

# 7

## Inner Products

Copyright © 2022, Gregory G. Smith  
Last updated: 2022-03-13

Compared to the coordinate space  $\mathbb{R}^n$ , an abstract vector space lacks geometry. For example, it does not have the concept of the length of a vector. To regain some geometry, this chapter introduces an additional structure that associates a scalar to each pair of vectors.

### 7.0 Inner Products

HOW DO WE CONCOCT AN ANALOGUE OF THE DOT PRODUCT? The desired structure builds on the properties of the dot product.

**7.0.0 Definition.** An *inner product* on  $\mathbb{C}$ -vector space  $V$  is a function from  $V \times V$  to  $\mathbb{C}$ , sending the pair  $(v, w)$  of vectors in  $V \times V$  to the scalar  $\langle v, w \rangle$  in  $\mathbb{C}$ , that satisfies the following four properties:

The axiomatic definition of an inner product was first given by [Giuseppe Peano](#) in 1898.

(linearity)	$\langle c\mathbf{u} + d\mathbf{v}, \mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{w} \rangle + d\langle \mathbf{v}, \mathbf{w} \rangle$	for all $\mathbf{u}, \mathbf{v}, \mathbf{w}$ in $V$ and all $c, d$ in $\mathbb{C}$ .
(conjugate-symmetry)	$\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$	for all $\mathbf{v}, \mathbf{w}$ in $V$ .
(nonnegativity)	$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$	for all $\mathbf{v}$ in $V$ .
(positivity)	$\langle \mathbf{v}, \mathbf{v} \rangle = 0$ if and only if $\mathbf{v} = \mathbf{0}$ .	

An *inner product space* is a vector space together with a specified inner product.

#### 7.0.1 Remarks.

- For any fixed vector  $w$  in  $V$ , linearity asserts that the map from  $V$  to  $\mathbb{C}$  defined by  $v \mapsto \langle v, w \rangle$  is linear.
- Conjugate-symmetry guarantees that, for all vectors  $v$  in  $V$ , we have  $\langle v, v \rangle = \overline{\langle v, v \rangle}$  or equivalently  $\langle v, v \rangle \in \mathbb{R}$ , so the nonnegative property is well-defined.
- When working over  $\mathbb{R}$ -vector spaces, conjugate-symmetry is just symmetry:  $\langle v, w \rangle = \langle w, v \rangle$  for all  $v$  and  $w$  in  $V$ .

Conventions differ as to which argument should be linear. In quantum mechanics and mathematical physics, one traditionally defines linearity in the second argument.

**7.0.2 Notation.** For any nonnegative integers  $m$  and  $n$ , the conjugate-transpose of a complex  $(m \times n)$ -matrix  $\mathbf{A}$  is the  $(n \times m)$ -matrix  $\mathbf{A}^* := \overline{\mathbf{A}^T} = (\overline{\mathbf{A}})^T$ .

Since conjugation and taking the transpose are both involutions, we have  $(\mathbf{A}^*)^* = \mathbf{A}$  for all  $\mathbf{A}$ .

**7.0.3 Problem.** Let  $n$  be a positive integer and let  $\mathbf{A}$  be an invertible complex  $(n \times n)$ -matrix. Show that  $\langle v, w \rangle := w^* \mathbf{A}^* \mathbf{A} v$  defines an inner product on the coordinate space  $\mathbb{C}^n$ .

*Solution.* For all vectors  $u, v, w$  in  $\mathbb{C}^n$  and all scalars  $c, d$  in  $\mathbb{C}$ , the linearity of matrix multiplication gives

$$\begin{aligned}\langle cu + dv, w \rangle &= w^* \mathbf{A}^* \mathbf{A} (cu + dv) \\ &= c(w^* \mathbf{A}^* \mathbf{A} u) + d(w^* \mathbf{A}^* \mathbf{A} v) = c \langle u, w \rangle + d \langle v, w \rangle,\end{aligned}$$

establishing the linearity property. The properties of the transpose and conjugation also give

$$\begin{aligned}\langle v, w \rangle &= w^* \mathbf{A}^* \mathbf{A} v = (w^* \mathbf{A}^* \mathbf{A} v)^\top = v^\top \mathbf{A}^\top (\mathbf{A}^*)^\top (w^*)^\top \\ &= \overline{v^* \mathbf{A}^* (\mathbf{A}^*)^* (w^*)^*} = \overline{v^* \mathbf{A}^* \mathbf{A} w} = \overline{\langle w, v \rangle},\end{aligned}$$

