(JRA)

5.6 INNER-PRODUCT SPACES, ORTHOGONAL BASES, AND PROJECTIONS (OPTIONAL)

Up to now we have considered a vector space solely as an entity with an algebraic structure. We know, however, that R^n possesses more than just an algebraic structure; in particular, we know that we can measure the size or length of a vector $\mathbf{x}$ in R^n by the quantity $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T\mathbf{x}}$. Similarly, we can define the distance from $\mathbf{x}$ to $\mathbf{y}$ as $\|\mathbf{x} - \mathbf{y}\|$. The ability to measure distances means that R^n has a geometric structure, which supplements the algebraic structure. The geometric structure can be employed to study problems of convergence, continuity, and the like. In this section we briefly describe how a suitable measure of distance might be imposed on a general vector space. Our development will be brief, and we will leave most of the details to the reader; but the ideas parallel those in Sections 3.6 and 3.8–3.9.

Inner-Product Spaces

To begin, we observe that the geometric structure for R^n is based on the scalar product $\mathbf{x}^T\mathbf{y}$. Essentially the scalar product is a real-valued function of two vector variables: Given $\mathbf{x}$ and $\mathbf{y}$ in R^n, the scalar product produces a number $\mathbf{x}^T\mathbf{y}$. Thus to derive a geometric structure for a vector space V, we should look for a generalization of the scalar-product function. A consideration of the properties of the scalar-product function leads to the definition of an inner-product function for a vector space. (With reference to Definition 7, which follows, we note that the expression $\mathbf{u}^T\mathbf{v}$ does not make sense in a general vector space V. Thus not only does the nomenclature change—scalar product becomes inner product—but also the notation changes as well, with $\langle \mathbf{u}, \mathbf{v} \rangle$ denoting the inner product of $\mathbf{u}$ and $\mathbf{v}$.)

DEFINITION 7 An ***inner product*** on a real vector space V is a function that assigns a real number, $\langle \mathbf{u}, \mathbf{v} \rangle$, to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V, and that satisfies these properties:

1. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \boldsymbol{\theta}$.
2. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.
3. $\langle a\mathbf{u}, \mathbf{v} \rangle = a\langle \mathbf{u}, \mathbf{v} \rangle$.
4. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.

The usual scalar product in R^n is an inner product in the sense of Definition 7, where $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$. To illustrate the flexibility of Definition 7, we also note that there are other sorts of inner products for R^n. The following example gives another inner product for R^2.

EXAMPLE 1 Let V be the vector space R^2, and let A be the (2×2) matrix

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

Verify that the function $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^TA\mathbf{v}$ is an inner product for R^2.

Solution Let $\mathbf{u}$ be a vector in R^2:

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Then

$$\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^TA\mathbf{u} = [u_1, u_2]\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

so $\langle \mathbf{u}, \mathbf{u} \rangle = 3u_1^2 + 4u_1u_2 + 4u_2^2 = 2u_1^2 + (u_1 + 2u_2)^2$. Thus $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $u_1 = u_2 = 0$. This shows that property 1 of Definition 7 is satisfied.

To see that property 2 of Definition 7 holds, note that A is symmetric; that is, $A^T = A$. Also observe that if $\mathbf{u}$ and $\mathbf{v}$ are in R^2, then $\mathbf{u}^TA\mathbf{v}$ is a (1×1) matrix, so $(\mathbf{u}^TA\mathbf{v})^T = \mathbf{u}^TA\mathbf{v}$. It now follows that $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^TA\mathbf{v} = (\mathbf{u}^TA\mathbf{v})^T = \mathbf{v}^TA^T(\mathbf{u}^T)^T = \mathbf{v}^TA^T\mathbf{u} = \langle \mathbf{v}, \mathbf{u} \rangle$.

Properties 3 and 4 of Definition 7 follow easily from the properties of matrix multiplication, so $\langle \mathbf{u}, \mathbf{v} \rangle$ is an inner product for R^2. ∎

In Example 1, an inner product for R^2 was defined in terms of a matrix A:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^TA\mathbf{v}.$$

In general, we might ask the following question:

> *"For what $(n \times n)$ matrices, A, does the operation $\mathbf{u}^TA\mathbf{v}$ define an inner product on R^n?"*

The answer to this question is suggested by the solution to Example 1. In particular (see Exercises 3 and 32), the operation $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$ is an inner product for R^n if and only if A is a symmetric positive-definite matrix.

