

## Chapter 6

# Inner Product Spaces

### 6.A Inner Products and Norms

9/30: • **Norm** (of  $x \in \mathbb{R}^n$ ): The length of  $x$ . Denoted by  $\|x\|$ . Given by

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

• **Dot product** (of  $x, y \in \mathbb{R}^n$ ): The quantity

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

• Properties of the dot product:

- $x \cdot x = \|x\|^2$ .
- $x \cdot x \geq 0$ .
- $x \cdot x = 0$  iff  $x = 0$ .
- Let  $y \in \mathbb{R}^n$ . Then  $T : \mathbb{R}^n \rightarrow \mathbb{R}$  defined by  $Tx = x \cdot y$  is linear.
- $x \cdot y = y \cdot x$ .

• **Norm** (of  $z \in \mathbb{C}^n$ ): The quantity

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

– Note that  $\|z\|^2 = z \cdot \bar{z}$ .

• **Inner product** (on  $V$ ): A function that takes each ordered pair  $(u, v) \in V \times V$  to a number  $\langle u, v \rangle \in \mathbb{F}$  and has the following properties.

**positivity**

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

**definiteness**

$$\langle v, v \rangle = 0 \text{ iff } v = 0.$$

**additivity in first slot**

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

**homogeneity in first slot**

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

**conjugate symmetry**

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

- Since every real number equals its complex conjugate, if  $V$  is real, we can dispense with the conjugacy condition in the conjugate symmetry condition and just have  $\langle u, v \rangle = \langle v, u \rangle$ .
- “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” (Axler, 2015, p. 166).

- **Euclidean inner product** (on  $\mathbb{F}^n$ ): The function defined by

$$\langle w, z \rangle = w_1 \bar{z}_1 + \cdots + w_n \bar{z}_n$$

- **Inner product space:** A vector space  $V$  along with an inner product on  $V$ .

- When  $\mathbb{F}^n$  is referred to as an inner product space, assume that the inner product is the Euclidean inner product unless explicitly stated otherwise.

- Basic properties of an inner product.

**Theorem 6.1.**

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

*Proof.* Let  $u \in V$  be arbitrary, and let  $T : V \rightarrow \mathbb{F}$  be defined by  $Tv = \langle v, u \rangle$ . Let  $v, w \in V$  be arbitrary, and let  $\lambda \in \mathbb{F}$ . Then

$$\begin{aligned} T(v+w) &= \langle v+w, u \rangle & T(\lambda v) &= \langle \lambda v, u \rangle \\ &= \langle v, u \rangle + \langle w, u \rangle & &= \lambda \langle v, u \rangle \\ &= Tv + Tw & &= \lambda Tv \end{aligned}$$

as desired. ■

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

*Proof.* Let  $u \in V$  be arbitrary. Since  $T$  as defined above is linear, Theorem 3.2 implies that  $0 = T(0) = \langle 0, u \rangle$ , as desired. ■

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

*Proof.* Let  $u \in V$  be arbitrary. By the conjugate symmetry property and the above,  $\langle u, 0 \rangle = \overline{\langle 0, u \rangle} = \bar{0} = 0$ , as desired. ■

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

*Proof.* Let  $u, v, w \in V$  be arbitrary. Then

$$\begin{aligned} \langle u, v+w \rangle &= \overline{\langle v+w, u \rangle} \\ &= \overline{\langle v, u \rangle + \langle w, u \rangle} \\ &= \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle} && \text{Theorem 4.1} \\ &= \langle u, v \rangle + \langle u, w \rangle \end{aligned}$$

as desired. ■

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

*Proof.* Let  $u, v \in V$ , and let  $\lambda \in \mathbb{F}$ . Then

$$\begin{aligned} \langle u, \lambda v \rangle &= \overline{\langle \lambda v, u \rangle} \\ &= \overline{\lambda \langle v, u \rangle} \\ &= \bar{\lambda} \overline{\langle v, u \rangle} && \text{Theorem 4.1} \\ &= \bar{\lambda} \langle u, v \rangle \end{aligned}$$

- **Norm** (of  $v \in V$ ): The quantity

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

- Basic properties of the norm.

**Theorem 6.2.** *Suppose  $v \in V$ . Then*

- (a)  $\|v\| = 0$  iff  $v = 0$ .

