $\mathbf{w} = (4, 5, 6, 7)$

**9.** Find the orthogonal component of the vector $\langle 3, 2 - 4\rangle$ along the line $2x + y - 2z = 0$, $3x - 2y + z = 0$.

**10.** Find the orthogonal component of the vector $(1, -2, 4)$ in $\mathcal{R}^3$ along the subspace spanned by the vectors $(3, -1, 4)$ and $(1, -1, -1)$.

**11.** Find the orthogonal component of the vector $(2 - i, 1 + i)$ along the subspace of $\mathcal{C}^2$ spanned by the vector $(3 - 2i, 1 - 4i)$.

**Answers**

**1.** $(1, 13/5, 0, 2/5)$ **2.** $(2, 3/4, 1, 1/4)$ **3.** $(1, 1/4, 1/2, 1/8)$
**4.** $(3, 5, 1)/\sqrt{35}$, $(1, -1, 2)/\sqrt{6}$, $(11, -5, -8)/\sqrt{210}$ **5.** $(3, 1, 0, -1, 0)$, $(4, 3, 2, 0, -1)$
**6.** $(4/5, -3/5)$ **7.** $(-6/7, -4/7, 2/7)$ **8.** $(8/7, 10/7, 12/7, 14/7)$
**9.** $\langle -9/122, -24/122, -21/122\rangle$ **10.** $(163/78, -37/78, 278/78)$
**11.** $(1/30)(27 + 8i, 29 - 14i)$

## §5.6 Gram-Schmidt Process

In Section 5.5 we saw advantages in using orthonormal bases for inner product spaces. It is straightforward to convert an orthogonal basis into an orthonormal basis, divide each vector by its norm. Theorems 5.6 and 5.7 provide a way to develop an orthonormal basis from a basis that is not orthogonal. It is called the Gram-Schmidt process.

### The Gram-Schmidt Process

We illustrate the process geometrically in $\mathcal{E}^3$, and then develop it in an arbitrary inner product space. Suppose that we have a basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for $\mathcal{E}^3$ that we wish to convert into an orthonormal basis. First we find an orthogonal basis; it can always be normalized later to produce an orthonormal basis. Take $\mathbf{w}_1 = \mathbf{v_1}$ as the first vector in an orthogonal basis. We use $\mathbf{v}_2$ to find a vector $\mathbf{w}_2$ orthogonal to $\mathbf{w}_1$. With equation 5.34b, we remove the orthogonal component of $\mathbf{v}_2$ along $\mathbf{w}_1$,

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{(\mathbf{w}_1, \mathbf{v}_2)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1,$$

(see Figure 5.3). Now, we use $\mathbf{v}_3$ to find a vector $\mathbf{w}_3$ orthogonal to $\mathbf{w}_1$ and $\mathbf{w}_2$. With equation 5.35b, we remove the component of $\mathbf{v}_3$ along the subspace spanned by $\mathbf{w}_1$ and $\mathbf{w}_2$,

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{(\mathbf{w}_1, \mathbf{v}_3)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{(\mathbf{w}_2, \mathbf{v}_3)}{\|\mathbf{w}_2\|^2}\mathbf{w}_2,$$

(see Figure 5.4). The three vectors $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$ constitute an orthogonal basis. An orthonormal basis can be obtained by normalizing these vectors.

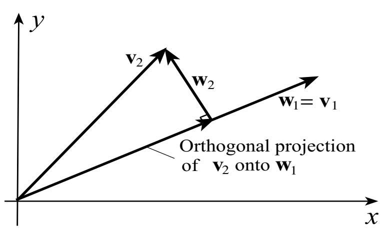


**Figure 5.3**

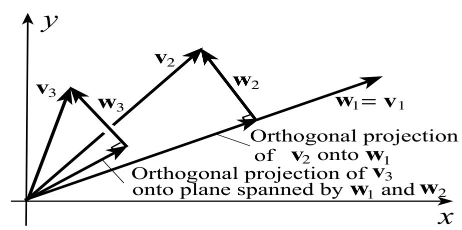


**Figure 5.4**

We do not normally use the Gram-Schmidt process to find an orthonormal basis for an inner product space (as was done above); there are easier ways. It is useful when we require an orthonormal basis of a subspace of an inner product space that is spanned by a given set of vectors that are not orthogonal. Suppose then, that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ is a basis for an $m$-dimensional subspace of an $n$-dimensional inner product space $V$, where $n > m$. From this basis, we will construct an orthogonal basis for the subspace. It is then a simple matter of rescaling to obtain an orthonormal basis.