establishing conjugate-symmetry. Setting  $w := \mathbf{A} v$ , we obtain

$$\begin{aligned}\langle v, v \rangle &= v^* \mathbf{A}^* \mathbf{A} v = (\mathbf{A} v)^* (\mathbf{A} v) = w^* w \\ &= \overline{w_1} w_1 + \overline{w_2} w_2 + \cdots + \overline{w_n} w_n = |w_1|^2 + |w_2|^2 + \cdots + |w_n|^2.\end{aligned}$$

We deduce the nonnegativity property, because the absolute value of any nonzero complex number is a positive real number. Since  $\mathbf{A}$  is invertible, we see that  $w = \mathbf{A} v = \mathbf{0}$  if and only if  $v = \mathbf{0}$ , which establishes the positivity property.  $\square$

When the invertible matrix is just the identity matrix, we obtain the canonical inner product on the coordinate space  $\mathbb{C}^n$ . Similar constructions give the canonical inner products on the vector space of matrices and the vector space of continuous functions.

**7.0.4 Definition.** For any nonnegative integer  $n$ , the *standard inner product* on  $\mathbb{C}^n$  is defined, for all  $v$  and  $w$  in  $\mathbb{C}^n$ , by

$$\langle v, w \rangle := w^* v = \sum_{k=1}^n \overline{w_k} v_k = \overline{w_1} v_1 + \overline{w_2} v_2 + \cdots + \overline{w_n} v_n.$$

**7.0.5 Definition.** For any two nonnegative integers  $m$  and  $n$ , the *Frobenius inner product* on the  $\mathbb{C}$ -vector space  $\mathbb{C}^{m \times n}$  of all complex  $(m \times n)$ -matrices is defined, for all matrices  $\mathbf{A}$  and  $\mathbf{B}$  in  $\mathbb{C}^{m \times n}$ , by

$$\langle \mathbf{A}, \mathbf{B} \rangle_{\text{F}} := \text{tr}(\mathbf{B}^* \mathbf{A}) = \sum_{j=1}^m \sum_{k=1}^n \overline{b_{j,k}} a_{j,k}.$$

If the matrices are converted into column vectors, then this inner product coincides with the standard inner product.

**7.0.6 Definition.** The *canonical inner product* on the  $\mathbb{R}$ -vector space of all real-valued continuous functions on the interval  $[a, b] \subset \mathbb{R}$  is defined, for all functions  $f$  and  $g$ , by  $\langle f, g \rangle := \int_a^b f(t) g(t) dt$ .

This inner product is named after [Georg Frobenius](#), seemingly inspired by his 1909 work on nonnegative matrices.

**7.0.7 Remarks.** For any continuous functions  $f, g, h$  on the interval  $[a, b]$  and any scalars  $c, e$  in  $\mathbb{R}$ , the properties of definite integrals give

$$\begin{aligned}\langle cf + eg, h \rangle &= \int_a^b (cf + eg)(t)h(t) dt = c \int_a^b f(t)h(t) dt + e \int_a^b g(t)h(t) dt = c \langle f, h \rangle + e \langle g, h \rangle, \\ \langle f, g \rangle &= \int_a^b f(t)g(t) dt = \int_a^b g(t)f(t) dt = \langle g, f \rangle, \\ \langle f, f \rangle &= \int_a^b f(t)f(t) dt = \int_a^b |f(t)|^2 dt \geq 0.\end{aligned}$$

For any continuous nonnegative function  $h$  such that  $\int_a^b h(t) dt = 0$ , it follows that  $h(t) = 0$  for all  $t \in [a, b]$ . Hence, the equation  $\langle f, f \rangle = 0$  implies that  $|f(t)|^2 = 0$  and  $f(t) = 0$  for all  $t \in [a, b]$ .

**7.0.8 Definition.** Two vectors  $v$  and  $w$  in an inner product space are *orthogonal* if  $\langle v, w \rangle = 0$ .

