

## 7 Inner Product Spaces

In this section we will generalize the work from section 6 to work in arbitrary vector spaces.

### 7.1 Inner Products

We first need to define an analogue of the dot product for an arbitrary vector space.

**Definition 7.1.** An *inner product* on a vector space  $V$  is an operation that takes in two vectors  $\mathbf{u}, \mathbf{v} \in V$  and returns a real number  $\langle \mathbf{u}, \mathbf{v} \rangle$ , satisfying the following conditions:

1. (Positive Definite)  $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ , and  $\langle \mathbf{u}, \mathbf{u} \rangle = 0$  if and only if  $\mathbf{u} = \mathbf{0}$ .
2. (Symmetric)  $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
3. (Bilinear)  $\langle \alpha \mathbf{u} + \beta \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{w} \rangle + \beta \langle \mathbf{v}, \mathbf{w} \rangle$ , and  $\langle \mathbf{u}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle = \alpha \langle \mathbf{u}, \mathbf{v} \rangle + \beta \langle \mathbf{u}, \mathbf{w} \rangle$ .

We write the *norm* of a vector  $\mathbf{v}$  as  $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ .

**Example 7.2.** Let  $V = \mathcal{C}([a, b], \mathbb{R})$  be the space of continuous functions on  $[a, b]$ , and define an inner product by  $\langle f, g \rangle = \int_a^b f(t)g(t) dt$ . Then

1.  $\langle f, f \rangle = \int_a^b f(t)^2 dt \geq 0$  since  $f(t)^2 \geq 0$ ; and further the integral is zero if and only if  $f(t)^2 = 0$  everywhere.
2.  $\langle f, g \rangle = \int_a^b f(t)g(t) dt = \int_a^b g(t)f(t) dt = \langle g, f \rangle$ .
3.  $\langle \alpha f + \beta g, h \rangle = \int_a^b (\alpha f(t) + \beta g(t))h(t) dt = \alpha \int_a^b f(t)h(t) dt + \beta \int_a^b g(t)h(t) dt = \alpha \langle f, h \rangle + \beta \langle g, h \rangle$ .

Thus this is an inner product on  $\mathcal{C}([a, b], \mathbb{R})$  by definition.

**Example 7.3.** Let  $V = \mathcal{P}_n(x)$  and fix real numbers  $x_0, x_1, \dots, x_n$  be distinct real numbers. For  $f, g \in V$ , define

$$\langle f, g \rangle = \sum_{i=0}^n f(x_i)g(x_i).$$

Then we can see  $\langle f, f \rangle = \sum_{i=0}^n f(x_i)^2 \geq 0$ , and the sum is equal to zero if and only if  $f(x_i) = 0$  for all  $i$ . But then  $f$  is a degree  $n$  polynomial with  $n + 1$  roots, and so must be constantly zero.

You will check the other two conditions on your homework.

We'd like to check that this inner product gives us all the things that the dot product did. In particular we want it to give us distance and angle and projections.

**Definition 7.4.** Let  $\mathbf{u}, \mathbf{v}$  be elements of an inner product space  $V$ . If  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ , we say that  $\mathbf{u}$  and  $\mathbf{v}$  are *orthogonal*.

We will eventually see that this corresponds to the two vectors being at a “right angle” to each other. But more immediately, we'll see that this means they are independent in a very specific way.

**Definition 7.5.** Suppose  $\mathbf{u}, \mathbf{v}$  are vectors in an inner product space  $V$ , and  $\mathbf{v} \neq 0$ . We define the projection of  $\mathbf{u}$  onto  $\mathbf{v}$  by

$$\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}.$$

**Proposition 7.6.** Let  $\mathbf{u}, \mathbf{v}$  be vectors in an inner product space  $V$ , with  $\mathbf{v} \neq 0$ . Let  $\mathbf{p} = \text{proj}_{\mathbf{v}} \mathbf{u}$ . Then:

1.  $\langle \mathbf{u} - \mathbf{p}, \mathbf{p} \rangle = 0$ —that is,  $\mathbf{u} - \mathbf{p}$  is orthogonal to  $\mathbf{p}$ .
2.  $\mathbf{u} = \mathbf{p}$  if and only if  $\mathbf{u}$  is a scalar multiple of  $\mathbf{v}$ .

*Proof.* 1. Exercise; see Homework 11.