There are a number of ways in which inner products can be defined on spaces of functions. For example, Exercise 6 will show that

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

defines one inner product for $\mathcal{P}_2$. The following example gives yet another inner product for $\mathcal{P}_2$.

EXAMPLE 2 For $p(t)$ and $q(t)$ in $\mathcal{P}_2$, verify that

$$\langle p, q \rangle = \int_0^1 p(t)q(t)\, dt$$

is an inner product.

Solution To check property 1 of Definition 7, note that

$$\langle p, p \rangle = \int_0^1 p(t)^2\, dt,$$

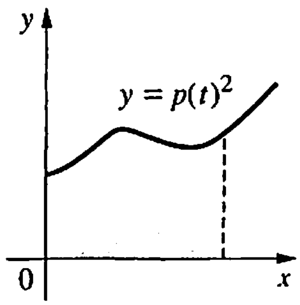

Figure 5.5
The value $\langle p, p \rangle$ is equal to the area under the graph of $y = p(t)^2$.

and $p(t)^2 \geq 0$ for $0 \leq t \leq 1$. Thus $\langle p, p \rangle$ is the area under the curve $p(t)^2, 0 \leq t \leq 1$. Hence $\langle p, p \rangle \geq 0$, and equality holds if and only if $p(t) = 0, 0 \leq t \leq 1$ (see Fig. 5.5).

Properties 2, 3, and 4 of Definition 7 are straightforward to verify, and we include here only the verification of property 4. If $p(t)$, $q(t)$, and $r(t)$ are in $\mathcal{P}_2$, then

$$\langle p, q + r \rangle = \int_0^1 p(t)[q(t) + r(t)]\, dt = \int_0^1 [p(t)q(t) + p(t)r(t)]\, dt$$

$$= \int_0^1 p(t)q(t)dt + \int_0^1 p(t)r(t)dt = \langle p, q \rangle + \langle p, r \rangle,$$

as required by property 4. ■

After the key step of defining a vector-space analog of the scalar product, the rest is routine. For purposes of reference we call a vector space with an inner product an ***inner-product space***. As in R^n, we can use the inner product as a measure of size: If V is an inner-product space, then for each $\mathbf{v}$ in V we define $\|\mathbf{v}\|$ (the ***norm*** of $\mathbf{v}$) as

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

Note that $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all $\mathbf{v}$ in V, so the norm function is always defined.

EXAMPLE 3 Use the inner product for $\mathcal{P}_2$ defined in Example 2 to determine $\|t^2\|$.

Solution By definition, $\|t^2\| = \sqrt{\langle t^2, t^2 \rangle}$. But $\langle t^2, t^2 \rangle = \int_0^1 t^2 t^2\, dt = \int_0^1 t^4\, dt = 1/5$, so $\|t^2\| = 1/\sqrt{5}$. ■

Before continuing, we pause to illustrate one way in which the inner-product space framework is used in practice. One of the many inner products for the vector space $C[0, 1]$ is

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\, dx.$$

If f is a relatively complicated function in $C[0, 1]$, we might wish to approximate f by a simpler function, say a polynomial. For definiteness suppose we want to find a polynomial p in $\mathcal{P}_2$ that is a good approximation to f. The phrase "good approximation" is too vague to be used in any calculation, but the inner-product space framework allows us to measure size and thus to pose some meaningful problems. In particular, we can ask for a polynomial p^* in $\mathcal{P}_2$ such that

$$\|f - p^*\| \leq \|f - p\|$$

for all p in $\mathcal{P}_2$. Finding such a polynomial p^* in this setting is equivalent to minimizing

$$\int_0^1 [f(x) - p(x)]^2 \, dx$$

among all p in $\mathcal{P}_2$. We will present a procedure for doing this shortly.

Orthogonal Bases

If $\mathbf{u}$ and $\mathbf{v}$ are vectors in an inner-product space V, we say that $\mathbf{u}$ and $\mathbf{v}$ are ***orthogonal*** if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Similarly, $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ is an ***orthogonal set*** in V if $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ when $i \neq j$. In addition, if an orthogonal set of vectors B is a basis for V, we call B an ***orthogonal basis***. The next two theorems correspond to their analogs in R^n, and we leave the proofs to the exercises. [See Eq. (5a), Eq. (5b), and Theorem 14 of Section 3.6.]

