0 Inner Product Spaces

An **inner product** on a vector space V is a function that associates with each (ordered) pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V a scalar $\langle \mathbf{u}, \mathbf{v} \rangle$ such that

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

(ii) $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$;

(iii) $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$;

(iv) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$; $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

An **inner product space** is a vector space V together with a specified inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ on V.

The Euclidean inner product—that is, the dot product $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$—is only one example of an inner product on the vector space $\mathbf{R}^n$ of n-tuples of real numbers. To see how other inner products on $\mathbf{R}^n$ can be defined, let A be a fixed $n \times n$ matrix. Given (column) vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbf{R}^n$, let us define the "product" $\langle \mathbf{u}, \mathbf{v} \rangle$ of these two vectors to be

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}. \tag{1}$$

Note that $\langle \mathbf{u}, \mathbf{v} \rangle$ is a 1×1 matrix—that is, $\langle \mathbf{u}, \mathbf{v} \rangle$ is a scalar. Then

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= \mathbf{u}^T A(\mathbf{v} + \mathbf{w}) \\ &= \mathbf{u}^T A\mathbf{v} + \mathbf{u}^T A\mathbf{w} \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle\end{aligned}$$

and

$$\begin{aligned}\langle c\mathbf{u}, \mathbf{v} \rangle &= (c\mathbf{u}^T)A\mathbf{v} \\ &= c\mathbf{u}^T A\mathbf{v} = c\langle \mathbf{u}, \mathbf{v} \rangle,\end{aligned}$$

so we see immediately that $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v}$ satisfies properties (ii) and (inner product.

In order to verify properties (i) and (iv) we must impose appropriate c on the matrix A. Suppose first that A is *symmetric:* $A = A^T$. Because **u** real number, it follows that $(\mathbf{u}^T A\mathbf{v})^T = \mathbf{u}^T A\mathbf{v}$. Consequently

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle &= \mathbf{u}^T A\mathbf{v} = (\mathbf{u}^T A\mathbf{v})^T \\ &= \mathbf{v}^T A^T u = \mathbf{v}^T A\mathbf{u} = \langle \mathbf{v}, \mathbf{u} \rangle .\end{aligned}$$

Thus the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v}$ satisfies property (i) provided that th A is symmetric.

The symmetric $n \times n$ matrix A is said to be **positive definite** if $\mathbf{u}^T A\mathbf{u}$ every nonzero n-vector **u**, in which case $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v}$ satisfies property (i innerproduct. Then our discussion shows that *if the $n \times n$ matrix A is symme positive definite, then*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v}$$

defines an inner product on $\mathbf{R}^n$. The familiar dot product $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} =$ simply the special case in which $A = I$, the $n \times n$ identity matrix.

Later we will state criteria for determining whether a given symmetri matrix A is positive definite, and hence whether $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v}$ defines a product on $\mathbf{R}^n$. In the case of a symmetric 2×2 matrix

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

this question can be answered by a simple technique of completing the squar Example 1 of this section. Note that if $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A\mathbf{v} = [u_1 \quad u_2] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

so that

$$\langle \mathbf{u}, \mathbf{v} \rangle = au_1v_1 + bu_1v_2 + bu_2v_1 + cu_2v_2.$$

Example 1 Consider the symmetric 2×2 matrix

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

Then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v} = 3u_1 v_1 + 2u_1 v_2 + 2u_2 v_1 + 4u_2 v_2$$

automatically satisfies properties (i)–(iii) of an inner product on $\mathbf{R}^2$. If $\mathbf{u} = (x,$
then (3) gives

$$\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^T A \mathbf{u} = 3x^2 + 4xy + 4y^2 = (x + 2y)^2 + 2x^2.$$

It is therefore clear that $\mathbf{u}^T A \mathbf{u} \geq 0$ and that $\mathbf{u}^T A \mathbf{v} = 0$ if and only if $x + 2y = 0 =$
that is, if and only if $x = y = 0$. Thus the symmetric matrix A is positive defin
and so $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$ defines an inner product on $\mathbf{R}^2$. Note that if $\mathbf{u} = (3, 1)$ a
$\mathbf{v} = (1, 4)$, then $\mathbf{u} \cdot \mathbf{v} = 7$, whereas

$$\langle \mathbf{u}, \mathbf{v} \rangle = [3 \quad 1] \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = 51.$$

Thus the inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$ is quite different from the Euclidean in
product on $\mathbf{R}^2$.