**Step 1** Take $\mathbf{w}_1 = \mathbf{v}_1$ as the first vector in the orthogonal basis.

**Step 2** To obtain the second vector in the orthogonal basis, subtract from $\mathbf{v}_2$ its orthogonal component along $\mathbf{w}_1$ (or $\mathbf{v}_1$). If we call the resulting vector $\mathbf{w}_2$, and use equation 5.34b to find the orthogonal component, we obtain

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{(\mathbf{w}_1, \mathbf{v}_2)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1.$$

This vector is orthogonal to $\mathbf{w}_1$. To confirm this, we note that

$$(\mathbf{w}_1, \mathbf{w}_2) = \left(\mathbf{w}_1, \mathbf{v}_2 - \frac{(\mathbf{w}_1, \mathbf{v}_2)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1\right) = (\mathbf{w}_1, \mathbf{v}_2) - \frac{(\mathbf{w}_1, \mathbf{v}_2)}{\|\mathbf{w}_1\|^2}(\mathbf{w}_1, \mathbf{w}_1) = 0.$$

**Step 3** To obtain the third vector in the orthogonal basis, subtract from $\mathbf{v}_3$ its orthogonal component along the subspace spanned by $\mathbf{w}_1$ and $\mathbf{w}_2$. If we call the resulting vector $\mathbf{w}_3$, and use 5.35b to find the orthogonal component, we get

$$\mathbf{w}_3 = \mathbf{v}_3 - \left[\frac{(\mathbf{w}_1, \mathbf{v}_3)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 + \frac{(\mathbf{w}_2, \mathbf{v}_3)}{\|\mathbf{w}_2\|^2}\mathbf{w}_2\right].$$

This vector is orthogonal to $\mathbf{w}_1$ and $\mathbf{w}_2$. We can confirm orthogonality to $\mathbf{w}_1$ as follows:

$$\begin{aligned}(\mathbf{w}_1, \mathbf{w}_3) &= \left(\mathbf{w}_1, \mathbf{v}_3 - \frac{(\mathbf{w}_1, \mathbf{v}_3)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 - \frac{(\mathbf{w}_2, \mathbf{v}_3)}{\|\mathbf{w}_2\|^2}\mathbf{w}_2\right)\\ &= (\mathbf{w}_1, \mathbf{v}_3) - \frac{(\mathbf{w}_1, \mathbf{v}_3)}{\|\mathbf{w}_1\|^2}(\mathbf{w}_1, \mathbf{w}_1) - \frac{(\mathbf{w}_2, \mathbf{v}_3)}{\|\mathbf{w}_2\|^2}(\mathbf{w}_1, \mathbf{w}_2)\\ &= (\mathbf{w}_1, \mathbf{v}_3) - (\mathbf{w}_1, \mathbf{v}_3) = 0.\end{aligned}$$

A similar calculation confirms orthogonality to $\mathbf{w}_2$.

**Step 4** To obtain the fourth vector in the orthogonal basis, subtract from $\mathbf{v}_4$ its orthogonal component along the subspace spanned by $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$. If we call the resulting vector $\mathbf{w}_4$, and once again use 5.35b for the orthogonal component, we obtain

$$\mathbf{w}_4 = \mathbf{v}_4 - \left[\frac{(\mathbf{w}_1, \mathbf{v}_4)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 + \frac{(\mathbf{w}_2, \mathbf{v}_4)}{\|\mathbf{w}_2\|^2}\mathbf{w}_2 + \frac{(\mathbf{w}_3, \mathbf{v}_4)}{\|\mathbf{w}_3\|^3}\mathbf{w}_3\right].$$

This vector is orthogonal to $\mathbf{w}_1$, $\mathbf{w}_2$, and $\mathbf{w}_3$.

Continuation of this process leads to $m$ orthogonal vectors $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$. Since these vectors are linear combinations of the vectors $\mathbf{v}_j$, they are in the subspace spanned by the $\mathbf{v}_j$. Because orthogonal vectors are linearly independent, these vectors also constitute a basis for the subspace. Division of each vector by its norm gives an orthonormal basis for the subspace.

Here is an example in $\mathcal{E}^4$ to illustrate.

**Example 5.24** Use the Gram-Schmidt process to construct an orthonormal basis for the subspace of $\mathcal{E}^4$ spanned by the linearly independent vectors $\mathbf{v}_1 = (1, 2, 2, 0)$, $\mathbf{v}_2 = (-2, 1, 4, 0)$, and $\mathbf{v}_3 = (3, 2, 0, -1)$.

