CHAPTER

6

Woman teaching geometry, from a fourteenth-century edition of Euclid's geometry book.

Inner Product Spaces

In making the definition of a vector space, we generalized the linear structure (addition and scalar multiplication) of $\mathbf{R}^2$ and $\mathbf{R}^3$. We ignored other important features, such as the notions of length and angle. These ideas are embedded in the concept we now investigate, inner products.

Our standing assumptions are as follows:

6.1 Notation $\mathbf{F}$, V

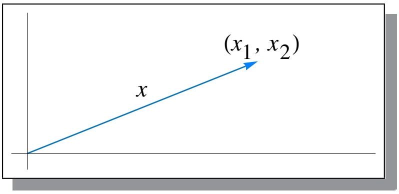
- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V denotes a vector space over $\mathbf{F}$.

LEARNING OBJECTIVES FOR THIS CHAPTER

- Cauchy–Schwarz Inequality
- Gram–Schmidt Procedure
- linear functionals on inner product spaces
- calculating minimum distance to a subspace

6.A *Inner Products and Norms*

Inner Products



The length of this vector x is $\sqrt{x_1{}^2 + x_2{}^2}$.

To motivate the concept of inner product, think of vectors in $\mathbf{R}^2$ and $\mathbf{R}^3$ as arrows with initial point at the origin. The length of a vector x in $\mathbf{R}^2$ or $\mathbf{R}^3$ is called the ***norm*** of x, denoted $\|x\|$. Thus for $x = (x_1, x_2) \in \mathbf{R}^2$, we have $\|x\| = \sqrt{x_1{}^2 + x_2{}^2}$.

Similarly, if $x = (x_1, x_2, x_3) \in \mathbf{R}^3$, then $\|x\| = \sqrt{x_1{}^2 + x_2{}^2 + x_3{}^2}$.

Even though we cannot draw pictures in higher dimensions, the generalization to $\mathbf{R}^n$ is obvious: we define the norm of $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ by

$$\|x\| = \sqrt{x_1{}^2 + \cdots + x_n{}^2}.$$

The norm is not linear on $\mathbf{R}^n$. To inject linearity into the discussion, we introduce the dot product.

6.2 Definition ***dot product***

For $x, y \in \mathbf{R}^n$, the ***dot product*** of x and y, denoted $x \cdot y$, is defined by

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n,$$

where $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$.

If we think of vectors as points instead of arrows, then $\|x\|$ should be interpreted as the distance from the origin to the point x.

Note that the dot product of two vectors in $\mathbf{R}^n$ is a number, not a vector. Obviously $x \cdot x = \|x\|^2$ for all $x \in \mathbf{R}^n$. The dot product on $\mathbf{R}^n$ has the following properties:

- $x \cdot x \geq 0$ for all $x \in \mathbf{R}^n$;
- $x \cdot x = 0$ if and only if $x = 0$;
- for $y \in \mathbf{R}^n$ fixed, the map from $\mathbf{R}^n$ to $\mathbf{R}$ that sends $x \in \mathbf{R}^n$ to $x \cdot y$ is linear;
- $x \cdot y = y \cdot x$ for all $x, y \in \mathbf{R}^n$.

An inner product is a generalization of the dot product. At this point you may be tempted to guess that an inner product is defined by abstracting the properties of the dot product discussed in the last paragraph. For real vector spaces, that guess is correct. However, so that we can make a definition that will be useful for both real and complex vector spaces, we need to examine the complex case before making the definition.

Recall that if $\lambda = a + bi$, where $a, b \in \mathbf{R}$, then

- the absolute value of λ, denoted $|\lambda|$, is defined by $|\lambda| = \sqrt{a^2 + b^2}$;
- the complex conjugate of λ, denoted $\bar{\lambda}$, is defined by $\bar{\lambda} = a - bi$;
- $|\lambda|^2 = \lambda\bar{\lambda}$.

See Chapter 4 for the definitions and the basic properties of the absolute value and complex conjugate.

For $z = (z_1, \ldots, z_n) \in \mathbf{C}^n$, we define the norm of z by

$$\|z\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

The absolute values are needed because we want $\|z\|$ to be a nonnegative number. Note that

$$\|z\|^2 = z_1\overline{z_1} + \cdots + z_n\overline{z_n}.$$

We want to think of $\|z\|^2$ as the inner product of z with itself, as we did in $\mathbf{R}^n$. The equation above thus suggests that the inner product of $w = (w_1, \ldots, w_n) \in \mathbf{C}^n$ with z should equal

$$w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

If the roles of the w and z were interchanged, the expression above would be replaced with its complex conjugate. In other words, we should expect that the inner product of w with z equals the complex conjugate of the inner product of z with w. With that motivation, we are now ready to define an inner product on V, which may be a real or a complex vector space.

Two comments about the notation used in the next definition:

- If λ is a complex number, then the notation $\lambda \geq 0$ means that λ is real and nonnegative.
- We use the common notation $\langle u, v\rangle$, with angle brackets denoting an inner product. Some people use parentheses instead, but then (u, v) becomes ambiguous because it could denote either an ordered pair or an inner product.

6.3 Definition ***inner product***

An ***inner product*** on V is a function that takes each ordered pair (u, v) of elements of V to a number $\langle u, v\rangle \in \mathbf{F}$ and has the following properties:

positivity
$\langle v, v\rangle \geq 0$ for all $v \in V$;

definiteness
$\langle v, v\rangle = 0$ if and only if $v = 0$;

additivity in first slot
$\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$ for all $u, v, w \in V$;

homogeneity in first slot
$\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$;

conjugate symmetry
$\langle u, v\rangle = \overline{\langle v, u\rangle}$ for all $u, v \in V$.

Although most mathematicians define an inner product as above, many physicists use a definition that requires homogeneity in the second slot instead of the first slot.

Every real number equals its complex conjugate. Thus if we are dealing with a real vector space, then in the last condition above we can dispense with the complex conjugate and simply state that $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$.

6.4 Example ***inner products***

(a) The ***Euclidean inner product*** on $\mathbf{F}^n$ is defined by

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = w_1\overline{z_1} + \cdots + w_n\overline{z_n}.$$

(b) If $c_1, \ldots, c_n$ are positive numbers, then an inner product can be defined on $\mathbf{F}^n$ by

$$\langle (w_1, \ldots, w_n), (z_1, \ldots, z_n)\rangle = c_1 w_1\overline{z_1} + \cdots + c_n w_n\overline{z_n}.$$

(c) An inner product can be defined on the vector space of continuous real-valued functions on the interval $[-1, 1]$ by

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx.$$

(d) An inner product can be defined on $\mathcal{P}(\mathbf{R})$ by

$$\langle p, q\rangle = \int_{0}^{\infty} p(x)q(x)e^{-x}\,dx.$$

6.5 Definition ***inner product space***

An ***inner product space*** is a vector space V along with an inner product on V.

The most important example of an inner product space is $\mathbf{F}^n$ with the Euclidean inner product given by part (a) of the last example. When $\mathbf{F}^n$ is referred to as an inner product space, you should assume that the inner product is the Euclidean inner product unless explicitly told otherwise.

So that we do not have to keep repeating the hypothesis that V is an inner product space, for the rest of this chapter we make the following assumption:

6.6 Notation V

For the rest of this chapter, V denotes an inner product space over $\mathbf{F}$.

Note the slight abuse of language here. An inner product space is a vector space along with an inner product on that vector space. When we say that a vector space V is an inner product space, we are also thinking that an inner product on V is lurking nearby or is obvious from the context (or is the Euclidean inner product if the vector space is $\mathbf{F}^n$).

6.7 Basic properties of an inner product

(a) For each fixed $u \in V$, the function that takes v to $\langle v, u\rangle$ is a linear map from V to $\mathbf{F}$.

(b) $\langle 0, u\rangle = 0$ for every $u \in V$.

(c) $\langle u, 0\rangle = 0$ for every $u \in V$.

(d) $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ for all $u, v, w \in V$.

(e) $\langle u, \lambda v\rangle = \bar{\lambda}\langle u, v\rangle$ for all $\lambda \in \mathbf{F}$ and $u, v \in V$.

Proof

(a) Part (a) follows from the conditions of additivity in the first slot and homogeneity in the first slot in the definition of an inner product.

(b) Part (b) follows from part (a) and the result that every linear map takes 0 to 0.

(c) Part (c) follows from part (b) and the conjugate symmetry property in the definition of an inner product.

(d) Suppose $u, v, w \in V$. Then

$$\begin{aligned}\langle u, v + w\rangle &= \overline{\langle v + w, u\rangle}\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle}\\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle}\\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

(e) Suppose $\lambda \in \mathbf{F}$ and $u, v \in V$. Then

$$\begin{aligned}\langle u, \lambda v\rangle &= \overline{\langle \lambda v, u\rangle}\\ &= \overline{\lambda\langle v, u\rangle}\\ &= \bar{\lambda}\overline{\langle v, u\rangle}\\ &= \bar{\lambda}\langle u, v\rangle,\end{aligned}$$

as desired. ■

Norms

Our motivation for defining inner products came initially from the norms of vectors on $\mathbf{R}^2$ and $\mathbf{R}^3$. Now we see that each inner product determines a norm.