THEOREM 10 Let $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthogonal basis for an inner-product space V. If $\mathbf{u}$ is any vector in V, then

$$\mathbf{u} = \frac{\langle \mathbf{v}_1, \mathbf{u} \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 + \frac{\langle \mathbf{v}_2, \mathbf{u} \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 + \cdots + \frac{\langle \mathbf{v}_n, \mathbf{u} \rangle}{\langle \mathbf{v}_n, \mathbf{v}_n \rangle}\mathbf{v}_n.$$

THEOREM 11 **Gram–Schmidt Orthogonalization** Let V be an inner-product space, and let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a basis for V. Let $\mathbf{v}_1 = \mathbf{u}_1$, and for $2 \leq k \leq n$ define $\mathbf{v}_k$ by

$$\mathbf{v}_k = \mathbf{u}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{u}_k, \mathbf{v}_j \rangle}{\langle \mathbf{v}_j, \mathbf{v}_j \rangle}\mathbf{v}_j.$$

Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for V.

EXAMPLE 4 Let the inner product on $\mathcal{P}_2$ be the one given in Example 2. Starting with the natural basis $\{1, x, x^2\}$, use Gram–Schmidt orthogonalization to obtain an orthogonal basis for $\mathcal{P}_2$.

Solution If we let $\{p_0, p_1, p_2\}$ denote the orthogonal basis, we have $p_0(x) = 1$ and find $p_1(x)$ from

$$p_1(x) = x - \frac{\langle p_0, x \rangle}{\langle p_0, p_0 \rangle} p_0(x).$$

We calculate

$$\langle p_0, x \rangle = \int_0^1 x \, dx = 1/2 \text{ and } \langle p_0, p_0 \rangle = \int_0^1 dx = 1;$$

so $p_1(x) = x - 1/2$. The next step of the Gram–Schmidt orthogonalization process is to form

$$p_2(x) = x^2 - \frac{\langle p_1, x^2\rangle}{\langle p_1, p_1\rangle} p_1(x) - \frac{\langle p_0, x^2\rangle}{\langle p_0, p_0\rangle} p_0(x).$$

The required constants are

$$\langle p_1, x^2\rangle = \int_0^1 (x^3 - x^2/2)\,dx = 1/12$$

$$\langle p_1, p_1\rangle = \int_0^1 (x^2 - x + 1/4)\,dx = 1/12$$

$$\langle p_0, x^2\rangle = \int_0^1 x^2\,dx = 1/3$$

$$\langle p_0, p_0\rangle = \int_0^1 dx = 1.$$

Therefore, $p_2(x) = x^2 - p_1(x) - p_0(x)/3 = x^2 - x + 1/6$, and $\{p_0, p_1, p_2\}$ is an orthogonal basis for $\mathcal{P}_2$ with respect to the inner product. ■

EXAMPLE 5 Let $B = \{p_0, p_1, p_2\}$ be the orthogonal basis for $\mathcal{P}_2$ obtained in Example 4. Find the coordinates of x^2 relative to B.

Solution By Theorem 10, $x^2 = a_0 p_0(x) + a_1 p_1(x) + a_2 p_2(x)$, where

$$\begin{aligned} a_0 &= \langle p_0, x^2\rangle/\langle p_0, p_0\rangle \\ a_1 &= \langle p_1, x^2\rangle/\langle p_1, p_1\rangle \\ a_2 &= \langle p_2, x^2\rangle/\langle p_2, p_2\rangle. \end{aligned}$$

The necessary calculations are

$$\langle p_0, x^2\rangle = \int_0^1 x^2\,dx = 1/3$$

$$\langle p_1, x^2\rangle = \int_0^1 [x^3 - (1/2)x^2]\,dx = 1/12$$

$$\langle p_2, x^2\rangle = \int_0^1 [x^4 - x^3 + (1/6)x^2]\,dx = 1/180$$

$$\langle p_0, p_0\rangle = \int_0^1 dx = 1$$

$$\langle p_1, p_1\rangle = \int_0^1 [x^2 - x + 1/4]\,dx = 1/12$$

$$\langle p_2, p_2\rangle = \int_0^1 [x^2 - x + 1/6]^2\,dx = 1/180.$$

Thus $a_0 = 1/3$, $a_1 = 1$, and $a_2 = 1$. We can easily check that $x^2 = (1/3)p_0(x) + p_1(x) + p_2(x)$. ■

