Chapter 10

Inner Product Spaces

10.1 Inner Products and Norms

The dot product was introduced in $\mathbb{R}^n$ to provide a natural generalization of the geometrical notions of length and orthogonality that were so important in Chapter 4. The plan in this chapter is to define an *inner product* on an arbitrary real vector space V (of which the dot product is an example in $\mathbb{R}^n$) and use it to introduce these concepts in V. While this causes some repetition of arguments in Chapter 8, it is well worth the effort because of the much wider scope of the results when stated in full generality.

Definition 10.1 Inner Product Spaces

An ***inner product*** *on a real vector space V is a function that assigns a real number $\langle \mathbf{v},\ \mathbf{w}\rangle$ to every pair $\mathbf{v}$, $\mathbf{w}$ of vectors in V in such a way that the following axioms are satisfied.*

P1. $\langle \mathbf{v},\ \mathbf{w}\rangle$ *is a real number for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P2. $\langle \mathbf{v},\ \mathbf{w}\rangle = \langle \mathbf{w},\ \mathbf{v}\rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.

P3. $\langle \mathbf{v}+\mathbf{w},\ \mathbf{u}\rangle = \langle \mathbf{v},\ \mathbf{u}\rangle + \langle \mathbf{w},\ \mathbf{u}\rangle$ *for all* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *in* V.

P4. $\langle r\mathbf{v},\ \mathbf{w}\rangle = r\langle \mathbf{v},\ \mathbf{w}\rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V *and all* r *in* $\mathbb{R}$.

P5. $\langle \mathbf{v},\ \mathbf{v}\rangle > 0$ *for all* $\mathbf{v} \neq \mathbf{0}$ *in* V.

A real vector space V with an inner product $\langle\ ,\ \rangle$ will be called an **inner product space**. Note that every subspace of an inner product space is again an inner product space using the same inner product.[1]

Example 10.1.1

$\mathbb{R}^n$ is an inner product space with the dot product as inner product:

$$\langle \mathbf{v},\ \mathbf{w}\rangle = \mathbf{v}\cdot\mathbf{w} \quad \text{for all } \mathbf{v},\ \mathbf{w} \in \mathbb{R}^n$$

See Theorem 5.3.1. This is also called the **euclidean** inner product, and $\mathbb{R}^n$, equipped with the dot product, is called **euclidean n-space**.

Example 10.1.2

If A and B are $m \times n$ matrices, define $\langle A,\ B\rangle = \operatorname{tr}(AB^T)$ where $\operatorname{tr}(X)$ is the trace of the square matrix X. Show that $\langle\ ,\ \rangle$ is an inner product in $\mathbf{M}_{mn}$.

[1]If we regard $\mathbb{C}^n$ as a vector space over the field $\mathbb{C}$ of complex numbers, then the "standard inner product" on $\mathbb{C}^n$ defined in Section 8.7 does not satisfy Axiom P4 (see Theorem 8.7.1(3)).

Solution. P1 is clear. Since $\text{tr}(P) = \text{tr}(P^T)$ for every square matrix P, we have P2:

$$\langle A, B\rangle = \text{tr}(AB^T) = \text{tr}[(AB^T)^T] = \text{tr}(BA^T) = \langle B, A\rangle$$

Next, P3 and P4 follow because trace is a linear transformation $\mathbf{M}_{mn} \to \mathbb{R}$ (Exercise 10.1.19). Turning to P5, let $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_m$ denote the rows of the matrix A. Then the (i, j)-entry of AA^T is $\mathbf{r}_i \cdot \mathbf{r}_j$, so

$$\langle A, A\rangle = \text{tr}(AA^T) = \mathbf{r}_1 \cdot \mathbf{r}_1 + \mathbf{r}_2 \cdot \mathbf{r}_2 + \cdots + \mathbf{r}_m \cdot \mathbf{r}_m$$

But $\mathbf{r}_j \cdot \mathbf{r}_j$ is the sum of the squares of the entries of $\mathbf{r}_j$, so this shows that $\langle A, A\rangle$ is the sum of the squares of all nm entries of A. Axiom P5 follows.

The importance of the next example in analysis is difficult to overstate.

Example 10.1.3: [2]

Let $\mathbf{C}[a, b]$ denote the vector space of **continuous functions** from $[a, b]$ to $\mathbb{R}$, a subspace of $\mathbf{F}[a, b]$. Show that

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

defines an inner product on $\mathbf{C}[a, b]$.

Solution. Axioms P1 and P2 are clear. As to axiom P4,

$$\langle rf, g\rangle = \int_a^b rf(x)g(x)dx = r\int_a^b f(x)g(x)dx = r\langle f, g\rangle$$

Axiom P3 is similar. Finally, theorems of calculus show that $\langle f, f\rangle = \int_a^b f(x)^2 dx \geq 0$ and, if f is continuous, that this is zero if and only if f is the zero function. This gives axiom P5.

If $\mathbf{v}$ is any vector, then, using axiom P3, we get

$$\langle \mathbf{0}, \mathbf{v}\rangle = \langle \mathbf{0}+\mathbf{0}, \mathbf{v}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle + \langle \mathbf{0}, \mathbf{v}\rangle$$

and it follows that the number $\langle \mathbf{0}, \mathbf{v}\rangle$ must be zero. This observation is recorded for reference in the following theorem, along with several other properties of inner products. The other proofs are left as Exercise 10.1.20.

Theorem 10.1.1

Let $\langle\ ,\ \rangle$ be an inner product on a space V; let $\mathbf{v}$, $\mathbf{u}$, and $\mathbf{w}$ denote vectors in V; and let r denote a real number.

1. $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$
2. $\langle \mathbf{v}, r\mathbf{w}\rangle = r\langle \mathbf{v}, \mathbf{w}\rangle = \langle r\mathbf{v}, \mathbf{w}\rangle$

[2] This example (and others later that refer to it) can be omitted with no loss of continuity by students with no calculus background.

3. $\langle \mathbf{v}, \mathbf{0} \rangle = 0 = \langle \mathbf{0}, \mathbf{v} \rangle$

4. $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$

If $\langle\ ,\ \rangle$ is an inner product on a space V, then, given $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in V,

$$\langle r\mathbf{u}+s\mathbf{v},\ \mathbf{w}\rangle = \langle r\mathbf{u},\ \mathbf{w}\rangle + \langle s\mathbf{v},\ \mathbf{w}\rangle = r\langle \mathbf{u},\ \mathbf{w}\rangle + s\langle \mathbf{v},\ \mathbf{w}\rangle$$

for all r and s in $\mathbb{R}$ by axioms P3 and P4. Moreover, there is nothing special about the fact that there are two terms in the linear combination or that it is in the first component:

$$\begin{aligned} \langle r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_n\mathbf{v}_n,\ \mathbf{w}\rangle &= r_1\langle \mathbf{v}_1,\ \mathbf{w}\rangle + r_2\langle \mathbf{v}_2,\ \mathbf{w}\rangle + \cdots + r_n\langle \mathbf{v}_n,\ \mathbf{w}\rangle, \text{ and} \\ \langle \mathbf{v},\ s_1\mathbf{w}_1 + s_2\mathbf{w}_2 + \cdots + s_m\mathbf{w}_m\rangle &= s_1\langle \mathbf{v},\ \mathbf{w}_1\rangle + s_2\langle \mathbf{v},\ \mathbf{w}_2\rangle + \cdots + s_m\langle \mathbf{v},\ \mathbf{w}_m\rangle \end{aligned}$$

hold for all r_i and s_i in $\mathbb{R}$ and all $\mathbf{v}$, $\mathbf{w}$, $\mathbf{v}_i$, and $\mathbf{w}_j$ in V. These results are described by saying that inner products "preserve" linear combinations. For example,

$$\begin{aligned} \langle 2\mathbf{u}-\mathbf{v},\ 3\mathbf{u}+2\mathbf{v}\rangle &= \langle 2\mathbf{u},\ 3\mathbf{u}\rangle + \langle 2\mathbf{u},\ 2\mathbf{v}\rangle + \langle -\mathbf{v},\ 3\mathbf{u}\rangle + \langle -\mathbf{v},\ 2\mathbf{v}\rangle \\ &= 6\langle \mathbf{u},\ \mathbf{u}\rangle + 4\langle \mathbf{u},\ \mathbf{v}\rangle - 3\langle \mathbf{v},\ \mathbf{u}\rangle - 2\langle \mathbf{v},\ \mathbf{v}\rangle \\ &= 6\langle \mathbf{u},\ \mathbf{u}\rangle + \langle \mathbf{u},\ \mathbf{v}\rangle - 2\langle \mathbf{v},\ \mathbf{v}\rangle \end{aligned}$$

If A is a symmetric $n \times n$ matrix and $\mathbf{x}$ and $\mathbf{y}$ are columns in $\mathbb{R}^n$, we regard the 1×1 matrix $\mathbf{x}^TA\mathbf{y}$ as a number. If we write

$$\langle \mathbf{x},\ \mathbf{y}\rangle = \mathbf{x}^TA\mathbf{y} \quad \text{for all columns } \mathbf{x},\ \mathbf{y} \text{ in } \mathbb{R}^n$$

then axioms P1–P4 follow from matrix arithmetic (only P2 requires that A is symmetric). Axiom P5 reads

$$\mathbf{x}^TA\mathbf{x} > 0 \quad \text{for all columns } \mathbf{x} \neq \mathbf{0} \text{ in } \mathbb{R}^n$$

and this condition characterizes the positive definite matrices (Theorem 8.3.2). This proves the first assertion in the next theorem.

Theorem 10.1.2

If A is any $n \times n$ positive definite matrix, then

$$\langle \mathbf{x},\ \mathbf{y}\rangle = \mathbf{x}^TA\mathbf{y} \text{ for all columns } \mathbf{x},\ \mathbf{y} \text{ in } \mathbb{R}^n$$

defines an inner product on $\mathbb{R}^n$, and every inner product on $\mathbb{R}^n$ arises in this way.

Proof. Given an inner product $\langle\ ,\ \rangle$ on $\mathbb{R}^n$, let $\{\mathbf{e}_1,\ \mathbf{e}_2,\ \dots,\ \mathbf{e}_n\}$ be the standard basis of $\mathbb{R}^n$. If $\mathbf{x} = \sum_{i=1}^{n} x_i\mathbf{e}_i$ and $\mathbf{y} = \sum_{j=1}^{n} y_j\mathbf{e}_j$ are two vectors in $\mathbb{R}^n$, compute $\langle \mathbf{x},\ \mathbf{y}\rangle$ by adding the inner product of each term $x_i\mathbf{e}_i$ to each term $y_j\mathbf{e}_j$. The result is a double sum.

$$\langle \mathbf{x},\ \mathbf{y}\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\langle x_i\mathbf{e}_i,\ y_j\mathbf{e}_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\langle \mathbf{e}_i,\ \mathbf{e}_j\rangle y_j$$

As the reader can verify, this is a matrix product:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \left[\begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array} \right] \left[\begin{array}{cccc} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_2, \mathbf{e}_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \langle \mathbf{e}_n, \mathbf{e}_2 \rangle & \cdots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{array} \right] \left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array} \right]$$

Hence $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$, where A is the $n \times n$ matrix whose (i, j)-entry is $\langle \mathbf{e}_i, \mathbf{e}_j \rangle$. The fact that

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \langle \mathbf{e}_j, \mathbf{e}_i \rangle$$

shows that A is symmetric. Finally, A is positive definite by Theorem 8.3.2. □

Thus, just as every linear operator $\mathbb{R}^n \to \mathbb{R}^n$ corresponds to an $n \times n$ matrix, every inner product on $\mathbb{R}^n$ corresponds to a positive definite $n \times n$ matrix. In particular, the dot product corresponds to the identity matrix I_n.

Remark

If we refer to the inner product space $\mathbb{R}^n$ without specifying the inner product, we mean that the dot product is to be used.

Example 10.1.4

Let the inner product $\langle \ , \ \rangle$ be defined on $\mathbb{R}^2$ by

$$\left\langle \left[\begin{array}{c} v_1 \\ v_2 \end{array} \right], \left[\begin{array}{c} w_1 \\ w_2 \end{array} \right] \right\rangle = 2v_1w_1 - v_1w_2 - v_2w_1 + v_2w_2$$

Find a symmetric 2×2 matrix A such that $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A \mathbf{y}$ for all $\mathbf{x}$, $\mathbf{y}$ in $\mathbb{R}^2$.

Solution. The (i, j)-entry of the matrix A is the coefficient of v_iw_j in the expression, so $A = \left[\begin{array}{rr} 2 & -1 \\ -1 & 1 \end{array} \right]$. Incidentally, if $\mathbf{x} = \left[\begin{array}{c} x \\ y \end{array} \right]$, then

$$\langle \mathbf{x}, \mathbf{x} \rangle = 2x^2 - 2xy + y^2 = x^2 + (x - y)^2 \geq 0$$

for all $\mathbf{x}$, so $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ implies $\mathbf{x} = \mathbf{0}$. Hence $\langle \ , \ \rangle$ is indeed an inner product, so A is positive definite.

Let $\langle \ , \ \rangle$ be an inner product on $\mathbb{R}^n$ given as in Theorem 10.1.2 by a positive definite matrix A. If $\mathbf{x} = \left[\begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array} \right]^T$, then $\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^T A \mathbf{x}$ is an expression in the variables $x_1, x_2, \dots, x_n$ called a **quadratic form**. These are studied in detail in Section 8.9.

Norm and Distance

Definition 10.2 Norm and Distance

As in $\mathbb{R}^n$, if $\langle\ ,\ \rangle$ is an inner product on a space V, the ***norm***[3] *$\|\mathbf{v}\|$ of a vector $\mathbf{v}$ in V is defined by*

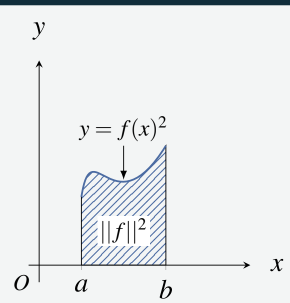
$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v},\ \mathbf{v}\rangle}$$

We define the ***distance*** *between vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V to be*

$$\mathrm{d}(\mathbf{v},\ \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\|$$

Note that axiom P5 guarantees that $\langle \mathbf{v},\ \mathbf{v}\rangle \geq 0$, so $\|\mathbf{v}\|$ is a real number.

Example 10.1.5



The norm of a continuous function $f = f(x)$ in $\mathbf{C}[a,\ b]$ (with the inner product from Example 10.1.3) is given by

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

Hence $\|f\|^2$ is the area beneath the graph of $y = f(x)^2$ between $x = a$ and $x = b$ (shaded in the diagram).

Example 10.1.6

Show that $\langle \mathbf{u}+\mathbf{v},\ \mathbf{u}-\mathbf{v}\rangle = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$ in any inner product space.

Solution.

$$\begin{aligned}\langle \mathbf{u}+\mathbf{v},\ \mathbf{u}-\mathbf{v}\rangle &= \langle \mathbf{u},\ \mathbf{u}\rangle - \langle \mathbf{u},\ \mathbf{v}\rangle + \langle \mathbf{v},\ \mathbf{u}\rangle - \langle \mathbf{v},\ \mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 - \langle \mathbf{u},\ \mathbf{v}\rangle + \langle \mathbf{u},\ \mathbf{v}\rangle - \|\mathbf{v}\|^2 \\ &= \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2\end{aligned}$$

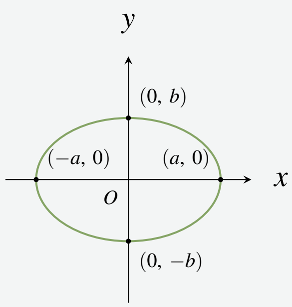
A vector $\mathbf{v}$ in an inner product space V is called a **unit vector** if $\|\mathbf{v}\| = 1$. The set of all unit vectors in V is called the **unit ball** in V. For example, if $V = \mathbb{R}^2$ (with the dot product) and $\mathbf{v} = (x,\ y)$, then

$$\|\mathbf{v}\|^2 = 1 \quad \text{if and only if} \quad x^2 + y^2 = 1$$

Hence the unit ball in $\mathbb{R}^2$ is the **unit circle** $x^2 + y^2 = 1$ with centre at the origin and radius 1. However, the shape of the unit ball varies with the choice of inner product.

[3]If the dot product is used in $\mathbb{R}^n$, the norm $\|\mathbf{x}\|$ of a vector $\mathbf{x}$ is usually called the **length** of $\mathbf{x}$.

Example 10.1.7



Let $a > 0$ and $b > 0$. If $\mathbf{v} = (x, y)$ and $\mathbf{w} = (x_1, y_1)$, define an inner product on $\mathbb{R}^2$ by

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{xx_1}{a^2} + \frac{yy_1}{b^2}$$

The reader can verify (Exercise 10.1.5) that this is indeed an inner product. In this case

$$\|\mathbf{v}\|^2 = 1 \quad \text{if and only if} \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

so the unit ball is the ellipse shown in the diagram.

Example 10.1.7 graphically illustrates the fact that norms and distances in an inner product space V vary with the choice of inner product in V.

Theorem 10.1.3

If $\mathbf{v} \neq \mathbf{0}$ is any vector in an inner product space V, then $\frac{1}{\|\mathbf{v}\|}\mathbf{v}$ is the unique unit vector that is a positive multiple of $\mathbf{v}$.

The next theorem reveals an important and useful fact about the relationship between norms and inner products, extending the Cauchy inequality for $\mathbb{R}^n$ (Theorem 5.3.2).

Theorem 10.1.4: Cauchy-Schwarz Inequality[4]

If $\mathbf{v}$ and $\mathbf{w}$ are two vectors in an inner product space V, then

$$\langle \mathbf{v}, \mathbf{w} \rangle^2 \leq \|\mathbf{v}\|^2 \|\mathbf{w}\|^2$$

Moreover, equality occurs if and only if one of $\mathbf{v}$ and $\mathbf{w}$ is a scalar multiple of the other.

Proof. Write $\|\mathbf{v}\| = a$ and $\|\mathbf{w}\| = b$. Using Theorem 10.1.1 we compute:

$$\begin{aligned} \|b\mathbf{v} - a\mathbf{w}\|^2 &= b^2\|\mathbf{v}\|^2 - 2ab\langle \mathbf{v}, \mathbf{w} \rangle + a^2\|\mathbf{w}\|^2 = 2ab(ab - \langle \mathbf{v}, \mathbf{w} \rangle) \\ \|b\mathbf{v} + a\mathbf{w}\|^2 &= b^2\|\mathbf{v}\|^2 + 2ab\langle \mathbf{v}, \mathbf{w} \rangle + a^2\|\mathbf{w}\|^2 = 2ab(ab + \langle \mathbf{v}, \mathbf{w} \rangle) \end{aligned} \tag{10.1}$$

It follows that $ab - \langle \mathbf{v}, \mathbf{w} \rangle \geq 0$ and $ab + \langle \mathbf{v}, \mathbf{w} \rangle \geq 0$, and hence that $-ab \leq \langle \mathbf{v}, \mathbf{w} \rangle \leq ab$. But then $|\langle \mathbf{v}, \mathbf{w} \rangle| \leq ab = \|\mathbf{v}\|\|\mathbf{w}\|$, as desired.