**Solution** If we choose $\mathbf{w}_1 = \mathbf{v}_1$ as the first vector in an orthogonal basis, and use $\mathbf{v}_2$ to construct a vector orthogonal to $\mathbf{w}_1$

$$\begin{aligned}\mathbf{w}_2 &= \mathbf{v}_2 - \frac{(\mathbf{w}_1, \mathbf{v}_2)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1\\ &= (-2, 1, 4, 0) - \frac{\big((1, 2, 2, 0), (-2, 1, 4, 0)\big)}{9}(1, 2, 2, 0)\\ &= (-2, 1, 4, 0) - \frac{8}{9}(1, 2, 2, 0) = \left(-\frac{26}{9}, -\frac{7}{9}, \frac{20}{9}, 0\right).\end{aligned}$$

Since the length of $\mathbf{w}_2$ at this point is irrelevant, we replace this $\mathbf{w}_2$ with $\mathbf{w}_2 = (26, 7, -20, 0)$. Let us check our calculations by seeing if $\mathbf{w}_2$ is indeed orthogonal to $\mathbf{w}_1$,

$$\big((1, 2, 2, 0), (26, 7, -20, 0)\big) = 26 + 14 - 40 + 0 = 0.$$

We now construct a vector orthogonal to $\mathbf{w}_1$ and $\mathbf{w}_2$,

$$\begin{aligned}\mathbf{w}_3 &= \mathbf{v}_3 - \left[\frac{(\mathbf{w}_1, \mathbf{v}_3)}{\|\mathbf{w}_1\|^2}\mathbf{w}_1 + \frac{(\mathbf{w}_2, \mathbf{v}_3)}{\|\mathbf{w}_2\|^2}\mathbf{w}_2\right]\\ &= (3, 2, 0, -1) - \frac{\big((1, 2, 2, 0), (3, 2, 0, -1)\big)}{9}(1, 2, 2, 0) - \frac{\big((26, 7, -20, 0), (3, 2, 0, -1)\big)}{1125}(26, 7, -20, 0)\\ &= (3, 2, 0, -1) - \frac{7}{9}(1, 2, 2, 0) - \frac{92}{1125}(26, 7, -20, 0) = \left(\frac{12}{125}, \frac{-16}{125}, \frac{2}{25}, 1\right).\end{aligned}$$

Once again lengths are irrelevant and we replace this $\mathbf{w}_3$ with $\mathbf{w}_3 = (12, -16, 10, 125)$, and check whether it is orthogonal to $\mathbf{w}_1$ and $\mathbf{w}_2$,

$$(\mathbf{w}_1, \mathbf{w}_3) = \big((1,2,2,0),(12,-16,10,125)\big) = 0, \qquad (\mathbf{w}_2, \mathbf{w}_3) = \big((26,7,-20,0),(12,-16,10,125)\big) = 0.$$

We now convert to an orthonormal basis by dividing each orthogonal vector by its norm,

$$\begin{aligned}\hat{\mathbf{w}}_1 &= \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} = \frac{1}{3}(1,2,2,0),\\ \hat{\mathbf{w}}_2 &= \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|} = \frac{1}{15\sqrt{5}}(26,7,-20,0),\\ \hat{\mathbf{w}}_3 &= \frac{\mathbf{w}_3}{\|\mathbf{w}_3\|} = \frac{1}{5\sqrt{645}}(12,-16,10,125).\bullet\end{aligned}$$

The order in which vectors are used to produce orthogonal vectors by the Gram-Schmidt process affects the outcome. Exercise 6 shows that a different set of orthonormal vectors is obtained in Example 5.24 if vector $(-2,1,4,0)$ is used first.

**Example 5.25** The vectors $f(x) = 1-x$ and $g(x) = 4x - x^2$ in the space $P_2(x)$ of real polynomials of degree less than or equal to two on the interval $0 \le x \le 1$ are not orthogonal with respect to the inner product of Example 5.9. Use them to find an orthogonal basis on this interval for the subspace spanned by the vectors.

**Solution** A vector (function) orthogonal to $f(x)$ is

$$h(x) = g(x) - \frac{\big(f(x), g(x)\big)}{\|f(x)\|^2} f(x),$$

where

$$\|f(x)\|^2 = \int_0^1 (1-x)^2\,dx = \frac{1}{3}, \quad \text{and} \quad (1-x, 4x-x^2) = \int_0^1 (1-x)(4x-x^2)\,dx = \frac{7}{12}.$$

Hence,

$$h(x) = 4x - x^2 - \frac{7/12}{1/3}(1-x) = -x^2 + \frac{23x}{4} - \frac{7}{4}.$$

An orthogonal basis for the subspace is therefore $\{1-x, 7-23x+4x^2\}$.●

Any set of vectors in a vector space spans a subspace. We can use the Gram-Schmidt process to replace the given vectors with an orthonormal basis for the subspace. As presented, the process requires a set of linearly independent vectors. This is not entirely true, but it is the best way to start. In other words, we would initially replace the given set of vectors with a linearly independent set. Here is an example to illustrate. When it is finished we discuss what happens when we start with a set of vectors that is linearly dependent.

