

Chapter 6

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 geometric features such as the notions of length and angle. These ideas are embedded in the concept of inner products, which we will investigate in this chapter.

Every inner product induces a norm, which you can think of as a length. This norm satisfies key properties such as the Pythagorean theorem, the triangle inequality, the parallelogram equality, and the Cauchy–Schwarz inequality.

The notion of perpendicular vectors in Euclidean geometry gets renamed to orthogonal vectors in the context of an inner product space. We will see that orthonormal bases are tremendously useful in inner product spaces. The Gram–Schmidt procedure constructs such bases. This chapter will conclude by putting together these tools to solve minimization problems.

standing assumptions for this chapter

- F denotes $\mathbf{R}$ or $\mathbf{C}$.
- V and W denote vector spaces over F .



Mathew Petroff CC BY-SA

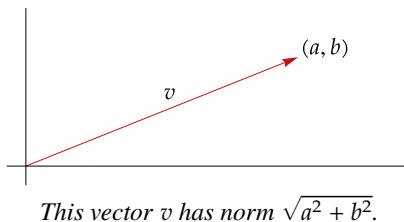
*The George Peabody Library, now part of Johns Hopkins University, opened while James Sylvester (1814–1897) was the university’s first mathematics professor. Sylvester’s publications include the first use of the word **matrix** in mathematics.*

6A Inner Products and Norms

Inner Products

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 v in $\mathbf{R}^2$ or $\mathbf{R}^3$ is called the *norm* of v and is denoted by $\|v\|$. Thus for $v = (a, b) \in \mathbf{R}^2$, we have

$$\|v\| = \sqrt{a^2 + b^2}.$$



Similarly, if $v = (a, b, c) \in \mathbf{R}^3$, then $\|v\| = \sqrt{a^2 + b^2 + c^2}$.

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

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

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

6.1 definition: dot product

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

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

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

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

If we think of a vector as a point instead of as an arrow, then $\|x\|$ should be interpreted to mean the distance from the origin to the point x .

- $x \cdot x \geq 0$ 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 by $|\lambda|$, is defined by $|\lambda| = \sqrt{a^2 + b^2}$;
- the complex conjugate of λ , denoted by $\bar{\lambda}$, 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 $z = (z_1, \dots, z_n) \in \mathbf{C}^n$, we define the norm of z by

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

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

$$\|z\|^2 = z_1\bar{z}_1 + \dots + z_n\bar{z}_n.$$

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 the vector $w = (w_1, \dots, w_n) \in \mathbf{C}^n$ with z should equal

$$w_1\bar{z}_1 + \dots + w_n\bar{z}_n.$$

If the roles of the w and z were interchanged, the expression above would be replaced with its complex conjugate. Thus 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.

One comment about the notation used in the next definition:

- For $\lambda \in \mathbf{C}$, the notation $\lambda \geq 0$ means λ is real and nonnegative.

6.2 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

$$\langle v, v \rangle \geq 0 \text{ for all } v \in V.$$

definiteness

$$\langle v, v \rangle = 0 \text{ if and only if } v = 0.$$

additivity in first slot

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle \text{ for all } u, v, w \in V.$$

homogeneity in first slot

$$\langle \lambda u, v \rangle = \lambda \langle u, v \rangle \text{ for all } \lambda \in \mathbf{F} \text{ and all } u, v \in V.$$

conjugate symmetry

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

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 $u, v \in V$.

Most mathematicians define inner products as above, but many physicists use a definition that requires homogeneity in the second slot instead of the first slot.

6.3 example: inner products

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

$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n) \rangle = w_1 \bar{z}_1 + \dots + w_n \bar{z}_n$$

for all $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

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

$$\langle (w_1, \dots, w_n), (z_1, \dots, z_n) \rangle = c_1 w_1 \bar{z}_1 + \dots + c_n w_n \bar{z}_n$$

for all $(w_1, \dots, w_n), (z_1, \dots, z_n) \in \mathbf{F}^n$.

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

$$\langle f, g \rangle = \int_{-1}^1 f g$$

for all f, g continuous real-valued functions on $[-1, 1]$.

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

$$\langle p, q \rangle = p(0)q(0) + \int_{-1}^1 p' q'$$

for all $p, q \in \mathcal{P}(\mathbf{R})$.

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

$$\langle p, q \rangle = \int_0^\infty p(x)q(x)e^{-x} dx$$

for all $p, q \in \mathcal{P}(\mathbf{R})$.

6.4 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 (a) in the example above. 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 and W are inner product spaces, we make the following assumption.