*Proof.* Suppose first that  $\|v\| = 0$ . Then  $0 = \sqrt{\langle v, v \rangle} = \langle v, v \rangle$ . Thus,  $v = 0$ . The proof is symmetric in the reverse direction. ■

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

*Proof.* Let  $\lambda \in \mathbb{F}$  be arbitrary. Then

$$\begin{aligned} \|\lambda v\|^2 &= \langle \lambda v, \lambda v \rangle \\ &= \lambda \bar{\lambda} \langle v, v \rangle \\ &= |\lambda|^2 \|v\|^2 \end{aligned} \quad \text{Theorem 4.1}$$

Taking square roots of the above gives the desired equality.<sup>[1]</sup> ■

- **Orthogonal** (vectors  $u, v \in V$ ): Two vectors  $u, v \in V$  such that  $\langle u, v \rangle = 0$ .<sup>[2]</sup>
- If  $u, v \in \mathbb{R}^2$  are nonzero, then  $\langle u, v \rangle = \|u\| \|v\| \cos \theta$ .
  - “Thus, two vectors in  $\mathbb{R}^2$  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*.” (Axler, 2015, p. 169).
- Orthogonality and zero.

**Theorem 6.3.**

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

*Proof.* Let  $u \in V$  be arbitrary. Then Theorem 6.1b implies that  $\langle 0, u \rangle = 0$ . Thus,  $u$  and  $0$  are orthogonal, as desired. ■

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

*Proof.* Let  $v \in V$  be such that  $v$  is orthogonal to itself. Then  $\langle v, v \rangle = 0$ . But by the property of definiteness, it follows that  $v = 0$ , as desired. ■

- The special case where  $V = \mathbb{R}^2$  of the following is over 2500 years old, although the following is not the original proof.

**Theorem 6.4** (Pythagorean Theorem). *Suppose  $u, v \in V$  are orthogonal. Then*

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

*Proof.* We have

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= (\langle u, u \rangle + \langle u, v \rangle) + (\langle v, u \rangle + \langle v, v \rangle) \\ &= \|u\|^2 + 0 + 0 + \|v\|^2 \\ &= \|u\|^2 + \|v\|^2 \end{aligned}$$

as desired. ■

<sup>1</sup>Notice this technique: Working with norms squared is usually easier than working directly with norms.

<sup>2</sup>The word *orthogonal* derives from the Greek word *orthogonios*, which means right-angled.

- Note that we can prove the converse of the Pythagorean Theorem in *real* inner product spaces as follows: If  $\|u + v\|^2 = \|u\|^2 + \|v\|^2$  for  $u, v$  real, then

$$\begin{aligned}
 \langle u, u \rangle + \langle v, v \rangle &= \langle u + v, u + v \rangle \\
 \langle u, u \rangle + \langle v, v \rangle &= \langle u, u \rangle + \langle v, u \rangle + \langle u, v \rangle + \langle v, v \rangle \\
 0 &= \langle u, v \rangle + \langle v, u \rangle \\
 &= \langle u, v \rangle + \overline{\langle u, v \rangle} \\
 &= 2 \operatorname{Re} \langle u, v \rangle \\
 &= 2 \langle u, v \rangle \\
 &= \langle u, v \rangle
 \end{aligned}$$

where the first equality holds by the definition of the norm, the second by the additivity of the inner product and Theorem 6.1d, the fifth by Theorem 4.1, and the sixth by the hypothesis that the inner product space is real, as desired.

- Let  $u, v \in V$  with  $v \neq 0$ . We are now equipped to consider how to write  $u$  as the sum of  $v$  plus a vector  $w$  orthogonal to  $v$ , as in Figure 6.1 below.

**Theorem 6.5.** 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  $\langle w, v \rangle = 0$  and  $u = cv + w$ .

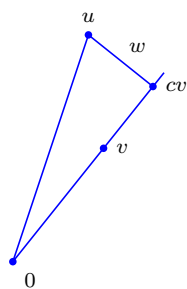


Figure 6.1: An orthogonal decomposition.

*Proof.* We want to write  $u$  in the form  $u = cv + w$  where  $w$  is orthogonal to  $v$ . We know that

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

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

$$\begin{aligned}
 0 &= \langle u - cv, v \rangle \\
 &= \langle u, v \rangle + \langle -cv, v \rangle \\
 &= \langle u, v \rangle - c \langle v, v \rangle \\
 &= \langle u, v \rangle - c \|v\|^2 \\
 c &= \frac{\langle u, v \rangle}{\|v\|^2}
 \end{aligned}$$

But this gives the values we want for  $c$  and  $w$ , as desired. ■

- This allows for the proof of a very important result.