**7.0.9 Lemma** (Properties of inner products). *Let  $V$  be a complex inner product space.*

- i. For all vectors  $v$  in  $V$ , we have  $\langle \mathbf{0}, v \rangle = \mathbf{0} = \langle v, \mathbf{0} \rangle$ .
- i. For all vectors  $u, v, w$  in  $V$  and all scalars  $c, d$  in  $\mathbb{C}$ , we have

$$\langle u, cv + dw \rangle = \bar{c} \langle u, v \rangle + \bar{d} \langle u, w \rangle.$$

The zero vector is orthogonal to every vector. Moreover, the positivity of inner products implies that the zero vector is the only vector orthogonal to itself.

*Proof.* The properties of linear maps [3.0.3] and the linearity of inner products imply that  $\langle \mathbf{0}, v \rangle = \mathbf{0}$  and conjugate-symmetry implies that  $\langle v, \mathbf{0} \rangle = \langle \mathbf{0}, v \rangle = \mathbf{0}$ . For all vectors  $u, v, w$  in  $V$  and all scalars  $c, d$  in  $\mathbb{C}$ , the same two properties also give

$$\begin{aligned}\langle u, cv + dw \rangle &= \overline{\langle cv + dw, u \rangle} \\ &= \bar{c} \overline{\langle v, u \rangle} + \bar{d} \overline{\langle w, u \rangle} = \bar{c} \langle u, v \rangle + \bar{d} \langle u, w \rangle. \quad \square\end{aligned}$$

### Exercises

**7.0.10 Problem.** Determine which of the following statements are true. If a statement is false, then provide a counterexample.

- i. The inner product of any two vectors is a real number.
- ii. Inner products are not defined when the base field is  $\mathbb{R}$ .
- iii. The  $\mathbb{C}$ -vector space  $\mathbb{C}^n$  has only one inner product.
- iv. If  $\mathbb{C}^1$  is equipped with the standard inner product, then every nonzero vector is orthogonal to a unique vector.

**7.0.11 Problem.** Let  $V$  be a complex inner product space. For all vectors  $v$  and  $w$  in  $V$ , prove the following identities:

$$\begin{aligned}(\text{polar identity}) \quad \langle v, w \rangle &= \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2 + i \|v + iw\|^2 - i \|v - iw\|^2) \\ (\text{parallelogram identity}) \quad \|v + w\|^2 + \|v - w\|^2 &= 2 (\|v\|^2 + \|w\|^2)\end{aligned}$$

## 7.1 Norms

HOW DO WE MEASURE LENGTH IN AN INNER PRODUCT SPACE? An inner product induces a notion of length.

**7.1.0 Definition.** Let  $V$  be an inner product space. For any vector  $v$  in  $V$ , its *norm* (also known as its *length*) is  $\|v\| := \sqrt{\langle v, v \rangle} \in \mathbb{R}$ .

To define a normed linear space, one must assume that the underlying field of scalars is either  $\mathbb{R}$  or  $\mathbb{C}$ .

**7.1.1 Lemma** (Properties of norms). *Let  $V$  be an inner product space.*

- i. We have  $\|v\| = 0$  if and only if  $v = \mathbf{0}$ .
- ii. For all vectors  $v$  in  $V$  and all scalars  $c$ , we have  $\|c v\| = |c| \|v\|$ .

*Proof.*

- i. The positivity of an inner product [7.0.0] asserts that  $\langle v, v \rangle = 0$  if and only if  $v = \mathbf{0}$ . Since the real number 0 has a unique square root, we have  $\|v\| = 0$  if and only if  $v = \mathbf{0}$ .
- ii. Linearity and conjugate-symmetry of an inner product [7.0.0] show that  $\|c v\|^2 = \langle c v, c v \rangle = (c \bar{c}) \langle v, v \rangle = |c|^2 \|v\|^2$ . Since the norm and the absolute value of any scalar are a nonnegative real numbers, taking square roots gives  $\|c v\| = |c| \|v\|$ .  $\square$

Since the square root symbol represents the unique nonnegative square root, we have  $|c| = \sqrt{c^2}$  for all scalars  $c$  in  $\mathbb{R}$ .