Essentially everything that has been done with the Euclidean inner produc
$\mathbf{R}^n$ in the first two sections of this chapter can be done with an arbitrary inner proc
space V (with an occasional proviso that the vector space V be finite-dimension
Given an arbitrary inner product $\langle \mathbf{u}, \mathbf{v} \rangle$ on a vector space V the **length** (or **norm**
the vector $\mathbf{u}$ (with respect to this inner product) is defined to be

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

For instance, the length of $\mathbf{u} = (3, 1)$ with respect to the inner product of Exan
1 is given by

$$\|\mathbf{u}\|^2 = [3 \quad 1] \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 43.$$

Thus $\|\mathbf{u}\| = \sqrt{43}$, whereas the Euclidean length of $\mathbf{u} = (3, 1)$ is $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$
$\sqrt{10}$.

The proof of Theorem translates (see Problem 19) into a proof of the **Cau
Schwarz inequality**

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

for an arbitrary inner product on any vector space V. It follows that the ang
between the nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ can be defined in this way:

$$\cos \theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|}.$$

Consequently we say that the vectors **u** and **v** are **orthogonal** provided that ⟨**u**, 0. The **triangle inequality**

$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

for an arbitrary inner product space follows from the Cauchy-Schwarz inequ And it follows that any finite set of mutually orthogonal vectors in an inner pr space is a linearly independent set.

The techniques of Section 4.9 are of special interest in the more genera ting of inner product spaces. The Gram-Schmidt orthogonalization algorithm c used to convert a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for a finite-dimensional inner product V into an orthogonal basis $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$. The analogues for this purpose formulas in Equations (12) and (13) in Section 4.9 are

$$\mathbf{u}_1 = \mathbf{v}_1$$

and

$$\mathbf{u}_{k+1} = \mathbf{v}_{k+1} - \frac{\langle \mathbf{u}_1, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 - \frac{\langle \mathbf{u}_2, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}\mathbf{u}_2 - \cdots - \frac{\langle \mathbf{u}_k, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_k$$

for $k = 1, 2, \ldots, n-1$ in turn. Thus $\mathbf{u}_{k+1}$ is obtained by subtracting from $\mathbf{v}_{k+1}$ of its components parallel (with respect to the given inner product) to the previ constructed orthogonal vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$.

Now let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be an orthogonal basis for the (finite-dimensi subspace W of the inner product space V. Given any vector **b** in V, we defi analogy with the formula in Equation (6) of Section 4.9) the **orthogonal proje** **p** of **b** into the subspace W to be

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

It is readily verified (see Problem 20) that $\mathbf{q} = \mathbf{b} - \mathbf{p}$ is orthogonal to vector in W, and it follows that **p** and **q** are the unique vectors parallel t orthogonal to W (respectively) such that $\mathbf{b} = \mathbf{p} + \mathbf{q}$. Finally, the triangle ineq can be used (as in Theorem 1 of Section 4.8) to show that the orthogonal proje **p** of b into W is the point of the subspace W, closest to **b**. If **b** itself is a vec W then $\mathbf{p} = \mathbf{b}$, and the right-hand side in (10) expresses **b** as a linear combir of the orthogonal basis vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$.

INNER PRODUCTS AND FUNCTION SPACES*

Some of the most interesting and important applications involving orthogonal and projections are to vector spaces of functions. We've introduced the v space $\mathcal{F}$ of all real-valued functions on the real line **R** as well as various in dimensional subspaces of $\mathcal{F}$, including the space $\mathcal{P}$ of all polynomials and the of all continuous functions on **R**.

* The remainder of this section is for those readers who have studied elementary calculus.

Here we want to discuss the infinite-dimensional vector space $C\,[a, b]$ consi ing of all continuous functions defined on the closed interval $[a, b]$, with the us vector space operations

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (cf)(x) = cf(x).$$

When it is unnecessary to refer explicitly to the interval $[a, b]$, we will simply wr $C = C[a, b]$.