**Theorem 6.6** (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.* We divide into two cases ( $v = 0$  and  $v \neq 0$ ). If  $v = 0$ , then

$$|\langle u, v \rangle| = 0 \leq 0 = \|u\|\sqrt{0} = \|u\|\sqrt{\langle u, v \rangle} = \|u\|\|v\|$$

and we also have that the equality holds since  $v = 0 = 0u$ ,  $0 \in \mathbb{F}$ . Now let  $v \neq 0$ . Then by Theorem 6.5,

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

where  $\langle v, w \rangle = 0$ . It follows by the Pythagorean Theorem that

$$\begin{aligned} \|u\|^2 &= \left\| \frac{\langle u, v \rangle}{\|v\|^2}v \right\|^2 + \|w\|^2 \\ &= \frac{|\langle u, v \rangle|^2}{\|v\|^4} \|v\|^2 + \|w\|^2 && \text{Theorem 6.2b} \\ &= \frac{|\langle u, v \rangle|^2}{\|v\|^2} + \|w\|^2 \\ &\geq \frac{|\langle u, v \rangle|^2}{\|v\|^2} \end{aligned}$$

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

Also note that the Cauchy-Schwarz inequality is an equality iff the last line is an equality, which happens iff  $w = 0$ . But  $w = 0$  iff  $u$  is a scalar multiple of  $v$ , as desired. ■

- Note that the Cauchy-Schwarz is known as such because the French mathematician Augustin-Louis Cauchy proved the top inequality below in 1821, and the German mathematician Hermann Schwarz proved the bottom inequality below in 1886; both are special cases of the above.

$$\begin{aligned} |x_1 y_1 + \cdots + x_n y_n|^2 &\leq (x_1^2 + \cdots + x_n^2)(y_1^2 + \cdots + y_n^2) \\ \left| \int_{-1}^1 f(x)g(x) \, dx \right|^2 &\leq \left( \int_{-1}^1 (f(x))^2 \, dx \right) \left( \int_{-1}^1 (g(x))^2 \, dx \right) \end{aligned}$$

- For the top one, we let  $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$ .
- For the bottom one, we let  $f, g$  be continuous real-valued functions on  $[-1, 1]$ .

- We now prove another important inequality.

**Theorem 6.7** (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

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \langle v, u \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle + \langle u, v \rangle + \overline{\langle u, v \rangle} \\ &= \|u\|^2 + \|v\|^2 + 2 \operatorname{Re} \langle u, v \rangle \\ &\leq \|u\|^2 + \|v\|^2 + 2|\langle u, v \rangle| \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| && \text{Cauchy-Schwarz Inequality} \\ &= (\|u\| + \|v\|)^2 \end{aligned}$$

Taking square roots of both sides gives the desired inequality.

This inequality is an equality iff  $\langle u, v \rangle = \|u\|\|v\|$ . Now suppose  $u = cv$  where  $c \in \mathbb{F}$  is positive. Then

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

The proof is the same in the reverse direction. ■

- One last equality.

**Theorem 6.8** (Parallelogram Equality). *Suppose  $u, v \in V$ . Then*

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

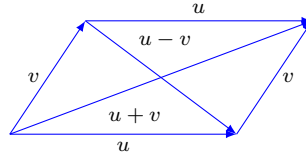


Figure 6.2: The parallelogram equality.

*Proof.* We have

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

as desired. ■

## Exercises

10/11: **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  $u, v \in V$ .

*Proof.* Let  $u, v \in V$  be arbitrary. Then

$$\begin{aligned}\frac{\|u + v\|^2 - \|u - v\|^2}{4} &= \frac{\langle u + v, u + v \rangle - \langle u - v, u - v \rangle}{4} \\ &= \frac{(\langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle) - (\langle u, u \rangle - \langle u, v \rangle - \langle v, u \rangle + \langle v, v \rangle)}{4} \\ &= \frac{2\langle u, v \rangle + 2\langle v, u \rangle}{4} \\ &= \frac{2\langle u, v \rangle + 2\langle u, v \rangle}{4} \\ &= \langle u, v \rangle\end{aligned}$$

as desired. ■

**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  $u, v \in V$ .

*Proof.* Let  $u, v \in V$  be arbitrary. Then

$$\begin{aligned} \frac{\|u + v\|^2 - \|u - v\|^2 + \|u + iv\|^2 i - \|u - iv\|^2 i}{4} &= \frac{\|u + v\|^2 - \|u - v\|^2}{4} + i \frac{\|u + iv\|^2 - \|u - iv\|^2}{4} \\ &= \frac{2\langle u, v \rangle + 2\langle v, u \rangle}{4} + i \frac{2\langle u, iv \rangle + 2\langle iv, u \rangle}{4} \\ &= \frac{2\langle u, v \rangle + 2\langle v, u \rangle + 2\langle u, v \rangle + 2\langle -v, u \rangle}{4} \\ &= \frac{4\langle u, v \rangle + 2\langle 0, u \rangle}{4} \\ &= \langle u, v \rangle \end{aligned}$$

as desired. ■

## 6.B Orthonormal Bases

10/3:

- **Orthonormal** (list of vectors): A list of vectors such that each vector in the list has norm 1 and is orthogonal to all the other vectors in the list.