Conversely, if $|\langle \mathbf{v}, \mathbf{w} \rangle| = \|\mathbf{v}\|\|\mathbf{w}\| = ab$ then $\langle \mathbf{v}, \mathbf{w} \rangle = \pm ab$. Hence (10.1) shows that $b\mathbf{v} - a\mathbf{w} = \mathbf{0}$ or $b\mathbf{v} + a\mathbf{w} = \mathbf{0}$. It follows that one of $\mathbf{v}$ and $\mathbf{w}$ is a scalar multiple of the other, even if $a = 0$ or $b = 0$. □

[4]Hermann Amandus Schwarz (1843–1921) was a German mathematician at the University of Berlin. He had strong geometric intuition, which he applied with great ingenuity to particular problems. A version of the inequality appeared in 1885.

Example 10.1.8

If f and g are continuous functions on the interval $[a, b]$, then (see Example 10.1.3)

$$\left(\int_a^b f(x)g(x)dx\right)^2 \leq \int_a^b f(x)^2dx \int_a^b g(x)^2dx$$

Another famous inequality, the so-called *triangle inequality*, also comes from the Cauchy-Schwarz inequality. It is included in the following list of basic properties of the norm of a vector.

Theorem 10.1.5

If V is an inner product space, the norm $\|\cdot\|$ has the following properties.

1. $\|\mathbf{v}\| \geq 0$ *for every vector* $\mathbf{v}$ *in* V.
2. $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$.
3. $\|r\mathbf{v}\| = |r|\|\mathbf{v}\|$ *for every* $\mathbf{v}$ *in* V *and every* r *in* $\mathbb{R}$.
4. $\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V *(**triangle inequality**).*

Proof. Because $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$, properties (1) and (2) follow immediately from (3) and (4) of Theorem 10.1.1. As to (3), compute

$$\|r\mathbf{v}\|^2 = \langle r\mathbf{v}, r\mathbf{v}\rangle = r^2\langle \mathbf{v}, \mathbf{v}\rangle = r^2\|\mathbf{v}\|^2$$

Hence (3) follows by taking positive square roots. Finally, the fact that $\langle \mathbf{v}, \mathbf{w}\rangle \leq \|\mathbf{v}\|\|\mathbf{w}\|$ by the Cauchy-Schwarz inequality gives

$$\begin{aligned}\|\mathbf{v}+\mathbf{w}\|^2 = \langle \mathbf{v}+\mathbf{w}, \mathbf{v}+\mathbf{w}\rangle &= \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w}\rangle + \|\mathbf{w}\|^2 \\ &\leq \|\mathbf{v}\|^2 + 2\|\mathbf{v}\|\|\mathbf{w}\| + \|\mathbf{w}\|^2 \\ &= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2\end{aligned}$$

Hence (4) follows by taking positive square roots. □

It is worth noting that the usual triangle inequality for absolute values,

$$|r+s| \leq |r| + |s| \text{ for all real numbers } r \text{ and } s$$

is a special case of (4) where $V = \mathbb{R} = \mathbb{R}^1$ and the dot product $\langle r, s\rangle = rs$ is used.

In many calculations in an inner product space, it is required to show that some vector $\mathbf{v}$ is zero. This is often accomplished most easily by showing that its norm $\|\mathbf{v}\|$ is zero. Here is an example.

Example 10.1.9

Let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be a spanning set for an inner product space V. If $\mathbf{v}$ in V satisfies $\langle \mathbf{v}, \mathbf{v}_i\rangle = 0$ for each $i = 1, 2, \dots, n$, show that $\mathbf{v} = \mathbf{0}$.

Solution. Write $\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$, r_i in $\mathbb{R}$. To show that $\mathbf{v} = \mathbf{0}$, we show that $\|\mathbf{v}\|^2 = \langle \mathbf{v},\ \mathbf{v} \rangle = 0$. Compute:

$$\langle \mathbf{v},\ \mathbf{v} \rangle = \langle \mathbf{v},\ r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n \rangle = r_1\langle \mathbf{v},\ \mathbf{v}_1 \rangle + \cdots + r_n\langle \mathbf{v},\ \mathbf{v}_n \rangle = 0$$

by hypothesis, and the result follows.

The norm properties in Theorem 10.1.5 translate to the following properties of distance familiar from geometry. The proof is Exercise 10.1.21.

Theorem 10.1.6

Let V be an inner product space.

1. $\mathrm{d}(\mathbf{v},\ \mathbf{w}) \geq 0$ *for all* $\mathbf{v}$, $\mathbf{w}$ *in* V.
2. $\mathrm{d}(\mathbf{v},\ \mathbf{w}) = 0$ *if and only if* $\mathbf{v} = \mathbf{w}$.
3. $\mathrm{d}(\mathbf{v},\ \mathbf{w}) = \mathrm{d}(\mathbf{w},\ \mathbf{v})$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.
4. $\mathrm{d}(\mathbf{v},\ \mathbf{w}) \leq \mathrm{d}(\mathbf{v},\ \mathbf{u}) + \mathrm{d}(\mathbf{u},\ \mathbf{w})$ *for all* $\mathbf{v}$, $\mathbf{u}$, *and* $\mathbf{w}$ *in* V.

Exercises for 10.1

Exercise 10.1.1 In each case, determine which of axioms P1–P5 fail to hold.

a. $V = \mathbb{R}^2$, $\langle (x_1,\ y_1),\ (x_2,\ y_2) \rangle = x_1y_1x_2y_2$

b. $V = \mathbb{R}^3$,
$\langle (x_1,\ x_2,\ x_3),\ (y_1,\ y_2,\ y_3) \rangle = x_1y_1 - x_2y_2 + x_3y_3$

c. $V = \mathbb{C}$, $\langle z,\ w \rangle = z\overline{w}$, where $\overline{w}$ is complex conjugation

d. $V = \mathbf{P}_3$, $\langle p(x),\ q(x) \rangle = p(1)q(1)$

e. $V = \mathbf{M}_{22}$, $\langle A,\ B \rangle = \det(AB)$

f. $V = \mathbf{F}[0,\ 1]$, $\langle f,\ g \rangle = f(1)g(0) + f(0)g(1)$

Exercise 10.1.2 Let V be an inner product space. If $U \subseteq V$ is a subspace, show that U is an inner product space using the same inner product.

Exercise 10.1.3 In each case, find a scalar multiple of $\mathbf{v}$ that is a unit vector.

a. $\mathbf{v} = f$ in $\mathbf{C}[0,\ 1]$ where $f(x) = x^2$
$\langle f,\ g \rangle \int_0^1 f(x)g(x)dx$

b. $\mathbf{v} = f$ in $\mathbf{C}[-\pi,\ \pi]$ where $f(x) = \cos x$
$\langle f,\ g \rangle \int_{-\pi}^{\pi} f(x)g(x)dx$

c. $\mathbf{v} = \left[\begin{array}{c} 1 \\ 3 \end{array} \right]$ in $\mathbb{R}^2$ where $\langle \mathbf{v},\ \mathbf{w} \rangle = \mathbf{v}^T \left[\begin{array}{cc} 1 & 1 \\ 1 & 2 \end{array} \right] \mathbf{w}$

d. $\mathbf{v} = \left[\begin{array}{c} 3 \\ -1 \end{array} \right]$ in $\mathbb{R}^2$, $\langle \mathbf{v},\ \mathbf{w} \rangle = \mathbf{v}^T \left[\begin{array}{cc} 1 & -1 \\ -1 & 2 \end{array} \right] \mathbf{w}$

Exercise 10.1.4 In each case, find the distance between $\mathbf{u}$ and $\mathbf{v}$.

a. $\mathbf{u} = (3,\ -1,\ 2,\ 0)$, $\mathbf{v} = (1,\ 1,\ 1,\ 3)$; $\langle \mathbf{u},\ \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$

b. $\mathbf{u} = (1,\ 2,\ -1,\ 2)$, $\mathbf{v} = (2,\ 1,\ -1,\ 3)$; $\langle \mathbf{u},\ \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$

c. $\mathbf{u} = f$, $\mathbf{v} = g$ in $\mathbf{C}[0,\ 1]$ where $f(x) = x^2$ and $g(x) = 1 - x$; $\langle f,\ g \rangle = \int_0^1 f(x)g(x)dx$

d. $\mathbf{u} = f$, $\mathbf{v} = g$ in $\mathbf{C}[-\pi,\ \pi]$ where $f(x) = 1$ and $g(x) = \cos x$; $\langle f,\ g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$

Exercise 10.1.5 Let $a_1,\ a_2,\ \ldots,\ a_n$ be positive numbers. Given $\mathbf{v} = (v_1,\ v_2,\ \ldots,\ v_n)$ and $\mathbf{w} = (w_1,\ w_2,\ \ldots,\ w_n)$, define $\langle \mathbf{v},\ \mathbf{w} \rangle = a_1v_1w_1 + \cdots + a_nv_nw_n$. Show that this is an inner product on $\mathbb{R}^n$.

Exercise 10.1.6 If $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis of V and if $\mathbf{v} = v_1\mathbf{b}_1 + \cdots + v_n\mathbf{b}_n$ and $\mathbf{w} = w_1\mathbf{b}_1 + \cdots + w_n\mathbf{b}_n$ are vectors in V, define

$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1w_1 + \cdots + v_nw_n.$$

Show that this is an inner product on V.

Exercise 10.1.7 If $p = p(x)$ and $q = q(x)$ are polynomials in $\mathbf{P}_n$, define

$$\langle p, q \rangle = p(0)q(0) + p(1)q(1) + \cdots + p(n)q(n)$$

Show that this is an inner product on $\mathbf{P}_n$.
[*Hint for P5*: Theorem 6.5.4 or Appendix D.]

Exercise 10.1.8 Let $\mathbf{D}_n$ denote the space of all functions from the set $\{1, 2, 3, \dots, n\}$ to $\mathbb{R}$ with pointwise addition and scalar multiplication (see Exercise 6.3.35). Show that $\langle\ ,\ \rangle$ is an inner product on $\mathbf{D}_n$ if $\langle \mathbf{f}, \mathbf{g} \rangle = f(1)g(1) + f(2)g(2) + \cdots + f(n)g(n)$.

Exercise 10.1.9 Let $\text{re}\,(z)$ denote the real part of the complex number z. Show that $\langle\ ,\ \rangle$ is an inner product on $\mathbb{C}$ if $\langle \mathbf{z}, \mathbf{w} \rangle = \text{re}\,(z\overline{w})$.

Exercise 10.1.10 If $T : V \to V$ is an isomorphism of the inner product space V, show that

$$\langle \mathbf{v}, \mathbf{w} \rangle_1 = \langle T(\mathbf{v}), T(\mathbf{w}) \rangle$$

defines a new inner product $\langle\ ,\ \rangle_1$ on V.

Exercise 10.1.11 Show that every inner product $\langle\ ,\ \rangle$ on $\mathbb{R}^n$ has the form $\langle \mathbf{x}, \mathbf{y} \rangle = (U\mathbf{x}) \cdot (U\mathbf{y})$ for some upper triangular matrix U with positive diagonal entries. [*Hint*: Theorem 8.3.3.]

Exercise 10.1.12 In each case, show that $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T A\mathbf{w}$ defines an inner product on $\mathbb{R}^2$ and hence show that A is positive definite.

a. $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

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

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

d. $A = \begin{bmatrix} 3 & 4 \\ 4 & 6 \end{bmatrix}$

Exercise 10.1.13 In each case, find a symmetric matrix A such that $\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T A\mathbf{w}$.

a. $\left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\rangle = v_1w_1 + 2v_1w_2 + 2v_2w_1 + 5v_2w_2$

b. $\left\langle \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\rangle = v_1w_1 - v_1w_2 - v_2w_1 + 2v_2w_2$

c. $\left\langle \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \right\rangle = 2v_1w_1 + v_2w_2 + v_3w_3 - v_1w_2 - v_2w_1 + v_2w_3 + v_3w_2$

d. $\left\langle \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \right\rangle = v_1w_1 + 2v_2w_2 + 5v_3w_3 - 2v_1w_3 - 2v_3w_1$

Exercise 10.1.14 If A is symmetric and $\mathbf{x}^T A\mathbf{x} = 0$ for all columns $\mathbf{x}$ in $\mathbb{R}^n$, show that $A = 0$. [*Hint*: Consider $\langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y} \rangle$ where $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T A\mathbf{y}$.]

Exercise 10.1.15 Show that the sum of two inner products on V is again an inner product.

Exercise 10.1.16 Let $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$, $\|\mathbf{w}\| = \sqrt{3}$, $\langle \mathbf{u}, \mathbf{v} \rangle = -1$, $\langle \mathbf{u}, \mathbf{w} \rangle = 0$ and $\langle \mathbf{v}, \mathbf{w} \rangle = 3$. Compute:

a. $\langle \mathbf{v}+\mathbf{w}, 2\mathbf{u}-\mathbf{v} \rangle$

b. $\langle \mathbf{u}-2\mathbf{v}-\mathbf{w}, 3\mathbf{w}-\mathbf{v} \rangle$

Exercise 10.1.17 Given the data in Exercise 10.1.16, show that $\mathbf{u}+\mathbf{v} = \mathbf{w}$.

Exercise 10.1.18 Show that no vectors exist such that $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$, and $\langle \mathbf{u}, \mathbf{v} \rangle = -3$.

Exercise 10.1.19 Complete Example 10.1.2.

Exercise 10.1.20 Prove Theorem 10.1.1.

Exercise 10.1.21 Prove Theorem 10.1.6.

Exercise 10.1.22 Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in an inner product space V.

a. Expand $\langle 2\mathbf{u}-7\mathbf{v}, 3\mathbf{u}+5\mathbf{v} \rangle$.

b. Expand $\langle 3\mathbf{u}-4\mathbf{v}, 5\mathbf{u}+\mathbf{v} \rangle$.

c. Show that $\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2$.

d. Show that $\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2$.

Exercise 10.1.23 Show that

$$\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 = \frac{1}{2}\{\|\mathbf{v}+\mathbf{w}\|^2 + \|\mathbf{v}-\mathbf{w}\|^2\}$$

for any $\mathbf{v}$ and $\mathbf{w}$ in an inner product space.

Exercise 10.1.24 Let $\langle\ ,\ \rangle$ be an inner product on a vector space V. Show that the corresponding distance function is translation invariant. That is, show that $\text{d}\,(\mathbf{v}, \mathbf{w}) = \text{d}\,(\mathbf{v}+\mathbf{u}, \mathbf{w}+\mathbf{u})$ for all $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{u}$ in V.

Exercise 10.1.25

a. Show that $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}[\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2]$ for all $\mathbf{u}$, $\mathbf{v}$ in an inner product space V.

b. If $\langle\ ,\ \rangle$ and $\langle\ ,\ \rangle'$ are two inner products on V that have equal associated norm functions, show that $\langle \mathbf{u},\ \mathbf{v}\rangle = \langle \mathbf{u},\ \mathbf{v}\rangle'$ holds for all $\mathbf{u}$ and $\mathbf{v}$.

Exercise 10.1.26 Let $\mathbf{v}$ denote a vector in an inner product space V.

a. Show that $W = \{\mathbf{w} \mid \mathbf{w} \text{ in } V,\ \langle \mathbf{v},\ \mathbf{w} = 0\}$ is a subspace of V.

b. Let W be as in (a). If $V = \mathbb{R}^3$ with the dot product, and if $\mathbf{v} = (1,\ -1,\ 2)$, find a basis for W.

Exercise 10.1.27 Given vectors $\mathbf{w}_1,\ \mathbf{w}_2,\ \dots,\ \mathbf{w}_n$ and $\mathbf{v}$, assume that $\langle \mathbf{v},\ \mathbf{w}_i\rangle = 0$ for each i. Show that $\langle \mathbf{v},\ \mathbf{w}\rangle = 0$ for all $\mathbf{w}$ in span $\{\mathbf{w}_1,\ \mathbf{w}_2,\ \dots,\ \mathbf{w}_n\}$.

Exercise 10.1.28 If $V = \text{span}\,\{\mathbf{v}_1,\ \mathbf{v}_2,\ \dots,\ \mathbf{v}_n\}$ and $\langle \mathbf{v},\ \mathbf{v}_i\rangle = \langle \mathbf{w},\ \mathbf{v}_i\rangle$ holds for each i. Show that $\mathbf{v} = \mathbf{w}$.

Exercise 10.1.29 Use the Cauchy-Schwarz inequality in an inner product space to show that:

a. If $\|\mathbf{u}\| \leq 1$, then $\langle \mathbf{u},\ \mathbf{v}\rangle^2 \leq \|\mathbf{v}\|^2$ for all $\mathbf{v}$ in V.

b. $(x\cos\theta + y\sin\theta)^2 \leq x^2 + y^2$ for all real x, y, and θ.

c. $\|r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n\|^2 \leq [r_1\|\mathbf{v}_1\| + \cdots + r_n\|\mathbf{v}_n\|]^2$ for all vectors $\mathbf{v}_i$, and all $r_i > 0$ in $\mathbb{R}$.

Exercise 10.1.30 If A is a $2 \times n$ matrix, let $\mathbf{u}$ and $\mathbf{v}$ denote the rows of A.

a. Show that $AA^T = \left[\begin{array}{cc} \|\mathbf{u}\|^2 & \mathbf{u}\cdot\mathbf{v} \\ \mathbf{u}\cdot\mathbf{v} & \|\mathbf{v}\|^2 \end{array}\right]$.

b. Show that $\det(AA^T) \geq 0$.

Exercise 10.1.31

a. If $\mathbf{v}$ and $\mathbf{w}$ are nonzero vectors in an inner product space V, show that $-1 \leq \frac{\langle \mathbf{v},\ \mathbf{w}\rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} \leq 1$, and hence that a unique angle θ exists such that $\frac{\langle \mathbf{v},\ \mathbf{w}\rangle}{\|\mathbf{v}\|\|\mathbf{w}\|} = \cos\theta$ and $0 \leq \theta \leq \pi$. This angle θ is called the **angle between v** and **w**.

b. Find the angle between $\mathbf{v} = (1,\ 2,\ -1,\ 1\,3)$ and $\mathbf{w} = (2,\ 1,\ 0,\ 2,\ 0)$ in $\mathbb{R}^5$ with the dot product.

c. If θ is the angle between $\mathbf{v}$ and $\mathbf{w}$, show that the **law of cosines** is valid:

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

Exercise 10.1.32 If $V = \mathbb{R}^2$, define $\|(x,\ y)\| = |x| + |y|$.

a. Show that $\|\cdot\|$ satisfies the conditions in Theorem 10.1.5.

b. Show that $\|\cdot\|$ does not arise from an inner product on $\mathbb{R}^2$ given by a matrix A. [*Hint*: If it did, use Theorem 10.1.2 to find numbers a, b, and c such that $\|(x,\ y)\|^2 = ax^2 + bxy + cy^2$ for all x and y.]