6.8 **Definition** ***norm,*** $\|v\|$

For $v \in V$, the ***norm*** of v, denoted $\|v\|$, is defined by

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

6.9 **Example** ***norms***

(a) If $(z_1, \dots, z_n) \in \mathbf{F}^n$ (with the Euclidean inner product), then

$$\|(z_1, \dots, z_n)\| = \sqrt{|z_1|^2 + \cdots + |z_n|^2}.$$

(b) In the vector space of continuous real-valued functions on $[-1, 1]$ [with inner product given as in part (c) of 6.4], we have

$$\|f\| = \sqrt{\int_{-1}^{1} \bigl(f(x)\bigr)^2\, dx}.$$

6.10 Basic properties of the norm

Suppose $v \in V$.

(a) $\|v\| = 0$ if and only if $v = 0$.

(b) $\|\lambda v\| = |\lambda| \, \|v\|$ for all $\lambda \in \mathbf{F}$.

Proof

(a) The desired result holds because $\langle v, v \rangle = 0$ if and only if $v = 0$.

(b) Suppose $\lambda \in \mathbf{F}$. Then

$$\begin{aligned} \|\lambda v\|^2 &= \langle \lambda v, \lambda v \rangle \\ &= \lambda \langle v, \lambda v \rangle \\ &= \lambda \bar{\lambda} \langle v, v \rangle \\ &= |\lambda|^2 \|v\|^2. \end{aligned}$$

Taking square roots now gives the desired equality. ■

The proof above of part (b) illustrates a general principle: working with norms squared is usually easier than working directly with norms.

Now we come to a crucial definition.

6.11 Definition *orthogonal*

Two vectors $u, v \in V$ are called ***orthogonal*** if $\langle u, v \rangle = 0$.

In the definition above, the order of the vectors does not matter, because $\langle u, v \rangle = 0$ if and only if $\langle v, u \rangle = 0$. Instead of saying that u and v are orthogonal, sometimes we say that u is orthogonal to v.

Exercise 13 asks you to prove that if u, v are nonzero vectors in $\mathbf{R}^2$, then

$$\langle u, v \rangle = \|u\| \, \|v\| \cos \theta,$$

where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin). Thus two vectors in $\mathbf{R}^2$ are orthogonal (with respect to the usual Euclidean inner product) if and only if the cosine of the angle between them is 0, which happens if and only if the vectors are perpendicular in the usual sense of plane geometry. Thus you can think of the word *orthogonal* as a fancy word meaning *perpendicular*.

We begin our study of orthogonality with an easy result.

6.12 Orthogonality and 0

(a) 0 is orthogonal to every vector in V.

(b) 0 is the only vector in V that is orthogonal to itself.

Proof

(a) Part (b) of 6.7 states that $\langle 0, u\rangle = 0$ for every $u \in V$.

(b) If $v \in V$ and $\langle v, v\rangle = 0$, then $v = 0$ (by definition of inner product). ■

*The word **orthogonal** comes from the Greek word **orthogonios**, which means right-angled.*

For the special case $V = \mathbf{R}^2$, the next theorem is over 2,500 years old. Of course, the proof below is not the original proof.

6.13 Pythagorean Theorem

Suppose u and v are orthogonal vectors in V. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

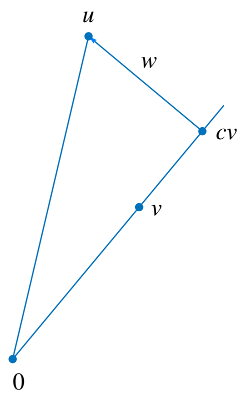
Proof We have

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

as desired.

The proof given above of the Pythagorean Theorem shows that the conclusion holds if and only if $\langle u, v\rangle + \langle v, u\rangle$, which equals $2\operatorname{Re}\langle u, v\rangle$, is 0. Thus the converse of the Pythagorean Theorem holds in real inner product spaces.

Suppose $u, v \in V$, with $v \neq 0$. We would like to write u as a scalar multiple of v plus a vector w orthogonal to v, as suggested in the next picture.



An orthogonal decomposition.

To discover how to write u as a scalar multiple of v plus a vector orthogonal to v, let $c \in \mathbf{F}$ denote a scalar. Then

$$u = cv + (u - cv).$$

Thus we need to choose c so that v is orthogonal to $(u - cv)$. In other words, we want

$$0 = \langle u - cv, v\rangle = \langle u, v\rangle - c\|v\|^2.$$

The equation above shows that we should choose c to be $\langle u, v\rangle / \|v\|^2$. Making this choice of c, we can write

$$u = \frac{\langle u, v\rangle}{\|v\|^2} v + \left(u - \frac{\langle u, v\rangle}{\|v\|^2} v\right).$$

As you should verify, the equation above writes u as a scalar multiple of v plus a vector orthogonal to v. In other words, we have proved the following result.

6.14 An orthogonal decomposition

Suppose $u, v \in V$, with $v \neq 0$. Set $c = \dfrac{\langle u, v\rangle}{\|v\|^2}$ and $w = u - \dfrac{\langle u, v\rangle}{\|v\|^2} v$. Then

$$\langle w, v\rangle = 0 \quad \text{and} \quad u = cv + w.$$

The orthogonal decomposition 6.14 will be used in the proof of the Cauchy–Schwarz Inequality, which is our next result and is one of the most important inequalities in mathematics.

French mathematician Augustin-Louis Cauchy (1789–1857) proved 6.17(a) in 1821. German mathematician Hermann Schwarz (1843–1921) proved 6.17(b) in 1886.

6.15 Cauchy–Schwarz Inequality

Suppose $u, v \in V$. Then

$$|\langle u, v\rangle| \leq \|u\|\,\|v\|.$$

This inequality is an equality if and only if one of u, v is a scalar multiple of the other.

Proof If $v = 0$, then both sides of the desired inequality equal 0. Thus we can assume that $v \neq 0$. Consider the orthogonal decomposition

$$u = \frac{\langle u, v\rangle}{\|v\|^2}v + w$$

given by 6.14, where w is orthogonal to v. By the Pythagorean Theorem,

$$\begin{aligned} \|u\|^2 &= \left\|\frac{\langle u, v\rangle}{\|v\|^2}v\right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v\rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, v\rangle|^2}{\|v\|^2}. \qquad \textbf{6.16} \end{aligned}$$

Multiplying both sides of this inequality by $\|v\|^2$ and then taking square roots gives the desired inequality.

Looking at the proof in the paragraph above, note that the Cauchy–Schwarz Inequality is an equality if and only if 6.16 is an equality. Obviously this happens if and only if $w = 0$. But $w = 0$ if and only if u is a multiple of v (see 6.14). Thus the Cauchy–Schwarz Inequality is an equality if and only if u is a scalar multiple of v or v is a scalar multiple of u (or both; the phrasing has been chosen to cover cases in which either u or v equals 0). ■

6.17 Example ***examples of the Cauchy–Schwarz Inequality***

(a) If $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbf{R}$, then

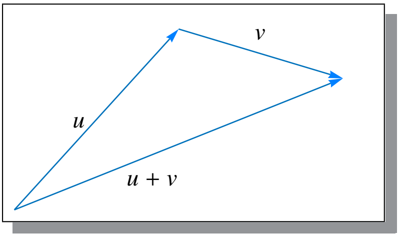
$$|x_1 y_1 + \cdots + x_n y_n|^2 \leq ({x_1}^2 + \cdots + {x_n}^2)({y_1}^2 + \cdots + {y_n}^2).$$

(b) If f, g are continuous real-valued functions on $[-1, 1]$, then

$$\left|\int_{-1}^{1} f(x)g(x)\,dx\right|^2 \leq \left(\int_{-1}^{1} \bigl(f(x)\bigr)^2\,dx\right)\left(\int_{-1}^{1} \bigl(g(x)\bigr)^2\,dx\right).$$

The next result, called the Triangle Inequality, has the geometric interpretation that the length of each side of a triangle is less than the sum of the lengths of the other two sides.

Note that the Triangle Inequality implies that the shortest path between two points is a line segment.



6.18 Triangle Inequality

Suppose $u, v \in V$. Then

$$\|u + v\| \leq \|u\| + \|v\|.$$

This inequality is an equality if and only if one of u, v is a nonnegative multiple of the other.