– In other words, a list  $e_1, \dots, e_m$  is orthonormal if

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

- Orthonormal lists are particularly easy to work with.

**Theorem 6.9.** If  $e_1, \dots, e_m$  is an orthonormal list of vectors in  $V$ , then

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

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

*Proof.* We have that

$$\begin{aligned} \|a_1 e_1 + \dots + a_m e_m\|^2 &= \|a_1 e_1\|^2 + \dots + \|a_m e_m\|^2 && \text{Pythagorean Theorem} \\ &= |a_1|^2 \|e_1\|^2 + \dots + |a_m|^2 \|e_m\|^2 && \text{Theorem 6.2b} \\ &= |a_1|^2 + \dots + |a_m|^2 \end{aligned}$$

as desired. ■

- The next result directly follows from the previous one.

**Theorem 6.10.** Every orthonormal list of vectors is linearly independent.

*Proof.* Let  $e_1, \dots, e_m$  be an orthonormal list of vectors, and let  $a_1, \dots, a_m \in \mathbb{F}$  be such that  $a_1 e_1 + \dots + a_m e_m = 0$ . Then

$$\begin{aligned} 0 &= \|0\| \\ &= \|a_1 e_1 + \dots + a_m e_m\| \\ &= |a_1|^2 + \dots + |a_m|^2 && \text{Theorem 6.9} \end{aligned}$$

But this implies that each  $a_j = 0$ , as desired. ■

- **Orthonormal basis** (of  $V$ ): An orthonormal list of vectors in  $V$  that is also a basis of  $V$ .
- We now prove an easy condition for identifying orthonormal bases.

**Theorem 6.11.** *Every orthonormal list of vectors in  $V$  with length  $\dim V$  is an orthonormal basis of  $V$ .*

*Proof.* Let  $e_1, \dots, e_m$  be an orthonormal list of vectors in  $V$  with length  $\dim V$ . It follows by Theorem 6.10 that  $e_1, \dots, e_m$  is linearly independent. Therefore, by Theorem 2.12,  $e_1, \dots, e_m$  is a basis of  $V$ . ■

10/6:

- For an arbitrary basis  $e_1, \dots, e_n$  of  $V$ , it can be quite troublesome to compute that  $a_1, \dots, a_n \in \mathbb{F}$  that make  $a_1 e_1 + \dots + a_n e_n = v$  for an arbitrary  $v \in V$ .
- However, it is quite simple for an orthonormal basis:

**Theorem 6.12.** *Suppose  $e_1, \dots, e_n$  is an orthonormal basis of  $V$  and  $v \in V$ . Then*

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

and

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

*Proof.* Because  $e_1, \dots, e_n$  is a basis of  $V$ , we have that

$$v = a_1 e_1 + \dots + a_n e_n$$

for some  $a_1, \dots, a_n \in \mathbb{F}$ . It follows by taking inner products with an arbitrary  $e_j$  that

$$\begin{aligned} \langle v, e_j \rangle &= \langle a_1 e_1 + \dots + a_n e_n, e_j \rangle \\ &= a_1 \langle e_1, e_j \rangle + \dots + a_n \langle e_n, e_j \rangle \\ &= a_1 \cdot 0 + \dots + a_{j-1} \cdot 0 + a_j \cdot 1 + a_{j+1} \cdot 0 + \dots + a_n \cdot 0 \\ &= a_j \end{aligned}$$

for each  $j = 1, \dots, n$ .

The second equation follows from the first equation by Theorem 6.9. ■

- The following algorithm<sup>[3]</sup> gives a method for turning a linearly independent list into an orthonormal list with the same span as the original list.

**Theorem 6.13** (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*

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

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

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

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

*Proof.* We induct on  $j$ . For the base case  $j = 1$ , we have that  $e_1$  is trivially orthogonal to all other vectors in the list and that

$$\|e_1\| = \left\| \frac{v_1}{\|v_1\|} \right\| = \frac{1}{\|v_1\|} \|v_1\| = 1$$

---

<sup>3</sup>Danish mathematician Jørgen Gram and German mathematician Erhard Schmidt popularized this algorithm.



proving that  $e_1$  is an orthonormal list of vectors in  $V$ . Additionally, since  $e_1$  is a scalar multiple of  $v_1$ , we naturally also have that

$$\text{span}(v_1) = \text{span}(e_1)$$

Now suppose inductively that we have proven that  $e_1, \dots, e_j$  is an orthonormal list of vectors such that  $\text{span}(v_1, \dots, v_j) = \text{span}(e_1, \dots, e_j)$ . We now wish to prove the claim for  $e_{j+1}$ . To do so, it will suffice to show that  $e_{j+1}$  is well defined, that  $e_1, \dots, e_{j+1}$  is an orthonormal list of vectors, and that  $\text{span}(v_1, \dots, v_{j+1}) = \text{span}(e_1, \dots, e_{j+1})$ . Let's begin.