Orthogonal Projections

We return now to the previously discussed problem of finding a polynomial p^* in $\mathcal{P}_2$ that is the best approximation of a function f in $C[0, 1]$. Note that the problem amounts to determining a vector p^* in a subspace of an inner-product space, where p^* is closer to f than any other vector in the subspace. The essential aspects of this problem can be stated formally as the following general problem:

Let V be an inner-product space and let W be a subspace of V. Given a vector $\mathbf{v}$ in V, find a vector $\mathbf{w}^*$ in W such that

$$\|\mathbf{v} - \mathbf{w}^*\| \leq \|\mathbf{v} - \mathbf{w}\| \quad \text{for all } \mathbf{w} \text{ in } W. \tag{1}$$

A vector $\mathbf{w}^*$ in W satisfying inequality (1) is called the ***projection of v*** onto W, or (frequently) the ***best least-squares approximation*** to $\mathbf{v}$. Intuitively $\mathbf{w}^*$ is the nearest vector in W to $\mathbf{v}$.

The solution process for this problem is almost exactly the same as that for the least-squares problem in R^n. One distinction in our general setting is that the subspace W might not be finite dimensional. If W is an infinite-dimensional subspace of V, then there may or may not be a projection of $\mathbf{v}$ onto W. If W is finite dimensional, then a projection always exists, is unique, and can be found explicitly. The next two theorems outline this concept, and again we leave the proofs to the reader since they parallel the proof of Theorem 18 of Section 3.9.

THEOREM 12 Let V be an inner-product space, and let W be a subspace of V. Let $\mathbf{v}$ be a vector in V, and suppose $\mathbf{w}^*$ is a vector in W such that

$$\langle \mathbf{v} - \mathbf{w}^*, \mathbf{w} \rangle = 0 \quad \text{for all } \mathbf{w} \text{ in } W.$$

Then $\|\mathbf{v} - \mathbf{w}^*\| \leq \|\mathbf{v} - \mathbf{w}\|$ for all $\mathbf{w}$ in W with equality holding only for $\mathbf{w} = \mathbf{w}^*$.

THEOREM 13 Let V be an inner-product space, and let $\mathbf{v}$ be a vector in V. Let W be an n-dimensional subspace of V, and let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal basis for W. Then

$$\|\mathbf{v} - \mathbf{w}^*\| \leq \|\mathbf{v} - \mathbf{w}\| \text{ for all } \mathbf{w} \text{ in } W$$

if and only if

$$\mathbf{w}^* = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\mathbf{u}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle}\mathbf{u}_n. \tag{2}$$

In view of Theorem 13, it follows that when W is a finite-dimensional subspace of an inner-product space V, we can always find projections by first finding an orthogonal basis for W (by using Theorem 11) and then calculating the projection $\mathbf{w}^*$ from Eq. (2).

To illustrate the process of finding a projection, we return to the inner-product space $C[0, 1]$ with the subspace $\mathcal{P}_2$. As a specific but rather unrealistic function, f, we choose $f(x) = \cos x$, x in radians. The inner product is

$$\langle f, g \rangle = \int_0^1 f(x)g(x)\,dx.$$

Example 6 In the vector space $C[0,1]$, let $f(x) = \cos x$. Find the projection of f onto the subspace $\mathcal{P}_2$.

Solution Let $\{p_0, p_1, p_2\}$ be the orthogonal basis for $\mathcal{P}_2$ found in Example 4. (Note that the inner product used in Example 4 coincides with the present inner product on $C[0,1]$. By Theorem 13, the projection of f onto $\mathcal{P}_2$ is the polynomial p^* defined by

$$p^*(x) = \frac{\langle f, p_0\rangle}{\langle p_0, p_0\rangle} p_0(x) + \frac{\langle f, p_1\rangle}{\langle p_1, p_1\rangle} p_1(x) + \frac{\langle f, p_2\rangle}{\langle p_2, p_2\rangle} p_2(x),$$

