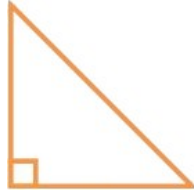


## Chapter 4: Orthogonality



$$\hat{e}_1 \perp \hat{e}_2$$

Orthogonality is the generalization of the idea of perpendicularity. Recall  $\vec{v}, \vec{w}$  are called orthogonal if  $\langle \vec{v}, \vec{w} \rangle = 0$ .

**Definition:** A basis  $\vec{u}_1, \dots, \vec{u}_n$  of an  $n$ -dimensional inner product space  $V$  is called an *orthogonal basis* if  $\langle \vec{u}_i, \vec{u}_j \rangle = 0$  for all  $i \neq j$ . It is called an *orthonormal basis* if, in addition, each vector has unit length:  $|\vec{u}_i| = 1$ , for all  $i = 1, \dots, n$ .

In application, and in theoretical research, choosing mutually orthogonal basis elements is extremely powerful.

**Definition:**  $A$  is called an *orthogonal matrix* if its columns form an orthonormal system.

In this chapter, we will learn how to adjust a basis to be an **orthonormal basis** (§4.2, Gram-Schmidt), to decompose matrices into **orthogonal matrices** (§4.3 QR decomp), and to **orthogonally project vectors** onto subspaces (§4.4).

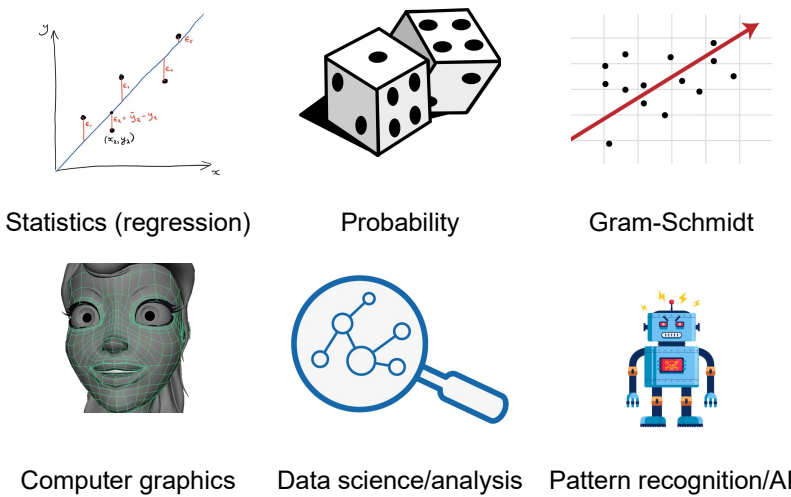
We will also learn about **orthogonal subspaces** which are orthogonal to each other (§4.4).

The four fundamental subspaces of a matrix ((co-)image, (co-)kernel) that were introduced in §2 come in mutually orthogonal pairs. This will lead to an intriguing result in §4.4 (Fredholm Alternative).

### Applications

Fourier analysis decomposes waves (music, or *any* signal) into sines/cosines which form orthogonal system of functions (CDs, DVDs, MP3s).

Orthogonal projections turn out to be incredibly important in:



Statistics (regression)

Probability

Gram-Schmidt

Computer graphics

Data science/analysis

Pattern recognition/AI

### 4.1 Orthogonal and Orthonormal Bases

**Lemma:** If  $\vec{v}_1, \dots, \vec{v}_n$  is an orthogonal basis of a vector space  $V$ , then the normalized vectors

$$\vec{u}_i = \frac{\vec{v}_i}{|\vec{v}_i|}, \quad i = 1, \dots, n \text{ form an orthonormal basis.}$$

**Specifically:**  $\vec{v}_1 = (1, 2, -1)$ ,  $\vec{v}_2 = (0, 1, 2)$ , and  $\vec{v}_3 = (5, -2, 1)$  are easily seen to form a basis of  $\mathbb{R}^3$ .

One can also check:  $\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \vec{v}_3 = \vec{v}_2 \cdot \vec{v}_3 = 0$ . So, form an orthogonal basis wrt standard dot product.

When we divide each orthogonal basis by its length, the result is the orthonormal basis:

$$\vec{u}_1 = \frac{1}{\sqrt{6}}(1, 2, -1) = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right), \quad \vec{u}_2 = \left(0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right), \quad \text{and} \quad \vec{u}_3 = \left(\frac{5}{\sqrt{30}}, -\frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}}\right).$$

Observe  $|\vec{u}_i| = 1$ , and  $\vec{u}_i \cdot \vec{u}_j = 0$ , when  $i \neq j$ .