To confirm that the  $e_{j+1}$  is well defined, we must confirm that the norm in the bottom is nonzero. But since  $v_1, \dots, v_m$  is linearly independent by hypothesis, we have that  $v_{j+1} \notin \text{span}(v_1, \dots, v_j) = \text{span}(e_1, \dots, e_j)$ , meaning that we must have

$$\begin{aligned} v_{j+1} &\neq \langle v_{j+1}, e_1 \rangle e_1 + \dots + \langle v_{j+1}, e_j \rangle e_j \\ v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j &\neq 0 \\ \|v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j\| &\neq 0 \end{aligned} \quad \text{Theorem 6.2a}$$

as desired.

To confirm that  $e_1, \dots, e_{j+1}$  is an orthonormal list of vectors, it will suffice to show that  $e_{j+1}$  has norm 1 (the inductive hypothesis guarantees that  $e_1, \dots, e_j$  have norm 1) and that each vector in  $e_1, \dots, e_j$  is orthogonal to  $e_{j+1}$  (again, we already know by the inductive hypothesis that  $e_1, \dots, e_j$  is mutually orthogonal). We can address the first part with an argument symmetric to that used in the base case (indeed, it is easily seen that any vector divided by its norm has norm 1). As to the second part of the argument, we have for any  $1 \leq k < j+1$  that

$$\begin{aligned} \langle e_{j+1}, e_k \rangle &= \left\langle \frac{v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j}{\|v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j\|}, e_k \right\rangle \\ &= \frac{1}{\|v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j\|} (\langle v_{j+1}, e_k \rangle - \langle v_{j+1}, e_1 \rangle \langle e_1, e_k \rangle - \dots - \langle v_{j+1}, e_j \rangle \langle e_j, e_k \rangle) \\ &= \frac{1}{\|v_{j+1} - \langle v_{j+1}, e_1 \rangle e_1 - \dots - \langle v_{j+1}, e_j \rangle e_j\|} (\langle v_{j+1}, e_k \rangle - \langle v_{j+1}, e_k \rangle \cdot 1) \\ &= 0 \end{aligned}$$

as desired.

To confirm that  $\text{span}(v_1, \dots, v_{j+1}) = \text{span}(e_1, \dots, e_{j+1})$ , Exercise 2.C.1 tells us that it will suffice to show that  $\text{span}(v_1, \dots, v_{j+1}) \subset \text{span}(e_1, \dots, e_{j+1})$  and that  $\dim \text{span}(v_1, \dots, v_{j+1}) = \dim \text{span}(e_1, \dots, e_{j+1})$ . Since  $\text{span}(v_1, \dots, v_j) = \text{span}(e_1, \dots, e_j)$  by hypothesis, we have that  $v_1, \dots, v_j \in \text{span}(e_1, \dots, e_j) \subset \text{span}(e_1, \dots, e_{j+1})$ . Additionally, by the definition of  $e_{j+1}$ , we have that  $v_{j+1} \in \text{span}(e_1, \dots, e_{j+1})$ . Therefore, we have that

$$\text{span}(v_1, \dots, v_{j+1}) \subset \text{span}(e_1, \dots, e_{j+1})$$

Additionally, since  $v_1, \dots, v_{j+1}$  is linearly independent by hypothesis, and  $e_1, \dots, e_{j+1}$  is linearly independent by Theorem 6.10, we have that both subspaces have dimension  $j+1$ , as desired. ■

- Existence of an orthonormal basis.

**Theorem 6.14.** *Every finite-dimensional inner product space has an orthonormal basis.*

*Proof.* Let  $V$  be a finite-dimensional inner product space. Choose a basis of  $V$ . Applying the Gram-Schmidt Procedure to this (by definition linearly independent) basis yields an orthonormal list of vectors in  $V$  of length  $\dim V$ . Therefore, by Theorem 6.11, the orthonormal list is an orthonormal basis of  $V$ . ■

- Extending an orthonormal list to an orthonormal basis.