To provide the vector space $C[a, b]$ with an inner product, we define

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

for any two functions f and g in $C[a, b]$. The fact that $\langle f, g \rangle$ satisfies propert (i)–(iii) of an inner product follows from familiar elementary facts about integr For instance,

$$\begin{aligned}\langle f, g + h \rangle &= \int_a^b f(x)\{g(x) + h(x)\}dx \\ &= \int_a^b f(x)g(x)dx + \int_a^b f(x)h(x)\,dx \\ &= \langle f, g \rangle + \langle f, h \rangle .\end{aligned}$$

It is also true (though perhaps not so obvious) that if f is a continuous function su that

$$\langle f, f \rangle = \int_a^b \{f(x)\}^2 dx = 0,$$

then it follows that $f(x) \equiv 0$ on $[a, b]$; that is, f is the zero function in $C\,[a,$ Therefore, $\langle f, g \rangle$ as defined in (11) also satisfies Property (iv) and hence is an in product on $C\,[a, b]$.

The **norm** $\|f\|$ of the function f in C is defined to be

$$\|f\| = \sqrt{\langle f, f \rangle} = \left(\int_a^b \{f(x)\}^2 dx\right)^{1/2}. \qquad (\,$$

Then the Cauchy-Schwarz and triangle inequalities for $C[a, b]$ take the forms

$$\left|\int_a^b f(x)g(x)dx\right| \leq \left(\int_a^b \{f(x)\}^2\,dx\right)^{1/2} \left(\int_a^b \{g(x)\}^2\,dx\right)^{1/2} \qquad (\,$$

and

$$\left(\int_a^b \{f(x) + g(x)\}^2\,dx\right)^{1/2} \leq \left(\int_a^b \{f(x)\}^2 dx\right)^{1/2} + \left(\int_a^b \{g(x)\}^2\,dx\right)^{1/2}. \qquad (\,$$

respectively. It may surprise you to observe that these inequalities involving i grals follow immediately from the general inequalities in (5) and (7), which do explicitly involve definite integrals.

Example 2 Let $\mathcal{P}_n$ denote the subspace of $C[-1, 1]$ consisting of all polynomials of d most n. $\mathcal{P}_n$ is an $(n+1)$-dimensional vector space, with basis elements

$$q_0(x) = 1, q_1(x) = x, q_2(x) = x^2, \ldots , q_n(x) = x^n.$$

We want to apply the Gram-Schmidt algorithm to convert $\{q_0, q_1, \ldots , q_n\}$ orthogonal basis $\{p_0, p_1, \ldots , p_n\}$ for $\mathcal{P}_n$. According to (8) and (9), we begi

$$p_0(x) = q_0(x) = 1,$$

and first calculate

$$\langle p_0, p_0\rangle = \int_{-1}^{1} 1 \cdot 1 \, dx = 2,$$

$$\langle p_0, q_1\rangle = \int_{-1}^{1} 1 \cdot x \, dx = 0.$$

Then

$$p_1 = q_1 - \frac{\langle p_0, q_1\rangle}{\langle p_0, p_0\rangle} p_0 = q_1 - \frac{0}{2} p_0 = q_1,$$

so

$$p_1(x) = x.$$

Next,

$$\langle p_1, p_1\rangle = \langle p_0, q_2\rangle = \int_{-1}^{1} x^2 \, dx = \frac{2}{3}$$

and

$$\langle p_1, q_2\rangle = \int_{-1}^{1} x^3 \, dx = 0,$$

so

$$\begin{aligned} p_2 &= q_2 - \frac{\langle p_0, q_2\rangle}{\langle p_0, p_0\rangle} p_0 - \frac{\langle p_1, q_2\rangle}{\langle p_1, p_1\rangle} p_1 \\ &= q_2 - \frac{\frac{2}{3}}{2} p_0 - \frac{0}{\frac{2}{3}} p_1 = q_2 - \frac{1}{3} p_0, \end{aligned}$$

and hence

$$p_2(x) = x^2 - \frac{1}{3} = \frac{1}{3}(3x^2 - 1).$$

To go one step further, we compute the integrals

$$\langle p_0, q_3 \rangle = \int_{-1}^{1} x^3 dx = 0,$$

$$\langle p_1, q_3 \rangle = \int_{-1}^{1} x^4 dx = \frac{2}{5},$$

$$\langle p_2, q_3 \rangle = \int_{-1}^{1} x^3 \left(x^2 - \frac{1}{3}\right) dx = 0,$$