10.2 Orthogonal Sets of Vectors

The idea that two lines can be perpendicular is fundamental in geometry, and this section is devoted to introducing this notion into a general inner product space V. To motivate the definition, recall that two nonzero geometric vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ are perpendicular (or orthogonal) if and only if $\mathbf{x}\cdot\mathbf{y} = 0$. In general, two vectors $\mathbf{v}$ and $\mathbf{w}$ in an inner product space V are said to be **orthogonal** if

$$\langle \mathbf{v},\ \mathbf{w}\rangle = 0$$

A set $\{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_n\}$ of vectors is called an **orthogonal set of vectors** if

1. *Each* $\mathbf{f}_i \neq \mathbf{0}$.
2. $\langle \mathbf{f}_i,\ \mathbf{f}_j\rangle = 0$ *for all* $i \neq j$.

If, in addition, $\|\mathbf{f}_i\| = 1$ for each i, the set $\{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_n\}$ is called an **orthonormal** set.

Example 10.2.1

$\{\sin x, \cos x\}$ is orthogonal in $\mathbf{C}[-\pi, \pi]$ because

$$\int_{-\pi}^{\pi} \sin x \cos x \, dx = \left[-\tfrac{1}{4}\cos 2x\right]_{-\pi}^{\pi} = 0$$

The first result about orthogonal sets extends Pythagoras' theorem in $\mathbb{R}^n$ (Theorem 5.3.4) and the same proof works.

Theorem 10.2.1: Pythagoras' Theorem

If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is an orthogonal set of vectors, then

$$\|\mathbf{f}_1 + \mathbf{f}_2 + \cdots + \mathbf{f}_n\|^2 = \|\mathbf{f}_1\|^2 + \|\mathbf{f}_2\|^2 + \cdots + \|\mathbf{f}_n\|^2$$

The proof of the next result is left to the reader.

Theorem 10.2.2

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ be an orthogonal set of vectors.

1. *$\{r_1\mathbf{f}_1, r_2\mathbf{f}_2, \dots, r_n\mathbf{f}_n\}$ is also orthogonal for any $r_i \neq 0$ in $\mathbb{R}$.*
2. *$\left\{\frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1, \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2, \dots, \frac{1}{\|\mathbf{f}_n\|}\mathbf{f}_n\right\}$ is an orthonormal set.*

As before, the process of passing from an orthogonal set to an orthonormal one is called **normalizing** the orthogonal set. The proof of Theorem 5.3.5 goes through to give

Theorem 10.2.3

Every orthogonal set of vectors is linearly independent.

Example 10.2.2

Show that $\left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \right\}$ is an orthogonal basis of $\mathbb{R}^3$ with inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T A \mathbf{w}, \quad \text{where } A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solution. We have

$$\left\langle \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\rangle = \begin{bmatrix} 2 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 0$$

and the reader can verify that the other pairs are orthogonal too. Hence the set is orthogonal, so it is linearly independent by Theorem 10.2.3. Because $\dim \mathbb{R}^3 = 3$, it is a basis.

The proof of Theorem 5.3.6 generalizes to give the following:

Theorem 10.2.4: Expansion Theorem

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ be an orthogonal basis of an inner product space V. If $\mathbf{v}$ is any vector in V, then

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}\mathbf{f}_n$$

is the expansion of $\mathbf{v}$ as a linear combination of the basis vectors.

The coefficients $\frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}, \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}, \dots, \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}$ in the expansion theorem are sometimes called the **Fourier coefficients** of $\mathbf{v}$ with respect to the orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$. This is in honour of the French mathematician J.B.J. Fourier (1768–1830). His original work was with a particular orthogonal set in the space $\mathbf{C}[a, b]$, about which there will be more to say in Section 10.5.

Example 10.2.3

If $a_0, a_1, \dots, a_n$ are distinct numbers and $p(x)$ and $q(x)$ are in $\mathbf{P}_n$, define

$$\langle p(x), q(x) \rangle = p(a_0)q(a_0) + p(a_1)q(a_1) + \cdots + p(a_n)q(a_n)$$

This is an inner product on $\mathbf{P}_n$. (Axioms P1–P4 are routinely verified, and P5 holds because 0 is the only polynomial of degree n with $n+1$ distinct roots. See Theorem 6.5.4 or Appendix D.) Recall that the **Lagrange polynomials** $\delta_0(x), \delta_1(x), \dots, \delta_n(x)$ relative to the numbers $a_0, a_1, \dots, a_n$ are defined as follows (see Section 6.5):

$$\delta_k(x) = \frac{\prod_{i \neq k}(x - a_i)}{\prod_{i \neq k}(a_k - a_i)} \quad k = 0, 1, 2, \dots, n$$

where $\prod_{i \neq k}(x - a_i)$ means the product of all the terms

$$(x - a_0), (x - a_1), (x - a_2), \dots, (x - a_n)$$

except that the kth term is omitted. Then $\{\delta_0(x), \delta_1(x), \dots, \delta_n(x)\}$ is orthonormal with respect to $\langle \ , \ \rangle$ because $\delta_k(a_i) = 0$ if $i \neq k$ and $\delta_k(a_k) = 1$. These facts also show that $\langle p(x), \delta_k(x) \rangle = p(a_k)$ so the expansion theorem gives

$$p(x) = p(a_0)\delta_0(x) + p(a_1)\delta_1(x) + \cdots + p(a_n)\delta_n(x)$$

for each $p(x)$ in $\mathbf{P}_n$. This is the **Lagrange interpolation expansion** of $p(x)$, Theorem 6.5.3, which is important in numerical integration.

Lemma 10.2.1: Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ *be an orthogonal set of vectors in an inner product space* V*, and let* $\mathbf{v}$ *be any vector not in* $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$*. Define*

$$\mathbf{f}_{m+1} = \mathbf{v} - \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ *is an orthogonal set of vectors.*

The proof of this result (and the next) is the same as for the dot product in $\mathbb{R}^n$ (Lemma 8.1.1 and Theorem 8.1.2).

Theorem 10.2.5: Gram-Schmidt Orthogonalization Algorithm

Let V *be an inner product space and let* $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ *be any basis of* V*. Define vectors* $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ *in* V *successively as follows:*

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{v}_1 \\
\mathbf{f}_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_3, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
\vdots &\qquad \vdots \\
\mathbf{f}_k &= \mathbf{v}_k - \frac{\langle \mathbf{v}_k, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle \mathbf{v}_k, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\langle \mathbf{v}_k, \mathbf{f}_{k-1} \rangle}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \dots, n$*. Then*

1. $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ *is an orthogonal basis of* V*.*
2. $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ *holds for each* $k = 1, 2, \dots, n$*.*

The purpose of the Gram-Schmidt algorithm is to convert a basis of an inner product space into an *orthogonal* basis. In particular, it shows that every finite dimensional inner product space *has* an orthogonal basis.

Example 10.2.4

Consider $V = \mathbf{P}_3$ with the inner product $\langle p, q \rangle = \int_{-1}^{1} p(x)q(x)dx$. If the Gram-Schmidt algorithm is applied to the basis $\{1, x, x^2, x^3\}$, show that the result is the orthogonal basis

$$\{1, x, \tfrac{1}{3}(3x^2 - 1), \tfrac{1}{5}(5x^3 - 3x)\}$$

Solution. Take $\mathbf{f}_1 = 1$. Then the algorithm gives

$$\begin{aligned}
\mathbf{f}_2 &= x - \frac{\langle x, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 = x - \frac{0}{2}\mathbf{f}_1 = x \\
\mathbf{f}_3 &= x^2 - \frac{\langle x^2, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\langle x^2, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&= x^2 - \frac{\frac{2}{3}}{2}1 - \frac{0}{\frac{2}{3}}x
\end{aligned}$$

$$= \frac{1}{3}(3x^2 - 1)$$

The verification that $\mathbf{f}_4 = \frac{1}{5}(5x^3 - 3x)$ is omitted.

The polynomials in Example 10.2.4 are such that the leading coefficient is 1 in each case. In other contexts (the study of differential equations, for example) it is customary to take multiples $p(x)$ of these polynomials such that $p(1) = 1$. The resulting orthogonal basis of $\mathbf{P}_3$ is

$$\{1,\ x,\ \frac{1}{3}(3x^2 - 1),\ \frac{1}{5}(5x^3 - 3x)\}$$

and these are the first four **Legendre polynomials**, so called to honour the French mathematician A. M. Legendre (1752–1833). They are important in the study of differential equations.

If V is an inner product space of dimension n, let $E = \{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_n\}$ be an orthonormal basis of V (by Theorem 10.2.5). If $\mathbf{v} = v_1\mathbf{f}_1 + v_2\mathbf{f}_2 + \cdots + v_n\mathbf{f}_n$ and $\mathbf{w} = w_1\mathbf{f}_1 + w_2\mathbf{f}_2 + \cdots + w_n\mathbf{f}_n$ are two vectors in V, we have $C_E(\mathbf{v}) = \left[\begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array}\right]^T$ and $C_E(\mathbf{w}) = \left[\begin{array}{cccc} w_1 & w_2 & \cdots & w_n \end{array}\right]^T$. Hence

$$\langle \mathbf{v},\ \mathbf{w} \rangle = \langle \sum_i v_i\mathbf{f}_i,\ \sum_j w_j\mathbf{f}_j \rangle = \sum_{i,\ j} v_i w_j \langle \mathbf{f}_i,\ \mathbf{f}_j \rangle = \sum_i v_i w_i = C_E(\mathbf{v}) \cdot C_E(\mathbf{w})$$

This shows that the coordinate isomorphism $C_E : V \to \mathbb{R}^n$ preserves inner products, and so proves

Corollary 10.2.1

If V is any n-dimensional inner product space, then V is isomorphic to $\mathbb{R}^n$ as inner product spaces. More precisely, if E is any orthonormal basis of V, the coordinate isomorphism

$$C_E : V \to \mathbb{R}^n \text{ satisfies } \langle \mathbf{v},\ \mathbf{w} \rangle = C_E(\mathbf{v}) \cdot C_E(\mathbf{w})$$

for all $\mathbf{v}$ and $\mathbf{w}$ in V.

The orthogonal complement of a subspace U of $\mathbb{R}^n$ was defined (in Chapter 8) to be the set of all vectors in $\mathbb{R}^n$ that are orthogonal to every vector in U. This notion has a natural extension in an arbitrary inner product space. Let U be a subspace of an inner product space V. As in $\mathbb{R}^n$, the **orthogonal complement** $U^\perp$ of U in V is defined by

$$U^\perp = \{\mathbf{v} \mid \mathbf{v} \in V,\ \langle \mathbf{v},\ \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \in U\}$$

Theorem 10.2.6

Let U be a finite dimensional subspace of an inner product space V.

1. *$U^\perp$ is a subspace of V and $V = U \oplus U^\perp$.*
2. *If $\dim V = n$, then $\dim U + \dim U^\perp = n$.*
3. *If $\dim V = n$, then $U^{\perp\perp} = U$.*

Proof.

1. $U^\perp$ is a subspace by Theorem 10.1.1. If $\mathbf{v}$ is in $U \cap U^\perp$, then $\langle \mathbf{v}, \mathbf{v} \rangle = 0$, so $\mathbf{v} = \mathbf{0}$ again by Theorem 10.1.1. Hence $U \cap U^\perp = \{\mathbf{0}\}$, and it remains to show that $U + U^\perp = V$. Given $\mathbf{v}$ in V, we must show that $\mathbf{v}$ is in $U + U^\perp$, and this is clear if $\mathbf{v}$ is in U. If $\mathbf{v}$ is not in U, let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ be an orthogonal basis of U. Then the orthogonal lemma shows that $\mathbf{v} - \left(\frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m \right)$ is in $U^\perp$, so $\mathbf{v}$ is in $U + U^\perp$ as required.

2. This follows from Theorem 9.3.6.

3. We have $\dim U^{\perp\perp} = n - \dim U^\perp = n - (n - \dim U) = \dim U$, using (2) twice. As $U \subseteq U^{\perp\perp}$ always holds (verify), (3) follows by Theorem 6.4.2. □

We digress briefly and consider a subspace U of an arbitrary vector space V. As in Section 9.3, if W is any complement of U in V, that is, $V = U \oplus W$, then each vector $\mathbf{v}$ in V has a *unique* representation as a sum $\mathbf{v} = \mathbf{u} + \mathbf{w}$ where $\mathbf{u}$ is in U and $\mathbf{w}$ is in W. Hence we may define a function $T : V \to V$ as follows:

$$T(\mathbf{v}) = \mathbf{u} \quad \text{where } \mathbf{v} = \mathbf{u} + \mathbf{w}, \ \mathbf{u} \text{ in } U, \ \mathbf{w} \text{ in } W$$

Thus, to compute $T(\mathbf{v})$, express $\mathbf{v}$ in any way at all as the sum of a vector $\mathbf{u}$ in U and a vector in W; then $T(\mathbf{v}) = \mathbf{u}$.

This function T is a linear operator on V. Indeed, if $\mathbf{v}_1 = \mathbf{u}_1 + \mathbf{w}_1$ where $\mathbf{u}_1$ is in U and $\mathbf{w}_1$ is in W, then $\mathbf{v} + \mathbf{v}_1 = (\mathbf{u} + \mathbf{u}_1) + (\mathbf{w} + \mathbf{w}_1)$ where $\mathbf{u} + \mathbf{u}_1$ is in U and $\mathbf{w} + \mathbf{w}_1$ is in W, so

$$T(\mathbf{v} + \mathbf{v}_1) = \mathbf{u} + \mathbf{u}_1 = T(\mathbf{v}) + T(\mathbf{v}_1)$$

Similarly, $T(a\mathbf{v}) = aT(\mathbf{v})$ for all a in $\mathbb{R}$, so T is a linear operator. Furthermore, $\text{im } T = U$ and $\ker T = W$ as the reader can verify, and T is called the **projection on U with kernel W**.

If U is a subspace of V, there are many projections on U, one for each complementary subspace W with $V = U \oplus W$. If V is an *inner product space*, we single out one for special attention. Let U be a finite dimensional subspace of an inner product space V.

Definition 10.3 Orthogonal Projection on a Subspace

The projection on U with kernel $U^\perp$ is called the ***orthogonal projection*** *on U (or simply the* ***projection*** *on U) and is denoted $\text{proj}_U : V \to V$.*

Theorem 10.2.7: Projection Theorem

Let U be a finite dimensional subspace of an inner product space V and let $\mathbf{v}$ be a vector in V.

1. $\text{proj}_U : V \to V$ *is a linear operator with image U and kernel $U^\perp$.*
2. $\text{proj}_U\ \mathbf{v}$ *is in U and $\mathbf{v} - \text{proj}_U\ \mathbf{v}$ is in $U^\perp$.*
3. *If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is any orthogonal basis of U, then*

$$\text{proj}_U\ \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle \mathbf{v}, \mathbf{f}_2 \rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

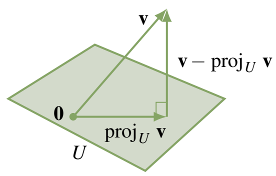
Proof. Only (3) remains to be proved. But since $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is an orthogonal basis of U and since $\text{proj}_U \mathbf{v}$ is in U, the result follows from the expansion theorem (Theorem 10.2.4) applied to the finite dimensional space U. □

Note that there is no requirement in Theorem 10.2.7 that V is finite dimensional.

Example 10.2.5

Let U be a subspace of the finite dimensional inner product space V. Show that $\text{proj}_{U^\perp} \mathbf{v} = \mathbf{v} - \text{proj}_U \mathbf{v}$ for all $\mathbf{v} \in V$.

Solution. We have $V = U^\perp \oplus U^{\perp\perp}$ by Theorem 10.2.6. If we write $\mathbf{p} = \text{proj}_U \mathbf{v}$, then $\mathbf{v} = (\mathbf{v} - \mathbf{p}) + \mathbf{p}$ where $\mathbf{v} - \mathbf{p}$ is in $U^\perp$ and $\mathbf{p}$ is in $U = U^{\perp\perp}$ by Theorem 10.2.7. Hence $\text{proj}_{U^\perp} \mathbf{v} = \mathbf{v} - \mathbf{p}$. See Exercise 8.1.7.



The vectors $\mathbf{v}$, $\text{proj}_U \mathbf{v}$, and $\mathbf{v} - \text{proj}_U \mathbf{v}$ in Theorem 10.2.7 can be visualized geometrically as in the diagram (where U is shaded and $\dim U = 2$). This suggests that $\text{proj}_U \mathbf{v}$ is the vector in U closest to $\mathbf{v}$. This is, in fact, the case.

Theorem 10.2.8: Approximation Theorem

*Let U be a finite dimensional subspace of an inner product space V. If $\mathbf{v}$ is any vector in V, then $\text{proj}_U \mathbf{v}$ is the vector in U that is closest to $\mathbf{v}$. Here **closest** means that*

$$\|\mathbf{v} - \text{proj}_U \mathbf{v}\| < \|\mathbf{v} - \mathbf{u}\|$$

for all $\mathbf{u}$ in U, $\mathbf{u} \neq \text{proj}_U \mathbf{v}$.

Proof. Write $\mathbf{p} = \text{proj}_U \mathbf{v}$, and consider $\mathbf{v} - \mathbf{u} = (\mathbf{v} - \mathbf{p}) + (\mathbf{p} - \mathbf{u})$. Because $\mathbf{v} - \mathbf{p}$ is in $U^\perp$ and $\mathbf{p} - \mathbf{u}$ is in U, Pythagoras' theorem gives

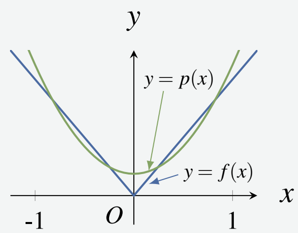
$$\|\mathbf{v} - \mathbf{u}\|^2 = \|\mathbf{v} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{u}\|^2 > \|\mathbf{v} - \mathbf{p}\|^2$$

because $\mathbf{p} - \mathbf{u} \neq 0$. The result follows. □

Example 10.2.6

Consider the space $\mathbf{C}[-1, 1]$ of real-valued continuous functions on the interval $[-1, 1]$ with inner product $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$. Find the polynomial $p = p(x)$ of degree at most 2 that best approximates the absolute-value function f given by $f(x) = |x|$.

Solution. Here we want the vector p in the subspace $U = \mathbf{P}_2$ of $\mathbf{C}[-1, 1]$ that is closest to f. In Example 10.2.4 the Gram-Schmidt algorithm was applied to give an orthogonal basis $\{\mathbf{f}_1 = 1, \mathbf{f}_2 = x, \mathbf{f}_3 = 3x^2 - 1\}$ of $\mathbf{P}_2$ (where, for convenience, we have changed $\mathbf{f}_3$ by a numerical factor). Hence the required polynomial is



$$\begin{aligned} p &= \text{proj}_{\mathbf{P}_2}\ f \\ &= \frac{\langle f,\ \mathbf{f}_1\rangle}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\langle f,\ \mathbf{f}_2\rangle}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \frac{\langle f,\ \mathbf{f}_3\rangle}{\|\mathbf{f}_3\|^2}\mathbf{f}_3 \\ &= \tfrac{1}{2}\mathbf{f}_1 + 0\mathbf{f}_2 + \tfrac{1/2}{8/5}\mathbf{f}_3 \\ &= \tfrac{3}{16}(5x^2+1) \end{aligned}$$

The graphs of $p(x)$ and $f(x)$ are given in the diagram.