**Example 5.26** Find an orthogonal basis for the subspace of $\mathcal{E}^5$ spanned by the vectors $(1,2,3,2,-5)$, $(3,-2,4,0,-1)$, and $(-1,6,2,4,-9)$.

**Solution** To determine whether the vectors are linearly independent, we row reduce

$$A = \begin{pmatrix} 1 & 2 & 3 & 2 & -5\\ 3 & -2 & 4 & 0 & -1\\ -1 & 6 & 2 & 4 & -9 \end{pmatrix} \qquad \text{to} \qquad A_{\text{rref}} = \begin{pmatrix} 1 & 0 & 7/4 & 1/2 & -3/2\\ 0 & 1 & 5/8 & 3/4 & -7/4\\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This shows that the vectors are linearly dependent, and a pair of vectors that span the same subspace is $(4,0,7,2,-6)$ and $(0,8,5,6,-14)$. To find an orthogonal basis, we can use $(4,0,7,2,-6)$ as one vector, and find a vector perpendicular to it using the Gram-Schmidt process

$$(0,8,5,6,-14) - \frac{\big((4,0,7,2,-6),(0,8,5,6,-14)\big)}{105}(4,0,7,2,-6) = \frac{4}{105}(-131,210,-98,92,-171).$$

We take the second vector in the orthogonal basis to be $(-131,210,-98,92,-171)$.•

Previous to this example, we suggested that it is not necessary to start with a set of linearly independent vectors, but it is always best to do so. If we don't, the Gram-Schmidt process leads to the zero vector when it encounters a vector that is a linear combination of previous vectors. We illustrate in the following very simple example, and discuss Example 5.26 in Exercise 25.

**Example 5.27** Show that when the Gram-Schmidt process is applied to the vectors $\{\hat{\mathbf{i}}, \hat{\mathbf{i}}+\hat{\mathbf{j}}, \hat{\mathbf{i}}+2\hat{\mathbf{j}}\}$, it leads to the zero vector at the third stage.

**Solution** The vector orthogonal to $\hat{\mathbf{i}}$ as produced by the Gram-Schmidt process is

$$\mathbf{v}_2 = (\hat{\mathbf{i}}+\hat{\mathbf{j}}) - \frac{(\hat{\mathbf{i}}, \hat{\mathbf{i}}+\hat{\mathbf{j}})}{1}\hat{\mathbf{i}} = \hat{\mathbf{j}},$$

as expected. The third vector should be the component of $\hat{\mathbf{i}}+2\hat{\mathbf{j}}$ orthogonal to $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$, which does not exist. The Gram-Schmidt process should therefore gives us the zero vector,

$$\mathbf{v}_3 = (\hat{\mathbf{i}}+2\hat{\mathbf{j}}) - \frac{(\hat{\mathbf{i}}, \hat{\mathbf{i}}+2\hat{\mathbf{j}})}{1}\hat{\mathbf{i}} - \frac{(\hat{\mathbf{i}}+2\hat{\mathbf{j}}, \hat{\mathbf{j}})}{1}\hat{\mathbf{j}} = \mathbf{0}.\bullet$$

**Example 5.28** Find an orthogonal basis for the kernel of the linear transformation with matrix

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

**Solution** We row reduce the matrix to find the kernel of the transformation,

$$A_{\text{rref}} = \begin{pmatrix} 1 & 0 & -3/2 & -5/2 \\ 0 & 1 & -5/4 & -5/4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

If we denote components of vectors in the space by $(v_1, v_2, v_3, v_4)$, then vectors in the kernel satisfy

$$v_1 - \frac{3v_3}{2} - \frac{5v_4}{2} = 0, \qquad v_2 - \frac{5v_3}{4} - \frac{5v_4}{4} = 0.$$

Thus, the kernel consists of vectors of the form

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} 3v_3/2 + 5v_4/2 \\ 5v_3/4 + 5v_4/4 \\ v_3 \\ v_4 \end{pmatrix} = \frac{v_3}{4}\begin{pmatrix} 6 \\ 5 \\ 4 \\ 0 \end{pmatrix} + \frac{v_4}{4}\begin{pmatrix} 10 \\ 5 \\ 0 \\ 4 \end{pmatrix}.$$

A basis for the kernel is $(6,5,4,0)$ and $(10,5,0,4)$. We can use the Gram-Schmidt process to replace the second vector by a vector orthogonal to $\mathbf{v}_1 = (6,5,4,0)$,

$$\mathbf{v}_2 = (10,5,0,4) - \frac{\big((6,5,4,0),(10,5,0,4)\big)}{36+25+16}(6,5,4,0) = (10,5,0,4) - \frac{85}{77}(6,5,4,0)$$
$$= \frac{4}{77}(85,-10,-85,77).$$