2. If  $\mathbf{u} = \beta \mathbf{v}$ , then

$$\text{proj}_{\mathbf{v}} \mathbf{u} = \frac{\langle \beta \mathbf{v}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} = \beta \mathbf{v} = \mathbf{u}.$$

Conversely, suppose  $\mathbf{u} = \text{proj}_{\mathbf{v}} \mathbf{u}$ . Then by definition

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v},$$

so set  $\beta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle}$  and we have  $\mathbf{u} = \beta \mathbf{v}$ . □

**Example 7.7.** Let  $V = \mathcal{C}([-1, 1], \mathbb{R})$  be the space of continuous functions on the closed interval  $[-1, 1]$ , with the inner product given as above. Consider the vectors  $1, x$ . We compute:

$$\begin{aligned} \|1\| &= \sqrt{\int_{-1}^1 1 \, dx} = \sqrt{x|_{-1}^1} = \sqrt{2} \\ \|x\| &= \sqrt{\int_{-1}^1 x^2 \, dx} = \sqrt{x^3/3|_{-1}^1} = \sqrt{2/3} \\ \langle 1, x \rangle &= \int_{-1}^1 x \, dx = x^2/2|_{-1}^1 = 0 \end{aligned}$$

so  $1$  and  $x$  are orthogonal. Thus the projection of  $x$  onto  $1$  will give the zero vector: the two vectors have no “direction” in common.

Let’s consider now the vector  $1 + x$ . We have

$$\begin{aligned}\langle 1 + x, 1 \rangle &= \int_{-1}^1 1 + x \, dx = x + x^2/2 \Big|_{-1}^1 = 2 \\ \langle 1 + x, x \rangle &= \int_{-1}^1 x + x^2 \, dx = x^2/2 + x^3/3 \Big|_{-1}^1 = 2/3.\end{aligned}$$

Now we compute

$$\begin{aligned}\text{proj}_1 1 + x &= \frac{\langle 1 + x, 1 \rangle}{\langle 1, 1 \rangle} 1 = \frac{2}{2} 1 = 1 \\ \text{proj}_x 1 + x &= \frac{\langle 1 + x, x \rangle}{\langle x, x \rangle} x = \frac{2/3}{2/3} x = x.\end{aligned}$$

Thus we can use the inner product to decompose  $1 + x$  into its  $1$  component and its  $x$  component (and the remainder, if there were any).

If two vectors are orthogonal, then they are independent; they don’t have anything pointing in the same direction. This means their lengths are in some sense independent.

**Proposition 7.8** (Pythagorean Law). *If  $\mathbf{u}, \mathbf{v}$  are orthogonal, then*

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

*Proof.* Exercise: See homework 11. □

**Example 7.9.** Returning to our previous example, we can compute that

$$\|1 + x\| = \sqrt{\int_{-1}^1 1 + 2x + x^2 \, dx} = \sqrt{x + x^2 + x^3/3 \Big|_{-1}^1} = \sqrt{8/3}.$$

We can confirm that indeed,

$$\|1 + x\|^2 = 8/3 = 2 + 2/3 = \|1\|^2 + \|x\|^2.$$

Using projections we can prove a very important result called the Cauchy-Schwarz inequality.

**Theorem 7.10** (Cauchy-Schwarz Inequality). *If  $\mathbf{u}, \mathbf{v}$  are in an inner product space  $V$ , then*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|. \tag{3}$$

*Equality holds if and only if  $\mathbf{u}$  and  $\mathbf{v}$  are linearly dependent.*

*Proof.* If  $\mathbf{v} = \mathbf{0}$ , both sides are zero. So assume  $\mathbf{v} \neq \mathbf{0}$ .

Let  $\mathbf{p} = \text{proj}_{\mathbf{v}}\mathbf{u}$ . Then we can compute By the Pythagorean law 7.8, we know that

$$\|\mathbf{u}\|^2 = \|\mathbf{p}\|^2 + \|\mathbf{u} - \mathbf{p}\|^2.$$

But we know that

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

Thus we have

$$\begin{aligned} \frac{\langle \mathbf{u}, \mathbf{v} \rangle^2}{\langle \mathbf{v}, \mathbf{v} \rangle} &= \|\mathbf{u}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2 \\ \langle \mathbf{u}, \mathbf{v} \rangle^2 &= \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{p}\|^2 \|\mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \\ |\langle \mathbf{u}, \mathbf{v} \rangle| &\leq \|\mathbf{u}\| \|\mathbf{v}\|. \end{aligned}$$

Further, we can easily see that we get equality if and only if  $\mathbf{u} - \mathbf{p} = \mathbf{0}$ , if and only if  $\mathbf{u} = \mathbf{p}$ , if and only if  $\mathbf{u}$  is a scalar multiple of  $\mathbf{v}$ .  $\square$

Notice that this allows us to define an angle. Cauchy-Schwarz tells us that

$$-1 \leq \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \leq 1,$$

so we can coherently define

**Definition 7.11.** If  $\mathbf{u}, \mathbf{v}$  are non-zero vectors in an inner product space, we define the angle between them to be

$$\theta = \arccos \left( \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \|\mathbf{v}\|} \right).$$

Finally we'd like to return to the idea of distance, by thinking about what properties a "distance" function *should* have. We get

**Definition 7.12.** A vector space  $V$  together with an operation  $\|\cdot\| : V \rightarrow \mathbb{R}$  is said to be a *normed linear space* if:

1.  $\|\mathbf{v}\| \geq 0$  for any  $\mathbf{v} \in V$ , with equality if and only if  $\mathbf{v} = \mathbf{0}$ .
2.  $\|\alpha\mathbf{v}\| = |\alpha| \|\mathbf{v}\|$  for any  $\alpha \in \mathbb{R}, \mathbf{v} \in V$ .
3. (Triangle inequality)  $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$  for all  $\mathbf{v}, \mathbf{w} \in V$ .