If polynomials of degree at most n are allowed in Example 10.2.6, the polynomial in $\mathbf{P}_n$ is $\text{proj}_{\mathbf{P}_n}\ f$, and it is calculated in the same way. Because the subspaces $\mathbf{P}_n$ get larger as n increases, it turns out that the approximating polynomials $\text{proj}_{\mathbf{P}_n}\ f$ get closer and closer to f. In fact, solving many practical problems comes down to approximating some interesting vector $\mathbf{v}$ (often a function) in an infinite dimensional inner product space V by vectors in finite dimensional subspaces (which can be computed). If $U_1 \subseteq U_2$ are finite dimensional subspaces of V, then

$$\|\mathbf{v} - \text{proj}_{U_2}\ \mathbf{v}\| \leq \|\mathbf{v} - \text{proj}_{U_1}\ \mathbf{v}\|$$

by Theorem 10.2.8 (because $\text{proj}_{U_1}\ \mathbf{v}$ lies in U_1 and hence in U_2). Thus $\text{proj}_{U_2}\ \mathbf{v}$ is a better approximation to $\mathbf{v}$ than $\text{proj}_{U_1}\ \mathbf{v}$. Hence a general method in approximation theory might be described as follows: Given $\mathbf{v}$, use it to construct a sequence of finite dimensional subspaces

$$U_1 \subseteq U_2 \subseteq U_3 \subseteq \cdots$$

of V in such a way that $\|\mathbf{v} - \text{proj}_{U_k}\ \mathbf{v}\|$ approaches zero as k increases. Then $\text{proj}_{U_k}\ \mathbf{v}$ is a suitable approximation to $\mathbf{v}$ if k is large enough. For more information, the interested reader may wish to consult *Interpolation and Approximation* by Philip J. Davis (New York: Blaisdell, 1963).

Exercises for 10.2

Use the dot product in $\mathbb{R}^n$ unless otherwise instructed.

Exercise 10.2.1 In each case, verify that B is an orthogonal basis of V with the given inner product and use the expansion theorem to express $\mathbf{v}$ as a linear combination of the basis vectors.

a. $\mathbf{v} = \left[\begin{array}{c} a \\ b \end{array}\right], B = \left\{\left[\begin{array}{c} 1 \\ -1 \end{array}\right], \left[\begin{array}{c} 1 \\ 0 \end{array}\right]\right\}, V = \mathbb{R}^2,$
$\langle \mathbf{v},\ \mathbf{w}\rangle = \mathbf{v}^T A\mathbf{w}$ where $A = \left[\begin{array}{cc} 2 & 2 \\ 2 & 5 \end{array}\right]$

b. $\mathbf{v} = \left[\begin{array}{c} a \\ b \\ c \end{array}\right], B = \left\{\left[\begin{array}{c} 1 \\ 1 \\ 1 \end{array}\right], \left[\begin{array}{c} -1 \\ 0 \\ 1 \end{array}\right], \left[\begin{array}{c} 1 \\ -6 \\ 1 \end{array}\right]\right\},$
$V = \mathbb{R}^3$, $\langle \mathbf{v},\ \mathbf{w}\rangle = \mathbf{v}^T A\mathbf{w}$ where $A = \left[\begin{array}{ccc} 2 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 2 \end{array}\right]$

c. $\mathbf{v} = a + bx + cx^2$, $B = \{1x,\ 2-3x^2\}$, $V = \mathbf{P}_2$,
$\langle p,\ q\rangle = p(0)q(0) + p(1)q(1) + p(-1)q(-1)$

d. $\mathbf{v} = \left[\begin{array}{cc} a & b \\ c & d \end{array}\right]$,
$B = \left\{\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right], \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array}\right], \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right]\right\},$
$V = \mathbf{M}_{22}$, $\langle X,\ Y\rangle = \text{tr}\,(XY^T)$

Exercise 10.2.2 Let $\mathbb{R}^3$ have the inner product $\langle (x,\ y,\ z),\ (x',\ y',\ z')\rangle = 2xx' + yy' + 3zz'$. In each case, use the Gram-Schmidt algorithm to transform B into an orthogonal basis.

a. $B = \{(1,\ 1,\ 0),\ (1,\ 0,\ 1),\ (0,\ 1,\ 1)\}$

b. $B = \{(1,\ 1,\ 1),\ (1,\ -1,\ 1),\ (1,\ 1,\ 0)\}$

Exercise 10.2.3 Let $\mathbf{M}_{22}$ have the inner product $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$. In each case, use the Gram-Schmidt algorithm to transform B into an orthogonal basis.

a. $B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

b. $B = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$

Exercise 10.2.4 In each case, use the Gram-Schmidt process to convert the basis $B = \{1,\ x,\ x^2\}$ into an orthogonal basis of $\mathbf{P}_2$.

a. $\langle p,\ q\rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$

b. $\langle p,\ q\rangle = \int_0^2 p(x)q(x)dx$

Exercise 10.2.5 Show that $\{1,\ x - \frac{1}{2},\ x^2 - x + \frac{1}{6}\}$, is an orthogonal basis of $\mathbf{P}_2$ with the inner product

$$\langle p,\ q\rangle = \int_0^1 p(x)q(x)dx$$

and find the corresponding orthonormal basis.

Exercise 10.2.6 In each case find $U^\perp$ and compute $\dim U$ and $\dim U^\perp$.

a. $U = \operatorname{span}\{(1,\ 1,\ 2,\ 0),\ (3,\ -1,\ 2,\ 1),\ (1,\ -3,\ -2,\ 1)\}$ in $\mathbb{R}^4$

b. $U = \operatorname{span}\{(1,\ 1,\ 0,\ 0)\}$ in $\mathbb{R}^4$

c. $U = \operatorname{span}\{1,\ x\}$ in $\mathbf{P}_2$ with $\langle p,\ q\rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$

d. $U = \operatorname{span}\{x\}$ in $\mathbf{P}_2$ with $\langle p,\ q\rangle = \int_0^1 p(x)q(x)dx$

e. $U = \operatorname{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right\}$ in $\mathbf{M}_{22}$ with $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$

f. $U = \operatorname{span}\left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\}$ in $\mathbf{M}_{22}$ with $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$

Exercise 10.2.7 Let $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$ in $\mathbf{M}_{22}$. In each case find the matrix in U closest to A.

a. $U = \operatorname{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$, $A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$

b. $U = \operatorname{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right\}$, $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$

Exercise 10.2.8 In $\mathbf{P}_2$, let

$$\langle p(x),\ q(x)\rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$$

In each case find the polynomial in U closest to $f(x)$.

a. $U = \operatorname{span}\{1 + x,\ x^2\}$, $f(x) = 1 + x^2$

b. $U = \operatorname{span}\{1,\ 1 + x^2\}$; $f(x) = x$

Exercise 10.2.9 Using the inner product given by $\langle p,\ q\rangle = \int_0^1 p(x)q(x)dx$ on $\mathbf{P}_2$, write $\mathbf{v}$ as the sum of a vector in U and a vector in $U^\perp$.

a. $\mathbf{v} = x^2$, $U = \operatorname{span}\{x + 1,\ 9x - 5\}$

b. $\mathbf{v} = x^2 + 1$, $U = \operatorname{span}\{1,\ 2x - 1\}$

Exercise 10.2.10

a. Show that $\{\mathbf{u},\ \mathbf{v}\}$ is orthogonal if and only if $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$.

b. If $\mathbf{u} = \mathbf{v} = (1,\ 1)$ and $\mathbf{w} = (-1,\ 0)$, show that $\|\mathbf{u} + \mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$ but $\{\mathbf{u},\ \mathbf{v},\ \mathbf{w}\}$ is *not* orthogonal. Hence the converse to Pythagoras' theorem need not hold for more than two vectors.

Exercise 10.2.11 Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in an inner product space V. Show that:

a. $\mathbf{v}$ is orthogonal to $\mathbf{w}$ if and only if $\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v} - \mathbf{w}\|$.

b. $\mathbf{v} + \mathbf{w}$ and $\mathbf{v} - \mathbf{w}$ are orthogonal if and only if $\|\mathbf{v}\| = \|\mathbf{w}\|$.

Exercise 10.2.12 Let U and W be subspaces of an n-dimensional inner product space V. Suppose $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{u} \in U$ and $\mathbf{w} \in W$ and $\dim U + \dim W = n$. Show that $U^{\perp} = W$.

Exercise 10.2.13 If U and W are subspaces of an inner product space, show that $(U+W)^{\perp} = U^{\perp} \cap W^{\perp}$.

Exercise 10.2.14 If X is any set of vectors in an inner product space V, define

$$X^{\perp} = \{\mathbf{v} \mid \mathbf{v} \text{ in } V,\ \langle \mathbf{v}, \mathbf{x} \rangle = 0 \text{ for all } \mathbf{x} \text{ in } X\}$$

a. Show that $X^{\perp}$ is a subspace of V.

b. If $U = \text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$, show that $U^{\perp} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}^{\perp}$.

c. If $X \subseteq Y$, show that $Y^{\perp} \subseteq X^{\perp}$.

d. Show that $X^{\perp} \cap Y^{\perp} = (X \cup Y)^{\perp}$.

Exercise 10.2.15 If $\dim V = n$ and $\mathbf{w} \neq \mathbf{0}$ in V, show that $\dim\{\mathbf{v} \mid \mathbf{v} \text{ in } V,\ \langle \mathbf{v}, \mathbf{w} \rangle = 0\} = n-1$.

Exercise 10.2.16 If the Gram-Schmidt process is used on an orthogonal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V, show that $\mathbf{f}_k = \mathbf{v}_k$ holds for each $k = 1, 2, \dots, n$. That is, show that the algorithm reproduces the same basis.

Exercise 10.2.17 If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_{n-1}\}$ is orthonormal in an inner product space of dimension n, prove that there are exactly two vectors $\mathbf{f}_n$ such that $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_{n-1}, \mathbf{f}_n\}$ is an orthonormal basis.

Exercise 10.2.18 Let U be a finite dimensional subspace of an inner product space V, and let $\mathbf{v}$ be a vector in V.

a. Show that $\mathbf{v}$ lies in U if and only if $\mathbf{v} = \text{proj}_U(\mathbf{v})$.

b. If $V = \mathbb{R}^3$, show that $(-5, 4, -3)$ lies in $\text{span}\{(3, -2, 5), (-1, 1, 1)\}$ but that $(-1, 0, 2)$ does not.

Exercise 10.2.19 Let $\mathbf{n} \neq \mathbf{0}$ and $\mathbf{w} \neq \mathbf{0}$ be nonparallel vectors in $\mathbb{R}^3$ (as in Chapter 4).

a. Show that $\left\{\mathbf{n},\ \mathbf{n} \times \mathbf{w},\ \mathbf{w} - \frac{\mathbf{n} \cdot \mathbf{w}}{\|\mathbf{n}\|^2}\mathbf{n}\right\}$ is an orthogonal basis of $\mathbb{R}^3$.

b. Show that $\text{span}\left\{\mathbf{n} \times \mathbf{w},\ \mathbf{w} - \frac{\mathbf{n} \cdot \mathbf{w}}{\|\mathbf{n}\|^2}\mathbf{n}\right\}$ is the plane through the origin with normal $\mathbf{n}$.

Exercise 10.2.20 Let $E = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ be an orthonormal basis of V.

a. Show that $\langle \mathbf{v}, \mathbf{w} \rangle = C_E(\mathbf{v}) \cdot C_E(\mathbf{w})$ for all $\langle \mathbf{v}, \mathbf{w} \rangle$ in V.

b. If $P = [p_{ij}]$ is an $n \times n$ matrix, define $\mathbf{b}_i = p_{i1}\mathbf{f}_1 + \cdots + p_{in}\mathbf{f}_n$ for each i. Show that $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ is an orthonormal basis if and only if P is an orthogonal matrix.

Exercise 10.2.21 Let $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be an orthogonal basis of V. If $\mathbf{v}$ and $\mathbf{w}$ are in V, show that

$$\langle \mathbf{v}, \mathbf{w} \rangle = \frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle \langle \mathbf{w}, \mathbf{f}_1 \rangle}{\|\mathbf{f}_1\|^2} + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_n \rangle \langle \mathbf{w}, \mathbf{f}_n \rangle}{\|\mathbf{f}_n\|^2}$$

Exercise 10.2.22 Let $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be an orthonormal basis of V, and let $\mathbf{v} = v_1\mathbf{f}_1 + \cdots + v_n\mathbf{f}_n$ and $\mathbf{w} = w_1\mathbf{f}_1 + \cdots + w_n\mathbf{f}_n$. Show that

$$\langle \mathbf{v}, \mathbf{w} \rangle = v_1w_1 + \cdots + v_nw_n$$

and

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

(**Parseval's formula**).

Exercise 10.2.23 Let $\mathbf{v}$ be a vector in an inner product space V.

a. Show that $\|\mathbf{v}\| \geq \|\text{proj}_U \mathbf{v}\|$ holds for all finite dimensional subspaces U. [*Hint*: Pythagoras' theorem.]

b. If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ is any orthogonal set in V, prove **Bessel's inequality**:

$$\frac{\langle \mathbf{v}, \mathbf{f}_1 \rangle^2}{\|\mathbf{f}_1\|^2} + \cdots + \frac{\langle \mathbf{v}, \mathbf{f}_m \rangle^2}{\|\mathbf{f}_m\|^2} \leq \|\mathbf{v}\|^2$$

Exercise 10.2.24 Let $B = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ be an orthogonal basis of an inner product space V. Given $\mathbf{v} \in V$, let θ_i be the angle between $\mathbf{v}$ and $\mathbf{f}_i$ for each i (see Exercise 10.1.31). Show that

$$\cos^2\theta_1 + \cos^2\theta_2 + \cdots + \cos^2\theta_n = 1$$

[The $\cos\theta_i$ are called **direction cosines** for $\mathbf{v}$ corresponding to B.]

Exercise 10.2.25

a. Let S denote a set of vectors in a finite dimensional inner product space V, and suppose that $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for all $\mathbf{u}$ in S implies $\mathbf{v} = \mathbf{0}$. Show that $V = \text{span } S$. [*Hint*: Write $U = \text{span } S$ and use Theorem 10.2.6.]

b. Let $A_1, A_2, \dots, A_k$ be $n \times n$ matrices. Show that the following are equivalent.

   i. If $A_i\mathbf{b} = \mathbf{0}$ for all i (where $\mathbf{b}$ is a column in $\mathbb{R}^n$), then $\mathbf{b} = \mathbf{0}$.

   ii. The set of all rows of the matrices A_i spans $\mathbb{R}^n$.

Exercise 10.2.26 Let $[x_i) = (x_1, x_2, \dots)$ denote a sequence of real numbers x_i, and let

$$V = \{[x_i) \mid \text{ only finitely many } x_i \neq 0\}$$

Define componentwise addition and scalar multiplication on V as follows:

$$[x_i) + [y_i) = [x_i + y_i), \text{ and } a[x_i) = [ax_i) \text{ for } a \text{ in } \mathbb{R}.$$

Given $[x_i)$ and $[y_i)$ in V, define $\langle [x_i), [y_i) \rangle = \sum_{i=0}^{\infty} x_i y_i$. (Note that this makes sense since only finitely many x_i and y_i are nonzero.) Finally define

$$U = \{[x_i) \text{ in } V \mid \sum_{i=0}^{\infty} x_i = 0\}$$

a. Show that V is a vector space and that U is a subspace.

b. Show that $\langle \ , \ \rangle$ is an inner product on V.

c. Show that $U^\perp = \{\mathbf{0}\}$.

d. Hence show that $U \oplus U^\perp \neq V$ and $U \neq U^{\perp\perp}$.

10.3 Orthogonal Diagonalization

There is a natural way to define a symmetric linear operator T on a finite dimensional inner product space V. If T is such an operator, it is shown in this section that V has an orthogonal basis consisting of eigenvectors of T. This yields another proof of the principal axes theorem in the context of inner product spaces.

Theorem 10.3.1

Let $T : V \to V$ be a linear operator on a finite dimensional space V. Then the following conditions are equivalent.

1. *V has a basis consisting of eigenvectors of T.*
2. *There exists a basis B of V such that $M_B(T)$ is diagonal.*

Proof. We have $M_B(T) = \left[\begin{array}{cccc} C_B[T(\mathbf{b}_1)] & C_B[T(\mathbf{b}_2)] & \cdots & C_B[T(\mathbf{b}_n)] \end{array} \right]$ where $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ is any basis of V. By comparing columns:

$$M_B(T) = \left[\begin{array}{cccc} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{array} \right] \text{ if and only if } T(\mathbf{b}_i) = \lambda_i \mathbf{b}_i \text{ for each } i$$

Theorem 10.3.1 follows. □

Definition 10.4 Diagonalizable Linear Operators

A linear operator T on a finite dimensional space V is called ***diagonalizable*** *if V has a basis consisting of eigenvectors of T.*

Example 10.3.1

Let $T : \mathbf{P}_2 \to \mathbf{P}_2$ be given by

$$T(a+bx+cx^2) = (a+4c) - 2bx + (3a+2c)x^2$$

Find the eigenspaces of T and hence find a basis of eigenvectors.

Solution. If $B_0 = \{1,\ x,\ x^2\}$, then

$$M_{B_0}(T) = \left[\begin{array}{rrr} 1 & 0 & 4 \\ 0 & -2 & 0 \\ 3 & 0 & 2 \end{array} \right]$$

so $c_T(x) = (x+2)^2(x-5)$, and the eigenvalues of T are $\lambda = -2$ and $\lambda = 5$. One sees that $\left\{ \left[\begin{array}{r} 0 \\ 1 \\ 0 \end{array} \right], \left[\begin{array}{r} 4 \\ 0 \\ -3 \end{array} \right], \left[\begin{array}{r} 1 \\ 0 \\ 1 \end{array} \right] \right\}$ is a basis of eigenvectors of $M_{B_0}(T)$, so $B = \{x,\ 4-3x^2,\ 1+x^2\}$ is a basis of $\mathbf{P}_2$ consisting of eigenvectors of T.

If V is an inner product space, the expansion theorem gives a simple formula for the matrix of a linear operator with respect to an orthogonal basis.

Theorem 10.3.2

Let $T : V \to V$ be a linear operator on an inner product space V. If $B = \{\mathbf{b}_1,\ \mathbf{b}_2,\ \dots,\ \mathbf{b}_n\}$ is an orthogonal basis of V, then

$$M_B(T) = \left[\frac{\langle \mathbf{b}_i,\ T(\mathbf{b}_j) \rangle}{\|\mathbf{b}_i\|^2} \right]$$

Proof. Write $M_B(T) = \left[a_{ij} \right]$. The jth column of $M_B(T)$ is $C_B[T(\mathbf{e}_j)]$, so

$$T(\mathbf{b}_j) = a_{1j}\mathbf{b}_1 + \cdots + a_{ij}\mathbf{b}_i + \cdots + a_{nj}\mathbf{b}_n$$

On the other hand, the expansion theorem (Theorem 10.2.4) gives

$$\mathbf{v} = \frac{\langle \mathbf{b}_1,\ \mathbf{v} \rangle}{\|\mathbf{b}_1\|^2}\mathbf{b}_1 + \cdots + \frac{\langle \mathbf{b}_i,\ \mathbf{v} \rangle}{\|\mathbf{b}_i\|^2}\mathbf{b}_i + \cdots + \frac{\langle \mathbf{b}_n,\ \mathbf{v} \rangle}{\|\mathbf{b}_n\|^2}\mathbf{b}_n$$

for any $\mathbf{v}$ in V. The result follows by taking $\mathbf{v} = T(\mathbf{b}_j)$. □

Example 10.3.2

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$T(a, b, c) = (a + 2b - c, 2a + 3c, -a + 3b + 2c)$$

If the dot product in $\mathbb{R}^3$ is used, find the matrix of T with respect to the standard basis $B = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ where $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$.