Proof We have

$$\begin{aligned}
\|u + v\|^2 &= \langle u + v, u + v\rangle \\
&= \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \langle v, u\rangle \\
&= \langle u, u\rangle + \langle v, v\rangle + \langle u, v\rangle + \overline{\langle u, v\rangle} \\
&= \|u\|^2 + \|v\|^2 + 2\operatorname{Re}\langle u, v\rangle \\
&\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v\rangle| \qquad \textbf{6.19} \\
&\leq \|u\|^2 + \|v\|^2 + 2\|u\|\,\|v\| \qquad \textbf{6.20} \\
&= (\|u\| + \|v\|)^2,
\end{aligned}$$

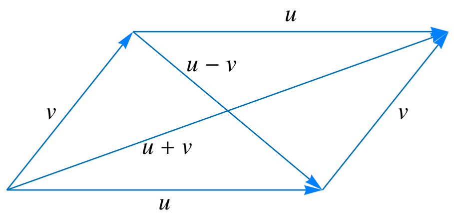
where 6.20 follows from the Cauchy–Schwarz Inequality (6.15). Taking square roots of both sides of the inequality above gives the desired inequality.

The proof above shows that the Triangle Inequality is an equality if and only if we have equality in 6.19 and 6.20. Thus we have equality in the Triangle Inequality if and only if

6.21 $$\langle u, v\rangle = \|u\|\|v\|.$$

If one of u, v is a nonnegative multiple of the other, then 6.21 holds, as you should verify. Conversely, suppose 6.21 holds. Then the condition for equality in the Cauchy–Schwarz Inequality (6.15) implies that one of u, v is a scalar multiple of the other. Clearly 6.21 forces the scalar in question to be nonnegative, as desired. ■

The next result is called the parallelogram equality because of its geometric interpretation: in every parallelogram, the sum of the squares of the lengths of the diagonals equals the sum of the squares of the lengths of the four sides.



The parallelogram equality.

6.22 Parallelogram Equality

Suppose $u, v \in V$. Then

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Proof We have

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

as desired. ■

Law professor Richard Friedman presenting a case before the U.S. Supreme Court in 2010:

Mr. Friedman: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging—
Chief Justice Roberts: I'm sorry. Entirely what?
Mr. Friedman: Orthogonal. Right angle. Unrelated. Irrelevant.
Chief Justice Roberts: Oh.
Justice Scalia: What was that adjective? I liked that.
Mr. Friedman: Orthogonal.
Chief Justice Roberts: Orthogonal.
Mr. Friedman: Right, right.
Justice Scalia: Orthogonal, ooh. (Laughter.)
Justice Kennedy: I knew this case presented us a problem. (Laughter.)

EXERCISES 6.A

1 Show that the function that takes $\big((x_1, x_2), (y_1, y_2)\big) \in \mathbf{R}^2 \times \mathbf{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbf{R}^2$.

2 Show that the function that takes $\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) \in \mathbf{R}^3 \times \mathbf{R}^3$ to $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbf{R}^3$.

3 Suppose $\mathbf{F} = \mathbf{R}$ and $V \neq \{0\}$. Replace the positivity condition (which states that $\langle v, v\rangle \geq 0$ for all $v \in V$) in the definition of an inner product (6.3) with the condition that $\langle v, v\rangle > 0$ for some $v \in V$. Show that this change in the definition does not change the set of functions from $V \times V$ to $\mathbf{R}$ that are inner products on V.

4 Suppose V is a real inner product space.

(a) Show that $\langle u + v, u - v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.

(b) Show that if $u, v \in V$ have the same norm, then $u+v$ is orthogonal to $u - v$.

(c) Use part (b) to show that the diagonals of a rhombus are perpendicular to each other.

5 Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is invertible.

6 Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0$ if and only if

$$\|u\| \leq \|u + av\|$$

for all $a \in \mathbf{F}$.

7 Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbf{R}$ if and only if $\|u\| = \|v\|$.

8 Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

9 Suppose $u, v \in V$ and $\|u\| \leq 1$ and $\|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v\rangle|.$$

10 Find vectors $u, v \in \mathbf{R}^2$ such that u is a scalar multiple of $(1, 3)$, v is orthogonal to $(1, 3)$, and $(1, 2) = u + v$.

11 Prove that

$$16 \leq (a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)$$

for all positive numbers a, b, c, d.

12 Prove that

$$(x_1 + \cdots + x_n)^2 \leq n({x_1}^2 + \cdots + {x_n}^2)$$

for all positive integers n and all real numbers $x_1, \ldots, x_n$.

13 Suppose u, v are nonzero vectors in $\mathbf{R}^2$. Prove that

$$\langle u, v \rangle = \|u\| \|v\| \cos\theta,$$

where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin).

Hint: Draw the triangle formed by u, v, and $u - v$; then use the law of cosines.

14 The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $x, y \in \mathbf{R}^n$ is defined to be

$$\arccos \frac{\langle x, y \rangle}{\|x\| \|y\|},$$

where the motivation for this definition comes from the previous exercise. Explain why the Cauchy–Schwarz Inequality is needed to show that this definition makes sense.

15 Prove that

$$\Big(\sum_{j=1}^{n} a_j b_j\Big)^2 \leq \Big(\sum_{j=1}^{n} j {a_j}^2\Big)\Big(\sum_{j=1}^{n} \frac{{b_j}^2}{j}\Big)$$

for all real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$.

16 Suppose $u, v \in V$ are such that

$$\|u\| = 3, \quad \|u+v\| = 4, \quad \|u-v\| = 6.$$

What number does $\|v\|$ equal?

17 Prove or disprove: there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = \max\{|x|, |y|\}$$

for all $(x, y) \in \mathbf{R}^2$.

18 Suppose $p > 0$. Prove that there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p}$$

for all $(x, y) \in \mathbf{R}^2$ if and only if $p = 2$.

19 Suppose V is a real inner product space. Prove that

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

for all $u, v \in V$.

20 Suppose V is a complex inner product space. Prove that

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2 + \|u + iv\|^2 i - \|u - iv\|^2 i}{4}$$

for all $u, v \in V$.

21 A norm on a vector space U is a function $\| \ \| \colon U \to [0, \infty)$ such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha| \|u\|$ for all $\alpha \in \mathbf{F}$ and all $u \in U$, and $\|u + v\| \leq \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $\| \ \|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle \ , \ \rangle$ on U such that $\|u\| = \langle u, u \rangle^{1/2}$ for all $u \in U$).

22 Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if $a_1, \dots, a_n \in \mathbf{R}$, then the square of the average of $a_1, \dots, a_n$ is less than or equal to the average of ${a_1}^2, \dots, {a_n}^2$.

23 Suppose $V_1, \dots, V_m$ are inner product spaces. Show that the equation

$$\langle (u_1, \dots, u_m), (v_1, \dots, v_m) \rangle = \langle u_1, v_1 \rangle + \dots + \langle u_m, v_m \rangle$$

defines an inner product on $V_1 \times \dots \times V_m$.
[*In the expression above on the right, $\langle u_1, v_1 \rangle$ denotes the inner product on V_1, ..., $\langle u_m, v_m \rangle$ denotes the inner product on V_m. Each of the spaces $V_1, \dots, V_m$ may have a different inner product, even though the same notation is used here.*]

24 Suppose $S \in \mathcal{L}(V)$ is an injective operator on V. Define $\langle \cdot, \cdot \rangle_1$ by

$$\langle u, v \rangle_1 = \langle Su, Sv \rangle$$

for $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V.

25 Suppose $S \in \mathcal{L}(V)$ is not injective. Define $\langle \cdot, \cdot \rangle_1$ as in the exercise above. Explain why $\langle \cdot, \cdot \rangle_1$ is not an inner product on V.

26 Suppose f, g are differentiable functions from $\mathbf{R}$ to $\mathbf{R}^n$.

(a) Show that

$$\langle f(t), g(t) \rangle' = \langle f'(t), g(t) \rangle + \langle f(t), g'(t) \rangle.$$

(b) Suppose $c > 0$ and $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t), f(t) \rangle = 0$ for every $t \in \mathbf{R}$.

(c) Interpret the result in part (b) geometrically in terms of the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin.

[*For the exercise above, a function* $f : \mathbf{R} \to \mathbf{R}^n$ *is called differentiable if there exist differentiable functions* $f_1, \dots, f_n$ *from* $\mathbf{R}$ *to* $\mathbf{R}$ *such that* $f(t) = \big(f_1(t), \dots, f_n(t)\big)$ *for each* $t \in \mathbf{R}$. *Furthermore, for each* $t \in \mathbf{R}$, *the derivative* $f'(t) \in \mathbf{R}^n$ *is defined by* $f'(t) = \big({f_1}'(t), \dots, {f_n}'(t)\big)$.]