*Remark 7.13.* These three conditions are equivalent to:

1. Nothing has a negative length, and only the zero vector has zero length.
2. Stretching a vector by a scalar multiplies its length by that scalar.
3. The sum of the lengths of two sides of a triangle is greater than the length of the third side. Alternatively, you can't get somewhere faster by adding a detour in the middle.

A normed linear space is in some sense the right setting in which to do calculus.

**Proposition 7.14.** *Let  $V$  be an inner product space. Then  $V$  is a normed linear space with norm given by  $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ .*

*Proof.* The first two conditions are easy to prove, so we'll just check the triangle inequality.

We compute

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

Taking the square root of both sides gives

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

as desired. □

*Remark 7.15.* There are norms that do not come from inner products. A good example is the norm on  $\mathbb{R}^n$  given by  $\|(a_1, a_2, \dots, a_n)\|_1 = |a_1| + |a_2| + \dots + |a_n|$ . We won't worry too much about those right now, though.

## 7.2 Orthonormal Bases

Throughout the course, we've been suggesting that we would often like to change from one coordinate system into another which is easier to work with. In this section we'll discuss one particular type of nice basis: one in which all the basis elements are orthogonal.

**Definition 7.16.** A set  $S = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$  is said to be *orthogonal* if  $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$  whenever  $i \neq j$ . We say it is *orthonormal* if every vector has magnitude 1.

**Proposition 7.17.** *Any orthogonal set of non-zero vectors is linearly independent.*

*Proof.* Suppose

$$a_1 \mathbf{u}_1 + \cdots + a_n \mathbf{u}_n = \mathbf{0}.$$

Then dotting the equation with itself, we get

$$\begin{aligned} \langle a_1 \mathbf{u}_1 + \cdots + a_n \mathbf{u}_n, a_1 \mathbf{u}_1 + \cdots + a_n \mathbf{u}_n \rangle &= 0 \\ \sum_{i,j=1}^n a_i a_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle &= 0 \end{aligned}$$

But since the  $\mathbf{u}_i$  are orthogonal,  $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$  when  $i \neq j$ , so this just gives us

$$\begin{aligned} \sum_{i=1}^n a_i^2 \langle \mathbf{u}_i, \mathbf{u}_i \rangle &= 0 \\ a_1^2 \|\mathbf{u}_1\|^2 + \cdots + a_n^2 \|\mathbf{u}_n\|^2 &= 0. \end{aligned}$$

And thus, since  $\|\mathbf{u}_i\| > 0$ , we must have  $a_i = 0$  for each  $i$ . □

Thus every orthogonal set is a basis for its span.

**Definition 7.18.** Let  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  be a basis for  $V$ . We say that  $E$  is an *orthogonal basis* if  $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$  whenever  $i \neq j$ .

We say that  $E$  is an *orthonormal basis* if, furthermore,  $\|\mathbf{e}_i\| = 1$ . Thus  $E$  is an orthonormal basis if and only if

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

**Example 7.19.** • The standard basis for  $\mathbb{R}^3$  is orthonormal.

- The basis  $\{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$  for  $\mathbb{R}^3$  is orthogonal but not orthonormal.

But  $\{(\sqrt{2}/2, \sqrt{2}/2, 0), (\sqrt{2}/2, -\sqrt{2}/2, 0), (0, 0, 1)\}$  is orthonormal.

- Let  $V = \mathcal{P}_2(x)$  with inner product given by  $\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt$ . The basis  $E = \{1, x, 3x^2 - 1\}$  is an orthogonal basis for  $V$ , but not orthonormal.

The basis  $F = \left\{ \frac{1}{\sqrt{2}}, \frac{x\sqrt{3}}{\sqrt{2}}, \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1) \right\}$  is orthonormal.

- Let  $V = \mathcal{P}_2(x)$  with inner product given by  $\langle f, g \rangle = f(-1)g(-1) + f(0)g(0) + f(1)g(1)$ . Then  $E = \{1, x, x^2 - 2/3\}$  is an orthogonal basis for  $V$ .

An orthonormal basis is  $F = \left\{ \frac{\sqrt{3}}{3}, \frac{x\sqrt{2}}{2}, \frac{\sqrt{3}}{\sqrt{2}} \left(x^2 - \frac{2}{3}\right) \right\}$ .

Orthonormal bases are particularly nice, for a few reasons.