Solution. The basis B is orthonormal, so Theorem 10.3.2 gives

$$M_B(T) = \left[\begin{array}{ccc} \mathbf{e}_1 \cdot T(\mathbf{e}_1) & \mathbf{e}_1 \cdot T(\mathbf{e}_2) & \mathbf{e}_1 \cdot T(\mathbf{e}_3) \\ \mathbf{e}_2 \cdot T(\mathbf{e}_1) & \mathbf{e}_2 \cdot T(\mathbf{e}_2) & \mathbf{e}_2 \cdot T(\mathbf{e}_3) \\ \mathbf{e}_3 \cdot T(\mathbf{e}_1) & \mathbf{e}_3 \cdot T(\mathbf{e}_2) & \mathbf{e}_3 \cdot T(\mathbf{e}_3) \end{array} \right] = \left[\begin{array}{rrr} 1 & 2 & -1 \\ 2 & 0 & 3 \\ -1 & 3 & 2 \end{array} \right]$$

Of course, this can also be found in the usual way.

It is not difficult to verify that an $n \times n$ matrix A is symmetric if and only if $\mathbf{x} \cdot (A\mathbf{y}) = (A\mathbf{x}) \cdot \mathbf{y}$ holds for all columns $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. The analog for operators is as follows:

Theorem 10.3.3

Let V be a finite dimensional inner product space. The following conditions are equivalent for a linear operator $T : V \to V$.

1. $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V.
2. *The matrix of T is symmetric with respect to every orthonormal basis of V.*
3. *The matrix of T is symmetric with respect to some orthonormal basis of V.*
4. *There is an orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ of V such that $\langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle = \langle T(\mathbf{f}_i), \mathbf{f}_j \rangle$ holds for all i and j.*

Proof. (1) $\Rightarrow$ (2). Let $B = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be an orthonormal basis of V, and write $M_B(T) = \left[a_{ij} \right]$. Then $a_{ij} = \langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle$ by Theorem 10.3.2. Hence (1) and axiom P2 give

$$a_{ij} = \langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle = \langle T(\mathbf{f}_i), \mathbf{f}_j \rangle = \langle \mathbf{f}_j, T(\mathbf{f}_i) \rangle = a_{ji}$$

for all i and j. This shows that $M_B(T)$ is symmetric.

(2) $\Rightarrow$ (3). This is clear.

(3) $\Rightarrow$ (4). Let $B = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be an orthonormal basis of V such that $M_B(T)$ is symmetric. By (3) and Theorem 10.3.2, $\langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle = \langle \mathbf{f}_j, T(\mathbf{f}_i) \rangle$ for all i and j, so (4) follows from axiom P2.

(4) $\Rightarrow$ (1). Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in V and write them as $\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{f}_i$ and $\mathbf{w} = \sum_{j=1}^{n} w_j \mathbf{f}_j$. Then

$$\langle \mathbf{v}, T(\mathbf{w}) \rangle = \left\langle \sum_i v_i \mathbf{f}_i, \sum_j w_j T\mathbf{f}_j \right\rangle = \sum_i \sum_j v_i w_j \langle \mathbf{f}_i, T(\mathbf{f}_j) \rangle$$

$$= \sum_i \sum_j v_i w_j \langle T(\mathbf{f}_i), \mathbf{f}_j \rangle$$
$$= \left\langle \sum_i v_i T(\mathbf{f}_i), \sum_j w_j \mathbf{f}_j \right\rangle$$
$$= \langle T(\mathbf{v}), \mathbf{w} \rangle$$

where we used (4) at the third stage. This proves (1). □

A linear operator T on an inner product space V is called **symmetric** if $\langle \mathbf{v}, T(\mathbf{w}) \rangle = \langle T(\mathbf{v}), \mathbf{w} \rangle$ holds for all $\mathbf{v}$ and $\mathbf{w}$ in V.

Example 10.3.3

If A is an $n \times n$ matrix, let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the matrix operator given by $T_A(\mathbf{v}) = A\mathbf{v}$ for all columns $\mathbf{v}$. If the dot product is used in $\mathbb{R}^n$, then T_A is a symmetric operator if and only if A is a symmetric matrix.

Solution. If E is the standard basis of $\mathbb{R}^n$, then E is orthonormal when the dot product is used. We have $M_E(T_A) = A$ (by Example 9.1.4), so the result follows immediately from part (3) of Theorem 10.3.3.

It is important to note that whether an operator is symmetric depends on which inner product is being used (see Exercise 10.3.2).

If V is a finite dimensional inner product space, the eigenvalues of an operator $T : V \to V$ are the same as those of $M_B(T)$ for any orthonormal basis B (see Theorem 9.3.3). If T is symmetric, $M_B(T)$ is a symmetric matrix and so has real eigenvalues by Theorem 5.5.7. Hence we have the following:

Theorem 10.3.4

A symmetric linear operator on a finite dimensional inner product space has real eigenvalues.

If U is a subspace of an inner product space V, recall that its orthogonal complement is the subspace $U^\perp$ of V defined by

$$U^\perp = \{\mathbf{v} \text{ in } V \mid \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ for all } \mathbf{u} \text{ in } U\}$$

Theorem 10.3.5

Let $T : V \to V$ be a symmetric linear operator on an inner product space V, and let U be a T-invariant subspace of V. Then:

1. *The restriction of T to U is a symmetric linear operator on U.*
2. *$U^\perp$ is also T-invariant.*

Proof.

1. U is itself an inner product space using the same inner product, and condition 1 in Theorem 10.3.3 that T is symmetric is clearly preserved.

2. If $\mathbf{v}$ is in $U^{\perp}$, our task is to show that $T(\mathbf{v})$ is also in $U^{\perp}$; that is, $\langle T(\mathbf{v}), \mathbf{u}\rangle = 0$ for all $\mathbf{u}$ in U. But if $\mathbf{u}$ is in U, then $T(\mathbf{u})$ also lies in U because U is T-invariant, so

$$\langle T(\mathbf{v}), \mathbf{u}\rangle = \langle \mathbf{v}, T(\mathbf{u})\rangle$$

using the symmetry of T and the definition of $U^{\perp}$. □

The principal axes theorem (Theorem 8.2.2) asserts that an $n \times n$ matrix A is symmetric if and only if $\mathbb{R}^n$ has an orthogonal basis of eigenvectors of A. The following result not only extends this theorem to an arbitrary n-dimensional inner product space, but the proof is much more intuitive.

Theorem 10.3.6: Principal Axes Theorem

The following conditions are equivalent for a linear operator T on a finite dimensional inner product space V.

1. *T is symmetric.*
2. *V has an orthogonal basis consisting of eigenvectors of T.*

Proof. (1) $\Rightarrow$ (2). Assume that T is symmetric and proceed by induction on $n = \dim V$. If $n = 1$, *every* nonzero vector in V is an eigenvector of T, so there is nothing to prove. If $n \geq 2$, assume inductively that the theorem holds for spaces of dimension less than n. Let λ_1 be a real eigenvalue of T (by Theorem 10.3.4) and choose an eigenvector $\mathbf{f}_1$ corresponding to λ_1. Then $U = \mathbb{R}\mathbf{f}_1$ is T-invariant, so $U^{\perp}$ is also T-invariant by Theorem 10.3.5 (T is symmetric). Because $\dim U^{\perp} = n-1$ (Theorem 10.2.6), and because the restriction of T to $U^{\perp}$ is a symmetric operator (Theorem 10.3.5), it follows by induction that $U^{\perp}$ has an orthogonal basis $\{\mathbf{f}_2, \dots, \mathbf{f}_n\}$ of eigenvectors of T. Hence $B = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is an orthogonal basis of V, which proves (2).

(2) $\Rightarrow$ (1). If $B = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ is a basis as in (2), then $M_B(T)$ is symmetric (indeed diagonal), so T is symmetric by Theorem 10.3.3. □

The matrix version of the principal axes theorem is an immediate consequence of Theorem 10.3.6. If A is an $n \times n$ symmetric matrix, then $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is a symmetric operator, so let B be an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of T_A (and hence of A). Then P^TAP is diagonal where P is the orthogonal matrix whose columns are the vectors in B (see Theorem 9.2.4).

Similarly, let $T : V \to V$ be a symmetric linear operator on the n-dimensional inner product space V and let B_0 be any convenient orthonormal basis of V. Then an orthonormal basis of eigenvectors of T can be computed from $M_{B_0}(T)$. In fact, if $P^T M_{B_0}(T)P$ is diagonal where P is orthogonal, let $B = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be the vectors in V such that $C_{B_0}(\mathbf{f}_j)$ is column j of P for each j. Then B consists of eigenvectors of T by Theorem 9.3.3, and they are orthonormal because B_0 is orthonormal. Indeed

$$\langle \mathbf{f}_i, \mathbf{f}_j\rangle = C_{B_0}(\mathbf{f}_i) \cdot C_{B_0}(\mathbf{f}_j)$$

holds for all i and j, as the reader can verify. Here is an example.

Example 10.3.4

Let $T : \mathbf{P}_2 \to \mathbf{P}_2$ be given by

$$T(a+bx+cx^2) = (8a-2b+2c) + (-2a+5b+4c)x + (2a+4b+5c)x^2$$

Using the inner product $\langle a+bx+cx^2,\ a'+b'x+c'x^2\rangle = aa'+bb'+cc'$, show that T is symmetric and find an orthonormal basis of $\mathbf{P}_2$ consisting of eigenvectors.

Solution. If $B_0 = \{1,\ x,\ x^2\}$, then $M_{B_0}(T) = \left[\begin{array}{rrr} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{array}\right]$ is symmetric, so T is symmetric. This matrix was analyzed in Example 8.2.5, where it was found that an *orthonormal* basis of eigenvectors is $\left\{\frac{1}{3}\left[\begin{array}{rrr} 1 & 2 & -2 \end{array}\right]^T,\ \frac{1}{3}\left[\begin{array}{rrr} 2 & 1 & 2 \end{array}\right]^T,\ \frac{1}{3}\left[\begin{array}{rrr} -2 & 2 & 1 \end{array}\right]^T\right\}$. Because B_0 is orthonormal, the corresponding orthonormal basis of $\mathbf{P}_2$ is

$$B = \left\{\tfrac{1}{3}(1+2x-2x^2),\ \tfrac{1}{3}(2+x+2x^2),\ \tfrac{1}{3}(-2+2x+x^2)\right\}$$

Exercises for 10.3

Exercise 10.3.1 In each case, show that T is symmetric by calculating $M_B(T)$ for some orthonormal basis B.

a. $T : \mathbb{R}^3 \to \mathbb{R}^3$;
$T(a,\ b,\ c) = (a-2b,\ -2a+2b+2c,\ 2b-c)$; dot product

b. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$;
$T\left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = \left[\begin{array}{cc} c-a & d-b \\ a+2c & b+2d \end{array}\right]$;
inner product:
$\left\langle \left[\begin{array}{cc} x & y \\ z & w \end{array}\right], \left[\begin{array}{cc} x' & y' \\ z' & w' \end{array}\right] \right\rangle = xx'+yy'+zz'+ww'$

c. $T : \mathbf{P}_2 \to \mathbf{P}_2$;
$T(a+bx+cx^2) = (b+c)+(a+c)x+(a+b)x^2$;
inner product:
$\langle a+bx+cx^2,\ a'+b'x+c'x^2\rangle = aa'+bb'+cc'$

Exercise 10.3.2 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$T(a,\ b) = (2a+b,\ a-b).$$

a. Show that T is symmetric if the dot product is used.

b. Show that T is *not* symmetric if $\langle \mathbf{x},\ \mathbf{y}\rangle = \mathbf{x}A\mathbf{y}^T$, where $A = \left[\begin{array}{cc} 1 & 1 \\ 1 & 2 \end{array}\right]$.

[*Hint*: Check that $B = \{(1,\ 0),\ (1,\ -1)\}$ is an orthonormal basis.]

Exercise 10.3.3 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$T(a,\ b) = (a-b,\ b-a)$$

Use the dot product in $\mathbb{R}^2$.

a. Show that T is symmetric.

b. Show that $M_B(T)$ is *not* symmetric if the orthogonal basis $B = \{(1,\ 0),\ (0,\ 2)\}$ is used. Why does this not contradict Theorem 10.3.3?

Exercise 10.3.4 Let V be an n-dimensional inner product space, and let T and S denote symmetric linear operators on V. Show that:

a. The identity operator is symmetric.

b. rT is symmetric for all r in $\mathbb{R}$.

c. $S+T$ is symmetric.

d. If T is invertible, then T^{-1} is symmetric.

e. If $ST = TS$, then ST is symmetric.

Exercise 10.3.5 In each case, show that T is symmetric and find an orthonormal basis of eigenvectors of T.

a. $T : \mathbb{R}^3 \to \mathbb{R}^3$;
$T(a,\ b,\ c) = (2a+2c,\ 3b,\ 2a+5c)$; use the dot product

b. $T : \mathbb{R}^3 \to \mathbb{R}^3$;
$T(a,\ b,\ c) = (7a-b,\ -a+7b,\ 2c)$; use the dot product

c. $T : \mathbf{P}_2 \to \mathbf{P}_2$;
$T(a+bx+cx^2) = 3b+(3a+4c)x+4bx^2$;
inner product
$\langle a+bx+cx^2,\ a'+b'x+c'x^2\rangle = aa'+bb'+cc'$

d. $T : \mathbf{P}_2 \to \mathbf{P}_2$;
$T(a+bx+cx^2) = (c-a)+3bx+(a-c)x^2$; inner product as in part (c)

Exercise 10.3.6 If A is any $n\times n$ matrix, let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be given by $T_A(\mathbf{x}) = A\mathbf{x}$. Suppose an inner product on $\mathbb{R}^n$ is given by $\langle \mathbf{x},\ \mathbf{y}\rangle = \mathbf{x}^T P\mathbf{y}$, where P is a positive definite matrix.

a. Show that T_A is symmetric if and only if $PA = A^T P$.

b. Use part (a) to deduce Example 10.3.3.

Exercise 10.3.7 Let $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ be given by $T(X) = AX$, where A is a fixed 2×2 matrix.

a. Compute $M_B(T)$, where
$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$.
Note the order!

b. Show that $c_T(x) = [c_A(x)]^2$.

c. If the inner product on $\mathbf{M}_{22}$ is $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$, show that T is symmetric if and only if A is a symmetric matrix.

Exercise 10.3.8 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by

$$T(a,\ b) = (b-a,\ a+2b)$$

Show that T is symmetric if the dot product is used in $\mathbb{R}^2$ but that it is not symmetric if the following inner product is used:

$$\langle \mathbf{x},\ \mathbf{y}\rangle = \mathbf{x}A\mathbf{y}^T,\ A = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

Exercise 10.3.9 If $T : V \to V$ is symmetric, write $T^{-1}(W) = \{\mathbf{v} \mid T(\mathbf{v}) \text{ is in } W\}$. Show that $T(U)^\perp = T^{-1}(U^\perp)$ holds for every subspace U of V.

Exercise 10.3.10 Let $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$ be defined by $T(X) = PXQ$, where P and Q are nonzero 2×2 matrices. Use the inner product $\langle X,\ Y\rangle = \operatorname{tr}(XY^T)$. Show that T is symmetric if and only if either P and Q are both symmetric or both are scalar multiples of $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. [*Hint*: If B is as in part (a) of Exercise 10.3.7, then $M_B(T) = \begin{bmatrix} aP & cP \\ bP & dP \end{bmatrix}$ in block form, where $Q = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
If $B_0 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$, then $M_B(T) = \begin{bmatrix} pQ^T & qQ^T \\ rQ^T & sQ^T \end{bmatrix}$, where $P = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$. Use the fact that $cP = bP^T \Rightarrow (c^2-b^2)P = 0$.]

Exercise 10.3.11 Let $T : V \to W$ be any linear transformation and let $B = \{\mathbf{b}_1,\ \dots,\ \mathbf{b}_n\}$ and $D = \{\mathbf{d}_1,\ \dots,\ \mathbf{d}_m\}$ be bases of V and W, respectively. If W is an inner product space and D is orthogonal, show that

$$M_{DB}(T) = \begin{bmatrix} \frac{\langle \mathbf{d}_i,\ T(\mathbf{b}_j)\rangle}{\|\mathbf{d}_i\|^2} \end{bmatrix}$$

This is a generalization of Theorem 10.3.2.

Exercise 10.3.12 Let $T : V \to V$ be a linear operator on an inner product space V of finite dimension. Show that the following are equivalent.

1. $\langle \mathbf{v},\ T(\mathbf{w})\rangle = -\langle T(\mathbf{v}),\ \mathbf{w}\rangle$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.
2. $M_B(T)$ is skew-symmetric for every orthonormal basis B.
3. $M_B(T)$ is skew-symmetric for some orthonormal basis B.

Such operators T are called **skew-symmetric** operators.

Exercise 10.3.13 Let $T : V \to V$ be a linear operator on an n-dimensional inner product space V.

a. Show that T is symmetric if and only if it satisfies the following two conditions.

   i. $c_T(x)$ factors completely over $\mathbb{R}$.

   ii. If U is a T-invariant subspace of V, then $U^\perp$ is also T-invariant.

b. Using the standard inner product on $\mathbb{R}^2$, show that $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T(a,\ b) = (a,\ a+b)$ satisfies condition (i) and that $S : \mathbb{R}^2 \to \mathbb{R}^2$ with $S(a,\ b) = (b,\ -a)$ satisfies condition (ii), but that neither is symmetric. (Example 9.3.4 is useful for S.)

[*Hint for part* (a): If conditions (i) and (ii) hold, proceed by induction on n. By condition (i), let $\mathbf{e}_1$ be an eigenvector of T. If $U = \mathbb{R}\mathbf{e}_1$, then $U^\perp$ is T-invariant by condition (ii), so show that the restriction of T to $U^\perp$ satisfies conditions (i) and (ii). (Theorem 9.3.1 is helpful for part (i)). Then apply induction to show that V has an orthogonal basis of eigenvectors (as in Theorem 10.3.6)].