**Proposition:** Let  $\vec{v}_1, \dots, \vec{v}_n \in V$  be nonzero, mutually orthogonal elements, so  $\vec{v}_i \neq \vec{0}$  and  $\langle \vec{v}_i, \vec{v}_j \rangle = 0$  for all  $i \neq j$ .

Then  $\vec{v}_1, \dots, \vec{v}_k$  are linearly independent.

**Proof:** Suppose  $c_1\vec{v}_1 + \dots + c_k\vec{v}_k = \vec{0}$ .

We must show that the  $c_i$  need be zero.

Let's take the inner product of this equation with any  $\vec{v}_i$ . Using linearity and orthogonality, we compute:

$$0 = \langle c_1\vec{v}_1 + \dots + c_k\vec{v}_k, \vec{v}_i \rangle = c_1\langle \vec{v}_1, \vec{v}_i \rangle + \dots + c_k\langle \vec{v}_k, \vec{v}_i \rangle = c_i\langle \vec{v}_i, \vec{v}_i \rangle = c_i|\vec{v}_i|^2.$$

Therefore, given that  $\vec{v}_i \neq \vec{0}$ , we conclude that  $c_i = 0$ .

Since this holds for all  $i = 1, \dots, k$ , the linear independence of  $\vec{v}_1, \dots, \vec{v}_n$  follows. ■

**Theorem:** Suppose  $\vec{v}_1, \dots, \vec{v}_n$  are nonzero, mutually orthogonal elements of an inner product space  $V$ .

Then,  $\vec{v}_1, \dots, \vec{v}_n$  form an orthogonal basis for their span  $W = \text{span}\{\vec{v}_1, \dots, \vec{v}_n\} \subset V$ , which is therefore a subspace of dimension  $n = \dim W$ . In particular, if  $\dim V = n$ , then  $\vec{v}_1, \dots, \vec{v}_n$  form an orthogonal basis for  $V$ .

**Specifically:** Consider  $P^{(2)}$ , consisting of  $p(x) = \alpha + \beta x + \gamma x^2$ , equipped with the  $L^2$  inner product and norm.

Standard monomials  $1, x, x^2$  do *not* form an orthogonal basis. As calculated previously:

$$\langle 1, x \rangle = \frac{1}{2}, \quad \langle 1, x^2 \rangle = \frac{1}{3}, \quad \langle x, x^2 \rangle = \frac{1}{4}.$$

One orthogonal basis is provided by:  $p_1(x) = 1$ ,  $p_2(x) = x - \frac{1}{2}$ ,  $p_3(x) = x^2 - x + \frac{1}{6}$ .

One easily verifies  $\langle p_1, p_2 \rangle = \langle p_1, p_3 \rangle = \langle p_2, p_3 \rangle = 0$ , while  $|p_1| = 1$ ,  $|p_2| = \frac{1}{2\sqrt{3}}$ ,  $|p_3| = \frac{1}{6\sqrt{5}}$ .

Forming the orthonormal basis:  $u_1(x) = 1$ ,  $u_2(x) = \sqrt{3}(2x - 1)$ ,  $u_3(x) = \sqrt{5}(6x^2 - 6x + 1)$ .

§4.5 will give us a systematic way to find the orthogonal basis yourself.

## Computations in Orthogonal Bases

In high dimensions, computations can take a long time. However, switching to an orthogonal or orthonormal system can dramatically speed up the computations. This has allowed for many advancements including MP3s, CDs, DVDs, YouTube, least-squares approximations, and the statistical analysis of large data sets (and much more).

**Theorem:** Let  $\vec{u}_1, \dots, \vec{u}_n$  be an orthonormal basis for an inner product space  $V$ . Then, one can write any element  $\vec{v} \in V$  as a linear combination:  $\vec{v} = c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$ , in which its coordinates can be calculated as  $c_i = \langle \vec{v}, \vec{u}_i \rangle$ .

Moreover,  $\vec{v}$ 's norm is given by the Pythagorean formula:  $|\vec{v}| = \sqrt{c_1^2 + \dots + c_n^2} = \sqrt{\sum_{i=1}^n \langle \vec{v}, \vec{u}_i \rangle^2}$ .

**Proof:** Let's compute the inner product of  $\vec{v}$  (or  $c_1 \vec{u}_1 + \dots + c_n \vec{u}_n$ ) with one of the basis vectors.