We therefore take $(85,-10,-85,77)$ as the second basis vector.•

There are many ways to produce an orthonormal basis for a vector space or subspace. It is a question of what other demands are to be imposed on the basis. We just noted that if we want a basis for a subspace that is spanned by a given set of vectors, then the Gram-Schmidt process replaces the given set with an orthonormal set that spans the same space (or subspace). Suppose we want an orthonormal basis for $\mathcal{E}^3$, but we demand that it contain a unit vector in the direction $\mathbf{v}_1 = (1, 2, 2)$. We could proceed in various ways. One way would be to introduce two additional vectors which along with $(1, 2, 2)$ form a linearly independent set, and then use the Gram-Schmidt process to construct an orthonormal set. Here is an alternative procedure. All vectors $(a, b, c)$ orthogonal to $\mathbf{v}_1$ must satisfy

$$0 = \big((a, b, c), (1, 2, 2)\big) = a + 2b + 2c \qquad \Longrightarrow \qquad a = -2b - 2c.$$

The are many choices for $b$ and $c$, a simple one being $b = 0$ and $c = 1$, in which case $a = -2$, and a vector orthogonal to $\mathbf{v}_1$ is $\mathbf{v}_2 = (-2, 1, 0)$. All vectors $(d, e, f)$ orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$ must satisfy

$$0 = \big((d, e, f), (1, 2, 2)\big) = d + 2e + 2f, \qquad 0 = \big((d, e, f), (-2, 1, 0)\big) = -2d + e.$$

All solutions of these can be expressed in the form $e = 2d$, $f = -5d/2$. If we choose $d = 2$, then $e = 4$ and $f = -5$. A vector orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$ is $\mathbf{v}_3 = (2, 4, -5)$. Now that we have an orthogonal set of vectors, we can re-scale for an orthonormal set,

$$\hat{\mathbf{v}}_1 = \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \frac{(1, 2, 2)}{\sqrt{6}}, \qquad \hat{\mathbf{v}}_2 = \frac{\mathbf{v}_2}{|\mathbf{v}_2|} = \frac{(-2, 1, 0)}{\sqrt{5}}, \qquad \hat{\mathbf{v}}_3 = \frac{\mathbf{v}_3}{|\mathbf{v}_3|} = \frac{(2, 4, -5)}{3\sqrt{5}}.$$

**Example 5.29** Show that with the inner product of Exercise 14 in Section 5.3, the Gram-Schmidt process applied to the polynomials 1, $x$, $x^2$, and $x^3$ leads to the normalized Legendre polynomials of Exercise 2 in Section 5.4.

**Solution** The norm of the function $f_0(x) = 1$ is given by

$$\|f_0(x)\|^2 = \int_{-1}^{1} 1^2\, dx = \{x\}_{-1}^{1} = 2.$$

A normalized function corresponding to 1 is therefore $\hat{f}_0(x) = \dfrac{1}{\sqrt{2}}$. Using the second polynomial, a function orthogonal to $f_0(x) = x$ is

$$f_1(x) = x - \frac{(1, x)}{2}(1) = x - \frac{1}{2}\int_{-1}^{1} x(1)\, dx = x - \frac{1}{2}\left\{\frac{x^2}{2}\right\}_{-1}^{1} = x.$$

Since the norm of this function is given by

$$\|f_1(x)\|^2 = \int_{-1}^{1} x^2\, dx = \left\{\frac{x^3}{3}\right\}_{-1}^{1} = \frac{2}{3},$$

a normalized function is $\hat{f}_1(x) = \dfrac{\sqrt{3}x}{\sqrt{2}}$.

Using $x^2$, a polynomial orthogonal to $f_0(x)$ and $f_1(x)$ is

$$\begin{aligned} f_2(x) &= x^2 - \frac{(1, x^2)}{2}(1) - \frac{(x, x^2)}{2/3}(x) = x^2 - \frac{1}{2}\int_{-1}^{1} x^2(1)\, dx - \frac{3x}{2}\int_{-1}^{1} x(x^2)\, dx \\ &= x^2 - \frac{1}{2}\left\{\frac{x^3}{3}\right\}_{-1}^{1} - \frac{3x}{2}\left\{\frac{x^4}{4}\right\}_{-1}^{1} = x^2 - \frac{1}{3}. \end{aligned}$$