**Proposition 7.20.** *Suppose  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  is an orthonormal basis for  $V$ . Then if  $\mathbf{u} = \sum_{i=1}^n a_i \mathbf{e}_i$  and  $\mathbf{v} = \sum_{i=1}^n b_i \mathbf{e}_i$ , then  $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{i=1}^n a_i b_i$ .*

*Consequently  $\|\mathbf{u}\| = |a_1|^2 + \dots + |a_n|^2$ .*

*Remark 7.21.* We use this all the time when we're computing the norm of vectors in  $\mathbb{R}^n$ . This also gives us our "normal" dot product.

More importantly, orthonormal bases make projection, coordinates, and changes of basis very easy.

**Proposition 7.22.** *Let  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  be a basis for  $V$ , with  $\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0$  when  $i \neq j$ . Then if  $\mathbf{v} \in V$ , we have*

$$\mathbf{v} = \sum_{i=1}^n (\text{proj}_{\mathbf{e}_i} \mathbf{v}) \mathbf{e}_i = (\text{proj}_{\mathbf{e}_1} \mathbf{v}) \mathbf{e}_1 + \dots + (\text{proj}_{\mathbf{e}_n} \mathbf{v}) \mathbf{e}_n.$$

*Proof.* Write  $\mathbf{v} = a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n$  and compute each projection. □

**Corollary 7.23.** *Let  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  be an orthonormal basis for  $V$ . Then*

$$\mathbf{v} = \sum_{i=1}^n \langle \mathbf{v}, \mathbf{e}_i \rangle \mathbf{e}_i = \langle \mathbf{v}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \dots + \langle \mathbf{v}, \mathbf{e}_n \rangle \mathbf{e}_n.$$

**Example 7.24.**  $E = \{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$  is orthogonal. Find the  $E$  coordinates of  $(6, 2, 1)$ .

We compute:

$$\begin{aligned} \text{proj}_{\mathbf{e}_1}(6, 2, 1) &= \frac{(6, 2, 1) \cdot (1, 1, 0)}{(1, 1, 0) \cdot (1, 1, 0)} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \frac{8}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ \text{proj}_{\mathbf{e}_2}(6, 2, 1) &= \frac{(6, 2, 1) \cdot (1, -1, 0)}{(1, -1, 0) \cdot (1, -1, 0)} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \frac{4}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \\ \text{proj}_{\mathbf{e}_3}(6, 2, 1) &= \frac{(6, 2, 1) \cdot (0, 0, 1)}{(0, 0, 1) \cdot (0, 0, 1)} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

$$[(6, 2, 1)]_E = (4, 2, 1)$$

**Example 7.25.** Let  $V = \mathcal{P}_2(x)$ , with inner product given by  $\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt$ . Then  $E = \{1, x, 3x^2 - 1\}$  is orthogonal. Write  $3x^2 - 6x + 4$  in  $E$ -coordinates.

We compute

$$\text{proj}_{\mathbf{e}_1} 3x^2 - 6x + 4 = \frac{\langle 3x^2 - 6x + 4, 1 \rangle}{\langle 1, 1 \rangle} (1) = \frac{1}{2} \int_{-1}^1 3x^2 - 6x + 4 dx (1)$$

$$= \frac{1}{2} (x^3 - 3x^2 + 4x \mid \mid_{-1}^1) (1) = 5(1)$$

$$\text{proj}_{\mathbf{e}_2} 3x^2 - 6x + 4 = \frac{\langle 3x^2 - 6x + 4, x \rangle}{\langle x, x \rangle} (x) = \frac{3}{2} \int_{-1}^1 3x^3 - 6x^2 + 4x dx (x)$$

$$= \frac{3}{2} \left( \frac{x^4}{4} - 2x^3 + 2x^2 \mid \mid_{-1}^1 \right) (x) = -6(x)$$

$$\text{proj}_{\mathbf{e}_3} 3x^2 - 6x + 4 = \frac{\langle 3x^2 - 6x + 4, 3x^2 - 1 \rangle}{\langle 3x^2 - 1, 3x^2 - 1 \rangle} (3x^2 - 1)$$

$$= \frac{5}{8} \int_{-1}^1 (3x^2 - 6x + 4)(3x^2 - 1) dx (3x^2 - 1)$$

$$= \frac{5}{8} (9x^5/5 - 9x^4/2 + 3x^3 + 3x^2 - 4x \mid \mid_{-1}^1) (3x^2 - 1) = 1(3x^2 - 1)$$

$$[3x^2 - 6x + 4]_E = (5, -6, 1).$$

**Example 7.26.** Let  $V = \mathbb{R}^3$  with the usual dot product. Then the standard basis  $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  is an orthonormal basis. Use the dot product to find the coordinates of  $(2, 3, 4)$ .