where

$$\langle f, p_0\rangle = \int_0^1 \cos(x)\,dx \simeq .841471$$

$$\langle f, p_1\rangle = \int_0^1 (x - 1/2)\cos(x)\,dx \simeq .038962$$

$$\langle f, p_2\rangle = \int_0^1 (x^2 - x + 1/6)\cos(x)\,dx \simeq -.002394.$$

From Example 5, we have $\langle p_0, p_0\rangle = 1$, $\langle p_1, p_1\rangle = 1/12$, and $\langle p_2, p_2\rangle = 1/180$. Therefore, $p^*(x)$ is given by

$$\begin{aligned} p^*(x) &= \langle f, p_0\rangle p_0(x) + 12\langle f, p_1\rangle p_1(x) + 180\langle f, p_2\rangle p_2(x) \\ &\simeq .841471 p_0(x) - .467544 p_1(x) - .430920 p_2(x). \end{aligned}$$

In order to assess how well $p^*(x)$ approximates $\cos x$ in the interval $[0, 1]$, we can tabulate $p^*(x)$ and $\cos x$ at various values of x (see Table 5.1). ■

Example 7 The function $\mathrm{Si}(x)$ (important in applications such as optics) is defined as follows:

$$\mathrm{Si}(x) = \int_0^x \frac{\sin u}{u}\,du, \text{ for } x \neq 0. \tag{3}$$

The integral in (3) is not an elementary one and so, for a given value of x, $\mathrm{Si}(x)$ must be evaluated using a numerical integration procedure. In this example, we approximate

Table 5.1

x	$p^*(x)$	$\cos x$	$p^*(x) - \cos x$
0.0	1.0034	1.000	.0034
0.2	.9789	.9801	−.0012
0.4	.9198	.9211	−.0013
0.6	.8263	.8253	.0010
0.8	.6983	.6967	.0016
1.0	.5359	.5403	−.0044

Si(x) by a cubic polynomial for $0 \le x \le 1$. In particular, it can be shown that if we define Si(0) = 0, then Si(x) is continuous for all x. Thus we can ask:

"What is the projection of Si(x) onto the subspace $\mathcal{P}_3$ of $C[0, 1]$?"

This projection will serve as an approximation to Si(x) for $0 \le x \le 1$.

Solution We used the computer algebra system Derive to carry out the calculations. Some of the steps are shown in Fig. 5.6. To begin, let $\{p_0, p_1, p_2, p_3\}$ be the orthogonal basis for $\mathcal{P}_3$ found by the Gram–Schmidt process. From Example 4, we already know that

6: $\displaystyle\int_0^1 x^3 \text{ P1 (x) dx}$

7: $\displaystyle\frac{3}{40}$

8: $\displaystyle\int_0^1 x^3 \text{ P2 (x) dx}$

9: $\displaystyle\frac{1}{120}$

15: $\text{P3 (x) :=}x^3 - \frac{1}{4}\text{ P0 (x)} - \frac{9}{10}\text{ P1 (x)} - \frac{3}{2}\text{ P2 (x)}$

16: $\text{P3 (x) :=}x^3 - \frac{3x^2}{2} + \frac{3x}{5} - \frac{1}{20}$

17: $\displaystyle\int_0^1 \text{P3 (x) P3 (x) dx}$

18: $\displaystyle\frac{1}{2800}$

49: $\displaystyle\int_0^1 180 \text{ P2 (x)} \int_0^x \frac{\text{SIN (u)}}{\text{u}} \text{ du dx}$

50: -0.0804033

51: $\displaystyle\int_0^1 2800 \text{ P3 (x)} \int_0^x \frac{\text{SIN (u)}}{\text{u}} \text{ du dx}$

52: -0.0510442

Figure 5.6 Some of the steps used by Derive to generate the projection of Si(x) onto $\mathcal{P}_3$ in Example 7

$p_0(x) = 1$, $p_1(x) = x - 1/2$, and $p_2(x) = x^2 - x + 1/6$. To find p_3, we first calculate the inner products

$$\langle p_0, x^3\rangle, \langle p_1, x^3\rangle, \langle p_2, x^3\rangle$$

(see steps 6–9 in Fig. 5.6 for $\langle p_1, x^3\rangle$ and $\langle p_2, x^3\rangle$).