Exercise 10.3.14 Let $B = \{\mathbf{f}_1,\ \mathbf{f}_2,\ \dots,\ \mathbf{f}_n\}$ be an orthonormal basis of an inner product space V. Given $T : V \to V$, define $T' : V \to V$ by

$$\begin{aligned} T'(\mathbf{v}) &= \langle \mathbf{v},\ T(\mathbf{f}_1)\rangle \mathbf{f}_1 + \langle \mathbf{v},\ T(\mathbf{f}_2)\rangle \mathbf{f}_2 + \cdots + \langle \mathbf{v},\ T(\mathbf{f}_n)\rangle \mathbf{f}_n \\ &= \sum_{i=1}^{n} \langle \mathbf{v},\ T(\mathbf{f}_i)\rangle \mathbf{f}_i \end{aligned}$$

a. Show that $(aT)' = aT'$.

b. Show that $(S+T)' = S' + T'$.

c. Show that $M_B(T')$ is the transpose of $M_B(T)$.

d. Show that $(T')' = T$, using part (c). [*Hint*: $M_B(S) = M_B(T)$ implies that $S = T$.]

e. Show that $(ST)' = T'S'$, using part (c).

f. Show that T is symmetric if and only if $T = T'$. [*Hint*: Use the expansion theorem and Theorem 10.3.3.]

g. Show that $T + T'$ and TT' are symmetric, using parts (b) through (e).

h. Show that $T'(\mathbf{v})$ is independent of the choice of orthonormal basis B. [*Hint*: If $D = \{\mathbf{g}_1,\ \dots,\ \mathbf{g}_n\}$ is also orthonormal, use the fact that $\mathbf{f}_i = \sum_{j=1}^{n} \langle \mathbf{f}_i,\ \mathbf{g}_j\rangle \mathbf{g}_j$ for each i.]

Exercise 10.3.15 Let V be a finite dimensional inner product space. Show that the following conditions are equivalent for a linear operator $T : V \to V$.

1. T is symmetric and $T^2 = T$.

2. $M_B(T) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ for some orthonormal basis B of V.

   An operator is called a **projection** if it satisfies these conditions. [*Hint*: If $T^2 = T$ and $T(\mathbf{v}) = \lambda \mathbf{v}$, apply T to get $\lambda \mathbf{v} = \lambda^2 \mathbf{v}$. Hence show that 0, 1 are the only eigenvalues of T.]

Exercise 10.3.16 Let V denote a finite dimensional inner product space. Given a subspace U, define $\text{proj}_U : V \to V$ as in Theorem 10.2.7.

a. Show that proj_U is a projection in the sense of Exercise 10.3.15.

b. If T is any projection, show that $T = \text{proj}_U$, where $U = \text{im}\, T$. [*Hint*: Use $T^2 = T$ to show that $V = \text{im}\, T \oplus \ker T$ and $T(\mathbf{u}) = \mathbf{u}$ for all $\mathbf{u}$ in $\text{im}\, T$. Use the fact that T is symmetric to show that $\ker T \subseteq (\text{im}\, T)^\perp$ and hence that these are equal because they have the same dimension.]

10.4 Isometries

We saw in Section 2.6 that rotations about the origin and reflections in a line through the origin are linear operators on $\mathbb{R}^2$. Similar geometric arguments (in Section 4.4) establish that, in $\mathbb{R}^3$, rotations about a line through the origin and reflections in a plane through the origin are linear. We are going to give an algebraic proof of these results that is valid in any inner product space. The key observation is that reflections and rotations are distance preserving in the following sense. If V is an inner product space, a transformation $S : V \to V$ (not necessarily linear) is said to be **distance preserving** if the distance between $S(\mathbf{v})$ and $S(\mathbf{w})$ is the same as the distance between $\mathbf{v}$ and $\mathbf{w}$ for all vectors $\mathbf{v}$ and $\mathbf{w}$; more formally, if

$$\|S(\mathbf{v}) - S(\mathbf{w})\| = \|\mathbf{v} - \mathbf{w}\| \quad \text{for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } V \tag{10.2}$$

Distance-preserving maps need not be linear. For example, if $\mathbf{u}$ is any vector in V, the transformation $S_{\mathbf{u}} : V \to V$ defined by $S_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} + \mathbf{u}$ for all $\mathbf{v}$ in V is called **translation** by $\mathbf{u}$, and it is routine to verify that $S_{\mathbf{u}}$ is distance preserving for any $\mathbf{u}$. However, $S_{\mathbf{u}}$ is linear only if $\mathbf{u} = \mathbf{0}$ (since then $S_{\mathbf{u}}(\mathbf{0}) = \mathbf{0}$). Remarkably, distance-preserving operators that do fix the origin are necessarily linear.

Lemma 10.4.1

Let V be an inner product space of dimension n, and consider a distance-preserving transformation $S : V \to V$. If $S(\mathbf{0}) = \mathbf{0}$, then S is linear.

Proof. We have $\|S(\mathbf{v}) - S(\mathbf{w})\|^2 = \|\mathbf{v} - \mathbf{w}\|^2$ for all $\mathbf{v}$ and $\mathbf{w}$ in V by (10.2), which gives

$$\langle S(\mathbf{v}),\, S(\mathbf{w})\rangle = \langle \mathbf{v},\, \mathbf{w}\rangle \quad \text{for all } \mathbf{v} \text{ and } \mathbf{w} \text{ in } V \tag{10.3}$$

Now let $\{\mathbf{f}_1,\, \mathbf{f}_2,\, \dots,\, \mathbf{f}_n\}$ be an orthonormal basis of V. Then $\{S(\mathbf{f}_1),\, S(\mathbf{f}_2),\, \dots,\, S(\mathbf{f}_n)\}$ is orthonormal by (10.3) and so is a basis because $\dim V = n$. Now compute:

$$\begin{aligned}
\langle S(\mathbf{v}+\mathbf{w}) - S(\mathbf{v}) - S(\mathbf{w}),\, S(\mathbf{f}_i)\rangle &= \langle S(\mathbf{v}+\mathbf{w}),\, S(\mathbf{f}_i)\rangle - \langle S(\mathbf{v}),\, S(\mathbf{f}_i)\rangle - \langle S(\mathbf{w}),\, S(\mathbf{f}_i)\rangle \\
&= \langle \mathbf{v}+\mathbf{w},\, \mathbf{f}_i\rangle - \langle \mathbf{v},\, \mathbf{f}_i\rangle - \langle \mathbf{w},\, \mathbf{f}_i\rangle \\
&= 0
\end{aligned}$$

for each i. It follows from the expansion theorem (Theorem 10.2.4) that $S(\mathbf{v}+\mathbf{w}) - S(\mathbf{v}) - S(\mathbf{w}) = 0$; that is, $S(\mathbf{v}+\mathbf{w}) = S(\mathbf{v}) + S(\mathbf{w})$. A similar argument shows that $S(a\mathbf{v}) = aS(\mathbf{v})$ holds for all a in $\mathbb{R}$ and $\mathbf{v}$ in V, so S is linear after all. □

Definition 10.5 Isometries

Distance-preserving linear operators are called ***isometries****.*

It is routine to verify that the composite of two distance-preserving transformations is again distance preserving. In particular the composite of a translation and an isometry is distance preserving. Surprisingly, the converse is true.

Theorem 10.4.1

If V is a finite dimensional inner product space, then every distance-preserving transformation $S : V \to V$ is the composite of a translation and an isometry.

Proof. If $S : V \to V$ is distance preserving, write $S(\mathbf{0}) = \mathbf{u}$ and define $T : V \to V$ by $T(\mathbf{v}) = S(\mathbf{v}) - \mathbf{u}$ for all $\mathbf{v}$ in V. Then $\|T(\mathbf{v}) - T(\mathbf{w})\| = \|\mathbf{v} - \mathbf{w}\|$ for all vectors $\mathbf{v}$ and $\mathbf{w}$ in V as the reader can verify; that is, T is distance preserving. Clearly, $T(\mathbf{0}) = \mathbf{0}$, so it is an isometry by Lemma 10.4.1. Since

$$S(\mathbf{v}) = \mathbf{u} + T(\mathbf{v}) = (S_{\mathbf{u}} \circ T)(\mathbf{v}) \quad \text{for all } \mathbf{v} \text{ in } V$$

we have $S = S_{\mathbf{u}} \circ T$, and the theorem is proved. □

In Theorem 10.4.1, $S = S_{\mathbf{u}} \circ T$ factors as the composite of an isometry T followed by a translation $S_{\mathbf{u}}$. More is true: this factorization is unique in that $\mathbf{u}$ and T are uniquely determined by S; and $\mathbf{w} \in V$ exists such that $S = T \circ S_{\mathbf{w}}$ is uniquely the composite of translation by $\mathbf{w}$ followed by the same isometry T (Exercise 10.4.12).

Theorem 10.4.1 focuses our attention on the isometries, and the next theorem shows that, while they preserve distance, they are characterized as those operators that preserve other properties.

Theorem 10.4.2

Let $T : V \to V$ be a linear operator on a finite dimensional inner product space V.
The following conditions are equivalent:

1. *T is an isometry.* *(T preserves distance)*
2. *$\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v}$ in V.* *(T preserves norms)*
3. *$\langle T(\mathbf{v}), T(\mathbf{w})\rangle = \langle \mathbf{v}, \mathbf{w}\rangle$ for all $\mathbf{v}$ and $\mathbf{w}$ in V.* *(T preserves inner products)*
4. *If $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ is an orthonormal basis of V, then $\{T(\mathbf{f}_1), T(\mathbf{f}_2), \dots, T(\mathbf{f}_n)\}$ is also an orthonormal basis.* *(T preserves orthonormal bases)*
5. *T carries some orthonormal basis to an orthonormal basis.*

Proof. (1) $\Rightarrow$ (2). Take $\mathbf{w} = \mathbf{0}$ in (10.2).

(2) $\Rightarrow$ (3). Since T is linear, (2) gives $\|T(\mathbf{v}) - T(\mathbf{w})\|^2 = \|T(\mathbf{v} - \mathbf{w})\|^2 = \|\mathbf{v} - \mathbf{w}\|^2$. Now (3) follows.

(3) $\Rightarrow$ (4). By (3), $\{T(\mathbf{f}_1), T(\mathbf{f}_2), \dots, T(\mathbf{f}_n)\}$ is orthogonal and $\|T(\mathbf{f}_i)\|^2 = \|\mathbf{f}_i\|^2 = 1$. Hence it is a basis because $\dim V = n$.

(4) $\Rightarrow$ (5). This needs no proof.

(5) $\Rightarrow$ (1). By (5), let $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ be an orthonormal basis of V such that$\{T(\mathbf{f}_1), \dots, T(\mathbf{f}_n)\}$ is also orthonormal. Given $\mathbf{v} = v_1\mathbf{f}_1 + \cdots + v_n\mathbf{f}_n$ in V, we have $T(\mathbf{v}) = v_1T(\mathbf{f}_1) + \cdots + v_nT(\mathbf{f}_n)$ so Pythagoras' theorem gives

$$\|T(\mathbf{v})\|^2 = v_1^2 + \cdots + v_n^2 = \|\mathbf{v}\|^2$$

Hence $\|T(\mathbf{v})\| = \|\mathbf{v}\|$ for all $\mathbf{v}$, and (1) follows by replacing $\mathbf{v}$ by $\mathbf{v} - \mathbf{w}$. □

Before giving examples, we note some consequences of Theorem 10.4.2.

Corollary 10.4.1

Let V be a finite dimensional inner product space.

1. *Every isometry of V is an isomorphism.*[5]
2. a. *$1_V : V \to V$ is an isometry.*
   b. *The composite of two isometries of V is an isometry.*
   c. *The inverse of an isometry of V is an isometry.*

Proof. (1) is by (4) of Theorem 10.4.2 and Theorem 7.3.1. (2a) is clear, and (2b) is left to the reader. If $T : V \to V$ is an isometry and $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ is an orthonormal basis of V, then (2c) follows because T^{-1} carries the orthonormal basis $\{T(\mathbf{f}_1), \dots, T(\mathbf{f}_n)\}$ back to $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$. □

[5] V must be finite dimensional—see Exercise 10.4.13.

The conditions in part (2) of the corollary assert that the set of isometries of a finite dimensional inner product space forms an algebraic system called a **group**. The theory of groups is well developed, and groups of operators are important in geometry. In fact, geometry itself can be fruitfully viewed as the study of those properties of a vector space that are preserved by a group of invertible linear operators.

Example 10.4.1

Rotations of $\mathbb{R}^2$ about the origin are isometries, as are reflections in lines through the origin: They clearly preserve distance and so are linear by Lemma 10.4.1. Similarly, rotations about lines through the origin and reflections in planes through the origin are isometries of $\mathbb{R}^3$.

Example 10.4.2

Let $T : \mathbf{M}_{nn} \to \mathbf{M}_{nn}$ be the transposition operator: $T(A) = A^T$. Then T is an isometry if the inner product is $\langle A, B \rangle = \operatorname{tr}(AB^T) = \sum_{i, j} a_{ij}b_{ij}$. In fact, T permutes the basis consisting of all matrices with one entry 1 and the other entries 0.

The proof of the next result requires the fact (see Theorem 10.4.2) that, if B is an orthonormal basis, then $\langle \mathbf{v}, \mathbf{w} \rangle = C_B(\mathbf{v}) \cdot C_B(\mathbf{w})$ for all vectors $\mathbf{v}$ and $\mathbf{w}$.

Theorem 10.4.3

Let $T : V \to V$ be an operator where V is a finite dimensional inner product space. The following conditions are equivalent.

1. *T is an isometry.*
2. *$M_B(T)$ is an orthogonal matrix for every orthonormal basis B.*
3. *$M_B(T)$ is an orthogonal matrix for some orthonormal basis B.*

Proof. (1) $\Rightarrow$ (2). Let $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be an orthonormal basis. Then the jth column of $M_B(T)$ is $C_B[T(\mathbf{e}_j)]$, and we have

$$C_B[T(\mathbf{e}_j)] \cdot C_B[T(\mathbf{e}_k)] = \langle T(\mathbf{e}_j), T(\mathbf{e}_k) \rangle = \langle \mathbf{e}_j, \mathbf{e}_k \rangle$$

using (1). Hence the columns of $M_B(T)$ are orthonormal in $\mathbb{R}^n$, which proves (2).

(2) $\Rightarrow$ (3). This is clear.

(3) $\Rightarrow$ (1). Let $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be as in (3). Then, as before,

$$\langle T(\mathbf{e}_j), T(\mathbf{e}_k) \rangle = C_B[T(\mathbf{e}_j)] \cdot C_B[T(\mathbf{e}_k)]$$

so $\{T(\mathbf{e}_1), \dots, T(\mathbf{e}_n)\}$ is orthonormal by (3). Hence Theorem 10.4.2 gives (1). □

It is important that B is *orthonormal* in Theorem 10.4.3. For example, $T : V \to V$ given by $T(\mathbf{v}) = 2\mathbf{v}$ preserves *orthogonal* sets but is not an isometry, as is easily checked.

If P is an orthogonal square matrix, then $P^{-1} = P^T$. Taking determinants yields $(\det P)^2 = 1$, so $\det P = \pm 1$. Hence:

Corollary 10.4.2

If $T : V \to V$ is an isometry where V is a finite dimensional inner product space, then $\det T = \pm 1$.

Example 10.4.3

If A is any $n \times n$ matrix, the matrix operator $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is an isometry if and only if A is orthogonal using the dot product in $\mathbb{R}^n$. Indeed, if E is the standard basis of $\mathbb{R}^n$, then $M_E(T_A) = A$ by Theorem 9.2.4.

Rotations and reflections that fix the origin are isometries in $\mathbb{R}^2$ and $\mathbb{R}^3$ (Example 10.4.1); we are going to show that these isometries (and compositions of them in $\mathbb{R}^3$) are the only possibilities. In fact, this will follow from a general structure theorem for isometries. Surprisingly enough, much of the work involves the two–dimensional case.

Theorem 10.4.4

Let $T : V \to V$ be an isometry on the two-dimensional inner product space V. Then there are two possibilities.
Either (1) *There is an orthonormal basis B of V such that*

$$M_B(T) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad 0 \leq \theta < 2\pi$$

or (2) *There is an orthonormal basis B of V such that*

$$M_B(T) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Furthermore, type (1) *occurs if and only if* $\det T = 1$, *and type* (2) *occurs if and only if* $\det T = -1$.

Proof. The final statement follows from the rest because $\det T = \det[M_B(T)]$ for any basis B. Let $B_0 = \{\mathbf{e}_1, \mathbf{e}_2\}$ be any ordered orthonormal basis of V and write

$$A = M_{B_0}(T) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}; \text{ that is, } \begin{array}{l} T(\mathbf{e}_1) = a\mathbf{e}_1 + c\mathbf{e}_2 \\ T(\mathbf{e}_2) = b\mathbf{e}_1 + d\mathbf{e}_2 \end{array}$$

Then A is orthogonal by Theorem 10.4.3, so its columns (and rows) are orthonormal. Hence

$$a^2 + c^2 = 1 = b^2 + d^2$$

so (a, c) and (d, b) lie on the unit circle. Thus angles θ and φ exist such that

$$\begin{array}{lll} a = \cos\theta, & c = \sin\theta & 0 \leq \theta < 2\pi \\ d = \cos\varphi, & b = \sin\varphi & 0 \leq \varphi < 2\pi \end{array}$$

Then $\sin(\theta+\varphi) = cd+ab = 0$ because the columns of A are orthogonal, so $\theta+\varphi = k\pi$ for some integer k. This gives $d = \cos(k\pi-\theta) = (-1)^k\cos\theta$ and $b = \sin(k\pi-\theta) = (-1)^{k+1}\sin\theta$. Finally

$$A = \left[\begin{array}{cc} \cos\theta & (-1)^{k+1}\sin\theta \\ \sin\theta & (-1)^k\cos\theta \end{array}\right]$$

If k is even we are in type (1) with $B = B_0$, so assume k is odd. Then $A = \left[\begin{array}{cc} a & c \\ c & -a \end{array}\right]$. If $a = -1$ and $c = 0$, we are in type (1) with $B = \{\mathbf{e}_2,\ \mathbf{e}_2\}$. Otherwise A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ with corresponding eigenvectors $\mathbf{x}_1 = \left[\begin{array}{c} 1+a \\ c \end{array}\right]$ and $\mathbf{x}_2 = \left[\begin{array}{c} -c \\ 1+a \end{array}\right]$ as the reader can verify. Write

$$\mathbf{f}_1 = (1+a)\mathbf{e}_1 + c\mathbf{e}_2 \quad \text{and} \quad \mathbf{f}_2 = -c\mathbf{e}_2 + (1+a)\mathbf{e}_2$$

Then $\mathbf{f}_1$ and $\mathbf{f}_2$ are orthogonal (verify) and $C_{B_0}(\mathbf{f}_i) = C_{B_0}(\lambda_i\mathbf{f}_i) = \mathbf{x}_i$ for each i. Moreover

$$C_{B_0}[T(\mathbf{f}_i)] = AC_{B_0}(\mathbf{f}_i) = A\mathbf{x}_i = \lambda_i\mathbf{x}_i = \lambda_i C_{B_0}(\mathbf{f}_i) = C_{B_0}(\lambda_i\mathbf{f}_i)$$

so $T(\mathbf{f}_i) = \lambda_i\mathbf{f}_i$ for each i. Hence $M_B(T) = \left[\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_2 \end{array}\right] = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]$ and we are in type (2) with $B = \left\{\frac{1}{\|\mathbf{f}_1\|}\mathbf{f}_1,\ \frac{1}{\|\mathbf{f}_2\|}\mathbf{f}_2\right\}$. □

Corollary 10.4.3

An operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ is an isometry if and only if T is a rotation or a reflection.