**Theorem 6.15.** *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 by Theorem 6.10,  $e_1, \dots, e_m$  is linearly independent. Hence this list can be extended to a basis  $e_1, \dots, e_m, v_1, \dots, v_n$  of  $V$  by Theorem 2.8. Applying the Gram-Schmidt Procedure to this basis yields an orthonormal list  $e_1, \dots, e_m, f_1, \dots, f_n$  of vectors in  $V$ . Note that the Gram-Schmidt Procedure doesn't change the first  $m$  vectors of the list: If  $1 \leq j \leq m$ , then

$$\begin{aligned} e'_j &= \frac{e_j - \langle e_j, e_1 \rangle e_1 - \dots - \langle e_j, e_{j-1} \rangle e_{j-1}}{\|e_j - \langle e_j, e_1 \rangle e_1 - \dots - \langle e_j, e_{j-1} \rangle e_{j-1}\|} \\ &= \frac{e_j}{\|e_j\|} \\ &= e_j \end{aligned}$$

Lastly, by Theorem 6.11, we know that  $e_1, \dots, e_m, f_1, \dots, f_n$  is a basis of  $V$ . ■

- The next result builds on Theorem 5.7.

**Theorem 6.16.** *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 the basis  $v_1, \dots, v_n$  of  $V$ . Then by Theorem 5.6,  $\text{span}(v_1, \dots, v_j)$  is invariant under  $T$  for each  $j = 1, \dots, n$ . Applying the Gram-Schmidt Procedure to  $v_1, \dots, v_n$ , we obtain an orthonormal basis  $e_1, \dots, e_n$  of  $V$  with the property that

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

for each  $j = 1, \dots, n$ . Combining the last two results, we have that  $\text{span}(e_1, \dots, e_j)$  is invariant under  $T$  for each  $j = 1, \dots, n$ . Therefore, by Theorem 5.6,  $T$  has an upper-triangular matrix with respect to the orthonormal basis  $e_1, \dots, e_n$ . ■

- We now prove an important application of the above.

**Theorem 6.17** (Schur's Theorem<sup>[4]</sup>). *Suppose  $V$  is a finite-dimensional complex vector space and  $T \in \mathcal{L}(V)$ . Then  $T$  has an upper-triangular matrix with respect to some orthonormal basis of  $V$ .*

*Proof.* By Theorem 5.7,  $T$  has an upper-triangular matrix with respect to some basis of  $V$ . Therefore, by Theorem 6.16,  $T$  has an upper-triangular matrix with respect to some orthonormal basis of  $V$ , as desired. ■

- We now show that every linear functional is equivalent to some inner product.

**Theorem 6.18** (Riesz Representation Theorem<sup>[5]</sup>). *Suppose  $V$  is finite-dimensional and  $\varphi$  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  $v \in V$ .*

<sup>4</sup>German mathematician Issai Schur published the first proof of this result in 1909.

<sup>5</sup>This result is named for Hungarian mathematician Frigyes Riesz, who proved several results in the early twentieth century very similar to this one.

*Proof.* We first find such a vector  $u$ . Let  $e_1, \dots, e_n$  be an orthonormal basis of  $V$ . Then

$$\begin{aligned}\varphi(v) &= \varphi(\langle v, e_1 \rangle e_1 + \dots + \langle v, e_n \rangle e_n) && \text{Theorem 6.12} \\ &= \langle v, e_1 \rangle \varphi(e_1) + \dots + \langle v, e_n \rangle \varphi(e_n) \\ &= \left\langle v, \overline{\varphi(e_1)} e_1 + \dots + \overline{\varphi(e_n)} e_n \right\rangle && \text{Theorem 6.1e}\end{aligned}$$

so we may take

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

Now suppose for the sake of contradiction that  $u_1, u_2 \in V$  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  $v \in V$ . In particular, if we let  $v = u_1 - u_2$ , we have that  $\langle u_1 - u_2, u_1 - u_2 \rangle = 0$ . Therefore, by the definiteness of the inner product,  $u_1 - u_2 = 0$ , i.e.,  $u_1 = u_2$ , a contradiction. ■

- Note that despite the fact that the expression for  $u$  derived in the proof of the Riesz Representation Theorem seems to depend on both  $\varphi$  and the chosen orthonormal basis, the statement of the theorem shows that the value of  $u$  depends only on  $\varphi$ , and is in truth the same for any orthonormal basis.

## 6.C Orthogonal Complements and Minimization Problems

- **Orthogonal complement** (of  $U \subset V$ ): The set of all vectors in  $V$  that are orthogonal to every vector in  $U$ . Denoted by  $U^\perp$ . Given by

$$U^\perp = \{v \in V : \langle v, u \rangle = 0 \ \forall \ u \in U\}$$

- Basic properties of the orthogonal complement.

**Theorem 6.19.**

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

*Proof.* Theorem 6.1b tells us that  $\langle 0, u \rangle = 0$  for every  $u \in U$ ; thus,  $0 \in U^\perp$ , as desired.