6.5 notation: V, W

For the rest of this chapter and the next chapter, V and W denote inner product spaces 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 clear from the context (or is the Euclidean inner product if the vector space is $\mathbf{F}^n$).

6.6 basic properties of an inner product

- (a) For each fixed $v \in V$, the function that takes $u \in V$ to $\langle u, v \rangle$ is a linear map from V to $\mathbf{F}$.
- (b) $\langle 0, v \rangle = 0$ for every $v \in V$.
- (c) $\langle v, 0 \rangle = 0$ for every $v \in V$.
- (d) $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$ for all $u, v, w \in V$.
- (e) $\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle$ for all $\lambda \in \mathbf{F}$ and all $u, v \in V$.

Proof

- (a) For $v \in V$, the linearity of $u \mapsto \langle u, v \rangle$ follows from the conditions of additivity and homogeneity in the first slot in the definition of an inner product.
- (b) Every linear map takes 0 to 0. Thus (b) follows from (a).
- (c) If $v \in V$, then the conjugate symmetry property in the definition of an inner product and (b) show that $\langle v, 0 \rangle = \overline{\langle 0, v \rangle} = \overline{0} = 0$.
- (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

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



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.

6.7 definition: *norm*, $\|v\|$

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

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

6.8 example: *norms*

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

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

(b) For f in the vector space of continuous real-valued functions on $[-1, 1]$ and with inner product given as in 6.3(c), we have

$$\|f\| = \sqrt{\int_{-1}^1 f^2}.$$

6.9 *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

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

Taking square roots now gives the desired equality. ■

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

Now we come to a crucial definition.

6.10 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 two vectors does not matter, because $\langle u, v \rangle = 0$ if and only if $\langle v, u \rangle = 0$. Instead of saying u and v are orthogonal, sometimes we say u is orthogonal to v .

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

Exercise 15 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 nonzero vectors in $\mathbf{R}^2$ are orthogonal (with respect to the 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.

6.11 *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) Recall that 6.6(b) states that $\langle 0, v \rangle = 0$ for every $v \in V$.
- (b) If $v \in V$ and $\langle v, v \rangle = 0$, then $v = 0$ (by definition of inner product). ■

For the special case $V = \mathbf{R}^2$, the next theorem was known over 3,500 years ago in Babylonia and then rediscovered and proved over 2,500 years ago in Greece. Of course, the proof below is not the original proof.

6.12 *Pythagorean theorem*

Suppose $u, v \in V$. If u and v are orthogonal, then

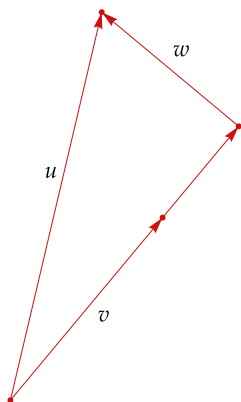
$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof Suppose $\langle u, v \rangle = 0$. Then

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

■

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 picture here.



*An orthogonal decomposition:
 u expressed as a scalar multiple of v plus a vector orthogonal to v .*

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 $(u - cv)$. Hence 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 displayed above explicitly writes u as a scalar multiple of v plus a vector orthogonal to v . Thus we have proved the following key result.

6.13 an orthogonal decomposition

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

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

The orthogonal decomposition 6.13 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.

6.14 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 6.13, where w is orthogonal to v . By the Pythagorean theorem,

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

6.15

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

The proof in the paragraph above shows that the Cauchy–Schwarz inequality is an equality if and only if 6.15 is an equality. This happens if and only if $w = 0$. But $w = 0$ if and only if u is a multiple of v (see 6.13). 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).

Augustin-Louis Cauchy (1789–1857) proved 6.16(a) in 1821. In 1859, Cauchy's student Viktor Bunyakovsky (1804–1889) proved integral inequalities like the one in 6.16(b). A few decades later, similar discoveries by Hermann Schwarz (1843–1921) attracted more attention and led to the name of this inequality.

6.16 example: Cauchy–Schwarz inequality

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

$$(x_1 y_1 + \dots + x_n y_n)^2 \leq (x_1^2 + \dots + x_n^2)(y_1^2 + \dots + y_n^2),$$

as follows from applying the Cauchy–Schwarz inequality to the vectors $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbf{R}^n$, using the usual Euclidean inner product.

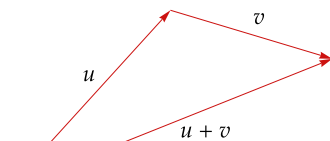
(b) If f, g are continuous real-valued functions on $[-1, 1]$, then

$$\left| \int_{-1}^1 fg \right|^2 \leq \left(\int_{-1}^1 f^2 \right) \left(\int_{-1}^1 g^2 \right),$$

as follows from applying the Cauchy–Schwarz inequality to Example 6.3(c).