**7.1.2 Theorem** (Pythagorean). *For any two orthogonal vectors  $v$  and  $w$  in an inner product space, we have  $\|v + w\|^2 = \|v\|^2 + \|w\|^2$ .*

*Proof.* Since orthogonality [7.0.8] implies that  $\langle v, w \rangle = 0$ , the linearity and conjugate-symmetry of an inner product [7.0.0] establish that

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\ &= \|v\|^2 + \langle v, w \rangle + \overline{\langle v, w \rangle} + \|w\|^2 = \|v\|^2 + \|w\|^2. \quad \square \end{aligned}$$

**7.1.3 Definition.** A set of vectors is *orthonormal* if the vectors are pairwise orthogonal and each vector has norm 1.

**7.1.4 Problem.** Let  $V := C([-1, 1])$  be the real inner product space consisting of continuous functions on the interval  $[-1, 1] \subset \mathbb{R}$  where  $\langle f, g \rangle := \int_{-1}^1 f(x) g(x) dx$ . Demonstrate that the three polynomial functions  $\frac{1}{\sqrt{2}}$ ,  $\frac{\sqrt{3}}{\sqrt{2}} x$ , and  $\frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1)$  form an orthonormal set.

*Solution.* Since

$$\begin{aligned} \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle &= \int_{-1}^1 \frac{1}{2} dx = \frac{1}{2}(1 - (-1)) = 1, \\ \left\langle \frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}} x \right\rangle &= \int_{-1}^1 \frac{\sqrt{3}}{2} x dx = \frac{\sqrt{3}}{2} \left[ \frac{1}{2} x^2 \right]_{x=-1}^{x=1} = 0, \\ \left\langle \frac{1}{\sqrt{2}}, \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1) \right\rangle &= \int_{-1}^1 \frac{\sqrt{5}}{4}(3x^2 - 1) dx = \frac{\sqrt{5}}{4} \left[ x^3 - x \right]_{x=-1}^{x=1} = \frac{\sqrt{5}}{4} \left[ x(x-1)(x+1) \right]_{x=-1}^{x=1} = 0, \\ \left\langle \frac{\sqrt{3}}{\sqrt{2}} x, \frac{\sqrt{3}}{\sqrt{2}} x \right\rangle &= \int_{-1}^1 \frac{3}{2} x^2 dx = \frac{3}{2} \left[ \frac{1}{3} x^3 \right]_{x=-1}^{x=1} = \frac{1}{2}(1 - (-1)) = 1, \\ \left\langle \frac{\sqrt{3}}{\sqrt{2}} x, \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1) \right\rangle &= \int_{-1}^1 \frac{\sqrt{15}}{4}(3x^3 - x) dx = \frac{\sqrt{15}}{4} \left[ \frac{3}{4} x^4 - \frac{1}{2} x^2 \right]_{x=-1}^{x=1} = 0, \\ \left\langle \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1), \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1) \right\rangle &= \int_{-1}^1 \frac{5}{8}(3x^2 - 1)^2 dx = \frac{5}{8} \left[ \frac{9}{5} x^5 - 2x^3 + x \right]_{x=-1}^{x=1} = \frac{1}{4}(9 - 10 + 5) = 1, \end{aligned}$$

these functions form an orthonormal set.  $\square$

**7.1.5 Corollary** (Parseval identity). *Let  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  be an orthonormal set of vectors in an inner product space. For all scalars  $c_1, c_2, \dots, c_m$ , we have  $\|c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_m \mathbf{u}_m\|^2 = |c_1|^2 + |c_2|^2 + \dots + |c_m|^2$ .*

This result is named after [Marc-Antoine Parseval](#) who described a similar result in an infinite-dimensional vector space.

*Induction proof.* We proceed by induction on the cardinality  $m$  of the orthonormal set. The base case  $m = 0$  is true because the empty sum of vectors is  $\mathbf{0}$ ,  $\|\mathbf{0}\| = 0$ , and the empty sum of scalars is 0. Suppose that  $m > 0$ . Since  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$  is an orthonormal set, the properties of an inner product [7.0.9] give

$$\begin{aligned} \langle c_1 \mathbf{u}_1, c_2 \mathbf{u}_2 + c_3 \mathbf{u}_3 + \dots + c_m \mathbf{u}_m \rangle &= c_1 \langle \mathbf{u}_1, c_2 \mathbf{u}_2 + c_3 \mathbf{u}_3 + \dots + c_m \mathbf{u}_m \rangle \\ &= c_1 \overline{c_2} \langle \mathbf{u}_1, \mathbf{u}_2 \rangle + c_1 \overline{c_3} \langle \mathbf{u}_1, \mathbf{u}_3 \rangle + \dots + c_1 \overline{c_m} \langle \mathbf{u}_1, \mathbf{u}_m \rangle \\ &= c_1 \overline{c_2} 0 + c_1 \overline{c_3} 0 + \dots + c_1 \overline{c_m} 0 = 0. \end{aligned}$$