Let  $v, w \in U^\perp$ . Then

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

for all  $u \in U$ . Thus,  $v + w \in U^\perp$ , as desired.

Let  $v \in U^\perp$  and  $\lambda \in \mathbb{F}$ . Then

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

for all  $u \in U$ . Thus,  $\lambda v \in U^\perp$ , as desired. ■

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

*Proof.* Let  $v \in V$  be arbitrary. Then since  $\langle v, 0 \rangle = 0$  by Theorem 6.1c, we have that  $v \in \{0\}^\perp$ . Thus,  $\{0\}^\perp = V$ . ■

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

*Proof.* Suppose  $v \in V^\perp$ . Then  $\langle v, u \rangle = 0$  for all  $u \in V$ . In particular,  $\langle v, v \rangle = 0$ . But this implies that  $v = 0$  by the definiteness of the inner product. Therefore,  $V^\perp = \{0\}$ . ■

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

*Proof.* Let  $v \in U \cap U^\perp$ . Then  $\langle v, v \rangle = 0$ . Therefore,  $U \cap U^\perp \subset \{0\}$ . ■

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

*Proof.* Let  $v \in W^\perp$ . Then  $\langle v, u \rangle = 0$  for all  $u \in W$ . But since  $U \subset W$ , this implies that  $\langle v, u \rangle = 0$  for all  $u \in U$ . Thus,  $v \in U^\perp$ . Therefore,  $W^\perp \subset U^\perp$ . ■

- We can decompose  $V$  into the direct sum of any subspace  $U$  and another subspace.

**Theorem 6.20.** Suppose  $U$  is a finite-dimensional subspace of  $V$ . Then

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

*Proof.* First, we will show that any  $v \in V$  can be written as the sum of an element of  $U$  and an element of  $U^\perp$ . Then, we will show that this decomposition is unique. Let's begin.

Let  $v \in V$  be arbitrary, and let  $e_1, \dots, e_m$  be an orthonormal basis of  $U$ . Then we may let

$$v = \underbrace{\langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m}_u + \underbrace{v - \langle v, e_1 \rangle e_1 - \dots - \langle v, e_m \rangle e_m}_w$$

Since  $e_1, \dots, e_m$  is a basis of  $U$ ,  $u \in U$ . Additionally, we have that

$$\begin{aligned} \langle w, e_j \rangle &= \langle v - \langle v, e_1 \rangle e_1 - \dots - \langle v, e_m \rangle e_m, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_1 \rangle \langle e_1, e_j \rangle - \dots - \langle v, e_m \rangle \langle e_m, e_j \rangle \\ &= \langle v, e_j \rangle - \langle v, e_j \rangle \cdot 1 \\ &= 0 \end{aligned}$$

for all  $e_j$ . Thus,  $w$  is orthogonal to every vector in  $\text{span}(e_1, \dots, e_m) = U$ . It follows that  $w \in U^\perp$ . Therefore,  $v = u + w$  for  $u \in U$  and  $w \in U^\perp$ , as desired.

This combined with the fact that  $U \cap U^\perp = \{0\}$  (by Theorem 6.19b) implies that  $V = U \oplus U^\perp$ , as desired. ■

- Dimension of the orthogonal complement.

**Theorem 6.21.** Suppose  $V$  is finite-dimensional and  $U$  is a subspace of  $V$ . Then

$$\dim U^\perp = \dim V - \dim U$$

*Proof.* By Theorem 6.20, we have that  $V = U \oplus U^\perp$ . Therefore, by Theorem 3.21,

$$\begin{aligned} \dim V &= \dim U + \dim U^\perp \\ \dim U^\perp &= \dim V - \dim U \end{aligned}$$

as desired. ■

- We now have an important corollary to Theorem 6.20.

**Theorem 6.22.** Suppose  $U$  is a finite-dimensional subspace of  $V$ . Then

$$U = (U^\perp)^\perp$$

*Proof.* Suppose first that  $u \in U$ . Then  $\langle u, v \rangle = 0$  for all  $v \in U^\perp$  by the definition of  $U^\perp$ . But this also means that  $u$  is orthogonal to every vector in  $U^\perp$ , i.e., that  $u \in (U^\perp)^\perp$ . Thus  $U \subset (U^\perp)^\perp$ .

Now suppose that  $v \in (U^\perp)^\perp$ . By Theorem 6.20, we have that  $v = u + w$ , where  $u \in U$  and  $w \in U^\perp$ . It follows that  $v - u = w \in U^\perp$  and, since  $v \in (U^\perp)^\perp$  by hypothesis and  $u \in (U^\perp)^\perp$  by the first part of the proof, that  $v - u \in (U^\perp)^\perp$ . Therefore,  $v - u \in U^\perp \cap (U^\perp)^\perp \subset \{0\}$ , so  $v - u = 0$  or  $v = u \in U$ , as desired. ■

- **Orthogonal projection** (of  $V$  onto  $U$ ): The operator  $P_U \in \mathcal{L}(V)$  defined by  $P_U v = u$  iff  $v = u + w$  where  $u \in U$  and  $w \in U^\perp$ .