Since the norm of this function is given by,

$$\|f_2(x)\|^2 = \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx = \int_{-1}^{1}\left(x^4 - \frac{2x^2}{3} + \frac{1}{9}\right) dx = \left\{\frac{x^5}{5} - \frac{2x^3}{9} + \frac{x}{9}\right\}_{-1}^{1} = \frac{8}{45},$$

a normalized function is $\hat{f}_2(x) = \sqrt{\dfrac{45}{8}}\left(x^2 - \dfrac{1}{3}\right) = \dfrac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1)$.
Using $x^3$, a polynomial orthogonal to $f_0(x)$, $f_1(x)$, and $f_2(x)$ is

$$\begin{aligned} f_3(x) &= x^3 - \frac{(1, x^3)}{2}(1) - \frac{(x, x^3)}{2/3}(x) - \frac{(x^2, x^3)}{8/45}(x^2) \\ &= x^3 - \frac{1}{2}\int_{-1}^{1} x^3(1)\,dx - \frac{3x}{2}\int_{-1}^{1} x(x^3)\,dx - \frac{45x^2}{8}\int_{-1}^{1} x^2(x^3)\,dx \\ &= x^3 - \frac{1}{2}\left\{\frac{x^4}{4}\right\}_{-1}^{1} - \frac{3x}{2}\left\{\frac{x^5}{5}\right\}_{-1}^{1} - \frac{45x^2}{8}\left\{\frac{x^6}{6}\right\}_{-1}^{1} = x^3 - \frac{3x}{5}. \end{aligned}$$

Since the norm of this function is given by,

$$\|f_3(x)\|^2 = \int_{-1}^{1}\left(x^3 - \frac{3x}{5}\right)^2 dx = \int_{-1}^{1}\left(x^6 - \frac{6x^4}{5} + \frac{9x^2}{25}\right) dx = \left\{\frac{x^7}{7} - \frac{6x^5}{25} + \frac{3x^3}{25}\right\}_{-1}^{1} = \frac{8}{175},$$

a normalized function is $\hat{f}_3(x) = \sqrt{\dfrac{175}{8}}\left(x^3 - \dfrac{3x}{5}\right) = \dfrac{\sqrt{7}}{2\sqrt{2}}(5x^3 - 3x)$.•

The Gram-Schmidt process can be applied to complex vectors as well as real ones, but, as we mentioned earlier, due to the fact that the inner product is not symmetric, we must be careful in the order of vectors in the inner product. Here is an example.

**Example 5.30** Use the Gram-Schmidt process to construct an orthonormal basis of $C^2$ using the vectors $\mathbf{v}_1 = (1 + i, i)$ and $\mathbf{v}_2 = (-i, 2 - i)$.

**Solution** A vector orthogonal to $\mathbf{v}_1$ is

$$\begin{aligned} \mathbf{w}_2 &= \mathbf{v}_2 - \frac{(\mathbf{v}_1, \mathbf{v}_2)}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \\ &= (-i, 2 - i) - \frac{\big((1 + i, i), (-i, 2 - i)\big)}{\|(1 + i, i)\|^2}(1 + i, i) \\ &= (-i, 2 - i) - \frac{(1 - i)(-i) - i(2 - i)}{|1 - i|^2 + |i|^2}(1 + i, i) \\ &= (-i, 2 - i) + \frac{2 + 3i}{3}(1 + i, i) \\ &= \frac{1}{3}(-1 + 2i, 3 - i). \end{aligned}$$

Since

$$\|(1 + i, i)\|^2 = 3, \qquad \|(-1 + 2i, 3 - i)\|^2 = |-1 + 2i|^2 + |3 - i|^2 = 15,$$

an orthonormal basis is

$$\hat{\mathbf{v}} = \frac{1}{\sqrt{3}}(1 + i, i), \qquad \hat{\mathbf{w}}_2 = \frac{1}{\sqrt{15}}(-1 + 2i, 3 - i).•$$

Theorems 5.6 and 5.7 gave instances when a vector space is the direct sum of orthogonal complements. The following theorem verifies that a vector space can be considered as the direct sum of any pair of orthogonal complements.

**Theorem 5.11** If $W$ is a subspace in a finite-dimensional inner product space $V$, then $V$ is the direct sum $W \oplus W^{\perp}$ of $W$ and and its orthogonal complement $W^{\perp}$.