Hence, the Pythagorean theorem [7.1.2], induction hypothesis, and properties of a norm [7.1.1] give

$$\begin{aligned} \|c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_m \mathbf{u}_m\|^2 &= \|c_1 \mathbf{u}_1\|^2 + \|c_2 \mathbf{u}_2 + c_3 \mathbf{u}_3 + \dots + c_m \mathbf{u}_m\|^2 \\ &= |c_1|^2 \|\mathbf{u}_1\|^2 + |c_2|^2 + |c_3|^2 + \dots + |c_m|^2 \\ &= |c_1|^2 + |c_2|^2 + \dots + |c_m|^2 \end{aligned}$$

because  $c_1 \mathbf{u}_1$  and  $c_2 \mathbf{u}_2 + c_3 \mathbf{u}_3 + \dots + c_m \mathbf{u}_m$  are orthogonal.  $\square$

**7.1.6 Corollary.** *An orthonormal set of vectors is linearly independent.*

*Proof.* Let  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  be an orthonormal set of vectors in an inner product space. Suppose that there are scalars  $c_1, c_2, \dots, c_m$  such that  $c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_m \mathbf{u}_m = \mathbf{0}$ . The Parseval identity [7.1.5] implies that  $0 = \|c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_m \mathbf{u}_m\|^2 = |c_1|^2 + |c_2|^2 + \dots + |c_m|^2$ . For all  $1 \leq k \leq m$ , the real number  $|c_k|^2$  is nonnegative, so we deduce that  $|c_1| = |c_2| = \dots = |c_m| = 0$  which implies that  $c_1 = c_2 = \dots = c_m = 0$ . Therefore, the vectors  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$  are linearly independent.  $\square$

**7.1.7 Corollary** (Orthonormal coordinates). *Let  $\mathcal{U} := (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$  be an ordered orthonormal basis for an inner product space  $V$ . For any vector  $\mathbf{v}$  in  $V$ , we have  $\mathbf{v} = \langle \mathbf{v}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{v}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \dots + \langle \mathbf{v}, \mathbf{u}_n \rangle \mathbf{u}_n$ . Equivalently, the coordinate vector of  $\mathbf{v}$  relative to the orthonormal basis  $\mathcal{U}$  is*

$$(\mathbf{v})_{\mathcal{U}} = [\langle \mathbf{v}, \mathbf{u}_1 \rangle \quad \langle \mathbf{v}, \mathbf{u}_2 \rangle \quad \dots \quad \langle \mathbf{v}, \mathbf{u}_n \rangle]^T.$$

*Proof.* Since  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  is a basis for the vector space  $V$ , there exists scalars  $c_1, c_2, \dots, c_n$  such that  $\mathbf{v} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n$ . For each  $1 \leq k \leq n$ , the linearity of an inner product [7.0.0] gives

$$\begin{aligned} \langle \mathbf{v}, \mathbf{u}_k \rangle &= \langle c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n, \mathbf{u}_k \rangle \\ &= c_1 \langle \mathbf{u}_1, \mathbf{u}_k \rangle + c_2 \langle \mathbf{u}_2, \mathbf{u}_k \rangle + \dots + c_{k-1} \langle \mathbf{u}_{k-1}, \mathbf{u}_k \rangle + c_k \langle \mathbf{u}_k, \mathbf{u}_k \rangle + c_{k+1} \langle \mathbf{u}_{k+1}, \mathbf{u}_k \rangle + \dots + c_n \langle \mathbf{u}_n, \mathbf{u}_k \rangle \\ &= c_1 0 + c_2 0 + \dots + c_{k-1} 0 + c_k 1 + c_{k+1} 0 + \dots + c_n 0 = c_k, \end{aligned}$$

because the vectors  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$  are orthonormal.  $\square$