– Since Theorem 6.20 implies that each  $v \in V$  can be written uniquely in the form  $u + w$  where  $u \in U$  and  $w \in U^\perp$  for any subspace  $U$  of  $V$ ,  $P_U v$  is well defined in general.

- Properties of the orthogonal projection.

**Theorem 6.23.** *Suppose  $U$  is a finite-dimensional subspace of  $V$  and  $v \in V$ . Then*

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

*Proof.* To prove that  $P_U$  is a linear map on  $V$ , suppose  $v_1, v_2 \in V$  and  $a_1, a_2 \in \mathbb{F}$ . Let

$$v_1 = u_1 + w_1 \qquad v_2 = u_2 + w_2$$

where  $u_1, u_2 \in U$  and  $w_1, w_2 \in U^\perp$ . Then

$$\begin{aligned} P_U(a_1 v_1 + a_2 v_2) &= P_U(a_1 u_1 + a_2 u_2 + a_1 w_1 + a_2 w_2) \\ &= a_1 u_1 + a_2 u_2 \\ &= a_1 P_U v_1 + a_2 P_U v_2 \end{aligned}$$

as desired. ■

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

*Proof.* Let  $u \in U$  be arbitrary. Then  $u = u + 0$  where  $u \in U$  and  $0 \in U^\perp$ . Thus,  $P_U u = u$ . ■

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

*Proof.* Let  $w \in U^\perp$ . Then  $w = 0 + w$  where  $0 \in U$  and  $w \in U^\perp$ . Thus,  $P_U w = 0$ . ■

- (d)  $\text{range } P_U = U$ .

*Proof.* By the definition of  $P_U$ ,  $\text{range } P_U \subset U$ . Additionally, by Theorem 6.23b,  $U \subset \text{range } P_U$ . Thus,  $\text{range } P_U = U$ , as desired. ■

- (e)  $P_U = U^\perp$ .

*Proof.* Theorem 6.23c implies that  $U^\perp \subset P_U$ . In the other direction, we have that if  $v \in P_U$ , then  $P_U v = 0$ . Thus, we must have  $v = 0 + v$ , where  $0 \in U$  and  $v \in U^\perp$ . Consequently,  $P_U \subset U^\perp$ . Therefore,  $P_U = U^\perp$ , as desired. ■

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

*Proof.* If  $v = u + w$  with  $u \in U$  and  $w \in U^\perp$ , then  $v - P_U v = v - u = w \in U^\perp$ , as desired. ■

- (g)  $P_U^2 = P_U$ .

*Proof.* If  $v = u + w$  with  $u \in U$  and  $w \in U^\perp$ , then  $(P_U^2)v = P_U(P_U v) = P_U u = u = P_U v$ , as desired. ■

- (h)  $\|P_U v\| \leq \|v\|$ .

*Proof.* If  $v = u + w$  with  $u \in U$  and  $w \in U^\perp$ , then

$$\begin{aligned} \|P_U v\| &= \|u\|^2 \\ &\leq \|u\|^2 + \|w\|^2 \\ &= \|v\|^2 \end{aligned} \qquad \text{Pythagorean Theorem}$$

Taking square roots gives the desired inequality. ■

(i) For every orthonormal basis  $e_1, \dots, e_m$  of  $U$ ,

$$P_U v = \langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m$$

*Proof.* If  $v = u + w$  with  $u \in U$  and  $w \in U^\perp$ , then

$$\begin{aligned} P_U v &= u \\ &= \langle v, e_1 \rangle e_1 + \dots + \langle v, e_m \rangle e_m \end{aligned} \quad \text{Theorem 6.12}$$

as desired. ■

- Minimizing the distance to a subspace.

**Theorem 6.24.** 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

$$\begin{aligned} \|v - P_U v\|^2 &\leq \|v - P_U v\|^2 + \|P_U v - u\|^2 \\ &= \|(v - P_U v) + (P_U v - u)\|^2 \\ &= \|v - u\|^2 \end{aligned} \quad \text{Pythagorean Theorem}$$

Taking square roots of the above give the desired inequality.

The inequality is an equality if and only if  $\|P_U v - u\|^2 = 0$ . But this happens if and only if  $u = P_U v$ , as desired. ■

- The above theorem is often combined with Theorem 6.23i to compute explicit solutions to minimization problems.