**Proof** Since $W$ and $W^{\perp}$ have only the zero vector in common, all that we need show is that every vector $\mathbf{v}$ in $V$ can be expressed in the form $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$, where $\mathbf{w}_1$ is in $W$, and

$\mathbf{w}_2$ is in $W^\perp$ (see the corollary to Theorem 1.18). Since $V$ is finite-dimensional, so also is $W$. Any basis of $W$ can be converted to an orthonormal one (by the Gram-Schmidt process, for instance). According to Theorem 5.7, $\mathbf{v}$ can be expressed in the form $\mathbf{v} = \mathbf{u} + \mathbf{r}$, where $\mathbf{u}$ is orthogonal to every vector in the basis of $W$, and $\mathbf{r}$ is the sum of the orthogonal components of $\mathbf{v}$ along the basis vectors of $W$; that is, $\mathbf{u}$ is in $W^\perp$ and $\mathbf{r}$ is in $W$.∎

---

## EXERCISES 5.6

**In Exercises 1–2 use the Gram-Schmidt process to construct an orthonormal basis for the subspace of $\mathcal{E}^3$ spanned by the given vectors.**

**1.** $(1, 2, -4)$, $(3, 1, 5)$

**2.** $(-1, 3, 7)$, $(2, -5, 6)$

**In Exercises 3–4 use the Gram-Schmidt process to construct an orthonormal basis for the subspace of $\mathcal{E}^4$ spanned by the given vectors.**

**3.** $(1, 0, 2, 2)$, $(-1, 1, 1, 2)$, $(0, 1, 1, -1)$

**4.** $(2, 2, -1.0)$, $(3, 1, 2, -4)$, $(3, 2, 1, 1)$

**5.** Find an orthonormal basis for the subspace of $\mathcal{E}^5$ spanned by the vectors $(-1, -1, 1, 0, 0)$, $(0, -1, 0, 0, 1)$, and $(1, -1, 0, 1, 0)$.

**6.** Repeat Example 5.24 beginning with the vector $(-2, 1, 4, 0)$, followed by $(1, 2, 2, 0)$.

**7.** Repeat Exercise 1, but use the inner product $\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) = x_1y_1 + 2x_2y_2 + 3x_3y_3$.

**8.** Use the inner product $\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) = x_1y_1 + 2x_2y_2 + x_3y_3/3$ to construct an orthonormal basis for the subspace of $\mathcal{E}^3$ spanned by the vectors $(1, 0, 1)$ and $(1, 1, 0)$.

**9.** Find an orthogonal basis for the subspace of $\mathcal{E}^5$ spanned by the vectors $(1, 0, -1, 1, 1)$, $(0, 1, -1, 1, 2)$, and $(0, -1, 2, 0, 4)$.

**10.** Find a basis for the orthogonal complement of the subspace spanned by the vectors $(1, -2, 3, -4)$ and $(2, 3, 4, -1)$ in $\mathcal{E}^4$.

**11.** Find an orthogonal basis for the subspace in Exercise 10.

**12.** Find an orthonormal basis for the subspace of $\mathcal{C}^4$ spanned by the vectors $(1, i, -1, -i)$, $(1, 2i, -3, -4i)$, and $(4, 0, -2, 0)$.

**In Exercises 13–20 find the orthogonal component of vector v along the subspace spanned by the given set of vectors. All vectors are in $\mathcal{E}^n$ for an appropriate value of $n$.**

**13.** $\mathbf{v} = (1, 2, 3)$; $\{(-2, 3, 7), (1, -4, 2)\}$

**14.** $\mathbf{v} = (1, 2, 3)$; $\{(-2, 3, 7), (1, -4, 3)\}$

**15.** $\mathbf{v} = (2, -2, 4, 1)$; $\{(1, 1, 2, 2), (-3, -3, 2, 1)\}$

**16.** $\mathbf{v} = (2, -2, 4, 1)$; $\{(1, 1, 2, 2), (-3, -3, 2, 2)\}$

**17.** $\mathbf{v} = (2, 3, 1, -4)$; $\{(1, 1, 1, 1), (-1, 2, 4, -5), (-5, -1, 3, 3)\}$

**18.** Find the orthogonal component of the vector $\langle 1, -2, 3\rangle$ along the plane $2x - 3y + z = 0$.

**19.** Show that with the inner product $\big(p_1(x), p_2(x)\big) = a_0b_0 + a_1b_1 + a_2b_2 + a_3b_3$ of two vectors $p_1(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ and $p_2(x) = b_0 + b_1x + b_2x^2 + b_3x^3$ in $P_2(x)$, the set $\{1, x, x^2, x^3\}$ is an orthonormal basis.

**20.** (a) Consider the space of polynomials $P_2(x)$ on the interval $-1 \le x \le 1$. Is the basis $\{1, x, x^2\}$ orthogonal with respect to the inner product

$$\big(p_1(x), p_2(x)\big) = \int_{-1}^{1} p_1(x)p_2(x)\,dx.$$

(b) Construct an orthonormal basis from the basis in part (a).

**21.** Repeat Exercise 20 but use the interval $0 \le x \le 1$.

**22.** Rows of the matrix

$$A = \begin{pmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{pmatrix}$$

are vectors in $\mathcal{E}^5$ and columns are vectors in $\mathcal{E}^4$. Find bases for: (a) the row space of $A$, (b) the null space of $A$, (c) the orthogonal complement of the row space of $A$, and (d) the column space of $A$.