In fact, if E is the standard basis of $\mathbb{R}^2$, then the clockwise rotation R_θ about the origin through an angle θ has matrix

$$M_E(R_\theta) = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array}\right]$$

(see Theorem 2.6.4). On the other hand, if $S : \mathbb{R}^2 \to \mathbb{R}^2$ is the reflection in a line through the origin (called the **fixed line** of the reflection), let $\mathbf{f}_1$ be a unit vector pointing along the fixed line and let $\mathbf{f}_2$ be a unit vector perpendicular to the fixed line. Then $B = \{\mathbf{f}_1,\ \mathbf{f}_2\}$ is an orthonormal basis, $S(\mathbf{f}_1) = \mathbf{f}_1$ and $S(\mathbf{f}_2) = -\mathbf{f}_2$, so

$$M_B(S) = \left[\begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array}\right]$$

Thus S is of type 2. Note that, in this case, 1 is an eigenvalue of S, and any eigenvector corresponding to 1 is a direction vector for the fixed line.

Example 10.4.4

In each case, determine whether $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ is a rotation or a reflection, and then find the angle or fixed line:

$$\text{(a) } A = \frac{1}{2}\left[\begin{array}{cc} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{array}\right] \qquad \text{(b) } A = \frac{1}{5}\left[\begin{array}{cc} -3 & 4 \\ 4 & 3 \end{array}\right]$$

Solution. Both matrices are orthogonal, so (because $M_E(T_A) = A$, where E is the standard basis) T_A is an isometry in both cases. In the first case, $\det A = 1$, so T_A is a counterclockwise rotation through θ, where $\cos\theta = \frac{1}{2}$ and $\sin\theta = -\frac{\sqrt{3}}{2}$. Thus $\theta = -\frac{\pi}{3}$. In (b), $\det A = -1$, so T_A is a reflection in this case. We verify that $\mathbf{d} = \left[\begin{array}{c} 1 \\ 2 \end{array}\right]$ is an eigenvector corresponding to the eigenvalue 1. Hence the fixed line $\mathbb{R}\mathbf{d}$ has equation $y = 2x$.

We now give a structure theorem for isometries. The proof requires three preliminary results, each of interest in its own right.

Lemma 10.4.2

Let $T : V \to V$ be an isometry of a finite dimensional inner product space V. If U is a T-invariant subspace of V, then $U^\perp$ is also T-invariant.

Proof. Let $\mathbf{w}$ lie in $U^\perp$. We are to prove that $T(\mathbf{w})$ is also in $U^\perp$; that is, $\langle T(\mathbf{w}), \mathbf{u}\rangle = 0$ for all $\mathbf{u}$ in U. At this point, observe that the restriction of T to U is an isometry $U \to U$ and so is an isomorphism by the corollary to Theorem 10.4.2. In particular, each $\mathbf{u}$ in U can be written in the form $\mathbf{u} = T(\mathbf{u}_1)$ for some $\mathbf{u}_1$ in U, so

$$\langle T(\mathbf{w}), \mathbf{u}\rangle = \langle T(\mathbf{w}), T(\mathbf{u}_1)\rangle = \langle \mathbf{w}, \mathbf{u}_1\rangle = 0$$

because $\mathbf{w}$ is in $U^\perp$. This is what we wanted. □

To employ Lemma 10.4.2 above to analyze an isometry $T : V \to V$ when $\dim V = n$, it is necessary to show that a T-invariant subspace U exists such that $U \neq 0$ and $U \neq V$. We will show, in fact, that such a subspace U can always be found of dimension 1 or 2. If T has a real eigenvalue λ then $\mathbb{R}\mathbf{u}$ is T-invariant where $\mathbf{u}$ is any λ-eigenvector. But, in case (1) of Theorem 10.4.4, the eigenvalues of T are $e^{i\theta}$ and $e^{-i\theta}$ (the reader should check this), and these are nonreal if $\theta \neq 0$ and $\theta \neq \pi$. It turns out that every complex eigenvalue λ of T has absolute value 1 (Lemma 10.4.3 below); and that U has a T-invariant subspace of dimension 2 if λ is not real (Lemma 10.4.4).

Lemma 10.4.3

Let $T : V \to V$ be an isometry of the finite dimensional inner product space V. If λ is a complex eigenvalue of T, then $|\lambda| = 1$.

Proof. Choose an orthonormal basis B of V, and let $A = M_B(T)$. Then A is a real orthogonal matrix so, using the standard inner product $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^T\overline{\mathbf{y}}$ in $\mathbb{C}$, we get

$$||A\mathbf{x}||^2 = (A\mathbf{x})^T(\overline{A\mathbf{x}}) = \mathbf{x}^T A^T \overline{A\mathbf{x}} = \mathbf{x}^T I\mathbf{x} = ||\mathbf{x}||^2$$

for all $\mathbf{x}$ in $\mathbb{C}^n$. But $A\mathbf{x} = \lambda\mathbf{x}$ for some $\mathbf{x} \neq \mathbf{0}$, whence $||\mathbf{x}||^2 = ||\lambda\mathbf{x}||^2 = |\lambda|^2||\mathbf{x}||^2$. This gives $|\lambda| = 1$, as required. □

Lemma 10.4.4

Let $T : V \to V$ be an isometry of the n-dimensional inner product space V. If T has a nonreal eigenvalue, then V has a two-dimensional T-invariant subspace.

Proof. Let B be an orthonormal basis of V, let $A = M_B(T)$, and (using Lemma 10.4.3) let $\lambda = e^{i\alpha}$ be a nonreal eigenvalue of A, say $A\mathbf{x} = \lambda\mathbf{x}$ where $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{C}^n$. Because A is real, complex conjugation gives $A\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$, so $\overline{\lambda}$ is also an eigenvalue. Moreover $\lambda \neq \overline{\lambda}$ (λ is nonreal), so $\{\mathbf{x},\ \overline{\mathbf{x}}\}$ is linearly independent in $\mathbb{C}^n$ (the argument in the proof of Theorem 5.5.4 works). Now define

$$\mathbf{z}_1 = \mathbf{x} + \overline{\mathbf{x}} \quad \text{and} \quad \mathbf{z}_2 = i(\mathbf{x} - \overline{\mathbf{x}})$$

Then $\mathbf{z}_1$ and $\mathbf{z}_2$ lie in $\mathbb{R}^n$, and $\{\mathbf{z}_1,\ \mathbf{z}_2\}$ is linearly independent over $\mathbb{R}$ because $\{\mathbf{x},\ \overline{\mathbf{x}}\}$ is linearly independent over $\mathbb{C}$. Moreover

$$\mathbf{x} = \tfrac{1}{2}(\mathbf{z}_1 - i\mathbf{z}_2) \quad \text{and} \quad \overline{\mathbf{x}} = \tfrac{1}{2}(\mathbf{z}_1 + i\mathbf{z}_2)$$

Now $\lambda + \overline{\lambda} = 2\cos\alpha$ and $\lambda - \overline{\lambda} = 2i\sin\alpha$, and a routine computation gives

$$\begin{aligned} A\mathbf{z}_1 &= \mathbf{z}_1\cos\alpha + \mathbf{z}_2\sin\alpha \\ A\mathbf{z}_2 &= -\mathbf{z}_1\sin\alpha + \mathbf{z}_2\cos\alpha \end{aligned}$$

Finally, let $\mathbf{e}_1$ and $\mathbf{e}_2$ in V be such that $\mathbf{z}_1 = C_B(\mathbf{e}_1)$ and $\mathbf{z}_2 = C_B(\mathbf{e}_2)$. Then

$$C_B[T(\mathbf{e}_1)] = AC_B(\mathbf{e}_1) = A\mathbf{z}_1 = C_B(\mathbf{e}_1\cos\alpha + \mathbf{e}_2\sin\alpha)$$

using Theorem 9.1.2. Because C_B is one-to-one, this gives the first of the following equations (the other is similar):

$$\begin{aligned} T(\mathbf{e}_1) &= \mathbf{e}_1\cos\alpha + \mathbf{e}_2\sin\alpha \\ T(\mathbf{e}_2) &= -\mathbf{e}_1\sin\alpha + \mathbf{e}_2\cos\alpha \end{aligned}$$

Thus $U = \text{span}\,\{\mathbf{e}_1,\ \mathbf{e}_2\}$ is T-invariant and two-dimensional. □

We can now prove the structure theorem for isometries.

Theorem 10.4.5

Let $T : V \to V$ be an isometry of the n-dimensional inner product space V. Given an angle θ, write $R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Then there exists an orthonormal basis B of V such that $M_B(T)$ has one of the following block diagonal forms, classified for convenience by whether n is odd or even:

$$n = 2k+1 \quad \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix}$$

$$n = 2k \quad \begin{bmatrix} R(\theta_1) & 0 & \cdots & 0 \\ 0 & R(\theta_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R(\theta_k) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R(\theta_{k-1}) \end{bmatrix}$$

Proof. We show first, by induction on n, that an orthonormal basis B of V can be found such that $M_B(T)$ is a block diagonal matrix of the following form:

$$M_B(T) = \begin{bmatrix} I_r & 0 & 0 & \cdots & 0 \\ 0 & -I_s & 0 & \cdots & 0 \\ 0 & 0 & R(\theta_1) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R(\theta_t) \end{bmatrix}$$

where the identity matrix I_r, the matrix $-I_s$, or the matrices $R(\theta_i)$ may be missing. If $n = 1$ and $V = \mathbb{R}\mathbf{v}$, this holds because $T(\mathbf{v}) = \lambda\mathbf{v}$ and $\lambda = \pm 1$ by Lemma 10.4.3. If $n = 2$, this follows from Theorem 10.4.4. If $n \geq 3$, either T has a real eigenvalue and therefore has a one-dimensional T-invariant subspace $U = \mathbb{R}\mathbf{u}$ for any eigenvector $\mathbf{u}$, or T has no real eigenvalue and therefore has a two-dimensional T-invariant subspace U by Lemma 10.4.4. In either case $U^\perp$ is T-invariant (Lemma 10.4.2) and $\dim U^\perp = n - \dim U < n$. Hence, by induction, let B_1 and B_2 be orthonormal bases of U and $U^\perp$ such that $M_{B_1}(T)$ and $M_{B_2}(T)$ have the form given. Then $B = B_1 \cup B_2$ is an orthonormal basis of V, and $M_B(T)$ has the desired form with a suitable ordering of the vectors in B.

Now observe that $R(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $R(\pi) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. It follows that an even number of 1s or -1s can be written as $R(\theta_1)$-blocks. Hence, with a suitable reordering of the basis B, the theorem follows. □

As in the dimension 2 situation, these possibilities can be given a geometric interpretation when $V = \mathbb{R}^3$ is taken as euclidean space. As before, this entails looking carefully at reflections and rotations in $\mathbb{R}^3$. If $Q : \mathbb{R}^3 \to \mathbb{R}^3$ is any reflection in a plane through the origin (called the **fixed plane** of the reflection), take $\{\mathbf{f}_2, \mathbf{f}_3\}$ to be any orthonormal basis of the fixed plane and take $\mathbf{f}_1$ to be a unit vector perpendicular to the fixed plane. Then $Q(\mathbf{f}_1) = -\mathbf{f}_1$, whereas $Q(\mathbf{f}_2) = \mathbf{f}_2$ and $Q(\mathbf{f}_3) = \mathbf{f}_3$. Hence $B = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is an orthonormal basis such that

$$M_B(Q) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Similarly, suppose that $R : \mathbb{R}^3 \to \mathbb{R}^3$ is any rotation about a line through the origin (called the **axis** of the rotation), and let $\mathbf{f}_1$ be a unit vector pointing along the axis, so $R(\mathbf{f}_1) = \mathbf{f}_1$. Now the plane through the origin perpendicular to the axis is an R-invariant subspace of $\mathbb{R}^2$ of dimension 2, and the restriction of R to this plane is a rotation. Hence, by Theorem 10.4.4, there is an orthonormal basis $B_1 = \{\mathbf{f}_2, \mathbf{f}_3\}$ of this plane such that $M_{B_1}(R) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. But then $B = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is an orthonormal basis of $\mathbb{R}^3$ such that the matrix of R is

$$M_B(R) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

However, Theorem 10.4.5 shows that there are isometries T in $\mathbb{R}^3$ of a third type: those with a matrix of the form

$$M_B(T) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

If $B = \{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$, let Q be the reflection in the plane spanned by $\mathbf{f}_2$ and $\mathbf{f}_3$, and let R be the rotation corresponding to θ about the line spanned by $\mathbf{f}_1$. Then $M_B(Q)$ and $M_B(R)$ are as above, and $M_B(Q)M_B(R) = M_B(T)$ as the reader can verify. This means that $M_B(QR) = M_B(T)$ by Theorem 9.2.1, and this in turn implies that $QR = T$ because M_B is one-to-one (see Exercise 9.1.26). A similar argument shows that $RQ = T$, and we have Theorem 10.4.6.

Theorem 10.4.6

If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is an isometry, there are three possibilities.

a. *T is a rotation, and* $M_B(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ *for some orthonormal basis B.*

b. *T is a reflection, and* $M_B(T) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *for some orthonormal basis B.*

c. *$T = QR = RQ$ where Q is a reflection, R is a rotation about an axis perpendicular to the fixed plane of Q and* $M_B(T) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$ *for some orthonormal basis B.*

Hence T is a rotation if and only if $\det T = 1$.

Proof. It remains only to verify the final observation that T is a rotation if and only if $\det T = 1$. But clearly $\det T = -1$ in parts (b) and (c). □

A useful way of analyzing a given isometry $T : \mathbb{R}^3 \to \mathbb{R}^3$ comes from computing the eigenvalues of T. Because the characteristic polynomial of T has degree 3, it must have a real root. Hence, there must be at least one real eigenvalue, and the only possible real eigenvalues are ± 1 by Lemma 10.4.3. Thus Table 10.1 includes all possibilities.

Table 10.1

Eigenvalues of T	**Action of T**
(1) 1, no other real eigenvalues	Rotation about the line $\mathbb{R}\mathbf{f}$ where $\mathbf{f}$ is an eigenvector corresponding to 1. [Case (a) of Theorem 10.4.6.]
(2) -1, no other real eigenvalues	Rotation about the line $\mathbb{R}\mathbf{f}$ followed by reflection in the plane $(\mathbb{R}\mathbf{f})^{\perp}$ where $\mathbf{f}$ is an eigenvector corresponding to -1. [Case (c) of Theorem 10.4.6.]
(3) $-1,\ 1,\ 1$	Reflection in the plane $(\mathbb{R}\mathbf{f})^{\perp}$ where $\mathbf{f}$ is an eigenvector corresponding to -1. [Case (b) of Theorem 10.4.6.]
(4) $1,\ -1,\ -1$	This is as in (1) with a rotation of π.
(5) $-1,\ -1,\ -1$	Here $T(\mathbf{x}) = -\mathbf{x}$ for all x. This is (2) with a rotation of π.
(6) $1,\ 1,\ 1$	Here T is the identity isometry.

Example 10.4.5

Analyze the isometry $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by $T \left[\begin{array}{c} x \\ y \\ z \end{array} \right] = \left[\begin{array}{c} y \\ z \\ -x \end{array} \right]$.

Solution. If B_0 is the standard basis of $\mathbb{R}^3$, then $M_{B_0}(T) = \left[\begin{array}{rrr} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{array} \right]$, so $c_T(x) = x^3 + 1 = (x+1)(x^2 - x + 1)$. This is (2) in Table 10.1. Write:

$$\mathbf{f}_1 = \frac{1}{\sqrt{3}} \left[\begin{array}{r} 1 \\ -1 \\ 1 \end{array} \right] \quad \mathbf{f}_2 = \frac{1}{\sqrt{6}} \left[\begin{array}{c} 1 \\ 2 \\ 1 \end{array} \right] \quad \mathbf{f}_3 = \frac{1}{\sqrt{2}} \left[\begin{array}{r} 1 \\ 0 \\ -1 \end{array} \right]$$

Here $\mathbf{f}_1$ is a unit eigenvector corresponding to $\lambda_1 = -1$, so T is a rotation (through an angle θ) about the line $L = \mathbb{R}\mathbf{f}_1$, followed by reflection in the plane U through the origin perpendicular to $\mathbf{f}_1$ (with equation $x - y + z = 0$). Then, $\{\mathbf{f}_1,\ \mathbf{f}_2\}$ is chosen as an orthonormal basis of U, so $B = \{\mathbf{f}_1,\ \mathbf{f}_2,\ \mathbf{f}_3\}$ is an orthonormal basis of $\mathbb{R}^3$ and

$$M_B(T) = \left[\begin{array}{ccc} -1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{array} \right]$$

Hence θ is given by $\cos\theta = \frac{1}{2}$, $\sin\theta = \frac{\sqrt{3}}{2}$, so $\theta = \frac{\pi}{3}$.

Let V be an n-dimensional inner product space. A subspace of V of dimension $n-1$ is called a **hyperplane** in V. Thus the hyperplanes in $\mathbb{R}^3$ and $\mathbb{R}^2$ are, respectively, the planes and lines through the origin. Let $Q : V \to V$ be an isometry with matrix

$$M_B(Q) = \left[\begin{array}{cc} -1 & 0 \\ 0 & I_{n-1} \end{array} \right]$$

for some orthonormal basis $B = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$. Then $Q(\mathbf{f}_1) = -\mathbf{f}_1$ whereas $Q(\mathbf{u}) = \mathbf{u}$ for each $\mathbf{u}$ in $U = \text{span}\,\{\mathbf{f}_2, \dots, \mathbf{f}_n\}$. Hence U is called the **fixed hyperplane** of Q, and Q is called **reflection** in U. Note that each hyperplane in V is the fixed hyperplane of a (unique) reflection of V. Clearly, reflections in $\mathbb{R}^2$ and $\mathbb{R}^3$ are reflections in this more general sense.

Continuing the analogy with $\mathbb{R}^2$ and $\mathbb{R}^3$, an isometry $T : V \to V$ is called a **rotation** if there exists an orthonormal basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ such that

$$M_B(T) = \left[\begin{array}{ccc} I_r & 0 & 0 \\ 0 & R(\theta) & 0 \\ 0 & 0 & I_s \end{array} \right]$$

in block form, where $R(\theta) = \left[\begin{array}{cc} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{array} \right]$, and where either I_r or I_s (or both) may be missing. If $R(\theta)$ occupies columns i and $i+1$ of $M_B(T)$, and if $W = \text{span}\,\{\mathbf{f}_i, \mathbf{f}_{i+1}\}$, then W is T-invariant and the matrix of $T : W \to W$ with respect to $\{\mathbf{f}_i, \mathbf{f}_{i+1}\}$ is $R(\theta)$. Clearly, if W is viewed as a copy of $\mathbb{R}^2$, then T is a rotation in W. Moreover, $T(\mathbf{u}) = \mathbf{u}$ holds for all vectors $\mathbf{u}$ in the $(n-2)$-dimensional subspace $U = \text{span}\,\{\mathbf{f}_1, \dots, \mathbf{f}_{i-1}, \mathbf{f}_{i+1}, \dots, \mathbf{f}_n\}$, and U is called the **fixed axis** of the rotation T. In $\mathbb{R}^3$, the axis of any rotation is a line (one-dimensional), whereas in $\mathbb{R}^2$ the axis is $U = \{\mathbf{0}\}$.