Using Theorem 11, we find p_3 and, for later use, $\langle p_3, p_3\rangle$:

$$\begin{aligned} p_3(x) &= x^3 - (3/2)x^2 + (3/5)x - 1/20 \\ \langle p_3, p_3\rangle &= 1/2800 \end{aligned}$$

(see steps 15–18 in Fig. 5.6). Finally, by Theorem 13, the projection of $\mathrm{Si}(x)$ onto $\mathcal{P}_3$ is the polynomial p^* defined by

$$\begin{aligned} p^*(x) &= \frac{\langle \mathrm{Si}, p_0\rangle}{\langle p_0, p_0\rangle} p_0(x) + \frac{\langle \mathrm{Si}, p_1\rangle}{\langle p_1, p_1\rangle} p_1(x) + \frac{\langle \mathrm{Si}, p_2\rangle}{\langle p_2, p_2\rangle} p_2(x) + \frac{\langle \mathrm{Si}, p_3\rangle}{\langle p_3, p_3\rangle} p_3(x) \\ &= \langle \mathrm{Si}, p_0\rangle p_0(x) + 12\langle \mathrm{Si}, p_1\rangle p_1(x) + 180\langle \mathrm{Si}, p_2\rangle p_2(x) + 2800\langle \mathrm{Si}, p_3\rangle p_3(x). \end{aligned}$$

In the expression above for p^*, the inner products $\langle \mathrm{Si}, p_k\rangle$ for $k = 0, 1, 2$, and 3 are given by

$$\langle \mathrm{Si}, p_k\rangle = \int_0^1 p_k(x)\mathrm{Si}(x)\, dx = \int_0^1 p_k(x) \left\{ \int_0^x \frac{\sin u}{u} du \right\} dx$$

(see steps 49–52 in Fig. 5.6 for $180\langle \mathrm{Si}, p_2\rangle$ and $2800\langle \mathrm{Si}, p_3\rangle$).

Now, since $\mathrm{Si}(x)$ must be estimated numerically, it follows that the inner products $\langle \mathrm{Si}, p_k\rangle$ must be estimated as well. Using Derive to approximate the inner products, we obtain the projection (or best least-squares approximation)

$$p^*(x) = .486385p_0(x) + .951172p_1(x) - .0804033p_2(x) - .0510442p_3(x).$$

To assess how well $p^*(x)$ approximates $\mathrm{Si}(x)$ in $[0, 1]$, we tabulate each function at a few selected points (see Table 5.2). As can be seen from Table 5.2, it appears that $p^*(x)$ is a very good approximation to $\mathrm{Si}(x)$. ∎

Table 5.2

x	$p^*(x)$	$\mathrm{Si}(x)$	$p^*(x) - \mathrm{Si}(x)$
0.0	.000049	.000000	.000049
0.2	.199578	.199556	.000028
0.4	.396449	.396461	−.000012
0.6	.588113	.588128	−.000015
0.8	.772119	.772095	.000024
1.0	.946018	.946083	−.000065

5.6 EXERCISES

1. Prove that $\langle \mathbf{x}, \mathbf{y} \rangle = 4x_1y_1 + x_2y_2$ is an inner product on R^2, where

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

2. Prove that $\langle \mathbf{x}, \mathbf{y} \rangle = a_1x_1y_1 + a_2x_2y_2 + \cdots + a_nx_ny_n$ is an inner product on R^n, where $a_1, a_2, \ldots, a_n$ are positive real numbers and where

$$\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \quad \text{and}$$
$$\mathbf{y} = [y_1, y_2, \ldots, y_n]^T.$$

3. A real $(n \times n)$ symmetric matrix A is called ***positive definite*** if $\mathbf{x}^TA\mathbf{x} > 0$ for all $\mathbf{x}$ in R^n, $\mathbf{x} \neq \boldsymbol{\theta}$. Let A be a symmetric positive-definite matrix, and verify that

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^TA\mathbf{y}$$

defines an inner product on R^n; that is, verify that the four properties of Definition 7 are satisfied.

4. Prove that the following symmetric matrix A is positive definite. Prove this by choosing an arbitrary vector $\mathbf{x}$ in R^2, $\mathbf{x} \neq \boldsymbol{\theta}$, and calculating $\mathbf{x}^TA\mathbf{x}$.

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

5. In $\mathcal{P}_2$ let $p(x) = a_0 + a_1x + a_2x^2$ and $q(x) = b_0 + b_1x + b_2x^2$. Prove that $\langle p, q \rangle = a_0b_0 + a_1b_1 + a_2b_2$ is an inner product on $\mathcal{P}_2$.