and

$$\langle p_2, p_2 \rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}.$$

Then

$$p_3 = q_3 - \frac{\langle p_0, q_3 \rangle}{\langle p_0, p_0 \rangle} p_0 - \frac{\langle p_1, q_3 \rangle}{\langle p_1, p_1 \rangle} p_1 - \frac{\langle p_2, q_3 \rangle}{\langle p_2, p_2 \rangle} p_2$$

$$= q_3 - \frac{0}{2} p_0 - \frac{\frac{2}{5}}{\frac{2}{3}} p_1 - \frac{0}{\frac{8}{45}} p_2 = q_3 - \frac{3}{5} p_1,$$

so

$$p_3(x) = x^3 - \frac{3}{5} x = \frac{1}{5}(5x^3 - 3x).$$

The orthogonal polynomials in (15)–(18) are constant multiples of the fa
Legendre polynomials. The first six Legendre polynomials are

$$P_0(x) = 1,$$

$$P_1(x) = x,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1),$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x),$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3),$$

$$P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x).$$

For reasons that need not concern us here, the constant multipliers are cho
that

$$P_0(1) = P_1(1) = P_2(1) = \cdots = 1.$$

Given a function f in $\mathcal{C}[-1, 1]$, the orthogonal projection p of f int
given (see the formula in (10)) in terms of Legendre polynomials by

$$p(x) = \frac{\langle P_0, f \rangle}{\langle P_0, p_0 \rangle} P_0(x) + \frac{\langle P_1, f \rangle}{\langle P_1, P_1 \rangle} P_1(x) + \cdots + \frac{\langle P_n, f \rangle}{\langle P_n, P_n \rangle} P_n(x).$$

Then $p(x)$ is the nth degree **least squares polynomial** approximation to f [on]
$[-1, 1]$. It is the nth degree polynomial that minimizes the **mean square err[or]**

$$\|f - p\|^2 = \int_{-1}^{1} \{f(x) - p(x)\}^2 dx.$$

Example 3 Let $\mathcal{T}_N$ denote the subspace of $\mathcal{C}[-\pi, \pi]$ that consists of all "trigonometric p[oly]mials" of the form

$$a_0 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx).$$

Then $\mathcal{T}_N$ is spanned by the $2N + 1$ functions

$$1, \cos x, \sin x, \cos 2x, \sin 2x, \ldots, \cos Nx, \sin Nx.$$

By standard techniques of integral calculus we find that

$$\langle 1, \cos nx \rangle = \int_{-\pi}^{\pi} \cos nx \, dx = 0,$$
$$\langle 1, \sin nx \rangle = \int_{-\pi}^{\pi} \sin nx \, dx = 0,$$
$$\langle \cos mx, \sin nx \rangle = \int_{-\pi}^{\pi} \cos mx \sin nx \, dx = 0$$

for all positive integers m and n, and that

$$\langle \sin mx, \sin nx \rangle = \int_{-\pi}^{\pi} \sin mx \sin nx \, dx = 0,$$
$$\langle \cos mx, \cos nx \rangle = \int_{-\pi}^{\pi} \cos mx \cos nx \, dx = 0$$

if $m \neq n$. Thus the $2N + 1$ nonzero functions in (21) are mutually orth[ogonal] and hence are linearly independent. It follows that $\mathcal{T}_N$ is a $(2N + 1)$-dimen[sional] subspace of $\mathcal{T}[-\pi, \pi]$ with the functions in (21) constituting an orthogonal [basis].

To find the norms of these basis functions, we calculate the integrals

$$\langle 1, 1 \rangle = \int_{-\pi}^{\pi} 1 \, dx = 2\pi,$$
$$\langle \cos nx, \cos nx \rangle = \int_{-\pi}^{\pi} \cos^2 nx \, dx$$
$$= \int_{-\pi}^{\pi} \frac{1}{2}(1 + \cos 2nx) \, dx$$
$$= \frac{1}{2}\left[x + \frac{1}{2n} \sin 2nx\right]_{-\pi}^{\pi} = \pi$$

and, similarly,

$$\langle \sin nx, \sin nx \rangle = \int_{-\pi}^{\pi} \sin^2 nx \, dx = \pi.$$

Thus

$$\|1\| = \sqrt{2\pi} \quad \text{and} \quad \|\cos nx\| = \|\sin nx\| = \sqrt{\pi} \qquad ($$

for all n.