We don't need to use the full projection operator; we just need to compute the inner products, since our basis is orthonormal and not just orthogonal.

$$\text{proj}_{\mathbf{e}_1} (2, 3, 4) = (2, 3, 4) \cdot (1, 0, 0) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{proj}_{\mathbf{e}_2} (2, 3, 4) = (2, 3, 4) \cdot (0, 1, 0) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\text{proj}_{\mathbf{e}_3} (2, 3, 4) = (2, 3, 4) \cdot (0, 0, 1) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$[(2, 3, 4)]_E = (2, 3, 4).$$

This isn't a surprise because it was already in coordinates with respect to the standard basis. But this also illustrates a more general principle: if your vector is already written in orthonormal coordinates, your inner product just becomes a dot product.



We'd like a way to generate an orthonormal basis if we don't already have one. This turns out to be straightforward; start with any basis, and one-by-one "fix" elements so that they're orthogonal to all the others.

**Proposition 7.27** (Gram-Schmidt Process). *Let  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  be a basis for  $V$ . Then there is an orthonormal basis  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ , where we set:*

$$\begin{array}{ll} \mathbf{f}_1 = \mathbf{e}_1 & \mathbf{u}_1 = \frac{\mathbf{f}_1}{\|\mathbf{f}_1\|} \\ \mathbf{f}_2 = \mathbf{e}_2 - \text{proj}_{\mathbf{u}_1} \mathbf{e}_2 & \mathbf{u}_2 = \frac{\mathbf{f}_2}{\|\mathbf{f}_2\|} \\ \mathbf{f}_3 = \mathbf{e}_3 - (\text{proj}_{\mathbf{u}_1} \mathbf{e}_3 + \text{proj}_{\mathbf{u}_2} \mathbf{e}_3) & \mathbf{u}_3 = \frac{\mathbf{f}_3}{\|\mathbf{f}_3\|} \\ \vdots & \vdots \\ \mathbf{f}_n = \mathbf{e}_n - (\text{proj}_{\mathbf{u}_1} \mathbf{e}_n + \dots + \text{proj}_{\mathbf{u}_{n-1}} \mathbf{e}_n) & \mathbf{u}_n = \frac{\mathbf{f}_n}{\|\mathbf{f}_n\|}. \end{array}$$

*Proof.* It's clear that each  $\mathbf{u}_i$  has norm 1, so we just need to check that they are pairwise orthogonal, which is the same as checking that the  $\mathbf{f}_i$  are all orthogonal

But we have constructed the  $\mathbf{f}_i$  to be orthogonal by subtracting off the pieces they have in common. For instance, we see that

$$\begin{aligned} \langle \mathbf{f}_1, \mathbf{f}_2 \rangle &= \left\langle \mathbf{e}_1, \mathbf{e}_2 - \frac{\langle \mathbf{e}_2, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \mathbf{e}_1 \right\rangle = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle - \frac{\langle \mathbf{e}_2, \mathbf{e}_1 \rangle}{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle \\ &= \langle \mathbf{e}_1, \mathbf{e}_2 \rangle - \langle \mathbf{e}_2, \mathbf{e}_1 \rangle = 0. \end{aligned}$$

In general, we see that

$$\langle \mathbf{f}_j, \text{proj}_{\mathbf{f}_i} \mathbf{f}_j \rangle = \left\langle \mathbf{f}_i, \frac{\langle \mathbf{f}_j, \mathbf{f}_j \rangle}{\langle \mathbf{f}_j, \mathbf{f}_j \rangle} \mathbf{f}_j \right\rangle = \langle \mathbf{f}_i, \mathbf{f}_j \rangle$$

and all the other projections will be zero since the  $\mathbf{f}_i$  are orthogonal, so each  $\mathbf{f}_j$  is orthogonal to all the previous  $\mathbf{f}_i$ .  $\square$

**Example 7.28.** Let  $V = \mathbb{R}^3$  with the usual dot product, and let  $E = \{(1, 1, -1), (1, -1, 1), (-1, 1, 1)\}$ . Use Gram-Schmidt to orthonormalize this basis.

We take  $\mathbf{f}_1 = (1, 1, -1)$ , and then we compute  $\|\mathbf{f}_1\| = \sqrt{3}$  so  $\mathbf{u}_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}\right)$ .

Then we set

$$\begin{aligned} \mathbf{f}_2 &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \text{Proj}_{(1,1,-1)} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{(1, -1, 1) \cdot (1, 1, -1)}{(1, -1, 1) \cdot (1, -1, 1)} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \frac{-1}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -2/3 \\ 2/3 \end{bmatrix} \\ \mathbf{u}_2 &= \frac{\mathbf{f}_2}{\|\mathbf{f}_2\|} = \frac{(4/3, -2/3, 2/3)}{\sqrt{24/9}} = \frac{\sqrt{3}}{2\sqrt{2}} \begin{bmatrix} 4/3 \\ -2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} \sqrt{6}/3 \\ -\sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix}. \end{aligned}$$