27 Suppose $u, v, w \in V$. Prove that

$$\|w - \tfrac{1}{2}(u + v)\|^2 = \frac{\|w - u\|^2 + \|w - v\|^2}{2} - \frac{\|u - v\|^2}{4}.$$

28 Suppose C is a subset of V with the property that $u, v \in C$ implies $\frac{1}{2}(u + v) \in C$. Let $w \in V$. Show that there is at most one point in C that is closest to w. In other words, show that there is at most one $u \in C$ such that

$$\|w - u\| \le \|w - v\| \quad \text{for all } v \in C.$$

Hint: Use the previous exercise.

29 For $u, v \in V$, define $d(u, v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

30 Fix a positive integer n. The ***Laplacian*** Δp of a twice differentiable function p on $\mathbf{R}^n$ is the function on $\mathbf{R}^n$ defined by

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

The function p is called ***harmonic*** if $\Delta p = 0$.

A ***polynomial*** on $\mathbf{R}^n$ is a linear combination of functions of the form ${x_1}^{m_1} \cdots {x_n}^{m_n}$, where $m_1, \dots, m_n$ are nonnegative integers.

Suppose q is a polynomial on $\mathbf{R}^n$. Prove that there exists a harmonic polynomial p on $\mathbf{R}^n$ such that $p(x) = q(x)$ for every $x \in \mathbf{R}^n$ with $\|x\| = 1$.

[*The only fact about harmonic functions that you need for this exercise is that if p is a harmonic function on $\mathbf{R}^n$ and $p(x) = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$, then $p = 0$.*]

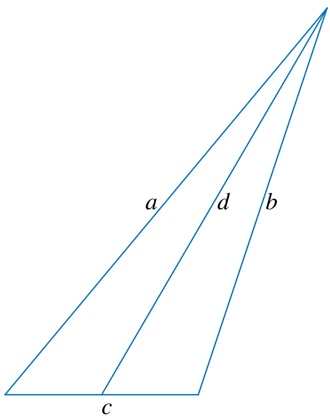
Hint: A reasonable guess is that the desired harmonic polynomial p is of the form $q + (1 - \|x\|^2)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbf{R}^n$ such that $q + (1 - \|x\|^2)r$ is harmonic by defining an operator T on a suitable vector space by

$$Tr = \Delta\big((1 - \|x\|^2)r\big)$$

and then showing that T is injective and hence surjective.

31 Use inner products to prove Apollonius's Identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$



6.B *Orthonormal Bases*

6.23 Definition ***orthonormal***

- A list of vectors is called ***orthonormal*** if each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.
- In other words, a list $e_1, \dots, e_m$ of vectors in V is orthonormal if

$$\langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

6.24 Example ***orthonormal lists***

(a) The standard basis in $\mathbf{F}^n$ is an orthonormal list.

(b) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right)$ is an orthonormal list in $\mathbf{F}^3$.

(c) $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}\right)$ is an orthonormal list in $\mathbf{F}^3$.

Orthonormal lists are particularly easy to work with, as illustrated by the next result.

6.25 The norm of an orthonormal linear combination

If $e_1, \dots, e_m$ is an orthonormal list of vectors in V, then

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \dots, a_m \in \mathbf{F}$.

Proof Because each e_j has norm 1, this follows easily from repeated applications of the Pythagorean Theorem (6.13). ■

The result above has the following important corollary.

6.26 An orthonormal list is linearly independent

Every orthonormal list of vectors is linearly independent.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V and $a_1, \dots, a_m \in \mathbf{F}$ are such that

$$a_1 e_1 + \cdots + a_m e_m = 0.$$

Then $|a_1|^2 + \cdots + |a_m|^2 = 0$ (by 6.25), which means that all the a_j's are 0. Thus $e_1, \dots, e_m$ is linearly independent. ■

6.27 Definition ***orthonormal basis***

An ***orthonormal basis*** of V is an orthonormal list of vectors in V that is also a basis of V.

For example, the standard basis is an orthonormal basis of $\mathbf{F}^n$.

6.28 An orthonormal list of the right length is an orthonormal basis

Every orthonormal list of vectors in V with length $\dim V$ is an orthonormal basis of V.

Proof By 6.26, any such list must be linearly independent; because it has the right length, it is a basis—see 2.39. ■

6.29 Example Show that

$$\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right), \left(\tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right), \left(-\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{1}{2}\right)$$

is an orthonormal basis of $\mathbf{F}^4$.

Solution We have

$$\left\|\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right)\right\| = \sqrt{\left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^2} = 1.$$

Similarly, the other three vectors in the list above also have norm 1.

We have

$$\left\langle\left(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{2}, -\tfrac{1}{2}\right)\right\rangle = \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \tfrac{1}{2} + \tfrac{1}{2} \cdot \left(-\tfrac{1}{2}\right) + \tfrac{1}{2} \cdot \left(-\tfrac{1}{2}\right) = 0.$$

Similarly, the inner product of any two distinct vectors in the list above also equals 0.

Thus the list above is orthonormal. Because we have an orthonormal list of length four in the four-dimensional vector space $\mathbf{F}^4$, this list is an orthonormal basis of $\mathbf{F}^4$ (by 6.28).

In general, given a basis $e_1, \dots, e_n$ of V and a vector $v \in V$, we know that there is some choice of scalars $a_1, \dots, a_n \in \mathbf{F}$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

The importance of orthonormal bases stems mainly from the next result.

Computing the numbers $a_1, \dots, a_n$ that satisfy the equation above can be difficult for an arbitrary basis of V. The next result shows, however, that this is easy for an orthonormal basis—just take $a_j = \langle v, e_j \rangle$.

6.30 Writing a vector as linear combination of orthonormal basis

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $v \in V$. Then

$$v = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n$$

and

$$\|v\|^2 = |\langle v, e_1 \rangle|^2 + \cdots + |\langle v, e_n \rangle|^2.$$

Proof Because $e_1, \dots, e_n$ is a basis of V, there exist scalars $a_1, \dots, a_n$ such that

$$v = a_1 e_1 + \cdots + a_n e_n.$$

Because $e_1, \dots, e_n$ is orthonormal, taking the inner product of both sides of this equation with e_j gives $\langle v, e_j \rangle = a_j$. Thus the first equation in 6.30 holds.

The second equation in 6.30 follows immediately from the first equation and 6.25. ■

Now that we understand the usefulness of orthonormal bases, how do we go about finding them? For example, does $\mathcal{P}_m(\mathbf{R})$, with inner product given by integration on $[-1, 1]$ [see 6.4(c)], have an orthonormal basis? The next result will lead to answers to these questions.

Danish mathematician Jørgen Gram (1850–1916) and German mathematician Erhard Schmidt (1876–1959) popularized this algorithm that constructs orthonormal lists.

The algorithm used in the next proof is called the ***Gram–Schmidt Procedure***. It gives a method for turning a linearly independent list into an orthonormal list with the same span as the original list.

6.31 Gram–Schmidt Procedure

Suppose $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Let $e_1 = v_1/\|v_1\|$. For $j = 2, \dots, m$, define e_j inductively by

$$e_j = \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|}.$$

Then $e_1, \dots, e_m$ is an orthonormal list of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j)$$

for $j = 1, \dots, m$.

Proof We will show by induction on j that the desired conclusion holds. To get started with $j = 1$, note that $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ because v_1 is a positive multiple of e_1.

Suppose $1 < j \le m$ and we have verified that

6.32 $$\operatorname{span}(v_1, \dots, v_{j-1}) = \operatorname{span}(e_1, \dots, e_{j-1})$$

and $e_1, \dots, e_{j-1}$ is an orthonormal list. Note that $v_j \notin \operatorname{span}(v_1, \dots, v_{j-1})$ (because $v_1, \dots, v_m$ is linearly independent). Thus $v_j \notin \operatorname{span}(e_1, \dots, e_{j-1})$. Hence we are not dividing by 0 in the definition of e_j given in 6.31. Dividing a vector by its norm produces a new vector with norm 1; thus $\|e_j\| = 1$.

Let $1 \le k < j$. Then

$$\begin{aligned}\langle e_j, e_k\rangle &= \left\langle \frac{v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|}, e_k \right\rangle \\ &= \frac{\langle v_j, e_k\rangle - \langle v_j, e_k\rangle}{\|v_j - \langle v_j, e_1\rangle e_1 - \cdots - \langle v_j, e_{j-1}\rangle e_{j-1}\|} \\ &= 0.\end{aligned}$$

Thus $e_1, \dots, e_j$ is an orthonormal list.

From the definition of e_j given in 6.31, we see that $v_j \in \operatorname{span}(e_1, \dots, e_j)$. Combining this information with 6.32 shows that

$$\operatorname{span}(v_1, \dots, v_j) \subset \operatorname{span}(e_1, \dots, e_j).$$

Both lists above are linearly independent (the v's by hypothesis, the e's by orthonormality and 6.26). Thus both subspaces above have dimension j, and hence they are equal, completing the proof. ■

6.33 Example Find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$, where the inner product is given by $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,dx$.

Solution We will apply the Gram–Schmidt Procedure (6.31) to the basis $1, x, x^2$.