Now suppose that $f(x)$ is an arbitrary continuous function in $C[-\pi, \pi]$. . cording to the formula in (10), the orthogonal projection $p(x)$ of $f(x)$ into subspace T_N is the sum of the $2N + 1$ orthogonal projections of $f(x)$ onto orthogonal basis elements in (21). These orthogonal projections are given by

$$\frac{\langle f(x), 1\rangle}{\langle 1, 1\rangle} = a_0$$

where

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

$$\frac{\langle f(x), \cos nx\rangle}{\langle \cos nx, \cos nx\rangle}\cos nx = a_n \cos nx \qquad ($$

where

$$a_n = \frac{\langle f(x), \cos nx\rangle}{\langle \cos nx, \cos nx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx; \qquad ($$

and

$$\frac{\langle f(x), \sin nx\rangle}{\langle \sin nx, \sin nx\rangle}\sin nx = b_n \sin nx$$

where

$$b_n = \frac{\langle f(x), \sin nx\rangle}{\langle \sin nx, \sin nx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx.$$

Consequently the orthogonal projection $p(x)$ of the function $f(x)$ into T given by

$$p(x) = a_0 + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx),$$

where the coefficients are given by the formulas in Equations (23)–(25). These stants $a_0, a_1, b_1, a_2, b_2, \ldots$ are called the **Fourier coefficients** of the function on $[-\pi, \pi]$. The fact that the orthogonal projection p is the element of T_N clo to f means that the Fourier coefficients of f minimize the mean square error

$$\|f - p\|^2 = \int_{-\pi}^{\pi}\left\{f(x) - a_0 - \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)\right\}^2 dx.$$

This is the sense in which the trigonometric polynomial $p(x)$ is the "best squares approximation" (in T_N) to the given continuous function $f(x)$.

Finally, we remark that $p(x)$ in (26) is a (finite) partial sum of the series

$$a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx).$$

With the coefficients given in (23)–(25), this infinite series is known as the I **series** of f on $[-\pi, \pi]$.

Example 4 Given $f(x) = x$ on $[-\pi, \pi]$, find the orthogonal projection p of f into $\mathcal{T}_4$.

Solution The formula in (23) yields

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\, dx = \frac{1}{2\pi}\left[\frac{1}{2}x^2\right]_{-\pi}^{\pi} = 0.$$

To find a_n and b_n for $n > 0$ we need the integral formulas

$$\int u \cos u\, du = \cos u + u \sin u + C$$

and

$$\int u \sin u\, du = \sin u - u \cos u + C.$$

Then the formula in (24) yields

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x \cos nx\, dx = \frac{1}{n^2\pi}\int_{-n\pi}^{n\pi} u \cos u\, du \quad (u = nx) \\ &= \frac{1}{n^2\pi}\left[\cos u + u \sin u\right]_{-n\pi}^{n\pi} = 0 \end{aligned}$$

for all positive integers n. And the formula in (25) yields

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x \sin nx\, dx = \frac{1}{n^2\pi}\int_{-n\pi}^{n\pi} u \sin u\, du \quad (u = nx) \\ &= \frac{1}{n^2\pi}\left[\sin u - u \cos u\right]_{-n\pi}^{n\pi} = -\frac{2}{n}\cos n\pi = \frac{2}{n}(-1)^{n+1} \end{aligned}$$

for all positive integers n. Substituting these values for $n \leq 4$ in (26), we g desired orthogonal projection

$$p(x) = 2\left(\sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \frac{1}{4}\sin 4x\right).$$

This is the "trigonometric polynomial of degree 4" that (in the least squares best approximates $f(x) = x$ on the interval $[-\pi, \pi]$.

4.10 Problems

For each 2 × 2 matrix A given in Problems 1–6, show that $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$ *is an inner product on* $\mathbf{R}^2$. *Given* $\mathbf{u} = (x, y)$, *write* $\mathbf{u}^T A\mathbf{u}$ *as a sum of squares as in Example 1.*

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

2. $\begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$

3. $\begin{bmatrix} 2 & -2 \\ -2 & 3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 \\ 3 & 10 \end{bmatrix}$

5. $\begin{bmatrix} 4 & 6 \\ 6 & 11 \end{bmatrix}$

6. $\begin{bmatrix} 9 & -3 \\ -3 & 2 \end{bmatrix}$

In each of Problems 7–10, apply the Gram-Schmidt algorithm to the vectors $e_1 = (1, 0)$ *and* $e_2 = (0, 1)$ *to obtain vectors* u_1 *and* u_2 *that are orthogonal with respect to the inner product* $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$.