Finally we have

$$\begin{aligned} \mathbf{f}_3 &= \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} - \text{proj}_{(1,1,-1)} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} - \text{proj}_{(4,-2,2)} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} - \frac{(-1, 1, 1) \cdot (1, 1, -1)}{(1, 1, -1) \cdot (1, 1, -1)} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \frac{(-1, 1, 1) \cdot (4, -2, 2)}{(4, -2, 2) \cdot (4, -2, 2)} \begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} - \frac{-1}{3} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - \frac{-4}{24} \begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\ \mathbf{u}_3 &= \frac{\mathbf{f}_3}{\|\mathbf{f}_3\|} = \frac{(0, 1, 1)}{\sqrt{2}} = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}. \end{aligned}$$

Thus an orthonormal basis for  $\mathbb{R}^3$  is

$$\left\{ \begin{bmatrix} \sqrt{3}/3 \\ \sqrt{3}/3 \\ -\sqrt{3}/3 \end{bmatrix}, \begin{bmatrix} \sqrt{6}/3 \\ -\sqrt{6}/6 \\ \sqrt{6}/6 \end{bmatrix}, \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \right\}.$$

Notice that while this is an orthonormal basis for  $\mathbb{R}^3$ , it is not the usual one. We will get different orthonormal bases out of the end, depending on which vector we start with.

**Example 7.29.** Let  $V = \mathcal{P}_2(x)$  with the inner product given by  $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$ . (Note this is a *different* inner product from the one we've been using!) Let's form an orthonormal basis from the set  $\{1, x, x^2\}$ .

We set  $\mathbf{f}_1 = 1$ . We compute that

$$\|\mathbf{1}\|^2 = \langle \mathbf{1}, \mathbf{1} \rangle = \int_0^1 1 \, dx = 1$$

so this is already a unit vector; we set  $\mathbf{u}_1 = 1$ .

We take

$$\mathbf{f}_2 = x - \frac{\langle \mathbf{1}, x \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle} (\mathbf{1}) = x - \frac{1}{2} (\mathbf{1}) = x - 1/2.$$

We compute

$$\|\mathbf{f}_2\| = \sqrt{\int_0^1 (x - 1/2)^2 \, dx} = \sqrt{12} = 2\sqrt{3}$$

so we set

$$\mathbf{u}_2 = \frac{\mathbf{f}_2}{\|\mathbf{f}_2\|} = 2\sqrt{3}(x - 1/2) = \sqrt{3}(2x - 1).$$

Finally, we have

$$\begin{aligned} \mathbf{f}_3 &= x^2 - \frac{\langle \mathbf{1}, x^2 \rangle}{\langle \mathbf{1}, \mathbf{1} \rangle} (\mathbf{1}) - \frac{\langle \sqrt{3}(2x - 1), x^2 \rangle}{\langle \sqrt{3}(2x - 1), \sqrt{3}(2x - 1) \rangle} \sqrt{3}(2x - 1) \\ &= x^2 - \int_0^1 x^2 \, dx (\mathbf{1}) - \sqrt{3} \int_0^1 2x^3 - x^2 \, dx (\sqrt{3}(2x - 1)) \\ &= x^2 - \frac{1}{3} - \frac{1}{2}(2x - 1) = x^2 - x + \frac{1}{6}. \end{aligned}$$

Then we compute

$$\begin{aligned} \|\mathbf{f}_3\| &= \sqrt{\int_0^1 (x^2 - x + 1/6)^2 \, dx} = \sqrt{\frac{1}{180}} = \frac{1}{6\sqrt{5}} \\ \mathbf{u}_3 &= \frac{\mathbf{f}_3}{\|\mathbf{f}_3\|} = 6\sqrt{5}x^2 - 6\sqrt{5}x + \sqrt{5}. \end{aligned}$$

Thus an orthonormal basis for  $\mathcal{P}_2(x)$  with this inner product is

$$\{1, \sqrt{3}(2x - 1), \sqrt{5}(6x^2 - 6x + 1)\}.$$

### 7.3 Orthogonal Subspaces

We have used orthogonality to give a vector space a particularly nice basis. We can also break the vector space into two (or more) independent subspaces.

**Definition 7.30.** If  $V$  is an inner product space and  $U, W$  are subspaces, we say that  $U$  and  $W$  are *orthogonal* and write  $U \perp W$  if  $\langle \mathbf{u}, \mathbf{w} \rangle = 0$  for every  $\mathbf{u} \in U, \mathbf{w} \in W$ .

If  $U \subset V$ , we define the *orthogonal complement of  $U$*  to be the set of all vectors perpendicular to everything in  $U$ :

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

**Example 7.31.** • In  $\mathbb{R}^2$ , the orthogonal complement of a line is a line. The orthogonal complement to a set with two points in it is also a line.

- In  $\mathbb{R}^3$ , the orthogonal complement of a line is a plane, and the orthogonal complement of a plane is a line.

**Proposition 7.32.** *If  $U$  is a subset of  $V$ , then  $U^\perp$  is a subspace of  $V$ .*

*Proof.* 1.  $\mathbf{0}$  is orthogonal to everything, and thus is in  $U^\perp$ .