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 polygonal path between two points is a single line segment (a polygonal path consists of line segments).



In this triangle, the length of $u + v$ is less than the length of u plus the length of v .

6.17 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 real multiple of the other.

Proof We have

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \langle v, u \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \overline{\langle u, v \rangle} \\ &= \|u\|^2 + \|v\|^2 + 2 \operatorname{Re} \langle u, v \rangle \\ 6.18 \quad &\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v \rangle| \\ 6.19 \quad &\leq \|u\|^2 + \|v\|^2 + 2\|u\| \|v\| \\ &= (\|u\| + \|v\|)^2, \end{aligned}$$

where 6.19 follows from the Cauchy–Schwarz inequality (6.14). Taking square roots of both sides of the inequality above gives the desired inequality.

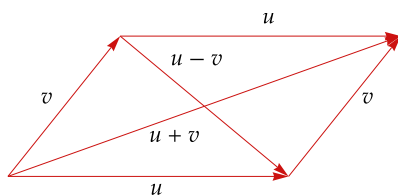
The proof above shows that the triangle inequality is an equality if and only if we have equality in 6.18 and 6.19. Thus we have equality in the triangle inequality if and only if

$$6.20 \quad \langle u, v \rangle = \|u\| \|v\|.$$

If one of u, v is a nonnegative real multiple of the other, then 6.20 holds. Conversely, suppose 6.20 holds. Then the condition for equality in the Cauchy–Schwarz inequality (6.14) implies that one of u, v is a scalar multiple of the other. This scalar must be a nonnegative real number, by 6.20, completing the proof. ■

For the reverse triangle inequality, see Exercise 20.

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. Note that the proof here is more straightforward than the usual proof in Euclidean geometry.



The diagonals of this parallelogram are $u + v$ and $u - v$.

6.21 parallelogram equality

Suppose $u, v \in V$. Then

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

Proof We have

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

as desired. ■

Exercises 6A

- 1 Prove or give a counterexample: If $v_1, \dots, v_m \in V$, then

$$\sum_{j=1}^m \sum_{k=1}^m \langle v_j, v_k \rangle \geq 0.$$

- 2 Suppose $S \in \mathcal{L}(V)$. Define $\langle \cdot, \cdot \rangle_1$ by

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

for all $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V if and only if S is injective.

- 3 (a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbf{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbf{R}^2$.
 (b) Show that the function taking an ordered pair $((x_1, x_2, x_3), (y_1, y_2, y_3))$ of elements of $\mathbf{R}^3$ to $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbf{R}^3$.
- 4 Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is injective.

5 Suppose V is a real inner product space.

- Show that $\langle u + v, u - v \rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.
- Show that if $u, v \in V$ have the same norm, then $u + v$ is orthogonal to $u - v$.
- Use (b) to show that the diagonals of a rhombus are perpendicular to each other.

6 Suppose $u, v \in V$. Prove that $\langle u, v \rangle = 0 \iff \|u\| \leq \|u + av\|$ 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 $a, b, c, x, y \in \mathbf{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \leq 1$. Prove that $a + b + c + 4x + 9y \leq 10$.

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

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

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

11 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 $(1, 2) = u + v$.

12 Suppose a, b, c, d are positive numbers.

- Prove that $(a + b + c + d) \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \right) \geq 16$.
- For which positive numbers a, b, c, d is the inequality above an equality?

13 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 $a_1^2, \dots, a_n^2$.

14 Suppose $v \in V$ and $v \neq 0$. Prove that $v/\|v\|$ is the unique closest element on the unit sphere of V to v . More precisely, prove that if $u \in V$ and $\|u\| = 1$, then

$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\|,$$

with equality only if $u = v/\|v\|$.

15 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: Use the law of cosines on the triangle formed by u , v , and $u - v$.

- 16 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 Exercise 15. Explain why the Cauchy–Schwarz inequality is needed to show that this definition makes sense.

- 17 Prove that

$$\left(\sum_{k=1}^n a_k b_k \right)^2 \leq \left(\sum_{k=1}^n k a_k^2 \right) \left(\sum_{k=1}^n \frac{b_k^2}{k} \right)$$

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

- 18 (a) Suppose $f: [1, \infty) \rightarrow [0, \infty)$ is continuous. Show that

$$\left(\int_1^\infty f \right)^2 \leq \int_1^\infty x^2 (f(x))^2 dx.$$

(b) For which continuous functions $f: [1, \infty) \rightarrow [0, \infty)$ is the inequality in (a) an equality with both sides finite?