To get started, with this inner product we have

$$\|1\|^2 = \int_{-1}^{1} 1^2\,dx = 2.$$

Thus $\|1\| = \sqrt{2}$, and hence $e_1 = \sqrt{\tfrac{1}{2}}$.

Now the numerator in the expression for e_2 is

$$x - \langle x, e_1\rangle e_1 = x - \Big(\int_{-1}^{1} x\sqrt{\tfrac{1}{2}}\,dx\Big)\sqrt{\tfrac{1}{2}} = x.$$

We have

$$\|x\|^2 = \int_{-1}^{1} x^2\,dx = \tfrac{2}{3}.$$

Thus $\|x\| = \sqrt{\tfrac{2}{3}}$, and hence $e_2 = \sqrt{\tfrac{3}{2}}x$.

Now the numerator in the expression for e_3 is

$$\begin{aligned} x^2 &- \langle x^2, e_1\rangle e_1 - \langle x^2, e_2\rangle e_2 \\ &= x^2 - \Big(\int_{-1}^{1} x^2\sqrt{\tfrac{1}{2}}\,dx\Big)\sqrt{\tfrac{1}{2}} - \Big(\int_{-1}^{1} x^2\sqrt{\tfrac{3}{2}}x\,dx\Big)\sqrt{\tfrac{3}{2}}x \\ &= x^2 - \tfrac{1}{3}. \end{aligned}$$

We have

$$\|x^2 - \tfrac{1}{3}\|^2 = \int_{-1}^{1} \big(x^4 - \tfrac{2}{3}x^2 + \tfrac{1}{9}\big)\,dx = \tfrac{8}{45}.$$

Thus $\|x^2 - \tfrac{1}{3}\| = \sqrt{\tfrac{8}{45}}$, and hence $e_3 = \sqrt{\tfrac{45}{8}}\big(x^2 - \tfrac{1}{3}\big)$.

Thus

$$\sqrt{\tfrac{1}{2}},\ \sqrt{\tfrac{3}{2}}x,\ \sqrt{\tfrac{45}{8}}\big(x^2 - \tfrac{1}{3}\big)$$

is an orthonormal list of length 3 in $\mathcal{P}_2(\mathbf{R})$. Hence this orthonormal list is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ by 6.28.

Now we can answer the question about the existence of orthonormal bases.

6.34 Existence of orthonormal basis

Every finite-dimensional inner product space has an orthonormal basis.

Proof Suppose V is finite-dimensional. Choose a basis of V. Apply the Gram–Schmidt Procedure (6.31) to it, producing an orthonormal list with length $\dim V$. By 6.28, this orthonormal list is an orthonormal basis of V. ■

Sometimes we need to know not only that an orthonormal basis exists, but also that every orthonormal list can be extended to an orthonormal basis. In the next corollary, the Gram–Schmidt Procedure shows that such an extension is always possible.

6.35 Orthonormal list extends to orthonormal basis

Suppose V is finite-dimensional. Then every orthonormal list of vectors in V can be extended to an orthonormal basis of V.

Proof Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Then $e_1, \dots, e_m$ is linearly independent (by 6.26). Hence this list can be extended to a basis $e_1, \dots, e_m, v_1, \dots, v_n$ of V (see 2.33). Now apply the Gram–Schmidt Procedure (6.31) to $e_1, \dots, e_m, v_1, \dots, v_n$, producing an orthonormal list

6.36 $$e_1, \dots, e_m, f_1, \dots, f_n;$$

here the formula given by the Gram–Schmidt Procedure leaves the first m vectors unchanged because they are already orthonormal. The list above is an orthonormal basis of V by 6.28. ■

Recall that a matrix is called upper triangular if all entries below the diagonal equal 0. In other words, an upper-triangular matrix looks like this:

$$\begin{pmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{pmatrix},$$

where the 0 in the matrix above indicates that all entries below the diagonal equal 0, and asterisks are used to denote entries on and above the diagonal.

In the last chapter we showed that if V is a finite-dimensional complex vector space, then for each operator on V there is a basis with respect to which the matrix of the operator is upper triangular (see 5.27). Now that we are dealing with inner product spaces, we would like to know whether there exists an *orthonormal* basis with respect to which we have an upper-triangular matrix.

The next result shows that the existence of a basis with respect to which T has an upper-triangular matrix implies the existence of an orthonormal basis with this property. This result is true on both real and complex vector spaces (although on a real vector space, the hypothesis holds only for some operators).

6.37 Upper-triangular matrix with respect to orthonormal basis

Suppose $T \in \mathcal{L}(V)$. If T has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Proof Suppose T has an upper-triangular matrix with respect to some basis $v_1, \dots, v_n$ of V. Thus $\operatorname{span}(v_1, \dots, v_j)$ is invariant under T for each $j = 1, \dots, n$ (see 5.26).

Apply the Gram–Schmidt Procedure to $v_1, \dots, v_n$, producing an orthonormal basis $e_1, \dots, e_n$ of V. Because

$$\operatorname{span}(e_1, \dots, e_j) = \operatorname{span}(v_1, \dots, v_j)$$

for each j (see 6.31), we conclude that $\operatorname{span}(e_1, \dots, e_j)$ is invariant under T for each $j = 1, \dots, n$. Thus, by 5.26, T has an upper-triangular matrix with respect to the orthonormal basis $e_1, \dots, e_n$. ■

German mathematician Issai Schur (1875–1941) published the first proof of the next result in 1909.

The next result is an important application of the result above.

6.38 Schur's Theorem

Suppose T is an operator on a finite-dimensional complex inner product space V. Then T has an upper-triangular matrix with respect to some orthonormal basis of V.

Proof Recall that T has an upper-triangular matrix with respect to some basis of V (see 5.27). Now apply 6.37. ■

Linear Functionals on Inner Product Spaces

Because linear maps into the scalar field $\mathbf{F}$ play a special role, we defined a special name for them in Section 3.F. That definition is repeated below in case you skipped Section 3.F.

6.39 Definition ***linear functional***

A ***linear functional*** on V is a linear map from V to $\mathbf{F}$. In other words, a linear functional is an element of $\mathcal{L}(V, \mathbf{F})$.

6.40 Example The function $\varphi\colon \mathbf{F}^3 \to \mathbf{F}$ defined by

$$\varphi(z_1, z_2, z_3) = 2z_1 - 5z_2 + z_3$$

is a linear functional on $\mathbf{F}^3$. We could write this linear functional in the form

$$\varphi(z) = \langle z, u \rangle$$

for every $z \in \mathbf{F}^3$, where $u = (2, -5, 1)$.

6.41 Example The function $\varphi\colon \mathcal{P}_2(\mathbf{R}) \to \mathbf{R}$ defined by

$$\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\,dt$$

is a linear functional on $\mathcal{P}_2(\mathbf{R})$ (here the inner product on $\mathcal{P}_2(\mathbf{R})$ is multiplication followed by integration on $[-1, 1]$; see 6.33). It is not obvious that there exists $u \in \mathcal{P}_2(\mathbf{R})$ such that

$$\varphi(p) = \langle p, u \rangle$$

for every $p \in \mathcal{P}_2(\mathbf{R})$ [we cannot take $u(t) = \cos(\pi t)$ because that function is not an element of $\mathcal{P}_2(\mathbf{R})$].

If $u \in V$, then the map that sends v to $\langle v, u \rangle$ is a linear functional on V. The next result shows that every linear functional on V is of this form. Example 6.41 above illustrates the power of the next result because for the linear functional in that example, there is no obvious candidate for u.

The next result is named in honor of Hungarian mathematician Frigyes Riesz (1880–1956), who proved several results early in the twentieth century that look very much like the result below.

6.42 Riesz Representation Theorem

Suppose V is finite-dimensional and φ is a linear functional on V. Then there is a unique vector $u \in V$ such that

$$\varphi(v) = \langle v, u \rangle$$

for every $v \in V$.

Proof First we show there exists a vector $u \in V$ such that $\varphi(v) = \langle v, u \rangle$ for every $v \in V$. Let $e_1, \dots, e_n$ be an orthonormal basis of V. Then

$$\begin{aligned}\varphi(v) &= \varphi(\langle v, e_1 \rangle e_1 + \cdots + \langle v, e_n \rangle e_n) \\ &= \langle v, e_1 \rangle \varphi(e_1) + \cdots + \langle v, e_n \rangle \varphi(e_n) \\ &= \langle v, \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n \rangle\end{aligned}$$

for every $v \in V$, where the first equality comes from 6.30. Thus setting

6.43
$$u = \overline{\varphi(e_1)} e_1 + \cdots + \overline{\varphi(e_n)} e_n,$$

we have $\varphi(v) = \langle v, u \rangle$ for every $v \in V$, as desired.