2. Suppose  $\mathbf{v} \in U^\perp$ , and  $r \in \mathbb{R}$ . Then for any  $\mathbf{u} \in U$  we have  $\langle r\mathbf{v}, \mathbf{u} \rangle = r\langle \mathbf{v}, \mathbf{u} \rangle = r \cdot 0 = 0$ , so  $r\mathbf{v} \in U^\perp$  by definition.

3. Suppose  $\mathbf{v}, \mathbf{w} \in U^\perp$ , and let  $\mathbf{u} \in U$ . Then

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

Thus  $\mathbf{v} + \mathbf{w}$  is orthogonal to  $\mathbf{u}$  for every  $\mathbf{u} \in U$ , and so  $\mathbf{v} + \mathbf{w} \in U^\perp$ .

Thus by the subspace theorem,  $U^\perp$  is a subspace of  $V$ . □

*Remark 7.33.* We will usually consider cases where  $U$  is also a subspace of  $V$ , but this isn't necessary; nothing above assumes anything about the structure of  $U$ .

A basic thing we want to do is, given a subspace, find a basis for the subspace and for its orthogonal complement. As with everything else, we can solve this problem by row-reducing matrices.

**Proposition 7.34.** *Let  $A$  be a matrix. Then  $\ker(A) = (\text{row}(A))^\perp$ .*

*Remark 7.35.* We used this earlier, when we were finding the normal vector to a plane.

*Proof.* If  $\mathbf{r}_i$  are the rows of the matrix  $A$ , then

$$A\mathbf{x} = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m \cdot \mathbf{x} \end{bmatrix}$$

and thus  $\mathbf{x} \in \ker(A)$  precisely if  $\mathbf{x}$  is orthogonal to every row of  $A$ . But if  $\mathbf{x}$  is orthogonal to every row vector of  $A$ , it is orthogonal to every linear combination of them, and thus is orthogonal to their span, which is the rowspace. □

**Example 7.36.** Suppose we want to find the orthogonal complement to  $U = \text{Span}\{(1, 4, 2), (1, 1, 1)\}$ . Then we write down the matrix

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 2/3 \\ 0 & 3 & 1 \end{bmatrix}$$

so  $U^\perp = \ker(A) = \{(-2\alpha, -\alpha, 3\alpha)\} = \text{Span}\{(2, 1, -3)\}$ . We can check that this is in fact orthogonal to the original two vectors.

There are a couple more useful facts we'd like to know about orthogonal complements, which show that they relate spaces in useful ways.

**Proposition 7.37.** *If  $U$  is a subspace of  $V$  and  $\mathbf{v} \in V$ , then there exist unique  $\mathbf{v}_U \in U, \mathbf{v}_{U^\perp} \in U^\perp$  such that  $\mathbf{v} = \mathbf{v}_U + \mathbf{v}_{U^\perp}$ .*

*We say that this is an orthogonal decomposition of  $\mathbf{v}$ .*

*Proof.* Let  $E = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  be an orthogonal basis for  $U$  and  $F = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$  an orthogonal basis for  $U^\perp$ .

We claim that  $E \cup F = \{\mathbf{e}_1, \dots, \mathbf{e}_n, \mathbf{f}_1, \dots, \mathbf{f}_m\}$  is an orthogonal basis for  $V$ . It must be orthogonal since  $E$  and  $F$  are orthogonal sets, and thus it is linearly independent. So we need to show that it spans  $V$ .

Suppose  $\mathbf{v} \in V$ , and consider the element

$$\mathbf{v}' = \mathbf{v} - \sum_{i=1}^n \text{proj}_{\mathbf{e}_i} \mathbf{v}.$$

This is an element of  $V$ , and by construction it is orthogonal to every  $\mathbf{e}_i$  and thus all of  $U$ , so  $\mathbf{v}' \in U^\perp$ . Thus  $\mathbf{v}' \in \text{Span}(F)$  and so  $\mathbf{v} \in \text{Span}(E \cup F)$ . Thus  $E \cup F$  spans  $V$ .

Then every element of  $V$  can be expressed uniquely as a linear combination of elements of  $E$  and  $F$ . This gives us a unique representation as a sum of an element of  $U$  and an element of  $U^\perp$ .

□

**Corollary 7.38.**  $\dim U + \dim U^\perp = \dim V$ .

**Example 7.39.** Give the orthogonal decomposition of  $(3, -1, 2)$  with respect to the subspace given by  $x - y + 2z = 0$  and its complement.

We need to find an orthonormal basis for either  $x - y + 2z = 0$  or its orthogonal complement. But we can see that the normal vector to this plane is in the orthogonal complement, so  $\{(1, -1, 2)\}$  is a basis for  $U^\perp$ .