- 19 Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T , then

$$|\lambda|^2 \leq \sum_{j=1}^n \sum_{k=1}^n |\mathcal{M}(T)_{j,k}|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j , column k of the matrix of T with respect to the basis $v_1, \dots, v_n$.

- 20 Prove that if $u, v \in V$, then $|\|u\| - \|v\|| \leq \|u - v\|$.

*The inequality above is called the **reverse triangle inequality**. For the reverse triangle inequality when $V = \mathbf{C}$, see Exercise 2 in Chapter 4.*

- 21 Suppose $u, v \in V$ are such that

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

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

- 22 Show that if $u, v \in V$, then

$$\|u + v\| \|u - v\| \leq \|u\|^2 + \|v\|^2.$$

- 23 Suppose $v_1, \dots, v_m \in V$ are such that $\|v_k\| \leq 1$ for each $k = 1, \dots, m$. Show that there exist $a_1, \dots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \dots + a_m v_m\| \leq \sqrt{m}.$$

- 24** Prove or give a counterexample: If $\|\cdot\|$ is the norm associated with an inner product on $\mathbf{R}^2$, then there exists $(x, y) \in \mathbf{R}^2$ such that $\|(x, y)\| \neq \max\{|x|, |y|\}$.
- 25** 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 $p = 2$.

- 26** Suppose V is a real inner product space. Prove that

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

for all $u, v \in V$.

- 27** Suppose V is a complex inner product space. Prove that

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

for all $u, v \in V$.

- 28** A norm on a vector space U is a function

$$\|\cdot\|: U \rightarrow [0, \infty)$$

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 $\|\cdot\|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle \cdot, \cdot \rangle$ on U such that $\|u\| = \langle u, u \rangle^{1/2}$ for all $u \in U$).

- 29** 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, for each $k = 1, \dots, m$, the inner product $\langle u_k, v_k \rangle$ denotes the inner product on V_k . Each of the spaces $V_1, \dots, V_m$ may have a different inner product, even though the same notation is used here.

- 30** Suppose V is a real inner product space. For $u, v, w, x \in V$, define

$$\langle u + iv, w + ix \rangle_{\mathbf{C}} = \langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i.$$

- (a) Show that $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ makes $V_{\mathbf{C}}$ into a complex inner product space.
 (b) Show that if $u, v \in V$, then

$$\langle u, v \rangle_{\mathbf{C}} = \langle u, v \rangle \quad \text{and} \quad \|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

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

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

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

$$\|w - u\| \leq \|w - x\|$$

for all $x \in E$.

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

(a) Show that

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

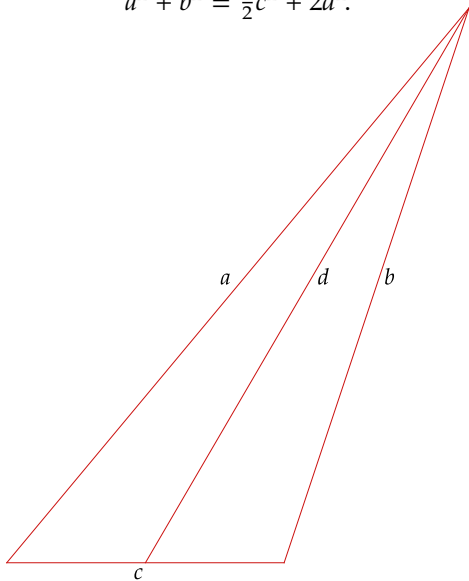
(b) Suppose c is a positive number and $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t), f(t) \rangle = 0$ for every $t \in \mathbf{R}$.

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

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

- 34 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

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



- 35 Fix a positive integer n . The *Laplacian* Δp of a twice differentiable real-valued 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 (with coefficients in $\mathbf{R}$) 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

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

and then showing that T is injective and hence surjective.

In realms of numbers, where the secrets lie,
A noble truth emerges from the deep,
Cauchy and Schwarz, their wisdom they apply,
An inequality for all to keep.

Two vectors, by this bond, are intertwined,
As inner products weave a gilded thread,
Their magnitude, by providence, confined,
A bound to which their destiny is wed.

Though shadows fall, and twilight dims the day,
This inequality will stand the test,
To guide us in our quest, to light the way,
And in its truth, our understanding rest.

So sing, ye muses, of this noble feat,
Cauchy–Schwarz, the bound that none can beat.

—written by ChatGPT with input *Shakespearean sonnet on Cauchy–Schwarz inequality*