Now we prove that only one vector $u \in V$ has the desired behavior. Suppose $u_1, u_2 \in V$ are such that

$$\varphi(v) = \langle v, u_1 \rangle = \langle v, u_2 \rangle$$

for every $v \in V$. Then

$$0 = \langle v, u_1 \rangle - \langle v, u_2 \rangle = \langle v, u_1 - u_2 \rangle$$

for every $v \in V$. Taking $v = u_1 - u_2$ shows that $u_1 - u_2 = 0$. In other words, $u_1 = u_2$, completing the proof of the uniqueness part of the result. ■

6.44 **Example** Find $u \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\, dt = \int_{-1}^{1} p(t) u(t)\, dt$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution Let $\varphi(p) = \int_{-1}^{1} p(t)\big(\cos(\pi t)\big)\,dt$. Applying formula 6.43 from the proof above, and using the orthonormal basis from Example 6.33, we have

$$u(x) = \Big(\int_{-1}^{1} \sqrt{\tfrac{1}{2}}\big(\cos(\pi t)\big)\,dt\Big)\sqrt{\tfrac{1}{2}} + \Big(\int_{-1}^{1} \sqrt{\tfrac{3}{2}}\,t\big(\cos(\pi t)\big)\,dt\Big)\sqrt{\tfrac{3}{2}}\,x$$
$$+ \Big(\int_{-1}^{1} \sqrt{\tfrac{45}{8}}\big(t^2 - \tfrac{1}{3}\big)\big(\cos(\pi t)\big)\,dt\Big)\sqrt{\tfrac{45}{8}}\big(x^2 - \tfrac{1}{3}\big).$$

A bit of calculus shows that

$$u(x) = -\tfrac{45}{2\pi^2}\big(x^2 - \tfrac{1}{3}\big).$$

Suppose V is finite-dimensional and φ a linear functional on V. Then 6.43 gives a formula for the vector u that satisfies $\varphi(v) = \langle v, u\rangle$ for all $v \in V$. Specifically, we have

$$u = \overline{\varphi(e_1)}e_1 + \cdots + \overline{\varphi(e_n)}e_n.$$

The right side of the equation above seems to depend on the orthonormal basis $e_1, \dots, e_n$ as well as on φ. However, 6.42 tells us that u is uniquely determined by φ. Thus the right side of the equation above is the same regardless of which orthonormal basis $e_1, \dots, e_n$ of V is chosen.

EXERCISES 6.B

1 (a) Suppose $\theta \in \mathbf{R}$. Show that $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ and $(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$ are orthonormal bases of $\mathbf{R}^2$.

(b) Show that each orthonormal basis of $\mathbf{R}^2$ is of the form given by one of the two possibilities of part (a).

2 Suppose $e_1, \dots, e_m$ is an orthonormal list of vectors in V. Let $v \in V$. Prove that
$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$$
if and only if $v \in \operatorname{span}(e_1, \dots, e_m)$.

3 Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ has an upper-triangular matrix with respect to the basis $(1, 0, 0)$, $(1, 1, 1)$, $(1, 1, 2)$. Find an orthonormal basis of $\mathbf{R}^3$ (use the usual inner product on $\mathbf{R}^3$) with respect to which T has an upper-triangular matrix.

4 Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

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

[*The orthonormal list above is often used for modeling periodic phenomena such as tides.*]

5 On $\mathcal{P}_2(\mathbf{R})$, consider the inner product given by

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

Apply the Gram–Schmidt Procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

6 Find an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ (with inner product as in Exercise 5) such that the differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to this basis.

7 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$p\left(\tfrac{1}{2}\right) = \int_0^1 p(x)q(x)\,dx$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

8 Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x)(\cos \pi x)\,dx = \int_0^1 p(x)q(x)\,dx$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

9 What happens if the Gram–Schmidt Procedure is applied to a list of vectors that is not linearly independent?

10 Suppose V is a real inner product space and $v_1, \dots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \dots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \dots, v_j) = \operatorname{span}(e_1, \dots, e_j)$$

for all $j \in \{1, \dots, m\}$.

11 Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle v, w \rangle_1 = 0$ if and only if $\langle v, w \rangle_2 = 0$. Prove that there is a positive number c such that $\langle v, w \rangle_1 = c \langle v, w \rangle_2$ for every $v, w \in V$.

12 Suppose V is finite-dimensional and $\langle \cdot, \cdot \rangle_1$, $\langle \cdot, \cdot \rangle_2$ are inner products on V with corresponding norms $\| \cdot \|_1$ and $\| \cdot \|_2$. Prove that there exists a positive number c such that

$$\|v\|_1 \leq c \|v\|_2$$

for every $v \in V$.

13 Suppose $v_1, \dots, v_m$ is a linearly independent list in V. Show that there exists $w \in V$ such that $\langle w, v_j \rangle > 0$ for all $j \in \{1, \dots, m\}$.

14 Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $v_1, \dots, v_n$ are vectors in V such that

$$\|e_j - v_j\| < \frac{1}{\sqrt{n}}$$

for each j. Prove that $v_1, \dots, v_n$ is a basis of V.

15 Suppose $C_{\mathbf{R}}([-1, 1])$ is the vector space of continuous real-valued functions on the interval $[-1, 1]$ with inner product given by

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

for $f, g \in C_{\mathbf{R}}([-1, 1])$. Let φ be the linear functional on $C_{\mathbf{R}}([-1, 1])$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C_{\mathbf{R}}([-1, 1])$ such that

$$\varphi(f) = \langle f, g \rangle$$

for every $f \in C_{\mathbf{R}}([-1, 1])$.
[*The exercise above shows that the Riesz Representation Theorem* (6.42) *does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.*]

16 Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, all the eigenvalues of T have absolute value less than 1, and $\epsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \epsilon \|v\|$ for every $v \in V$.

17 For $u \in V$, let Φu denote the linear functional on V defined by

$$(\Phi u)(v) = \langle v, u \rangle$$

for $v \in V$.

(a) Show that if $\mathbf{F} = \mathbf{R}$, then Φ is a linear map from V to V'. (Recall from Section 3.F that $V' = \mathcal{L}(V, \mathbf{F})$ and that V' is called the dual space of V.)

(b) Show that if $\mathbf{F} = \mathbf{C}$ and $V \neq \{0\}$, then Φ is not a linear map.

(c) Show that Φ is injective.

(d) Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. Use parts (a) and (c) and a dimension-counting argument (but without using 6.42) to show that Φ is an isomorphism from V onto V'.

[*Part* (*d*) *gives an alternative proof of the Riesz Representation Theorem* (*6.42*) *when* $\mathbf{F} = \mathbf{R}$. *Part* (*d*) *also gives a natural isomorphism* (*meaning that it does not depend on a choice of basis*) *from a finite-dimensional real inner product space onto its dual space.*]

6.C *Orthogonal Complements and Minimization Problems*

Orthogonal Complements

6.45 Definition ***orthogonal complement,*** $U^{\perp}$

If U is a subset of V, then the ***orthogonal complement*** of U, denoted $U^{\perp}$, is the set of all vectors in V that are orthogonal to every vector in U:

$$U^{\perp} = \{v \in V : \langle v, u \rangle = 0 \text{ for every } u \in U\}.$$

For example, if U is a line in $\mathbf{R}^3$ containing the origin, then $U^{\perp}$ is the plane containing the origin that is perpendicular to U. If U is a plane in $\mathbf{R}^3$ containing the origin, then $U^{\perp}$ is the line containing the origin that is perpendicular to U.

6.46 Basic properties of orthogonal complement

(a) If U is a subset of V, then $U^{\perp}$ is a subspace of V.

(b) $\{0\}^{\perp} = V$.

(c) $V^{\perp} = \{0\}$.

(d) If U is a subset of V, then $U \cap U^{\perp} \subset \{0\}$.

(e) If U and W are subsets of V and $U \subset W$, then $W^{\perp} \subset U^{\perp}$.

Proof

(a) Suppose U is a subset of V. Then $\langle 0, u \rangle = 0$ for every $u \in U$; thus $0 \in U^{\perp}$.

Suppose $v, w \in U^{\perp}$. If $u \in U$, then

$$\langle v + w, u \rangle = \langle v, u \rangle + \langle w, u \rangle = 0 + 0 = 0.$$

Thus $v + w \in U^{\perp}$. In other words, $U^{\perp}$ is closed under addition.

Similarly, suppose $\lambda \in \mathbf{F}$ and $v \in U^{\perp}$. If $u \in U$, then

$$\langle \lambda v, u \rangle = \lambda \langle v, u \rangle = \lambda \cdot 0 = 0.$$

Thus $\lambda v \in U^{\perp}$. In other words, $U^{\perp}$ is closed under scalar multiplication. Thus $U^{\perp}$ is a subspace of V.

(b) Suppose $v \in V$. Then $\langle v, 0\rangle = 0$, which implies that $v \in \{0\}^\perp$. Thus $\{0\}^\perp = V$.