We project  $(3, -1, 2)$  onto  $\text{Span}\{(1, -1, 2)\}$ . We have

$$\text{proj}_{(1,-1,2)} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \frac{(3, -1, 2) \cdot (1, -1, 2)}{(1, -1, 2) \cdot (1, -1, 2)} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \frac{8}{6} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -4/3 \\ 8/3 \end{bmatrix}$$

So this is the projection into  $U^\perp$ . The projection into  $U$  then is just what's left over: it's

$$\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} - \text{proj}_{(1,-1,2)} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 - 4/3 \\ -1 + 4/3 \\ 2 - 8/3 \end{bmatrix} = \begin{bmatrix} 5/3 \\ 1/3 \\ -2/3 \end{bmatrix}.$$

(We check that this vector is in fact in the plane  $U$ ). Then we have an orthogonal decomposition:  $(3, -1, 2) = (5/3, 1/3, -2/3) + (4/3, -4/3, 8/3)$ .

**Example 7.40.** Let  $V = \mathbb{R}^4$  and let  $U = \text{Span}\{(1, 2, 3, 4), (2, 1, -1, -2)\}$ . Find the orthogonal decomposition of  $(1, 1, 1, 1)$  into its components in  $U$  and  $U^\perp$ .

We write a matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 0 & -5 & -8 \\ 0 & 3 & 7 & 10 \end{bmatrix}$$

so  $\ker(A) = (5\alpha + 8\beta, -7\alpha - 10\beta, 3\alpha, 3\beta) = \text{Span}\{(5, -7, 3, 0), (8, -10, 0, 3)\}$ .

We need to find an orthogonal basis for either  $U$  or  $U^\perp$ . We compute

$$\begin{aligned} \mathbf{f}_1 &= \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \\ \mathbf{f}_2 &= \begin{bmatrix} 2 \\ 1 \\ -1 \\ -2 \end{bmatrix} - \text{proj}_{(1,2,3,4)} \begin{bmatrix} 2 \\ 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ -1 \\ -2 \end{bmatrix} - \frac{(2, 1, -1, -2) \cdot (1, 2, 3, 4)}{(1, 2, 3, 4) \cdot (1, 2, 3, 4)} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 1 \\ -1 \\ -2 \end{bmatrix} - \frac{-7}{30} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{30} \begin{bmatrix} 67 \\ 44 \\ -9 \\ -32 \end{bmatrix} \end{aligned}$$

We compute

$$\begin{aligned} \text{proj}_{\mathbf{f}_1} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} &= \frac{(1, 2, 3, 4) \cdot (1, 1, 1, 1)}{(1, 2, 3, 4) \cdot (1, 2, 3, 4)} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{10}{30} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 2/3 \\ 1 \\ 4/3 \end{bmatrix} \\ \text{proj}_{\mathbf{f}_2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} &= \frac{(1, 1, 1, 1) \cdot (67, 44, -9, -32)}{(67, 44, -9, -32) \cdot (67, 44, -9, -32)} \begin{bmatrix} 67 \\ 44 \\ -9 \\ -32 \end{bmatrix} = \frac{70}{7530} \begin{bmatrix} 67 \\ 44 \\ -9 \\ -32 \end{bmatrix} = \frac{7}{753} \begin{bmatrix} 67 \\ 44 \\ -9 \\ -32 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}_U &= \begin{bmatrix} 1/3 \\ 2/3 \\ 1 \\ 4/3 \end{bmatrix} + \frac{7}{753} \begin{bmatrix} 67 \\ 44 \\ -9 \\ -32 \end{bmatrix} = \frac{10}{251} \begin{bmatrix} 24 \\ 27 \\ 23 \\ 26 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}_{U^\perp} &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}_U = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{10}{251} \begin{bmatrix} 24 \\ 27 \\ 23 \\ 26 \end{bmatrix} = \frac{1}{251} \begin{bmatrix} 11 \\ -19 \\ 21 \\ -9 \end{bmatrix}. \end{aligned}$$

**Proposition 7.41.** *If  $U$  is a subspace of  $V$ , then  $(U^\perp)^\perp = U$ .*

*Proof.* If  $\mathbf{u} \in U$ , then  $\mathbf{u}$  is orthogonal to every  $\mathbf{w} \in U^\perp$  by definition. So  $U \subset (U^\perp)^\perp$ .

Conversely, suppose  $\mathbf{w} \in (U^\perp)^\perp$ . We can write  $\mathbf{w} = \mathbf{w}_U + \mathbf{w}_{U^\perp}$ . Then  $\mathbf{w} \in (U^\perp)^\perp$  so we know  $\langle \mathbf{w}, \mathbf{w}_{U^\perp} \rangle = \mathbf{0}$ .

But  $\langle \mathbf{w}, \mathbf{w}_{U^\perp} \rangle = \langle \mathbf{w}_{U^\perp}, \mathbf{w}_{U^\perp} \rangle = 0$  if and only if  $\mathbf{w}_{U^\perp} = \mathbf{0}$ . Thus  $\mathbf{w}_{U^\perp} = \mathbf{0}$ , and  $\mathbf{w} = \mathbf{w}_U \in U$ .  $\square$