7. A is the matrix of Problem 3.
8. A is the matrix of Problem 4.
9. A is the matrix of Problem 5.
10. A is the matrix of Problem 6.

For each 3×3 matrix A given in Problems 11 and 12, show that $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$ *is an inner product on* $\mathbf{R}^3$. *Given* $u = (x, y, z)$, *write* $u^T Au$ *as a sum of squares.*

11. $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 2 & 2 & 0 \\ 2 & 5 & 2 \\ 0 & 2 & 4 \end{bmatrix}$

In Problems 13 and 14, apply the Gram-Schmidt algorithm to the vectors $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, *and* $e_3 = (0, 0, 1)$ *to obtain vectors* u_1, $\mathbf{u}_2$, *and* u_3 *that are mutually orthogonal with respect to the inner product* $\langle \mathbf{u}, \mathbf{v}\rangle = \mathbf{u}^T A\mathbf{v}$.

13. A is the matrix of Problem 11.
14. A is the matrix of Problem 12.
15. Show that

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

defines an inner product on the space $\mathcal{P}_2$ of polynomials of degree at most 2.

16. Apply the Gram-Schmidt algorithm to the basis $\{1, x, x^2\}$ for $\mathcal{P}_2$ to construct a basis $\{p_0, p_1, p_2\}$ that is orthogonal with respect to the inner product of Problem 15.

17. Show that the symmetric 2 × 2 matrix

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

is positive definite if both $a > 0$ and $ac - b^2 > 0$. gestion: Write $ax^2 + 2bxy + cy^2$ as a sum of squares i form $a(x + \alpha y)^2 + \beta y^2$.

18. If the nonzero vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in an inner pro space V are mutually orthogonal, prove that they are early independent.

19. Translate the proof of Theorem 1 in Section 5.1 ir proof of the Cauchy-Schwarz inequality for an arbi inner product space.

20. Let $\mathbf{p}$ be the orthogonal projection (defined in Equ (10)) of $\mathbf{b}$ into the subspace W spanned by the orthog vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Show that $\mathbf{q} = \mathbf{b} - \mathbf{p}$ is orthog to W.

21. Let $\mathcal{W}$ be the subspace of $C[0, 1]$ consisting of all f tions of the form $f(x) = a + be^x$. Apply the G Schmidt algorithm to the basis $\{1, e^x\}$ to obtain th thogonal basis $\{p_1, p_2\}$, where

$$p_1(x) = 1 \quad \text{and} \quad p_2(x) = e^x - e + 1.$$

22. Show that the orthogonal projection of the fun $f(x) = x$ into the subspace $\mathcal{W}$ of Problem 21 is

$$p(x) = -\frac{1}{2} + \frac{e^x}{e - 1} \approx (0.5820)e^x - 0.5000.$$

This is the best (least squares) approximation to $f(x)$ by a function on $[0, 1]$ of the form $a + be^x$. *Sugges* The antiderivative of xe^x is $(x - 1)e^x + C$.

23. Continue the computations in Example 2 to derive the stant multiple

$$p_4(x) = \frac{1}{35}(35x^4 - 30x^2 + 3)$$

of the Legendre polynomial of degree 4.

24. The orthogonal projection of $f(x) = x^3$ into $\mathcal{P}_3$ i function f itself. Use this fact to express x^3 as a l combination of the Legendre polynomials $P_0(x)$, P $P_2(x)$, and $P_3(x)$ listed in (19).

25. This problem deals with orthogonal polynomials $[0, 1]$ rather than $C[-1, 1]$. Apply the Gram-Schmi gorithm to transform the basis $\{1, x, x^2\}$ for $\mathcal{P}_2$ int orthogonal basis $\{P_0, P_1, P_2\}$ where

$$p_0(x) = 1, \quad p_1(x) = \frac{1}{2}(2x - 1),$$

and

$$p_3(x) = \frac{1}{6}(6x^2 - 6x + 1).$$