6. Prove that $\langle p, q \rangle = p(0)q(0) + p(1)q(1) + p(2)q(2)$ is an inner product on $\mathcal{P}_2$.

7. Let $A = (a_{ij})$ and $B = (b_{ij})$ be (2×2) matrices. Show that $\langle A, B \rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$ is an inner product for the vector space of all (2×2) matrices.

8. For $\mathbf{x} = [1, -2]^T$ and $\mathbf{y} = [0, 1]^T$ in R^2, find $\langle \mathbf{x}, \mathbf{y} \rangle$, $\|\mathbf{x}\|$, $\|\mathbf{y}\|$, and $\|\mathbf{x} - \mathbf{y}\|$ using the inner product given in Exercise 1.

9. Repeat Exercise 8 with the inner product defined in Exercise 3 and the matrix A given in Exercise 4.

0. In $\mathcal{P}_2$ let $p(x) = -1 + 2x + x^2$ and $q(x) = 1 - x + 2x^2$. Using the inner product given in Exercise 5, find $\langle p, q \rangle$, $\|p\|$, $\|q\|$, and $\|p - q\|$.

11. Repeat Exercise 10 using the inner product defined in Exercise 6.

12. Show that $\{1, x, x^2\}$ is an orthogonal basis for $\mathcal{P}_2$ with the inner product defined in Exercise 5 but not with the inner product in Exercise 6.

13. In R^2 let $S = \{\mathbf{x}: \|\mathbf{x}\| = 1\}$. Sketch a graph of S if $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T\mathbf{y}$. Now graph S using the inner product given in Exercise 1.

14. Let A be the matrix given in Exercise 4, and for $\mathbf{x}, \mathbf{y}$ in R^2 define $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^TA\mathbf{y}$ (see Exercise 3). Starting with the natural basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, use Theorem 11 to obtain an orthogonal basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for R^2.

15. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be the orthogonal basis for R^2 obtained in Exercise 14 and let $\mathbf{v} = [3, 4]^T$. Use Theorem 10 to find scalars a_1, a_2 such that $\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$.

16. Use Theorem 11 to calculate an orthogonal basis $\{p_0, p_1, p_2\}$ for $\mathcal{P}_2$ with respect to the inner product in Exercise 6. Start with the natural basis $\{1, x, x^2\}$ for $\mathcal{P}_2$.

17. Use Theorem 10 to write $q(x) = 2 + 3x - 4x^2$ in terms of the orthogonal basis $\{p_0, p_1, p_2\}$ obtained in Exercise 16.

18. Show that the function defined in Exercise 6 is not an inner product for $\mathcal{P}_3$. [*Hint:* Find $p(x)$ in $\mathcal{P}_3$ such that $\langle p, p \rangle = 0$, but $p \neq \boldsymbol{\theta}$.]

19. Starting with the natural basis $\{1, x, x^2, x^3, x^4\}$, generate an orthogonal basis for $\mathcal{P}_4$ with respect to the inner product

$$\langle p, q \rangle = \sum_{i=-2}^{2} p(i)q(i).$$

20. If V is an inner-product space, show that $\langle \mathbf{v}, \boldsymbol{\theta} \rangle = 0$ for each vector $\mathbf{v}$ in V.

21. Let V be an inner-product space, and let $\mathbf{u}$ be a vector in V such that $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ for every vector $\mathbf{v}$ in V. Show that $\mathbf{u} = \boldsymbol{\theta}$.

22. Let a be a scalar and $\mathbf{v}$ a vector in an inner-product space V. Prove that $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$.

23. Prove that if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is an orthogonal set of nonzero vectors in an inner-product space, then this set is linearly independent.

24. Prove Theorem 10.

25. Approximate x^3 with a polynomial in $\mathcal{P}_2$. [*Hint:* Use the inner product
$$\langle p, q\rangle = \int_0^1 p(t)q(t)\,dt,$$
and let $\{p_0, p_1, p_2\}$ be the orthogonal basis for $\mathcal{P}_2$ obtained in Example 4. Now apply Theorem 13.]