**23.** The vectors $\mathbf{v}_1 = (1, 1, -1, 1)$, $\mathbf{v}_2 = (1, 0, 2, 1)$, and $\mathbf{v}_3 = (-1, 0, 0, 1)$ are orthogonal in $\mathcal{E}^4$. Show that if the Gram-Schmidt process is used to find a fourth orthogonal vector starting with any vector $(a, b, c, d)$ that is not in the span of $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, the same vector is obtained.

**24.** Find a simple orthonormal basis for the row space of the matrix

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

**25.** Show if the Gram-Schmidt process is applied to the three given vectors in Example 5.26, the zero vector is produced at the third stage.

**26.** Find the orthogonal component of the vector $(2 - i, 1 + i, i)$ along the subspace of $\mathcal{C}^3$ spanned by the vectors $(1, 2, i)$ and $(3 + i, 2 + 2i, 4)$.

**27.** Show that with the inner product of Exercise 15 in Section 5.3, the Gram-Schmidt process applied to the polynomials 1, $x$, $x^2$, and $x^3$ leads to the normalized Hermite polynomials of Exercise 3 in Section 5.4.

**28.** Show that with the inner product of Exercise 16 in Section 5.3, the Gram-Schmidt process applied to the polynomials 1, $x$, $x^2$, and $x^3$ leads to the normalized Chebyshev polynomials of Exercise 4 in Section 5.4.

### Answers

**1.** $(1, 2, -4)/\sqrt{21}$, $(26, 17, 15)/\sqrt{1190}$ **2.** $(-1, 3, 7)/\sqrt{59}$, $(143, -370, 179)/\sqrt{189\,390}$
**3.** $(1, 0, 2, 2)/3$, $(-14, 9, -1, 8)/\sqrt{342}$, $(0, 1, 1, -1)/\sqrt{3}$
**4.** $(2, 2, -1, 0)/3$, $(5, -1, 8, -12)/(3\sqrt{26})$, $(21, 1, 44, 38)/(7\sqrt{78})$
**5.** $(-1, -1, 1, 0, 0)\sqrt{3}$, $(1, -2, -1, 0, 3)/\sqrt{15}$, $(4, -3, 1, 5, -3)/(2\sqrt{15})$
**6.** $(-2, 1, 4, 0)/\sqrt{21}$, $(37, 34, 10, 0)/(5\sqrt{105})$, $(12, -16, 10, -125)/(5\sqrt{645})$
**7.** $(1, 2, -4)/\sqrt{57}$, $(224, 163, 73)/\sqrt{119\,301}$ **8.** $(\sqrt{3}/2, 0, \sqrt{3}/2)$, $(1/6, 2/3, -1/2)$
**9.** $(1, 0, -1, 1, 1)$, $(-1, 1, 0, 0, 1)$, $(1, -4, 5, -1, 5)$ **10.** $(2, -1, 0, 1)$, $(-17, 2, 7, 0)$
**11.** $(2, -1, 0, 1)$, $(-5, -4, 7, 6)$ **12.** $(1, i, -1, -i)/2$, $(3, i, 1, 3i)/(2\sqrt{5})$, $(1, -2i, -1, 0)/\sqrt{6}$
**13.** $(-1112, 1823, 3551)/1302$ **14.** $(-5773, 9152, 18629/6773$
**15.** $(-4, -4, 64, 55)/23$ **16.** $5(0, 0, 1, 1)2$ **17.** $(55, 51, 33, -93)/23$ **18.** $(-8, 5, 51)/14$
**20.**(a) No (b) $1/\sqrt{2}$, $\sqrt{3/2}x$, $\sqrt{5}(3x^2 - 1)/(2\sqrt{2})$
**21.**(a) No (b) 1, $\sqrt{3}(2x - 1)$, $\sqrt{5}(6x^2 - 6x + 1)$
**22.**(a) $(1, 0, 1, 0, -1)$, $(0, 1, 2, 0, 3)$, $(0, 0, 0, 1, 4)$ (b) $(-1, -2, 1, 0, 0)$, $(1, -3, 0, -4, 1)$
(c) $(-1, -2, 1, 0, 0)$, $(1, -3, 0, -4, 1)$ (d) $(1, 0, 0, 1)$, $(0, 1, 0, 6)$, $(0, 0, 1, 3)$
**23.** $(1, -3, -1, 1)$ **24.** $(1, 0, 0, -1)/\sqrt{2}$, $(5, 2, 0, 5)/(3\sqrt{6})$, $(1, -5, 9, 1)/(6\sqrt{3})$
**26.** $(1/68)(141 + 29i, 210 + 76i, 55 + 78i)$