(c) Suppose $v \in V^\perp$. Then $\langle v, v\rangle = 0$, which implies that $v = 0$. Thus $V^\perp = \{0\}$.

(d) Suppose U is a subset of V and $v \in U \cap U^\perp$. Then $\langle v, v\rangle = 0$, which implies that $v = 0$. Thus $U \cap U^\perp \subset \{0\}$.

(e) Suppose U and W are subsets of V and $U \subset W$. Suppose $v \in W^\perp$. Then $\langle v, u\rangle = 0$ for every $u \in W$, which implies that $\langle v, u\rangle = 0$ for every $u \in U$. Hence $v \in U^\perp$. Thus $W^\perp \subset U^\perp$. ■

Recall that if U, W are subspaces of V, then V is the direct sum of U and W (written $V = U \oplus W$) if each element of V can be written in exactly one way as a vector in U plus a vector in W (see 1.40).

The next result shows that every finite-dimensional subspace of V leads to a natural direct sum decomposition of V.

6.47 Direct sum of a subspace and its orthogonal complement

Suppose U is a finite-dimensional subspace of V. Then

$$V = U \oplus U^\perp.$$

Proof First we will show that

6.48
$$V = U + U^\perp.$$

To do this, suppose $v \in V$. Let $e_1, \dots, e_m$ be an orthonormal basis of U. Obviously

6.49
$$v = \underbrace{\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m}_{u} + \underbrace{v - \langle v, e_1\rangle e_1 - \cdots - \langle v, e_m\rangle e_m}_{w}.$$

Let u and w be defined as in the equation above. Clearly $u \in U$. Because $e_1, \dots, e_m$ is an orthonormal list, for each $j = 1, \dots, m$ we have

$$\begin{aligned} \langle w, e_j\rangle &= \langle v, e_j\rangle - \langle v, e_j\rangle \\ &= 0. \end{aligned}$$

Thus w is orthogonal to every vector in $\operatorname{span}(e_1, \dots, e_m)$. In other words, $w \in U^\perp$. Thus we have written $v = u + w$, where $u \in U$ and $w \in U^\perp$, completing the proof of 6.48.

From 6.46(d), we know that $U \cap U^\perp = \{0\}$. Along with 6.48, this implies that $V = U \oplus U^\perp$ (see 1.45). ■

Now we can see how to compute $\dim U^{\perp}$ from $\dim U$.

6.50 Dimension of the orthogonal complement

Suppose V is finite-dimensional and U is a subspace of V. Then

$$\dim U^{\perp} = \dim V - \dim U.$$

Proof: The formula for $\dim U^{\perp}$ follows immediately from 6.47 and 3.78. ■

The next result is an important consequence of 6.47.

6.51 The orthogonal complement of the orthogonal complement

Suppose U is a finite-dimensional subspace of V. Then

$$U = (U^{\perp})^{\perp}.$$

Proof: First we will show that

6.52
$$U \subset (U^{\perp})^{\perp}.$$

To do this, suppose $u \in U$. Then $\langle u, v\rangle = 0$ for every $v \in U^{\perp}$ (by the definition of $U^{\perp}$). Because u is orthogonal to every vector in $U^{\perp}$, we have $u \in (U^{\perp})^{\perp}$, completing the proof of 6.52.

To prove the inclusion in the other direction, suppose $v \in (U^{\perp})^{\perp}$. By 6.47, we can write $v = u + w$, where $u \in U$ and $w \in U^{\perp}$. We have $v - u = w \in U^{\perp}$. Because $v \in (U^{\perp})^{\perp}$ and $u \in (U^{\perp})^{\perp}$ (from 6.52), we have $v - u \in (U^{\perp})^{\perp}$. Thus $v - u \in U^{\perp} \cap (U^{\perp})^{\perp}$, which implies that $v - u$ is orthogonal to itself, which implies that $v - u = 0$, which implies that $v = u$, which implies that $v \in U$. Thus $(U^{\perp})^{\perp} \subset U$, which along with 6.52 completes the proof. ■

We now define an operator P_U for each finite-dimensional subspace of V.

6.53 Definition ***orthogonal projection,*** P_U

Suppose U is a finite-dimensional subspace of V. The ***orthogonal projection*** of V onto U is the operator $P_U \in \mathcal{L}(V)$ defined as follows: For $v \in V$, write $v = u + w$, where $u \in U$ and $w \in U^{\perp}$. Then $P_U v = u$.

The direct sum decomposition $V = U \oplus U^\perp$ given by 6.47 shows that each $v \in V$ can be uniquely written in the form $v = u + w$ with $u \in U$ and $w \in U^\perp$. Thus $P_U v$ is well defined.

6.54 **Example** Suppose $x \in V$ with $x \neq 0$ and $U = \operatorname{span}(x)$. Show that

$$P_U v = \frac{\langle v, x\rangle}{\|x\|^2} x$$

for every $v \in V$.

Solution Suppose $v \in V$. Then

$$v = \frac{\langle v, x\rangle}{\|x\|^2} x + \Bigl(v - \frac{\langle v, x\rangle}{\|x\|^2} x\Bigr),$$

where the first term on the right is in $\operatorname{span}(x)$ (and thus in U) and the second term on the right is orthogonal to x (and thus is in $U^\perp$). Thus $P_U v$ equals the first term on the right, as desired.

6.55 Properties of the orthogonal projection P_U

Suppose U is a finite-dimensional subspace of V and $v \in V$. Then

(a) $P_U \in \mathcal{L}(V)$;

(b) $P_U u = u$ for every $u \in U$;

(c) $P_U w = 0$ for every $w \in U^\perp$;

(d) range $P_U = U$;

(e) null $P_U = U^\perp$;

(f) $v - P_U v \in U^\perp$;

(g) $P_U{}^2 = P_U$;

(h) $\|P_U v\| \leq \|v\|$;

(i) for every orthonormal basis $e_1, \ldots, e_m$ of U,

$$P_U v = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m.$$

Proof

(a) To show that P_U is a linear map on V, suppose $v_1, v_2 \in V$. Write

$$v_1 = u_1 + w_1 \quad \text{and} \quad v_2 = u_2 + w_2$$

with $u_1, u_2 \in U$ and $w_1, w_2 \in U^\perp$. Thus $P_U v_1 = u_1$ and $P_U v_2 = u_2$. Now

$$v_1 + v_2 = (u_1 + u_2) + (w_1 + w_2),$$

where $u_1 + u_2 \in U$ and $w_1 + w_2 \in U^\perp$. Thus

$$P_U(v_1 + v_2) = u_1 + u_2 = P_U v_1 + P_U v_2.$$

Similarly, suppose $\lambda \in \mathbf{F}$. The equation $v = u + w$ with $u \in U$ and $w \in U^\perp$ implies that $\lambda v = \lambda u + \lambda w$ with $\lambda u \in U$ and $\lambda w \in U^\perp$. Thus $P_U(\lambda v) = \lambda u = \lambda P_U v$.

Hence P_U is a linear map from V to V.

(b) Suppose $u \in U$. We can write $u = u + 0$, where $u \in U$ and $0 \in U^\perp$. Thus $P_U u = u$.

(c) Suppose $w \in U^\perp$. We can write $w = 0 + w$, where $0 \in U$ and $w \in U^\perp$. Thus $P_U w = 0$.

(d) The definition of P_U implies that range $P_U \subset U$. Part (b) implies that $U \subset$ range P_U. Thus range $P_U = U$.

(e) Part (c) implies that $U^\perp \subset$ null P_U. To prove the inclusion in the other direction, note that if $v \in$ null P_U then the decomposition given by 6.47 must be $v = 0 + v$, where $0 \in U$ and $v \in U^\perp$. Thus null $P_U \subset U^\perp$.

(f) If $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$v - P_U v = v - u = w \in U^\perp.$$

(g) If $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$(P_U{}^2)v = P_U(P_U v) = P_U u = u = P_U v.$$

(h) If $v = u + w$ with $u \in U$ and $w \in U^\perp$, then

$$\|P_U v\|^2 = \|u\|^2 \leq \|u\|^2 + \|w\|^2 = \|v\|^2,$$

where the last equality comes from the Pythagorean Theorem.

(i) The formula for $P_U v$ follows from equation 6.49 in the proof of 6.47. ■

Minimization Problems

The remarkable simplicity of the solution to this minimization problem has led to many important applications of inner product spaces outside of pure mathematics.

The following problem often arises: given a subspace U of V and a point $v \in V$, find a point $u \in U$ such that $\|v - u\|$ is as small as possible. The next proposition shows that this minimization problem is solved by taking $u = P_U v$.

6.56 Minimizing the distance to a subspace

Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \leq \|v - u\|.$$

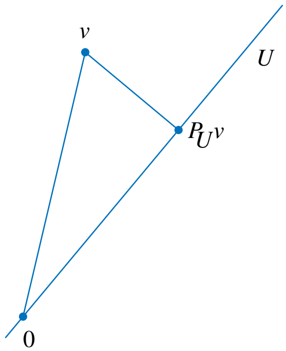
Furthermore, the inequality above is an equality if and only if $u = P_U v$.