With these definitions, the following theorem is an immediate consequence of Theorem 10.4.5 (the details are left to the reader).

Theorem 10.4.7

Let $T : V \to V$ be an isometry of a finite dimensional inner product space V. Then there exist isometries $T_1, \dots, T$ such that

$$T = T_k T_{k-1} \cdots T_2 T_1$$

where each T_i is either a rotation or a reflection, at most one is a reflection, and $T_i T_j = T_j T_i$ holds for all i and j. Furthermore, T is a composite of rotations if and only if $\det T = 1$.

Exercises for 10.4

Throughout these exercises, V denotes a finite dimensional inner product space.

Exercise 10.4.1 Show that the following linear operators are isometries.

a. $T : \mathbb{C} \to \mathbb{C}$; $T(z) = \overline{z}$; $\langle z, w\rangle = \text{re}\,(z\overline{w})$

b. $T : \mathbb{R}^n \to \mathbb{R}^n$; $T(a_1, a_2, \dots, a_n)$ $= (a_n, a_{n-1}, \dots, a_2, a_1)$; dot product

c. $T : \mathbf{M}_{22} \to \mathbf{M}_{22}$; $T \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = \left[\begin{array}{cc} c & d \\ b & a \end{array} \right]$; $\langle A, B\rangle = \text{tr}\,(AB^T)$

d. $T : \mathbb{R}^3 \to \mathbb{R}^3$; $T(a, b, c) = \frac{1}{9}(2a+2b-c, 2a+$

$2c - b,\ 2b + 2c - a$); dot product

Exercise 10.4.2 In each case, show that T is an isometry of $\mathbb{R}^2$, determine whether it is a rotation or a reflection, and find the angle or the fixed line. Use the dot product.

a. $T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a \\ b \end{bmatrix}$

b. $T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -a \\ -b \end{bmatrix}$

c. $T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} b \\ -a \end{bmatrix}$

d. $T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ -a \end{bmatrix}$

e. $T \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} a+b \\ b-a \end{bmatrix}$

f. $T \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} a-b \\ a+b \end{bmatrix}$

Exercise 10.4.3 In each case, show that T is an isometry of $\mathbb{R}^3$, determine the type (Theorem 10.4.6), and find the axis of any rotations and the fixed plane of any reflections involved.

a. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ -b \\ c \end{bmatrix}$

b. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{3}c - a \\ \sqrt{3}a + c \\ 2b \end{bmatrix}$

c. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} b \\ c \\ a \end{bmatrix}$

d. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a \\ -b \\ -c \end{bmatrix}$

e. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{2} \begin{bmatrix} a + \sqrt{3}b \\ b - \sqrt{3}a \\ 2c \end{bmatrix}$

f. $T \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} a + c \\ -\sqrt{2}b \\ c - a \end{bmatrix}$

Exercise 10.4.4 Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be an isometry. A vector $\mathbf{x}$ in $\mathbb{R}^2$ is said to be **fixed** by T if $T(\mathbf{x}) = \mathbf{x}$. Let E_1 denote the set of all vectors in $\mathbb{R}^2$ fixed by T. Show that:

a. E_1 is a subspace of $\mathbb{R}^2$.

b. $E_1 = \mathbb{R}^2$ if and only if $T = 1$ is the identity map.

c. $\dim E_1 = 1$ if and only if T is a reflection (about the line E_1).

d. $E_1 = \{0\}$ if and only if T is a rotation ($T \neq 1$).

Exercise 10.4.5 Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be an isometry, and let E_1 be the subspace of all fixed vectors in $\mathbb{R}^3$ (see Exercise 10.4.4). Show that:

a. $E_1 = \mathbb{R}^3$ if and only if $T = 1$.

b. $\dim E_1 = 2$ if and only if T is a reflection (about the plane E_1).

c. $\dim E_1 = 1$ if and only if T is a rotation ($T \neq 1$) (about the line E_1).

d. $\dim E_1 = 0$ if and only if T is a reflection followed by a (nonidentity) rotation.

Exercise 10.4.6 If T is an isometry, show that aT is an isometry if and only if $a = \pm 1$.

Exercise 10.4.7 Show that every isometry preserves the angle between any pair of nonzero vectors (see Exercise 10.1.31). Must an angle-preserving isomorphism be an isometry? Support your answer.

Exercise 10.4.8 If $T : V \to V$ is an isometry, show that $T^2 = 1_V$ if and only if the only complex eigenvalues of T are 1 and -1.

Exercise 10.4.9 Let $T : V \to V$ be a linear operator. Show that any two of the following conditions implies the third:

1. T is symmetric.
2. T is an involution ($T^2 = 1_V$).
3. T is an isometry.

   [*Hint*: In all cases, use the definition

   $$\langle \mathbf{v},\ T(\mathbf{w}) \rangle = \langle T(\mathbf{v}),\ \mathbf{w} \rangle$$

   of a symmetric operator. For (1) and (3) $\Rightarrow$ (2), use the fact that, if $\langle T^2(\mathbf{v}) - \mathbf{v},\ \mathbf{w} \rangle = 0$ for all $\mathbf{w}$, then $T^2(\mathbf{v}) = \mathbf{v}$.]

Exercise 10.4.10 If B and D are any orthonormal bases of V, show that there is an isometry $T : V \to V$ that carries B to D.

Exercise 10.4.11 Show that the following are equivalent for a linear transformation $S : V \to V$ where V is finite dimensional and $S \neq 0$:

1. $\langle S(\mathbf{v}),\ S(\mathbf{w}) \rangle = 0$ whenever $\langle \mathbf{v},\ \mathbf{w} \rangle = 0$;

2. $S = aT$ for some isometry $T : V \to V$ and some $a \neq 0$ in $\mathbb{R}$.

3. S is an isomorphism and preserves angles between nonzero vectors.

   [*Hint*: Given (1), show that $\|S(\mathbf{e})\| = \|S(\mathbf{f})\|$ for all unit vectors $\mathbf{e}$ and $\mathbf{f}$ in V.]

Exercise 10.4.12 Let $S : V \to V$ be a distance preserving transformation where V is finite dimensional.

a. Show that the factorization in the proof of Theorem 10.4.1 is unique. That is, if $S = S_{\mathbf{u}} \circ T$ and $S = S_{\mathbf{u}'} \circ T'$ where $\mathbf{u}, \mathbf{u}' \in V$ and $T, T' : V \to V$ are isometries, show that $\mathbf{u} = \mathbf{u}'$ and $T = T'$.

b. If $S = S_{\mathbf{u}} \circ T$, $\mathbf{u} \in V$, T an isometry, show that $\mathbf{w} \in V$ exists such that $S = T \circ S_{\mathbf{w}}$.

Exercise 10.4.13 Define $T : \mathbf{P} \to \mathbf{P}$ by $T(f) = xf(x)$ for all $f \in \mathbf{P}$, and define an inner product on $\mathbf{P}$ as follows: If $f = a_0 + a_1x + a_2x^2 + \cdots$ and $g = b_0 + b_1x + b_2x^2 + \cdots$ are in $\mathbf{P}$, define $\langle f, g\rangle = a_0b_0 + a_1b_1 + a_2b_2 + \cdots$.

a. Show that $\langle\ ,\ \rangle$ is an inner product on $\mathbf{P}$.

b. Show that T is an isometry of $\mathbf{P}$.

c. Show that T is one-to-one but not onto.

10.5 An Application to Fourier Approximation[6]

If U is an orthogonal basis of a vector space V, the expansion theorem (Theorem 10.2.4) presents a vector $v \in V$ as a linear combination of the vectors in U. Of course this requires that the set U is finite since otherwise the linear combination is an infinite sum and makes no sense in V.

However, given an infinite orthogonal set $U = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n, \dots\}$, we can use the expansion theorem for $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ for each n to get a series of "approximations" $\mathbf{v}_n$ for a given vector $\mathbf{v}$. A natural question is whether these $\mathbf{v}_n$ are getting closer and closer to $\mathbf{v}$ as n increases. This turns out to be a very fruitful idea.

In this section we shall investigate an important orthogonal set in the space $\mathbf{C}[-\pi, \pi]$ of continuous functions on the interval $[-\pi, \pi]$, using the inner product.

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$$

Of course, calculus will be needed. The orthogonal set in question is

$$\{1, \sin x, \cos x, \sin(2x), \cos(2x), \sin(3x), \cos(3x), \dots\}$$

Standard techniques of integration give

$$\begin{aligned} \|1\|^2 &= \int_{-\pi}^{\pi} 1^2 dx = 2\pi \\ \|\sin kx\|^2 &= \int_{-\pi}^{\pi} \sin^2(kx)dx = \pi \quad \text{for any } k = 1, 2, 3, \dots \\ \|\cos kx\|^2 &= \int_{-\pi}^{\pi} \cos^2(kx)dx = \pi \quad \text{for any } k = 1, 2, 3, \dots \end{aligned}$$

We leave the verifications to the reader, together with the task of showing that these functions are orthogonal:

$$\langle\sin(kx), \sin(mx)\rangle = 0 = \langle\cos(kx), \cos(mx)\rangle \quad \text{if } k \neq m$$

[6]The name honours the French mathematician J.B.J. Fourier (1768-1830) who used these techniques in 1822 to investigate heat conduction in solids.

and

$$\langle \sin(kx),\ \cos(mx) \rangle = 0 \quad \text{for all } k \geq 0 \text{ and } m \geq 0$$

(Note that $1 = \cos(0x)$, so the constant function 1 is included.)

Now define the following subspace of $\mathbf{C}[-\pi,\ \pi]$:

$$F_n = \text{span}\{1,\ \sin x,\ \cos x,\ \sin(2x),\ \cos(2x),\ \dots,\ \sin(nx),\ \cos(nx)\}$$

The aim is to use the approximation theorem (Theorem 10.2.8); so, given a function f in $\mathbf{C}[-\pi,\ \pi]$, define the **Fourier coefficients** of f by

$$a_0 = \frac{\langle f(x),\ 1\rangle}{\|1\|^2} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx$$

$$a_k = \frac{\langle f(x),\ \cos(kx)\rangle}{\|\cos(kx)\|^2} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)dx \quad k = 1,\ 2,\ \dots$$

$$b_k = \frac{\langle f(x),\ \sin(kx)\rangle}{\|\sin(kx)\|^2} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)dx \quad k = 1,\ 2,\ \dots$$

Then the approximation theorem (Theorem 10.2.8) gives Theorem 10.5.1.

Theorem 10.5.1

Let f be any continuous real-valued function defined on the interval $[-\pi,\ \pi]$. If a_0, a_1, $\dots$, and b_0, b_1, $\dots$ are the Fourier coefficients of f, then given $n \geq 0$,

$$f_n(x) = a_0 + a_1\cos x + b_1\sin x + a_2\cos(2x) + b_2\sin(2x) + \cdots + a_n\cos(nx) + b_n\sin(nx)$$

is a function in F_n that is closest to f in the sense that

$$\|f - f_n\| \leq \|f - g\|$$

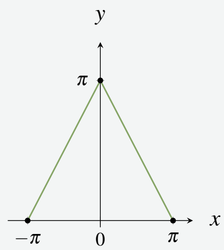
holds for all functions g in F_n.

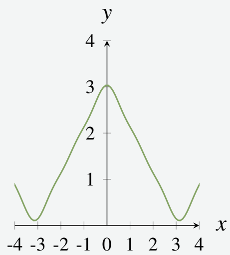
The function f_n is called the nth **Fourier approximation** to the function f.

Example 10.5.1

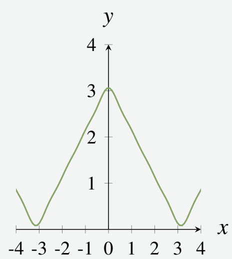
Find the fifth Fourier approximation to the function $f(x)$ defined on $[-\pi, \pi]$ as follows:

$$f(x) = \begin{cases} \pi + x & \text{if } -\pi \le x < 0 \\ \pi - x & \text{if } 0 \le x \le \pi \end{cases}$$





$f_5(x)$



$f_{13}(x)$

Solution. The graph of $y = f(x)$ appears in the top diagram. The Fourier coefficients are computed as follows. The details of the integrations (usually by parts) are omitted.

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx = \frac{\pi}{2}$$

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)dx = \frac{2}{\pi k^2}[1 - \cos(k\pi)] = \begin{cases} 0 & \text{if } k \text{ is even} \\ \frac{4}{\pi k^2} & \text{if } k \text{ is odd} \end{cases}$$

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)dx = 0 \quad \text{for all } k = 1, 2, \dots$$

Hence the fifth Fourier approximation is

$$f_5(x) = \frac{\pi}{2} + \frac{4}{\pi}\left\{\cos x + \frac{1}{3^2}\cos(3x) + \frac{1}{5^2}\cos(5x)\right\}$$

This is plotted in the middle diagram and is already a reasonable approximation to $f(x)$. By comparison, $f_{13}(x)$ is also plotted in the bottom diagram.

We say that a function f is an **even function** if $f(x) = f(-x)$ holds for all x; f is called an **odd function** if $f(-x) = -f(x)$ holds for all x. Examples of even functions are constant functions, the even powers x^2, x^4, $\dots$, and $\cos(kx)$; these functions are characterized by the fact that the graph of $y = f(x)$ is symmetric about the y axis. Examples of odd functions are the odd powers x, x^3, $\dots$, and $\sin(kx)$ where $k > 0$, and the graph of $y = f(x)$ is symmetric about the origin if f is odd. The usefulness of these functions stems from the fact that

$$\begin{array}{ll} \int_{-\pi}^{\pi} f(x)dx = 0 & \text{if } f \text{ is odd} \\ \int_{-\pi}^{\pi} f(x)dx = 2\int_0^{\pi} f(x)dx & \text{if } f \text{ is even} \end{array}$$

These facts often simplify the computations of the Fourier coefficients. For example:

1. The Fourier sine coefficients b_k all vanish if f is even.

2. The Fourier cosine coefficients a_k all vanish if f is odd.

This is because $f(x)\sin(kx)$ is odd in the first case and $f(x)\cos(kx)$ is odd in the second case.

The functions 1, $\cos(kx)$, and $\sin(kx)$ that occur in the Fourier approximation for $f(x)$ are all easy to generate as an electrical voltage (when x is time). By summing these signals (with the amplitudes given by the Fourier coefficients), it is possible to produce an electrical signal with (the approximation to) $f(x)$ as the voltage. Hence these Fourier approximations play a fundamental role in electronics.

Finally, the Fourier approximations $f_1, f_2, \dots$ of a function f get better and better as n increases. The reason is that the subspaces F_n increase:

$$F_1 \subseteq F_2 \subseteq F_3 \subseteq \cdots \subseteq F_n \subseteq \cdots$$

So, because $f_n = \text{proj}_{F_n} f$, we get (see the discussion following Example 10.2.6)

$$\|f - f_1\| \geq \|f - f_2\| \geq \cdots \geq \|f - f_n\| \geq \cdots$$

These numbers $\|f - f_n\|$ approach zero; in fact, we have the following fundamental theorem.

Theorem 10.5.2

Let f be any continuous function in $C[-\pi, \pi]$. Then

$$f_n(x) \text{ approaches } f(x) \text{ for all } x \text{ such that } -\pi < x < \pi.$$[7]

It shows that f has a representation as an infinite series, called the **Fourier series** of f:

$$f(x) = a_0 + a_1 \cos x + b_1 \sin x + a_2 \cos(2x) + b_2 \sin(2x) + \cdots$$

whenever $-\pi < x < \pi$. A full discussion of Theorem 10.5.2 is beyond the scope of this book. This subject had great historical impact on the development of mathematics, and has become one of the standard tools in science and engineering.

Thus the Fourier series for the function f in Example 10.5.1 is

$$f(x) = \frac{\pi}{2} + \frac{4}{\pi}\left\{\cos x + \frac{1}{3^2}\cos(3x) + \frac{1}{5^2}\cos(5x) + \frac{1}{7^2}\cos(7x) + \cdots\right\}$$

Since $f(0) = \pi$ and $\cos(0) = 1$, taking $x = 0$ leads to the series

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots$$

Example 10.5.2

Expand $f(x) = x$ on the interval $[-\pi, \pi]$ in a Fourier series, and so obtain a series expansion of $\frac{\pi}{4}$.

Solution. Here f is an odd function so all the Fourier cosine coefficients a_k are zero. As to the sine coefficients:

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin(kx)dx = \frac{2}{k}(-1)^{k+1} \quad \text{for } k \geq 1$$

where we omit the details of the integration by parts. Hence the Fourier series for x is

$$x = 2[\sin x - \tfrac{1}{2}\sin(2x) + \tfrac{1}{3}\sin(3x) - \tfrac{1}{4}\sin(4x) + \dots]$$

for $-\pi < x < \pi$. In particular, taking $x = \frac{\pi}{2}$ gives an infinite series for $\frac{\pi}{4}$.

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots$$

Many other such formulas can be proved using Theorem 10.5.2.

[7] We have to be careful at the end points $x = \pi$ or $x = -\pi$ because $\sin(k\pi) = \sin(-k\pi)$ and $\cos(k\pi) = \cos(-k\pi)$.

Exercises for 10.5

Exercise 10.5.1 In each case, find the Fourier approximation f_5 of the given function in $\mathbf{C}[-\pi, \pi]$.

a. $f(x) = \pi - x$

b. $f(x) = |x| = \begin{cases} x & \text{if } 0 \leq x \leq \pi \\ -x & \text{if } -\pi \leq x < 0 \end{cases}$

c. $f(x) = x^2$

d. $f(x) = \begin{cases} 0 & \text{if } -\pi \leq x < 0 \\ x & \text{if } 0 \leq x \leq \pi \end{cases}$

Exercise 10.5.2

a. Find f_5 for the even function f on $[-\pi, \pi]$ satisfying $f(x) = x$ for $0 \leq x \leq \pi$.

b. Find f_6 for the even function f on $[-\pi, \pi]$ satisfying $f(x) = \sin x$ for $0 \leq x \leq \pi$.

[*Hint*: If $k > 1$, $\int \sin x \cos(kx)$ $= \frac{1}{2}\left[\frac{\cos[(k-1)x]}{k-1} - \frac{\cos[(k+1)x]}{k+1}\right]$.]

Exercise 10.5.3

a. Prove that $\int_{-\pi}^{\pi} f(x)dx = 0$ if f is odd and that $\int_{-\pi}^{\pi} f(x)dx = 2\int_{0}^{\pi} f(x)dx$ if f is even.

b. Prove that $\frac{1}{2}[f(x) + f(-x)]$ is even and that $\frac{1}{2}[f(x) - f(-x)]$ is odd for any function f. Note that they sum to $f(x)$.

Exercise 10.5.4 Show that $\{1, \cos x, \cos(2x), \cos(3x), \dots\}$ is an orthogonal set in $\mathbf{C}[0, \pi]$ with respect to the inner product $\langle f, g\rangle = \int_0^{\pi} f(x)g(x)dx$.

Exercise 10.5.5

a. Show that $\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \cdots$ using Exercise 10.5.1(b).

b. Show that $\frac{\pi^2}{12} = 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \cdots$ using Exercise 10.5.1(c).