26. In Examples 4 and 7 we found $p_0(x), \ldots, p_3(x)$, which are orthogonal with respect to
$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx.$$
Continue the process, and find $p_4(x)$ so that $\{p_0, p_1, \ldots, p_4\}$ is an orthogonal basis for $\mathcal{P}_4$. (Clearly there is an infinite sequence of polynomials $p_0, p_1, \ldots, p_n, \ldots$ that satisfy
$$\int_0^1 p_i(x)p_j(x)\,dx = 0, \quad i \neq j.$$
These are called the ***Legendre polynomials***.)

27. With the orthogonal basis for $\mathcal{P}_3$ obtained in Example 7, use Theorem 13 to find the projection of $f(x) = \cos x$ in $\mathcal{P}_3$. Construct a table similar to Table 5.1 and note the improvement.

28. An inner product on $C[-1, 1]$ is
$$\langle f, g\rangle = \frac{2}{\pi}\int_{-1}^1 \frac{f(x)g(x)}{\sqrt{1-x^2}}\,dx.$$
Starting with the set $\{1, x, x^2, x^3, \ldots\}$, use the Gram–Schmidt process to find polynomials $T_0(x)$, $T_1(x)$, $T_2(x)$, and $T_3(x)$ such that $\langle T_i, T_j\rangle = 0$ when $i \neq j$. These polynomials are called the ***Chebyshev polynomials of the first kind***. [*Hint:* Make a change of variables $x = \cos\theta$.]

29. A sequence of orthogonal polynomials usually satisfies a three-term recurrence relation. For example, the Chebyshev polynomials are related by
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \ldots, \tag{R}$$
where $T_0(x) = 1$ and $T_1(x) = x$. Verify that the polynomials defined by the relation (R) above are indeed orthogonal in $C[-1, 1]$ with respect to the inner product in Exercise 28. Verify this as follows:

a) Make the change of variables $x = \cos\theta$, and use induction to show that $T_k(\cos\theta) = \cos k\theta$, $k = 0, 1, \ldots$, where $T_k(x)$ is defined by (R).

b) Using part a), show that $\langle T_i, T_j\rangle = 0$ when $i \neq j$.

c) Use induction to show that $T_k(x)$ is a polynomial of degree k, $k = 0, 1, \ldots$.

d) Use (R) to calculate T_2, T_3, T_4, and T_5.

30. Let $C[-1, 1]$ have the inner product of Exercise 28, and let f be in $C[-1, 1]$. Use Theorem 13 to prove that $\|f - p^*\| \leq \|f - p\|$ for all p in $\mathcal{P}_n$ if
$$p^*(x) = \frac{a_0}{2} + \sum_{j=1}^n a_j T_j(x),$$
where $a_j = \langle f, T_j\rangle$, $j = 0, 1, \ldots, n$.

31. The iterated trapezoid rule provides a good estimate of $\int_a^b f(x)\,dx$ when $f(x)$ is periodic in $[a, b]$. In particular, let N be a positive integer, and let $h = (b-a)/N$. Next, define x_i by $x_i = a + ih$, $i = 0, 1, \ldots, N$, and suppose $f(x)$ is in $C[a, b]$. If we define $A(f)$ by
$$A(f) = \frac{h}{2}f(x_0) + h\sum_{j=1}^{N-1} f(x_j) + \frac{h}{2}f(x_N),$$
then $A(f)$ is the iterated trapezoid rule applied to $f(x)$. Using the result in Exercise 30, write a computer program that generates a good approximation to $f(x)$ in $C[-1, 1]$. That is, for an input function $f(x)$ and a specified value of n, calculate estimates of $a_0, a_1, \ldots, a_n$, where
$$a_k = \langle f, T_k\rangle \simeq A(fT_k).$$
To do this calculation, make the usual change of variables $x = \cos\theta$ so that
$$a_k = \frac{2}{\pi}\int_0^\pi f(\cos\theta)\cos(k\theta)d\theta, \quad k = 0, 1, \ldots, n.$$
Use the iterated trapezoid rule to estimate each a_k. Test your program on $f(x) = e^{2x}$ and note that (R) can be used to evaluate $p^*(x)$ at any point x in $[-1, 1]$.

32. Show that if A is a real $(n \times n)$ matrix and if the expression $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$ defines an inner product on R^n, then A must be symmetric and positive definite (see Exercise 3 for the definition of positive definite). [*Hint:* Consider $\langle \mathbf{e}_i, \mathbf{e}_j\rangle$.]