Proof We have

6.57
$$\begin{aligned} \|v - P_U v\|^2 &\leq \|v - P_U v\|^2 + \|P_U v - u\|^2 \\ &= \|(v - P_U v) + (P_U v - u)\|^2 \\ &= \|v - u\|^2, \end{aligned}$$

where the first line above holds because $0 \leq \|P_U v - u\|^2$, the second line above comes from the Pythagorean Theorem [which applies because $v - P_U v \in U^\perp$ by 6.55(f), and $P_U v - u \in U$], and the third line above holds by simple algebra. Taking square roots gives the desired inequality.

Our inequality above is an equality if and only if 6.57 is an equality, which happens if and only if $\|P_U v - u\| = 0$, which happens if and only if $u = P_U v$.



$P_U v$ is the closest point in U to v.

The last result is often combined with the formula 6.55(i) to compute explicit solutions to minimization problems.

6.58 **Example** Find a polynomial u with real coefficients and degree at most 5 that approximates $\sin x$ as well as possible on the interval $[-\pi, \pi]$, in the sense that

$$\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$$

is as small as possible. Compare this result to the Taylor series approximation.

Solution Let $C_{\mathbf{R}}[-\pi, \pi]$ denote the real inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

6.59
$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x) \, dx.$$

Let $v \in C_{\mathbf{R}}[-\pi, \pi]$ be the function defined by $v(x) = \sin x$. Let U denote the subspace of $C_{\mathbf{R}}[-\pi, \pi]$ consisting of the polynomials with real coefficients and degree at most 5. Our problem can now be reformulated as follows:

Find $u \in U$ such that $\|v - u\|$ is as small as possible.

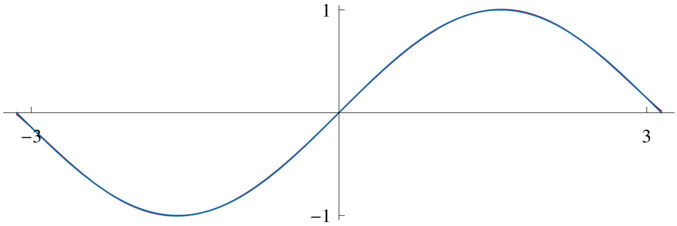
A computer that can perform integrations is useful here.

To compute the solution to our approximation problem, first apply the Gram–Schmidt Procedure (using the inner product given by 6.59) to the basis $1, x, x^2, x^3, x^4, x^5$ of U, producing an orthonormal basis $e_1, e_2, e_3, e_4, e_5, e_6$ of U. Then, again using the inner product given by 6.59, compute $P_U v$ using 6.55(i) (with $m = 6$). Doing this computation shows that $P_U v$ is the function u defined by

6.60
$$u(x) = 0.987862x - 0.155271x^3 + 0.00564312x^5,$$

where the π's that appear in the exact answer have been replaced with a good decimal approximation.

By 6.56, the polynomial u above is the best approximation to $\sin x$ on $[-\pi, \pi]$ using polynomials of degree at most 5 (here "best approximation" means in the sense of minimizing $\int_{-\pi}^{\pi} |\sin x - u(x)|^2 \, dx$). To see how good this approximation is, the next figure shows the graphs of both $\sin x$ and our approximation $u(x)$ given by 6.60 over the interval $[-\pi, \pi]$.



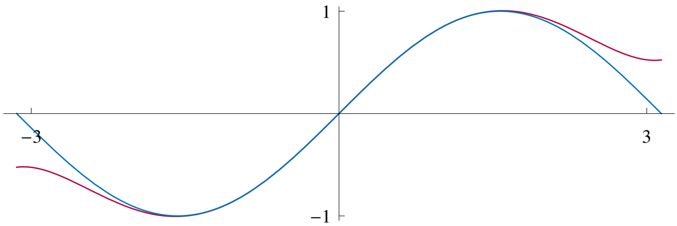
Graphs on $[-\pi, \pi]$ *of* $\sin x$ (*blue*) *and its approximation* $u(x)$ (*red*) *given by 6.60.*

Our approximation 6.60 is so accurate that the two graphs are almost identical—our eyes may see only one graph! Here the blue graph is placed almost exactly over the red graph. If you are viewing this on an electronic device, try enlarging the picture above, especially near 3 or -3, to see a small gap between the two graphs.

Another well-known approximation to $\sin x$ by a polynomial of degree 5 is given by the Taylor polynomial

6.61
$$x - \frac{x^3}{3!} + \frac{x^5}{5!}.$$

To see how good this approximation is, the next picture shows the graphs of both $\sin x$ and the Taylor polynomial 6.61 over the interval $[-\pi, \pi]$.



Graphs on $[-\pi, \pi]$ *of* $\sin x$ (*blue*) *and the Taylor polynomial 6.61* (*red*).

The Taylor polynomial is an excellent approximation to $\sin x$ for x near 0. But the picture above shows that for $|x| > 2$, the Taylor polynomial is not so accurate, especially compared to 6.60. For example, taking $x = 3$, our approximation 6.60 estimates $\sin 3$ with an error of about 0.001, but the Taylor series 6.61 estimates $\sin 3$ with an error of about 0.4. Thus at $x = 3$, the error in the Taylor series is hundreds of times larger than the error given by 6.60. Linear algebra has helped us discover an approximation to $\sin x$ that improves upon what we learned in calculus!

EXERCISES 6.C

1 Suppose $v_1, \dots, v_m \in V$. Prove that

$$\{v_1, \dots, v_m\}^{\perp} = \big(\operatorname{span}(v_1, \dots, v_m)\big)^{\perp}.$$

2 Suppose U is a finite-dimensional subspace of V. Prove that $U^{\perp} = \{0\}$ if and only if $U = V$.
[*Exercise 14(a) shows that the result above is not true without the hypothesis that U is finite-dimensional.*]

3 Suppose U is a subspace of V with basis $u_1, \dots, u_m$ and

$$u_1, \dots, u_m, w_1, \dots, w_n$$

is a basis of V. Prove that if the Gram–Schmidt Procedure is applied to the basis of V above, producing a list $e_1, \dots, e_m, f_1, \dots, f_n$, then $e_1, \dots, e_m$ is an orthonormal basis of U and $f_1, \dots, f_n$ is an orthonormal basis of $U^{\perp}$.

4 Suppose U is the subspace of $\mathbf{R}^4$ defined by

$$U = \operatorname{span}\big((1, 2, 3, -4), (-5, 4, 3, 2)\big).$$

Find an orthonormal basis of U and an orthonormal basis of $U^{\perp}$.

5 Suppose V is finite-dimensional and U is a subspace of V. Show that $P_{U^{\perp}} = I - P_U$, where I is the identity operator on V.

6 Suppose U and W are finite-dimensional subspaces of V. Prove that $P_U P_W = 0$ if and only if $\langle u, w \rangle = 0$ for all $u \in U$ and all $w \in W$.

7 Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in null P is orthogonal to every vector in range P. Prove that there exists a subspace U of V such that $P = P_U$.

8 Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \le \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

9 Suppose $T \in \mathcal{L}(V)$ and U is a finite-dimensional subspace of V. Prove that U is invariant under T if and only if $P_U T P_U = T P_U$.

10 Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V. Prove that U and $U^\perp$ are both invariant under T if and only if $P_U T = T P_U$.

11 In $\mathbf{R}^4$, let

$$U = \operatorname{span}\big((1,1,0,0),(1,1,1,2)\big).$$

Find $u \in U$ such that $\|u - (1,2,3,4)\|$ is as small as possible.

12 Find $p \in \mathcal{P}_3(\mathbf{R})$ such that $p(0) = 0$, $p'(0) = 0$, and

$$\int_0^1 |2 + 3x - p(x)|^2 \, dx$$

is as small as possible.

13 Find $p \in \mathcal{P}_5(\mathbf{R})$ that makes

$$\int_{-\pi}^{\pi} |\sin x - p(x)|^2 \, dx$$

as small as possible.

[*The polynomial 6.60 is an excellent approximation to the answer to this exercise, but here you are asked to find the exact solution, which involves powers of π. A computer that can perform symbolic integration will be useful.*]

14 Suppose $C_{\mathbf{R}}([-1,1])$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

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

for $f, g \in C_{\mathbf{R}}([-1,1])$. Let U be the subspace of $C_{\mathbf{R}}([-1,1])$ defined by

$$U = \{f \in C_{\mathbf{R}}([-1,1]) : f(0) = 0\}.$$

(a) Show that $U^\perp = \{0\}$.

(b) Show that 6.47 and 6.51 do not hold without the finite-dimensional hypothesis.