Using the orthonormality conditions:  $\langle \vec{u}_i, \vec{u}_j \rangle = \begin{cases} 0 & i \neq j, \\ 1 & i = j, \end{cases}$  and bilinearity of the inner product, we obtain:

$$\langle \vec{v}, \vec{u}_i \rangle = \left\langle \sum_{j=1}^n c_j \vec{u}_j, \vec{u}_i \right\rangle = \sum_{j=1}^n c_j \langle \vec{u}_j, \vec{u}_i \rangle = c_i |\vec{u}_i|^2 = c_i.$$

To prove the last formula in the theorem, we similarly expand

$$|\vec{v}|^2 = \langle \vec{v}, \vec{v} \rangle = \left\langle \sum_{j=1}^n c_j \vec{u}_j, \sum_{j=1}^n c_j \vec{u}_j \right\rangle = \sum_{i,j=1}^n c_i c_j \langle \vec{u}_i, \vec{u}_j \rangle = \sum_{i=1}^n c_i^2,$$

again making use of the orthonormality of the basis elements. ■

**Specifically:** let's rewrite  $\vec{v} = (1, 1, 1)$  in terms of the orthonormal basis:

$$\vec{u}_1 = \left( \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}} \right), \quad \vec{u}_2 = \left( 0, \frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right), \quad \text{and} \quad \vec{u}_3 = \left( \frac{5}{\sqrt{30}}, -\frac{2}{\sqrt{30}}, \frac{1}{\sqrt{30}} \right).$$

Computing the dot product:  $\vec{v} \cdot \vec{u}_1 = \frac{2}{\sqrt{6}}$ ,  $\vec{v} \cdot \vec{u}_2 = \frac{3}{\sqrt{5}}$ ,  $\vec{v} \cdot \vec{u}_3 = \frac{4}{\sqrt{30}}$ . And therefore:

$$\vec{v} = \frac{2}{\sqrt{6}} \vec{u}_1 + \frac{3}{\sqrt{5}} \vec{u}_2 + \frac{4}{\sqrt{30}} \vec{u}_3.$$

Observe how much easier this is than solving the linear system  $\vec{v} = [\vec{u}_1 \dots \vec{u}_n] \vec{c}$  (chapter 2).

**Theorem:** If  $\vec{v}_1, \dots, \vec{v}_n$  form an orthogonal basis, then the corresponding coordinates of  $\vec{v} := a_1 \vec{v}_1 + \dots + a_n \vec{v}_n$  are given by

$$a_i = \frac{\langle \vec{v}, \vec{v}_i \rangle}{|\vec{v}_i|^2}. \quad \text{In this case, its norm can be computed using the formula: } |\vec{v}|^2 = \sum_{i=1}^n a_i^2 |\vec{v}_i|^2 = \sum_{i=1}^n \left( \frac{\langle \vec{v}, \vec{v}_i \rangle}{|\vec{v}_i|^2} \right)^2.$$

**Example:** Let  $\mathbb{R}^2$  have the inner product defined by the positive definite matrix  $\mathbf{K} = \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$ .

a) Show that  $\vec{v}_1 = (1, 1)$ ,  $\vec{v}_2 = (-2, 1)$  form an orthogonal basis. ...

$$\langle \vec{v}_1, \vec{v}_2 \rangle = \vec{v}_1^T \mathbf{K} \vec{v}_2 = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -5 \\ 5 \end{bmatrix} = 0.$$

b) Write  $\vec{v} = (3, 2)$  as a linear combination of  $\vec{v}_1, \vec{v}_2$  using the orthogonality formul in the above theorem.

$$a_1 = \frac{\langle \vec{v}, \vec{v}_1 \rangle}{|\vec{v}_1|^2}, \quad a_2 = \frac{\langle \vec{v}, \vec{v}_2 \rangle}{|\vec{v}_2|^2}.$$

$$\langle \vec{v}, \vec{v}_1 \rangle = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 7. \quad \langle \vec{v}, \vec{v}_2 \rangle = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = -5.$$

$$|\vec{v}_1|^2 = \langle \vec{v}_1, \vec{v}_1 \rangle = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3. \quad |\vec{v}_2|^2 = \begin{bmatrix} -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = 15.$$

$$\text{So, } a_1 = \frac{7}{3}, \quad a_2 = \frac{-5}{15} \quad \text{and} \quad \vec{v} = \frac{7}{3} \vec{v}_1 - \frac{1}{3} \vec{v}_2.$$

c) Verify the norm formula in the above theorem for  $|\vec{v}|$ .

$$|\vec{v}|^2 = \sum_{i=1}^n a_i^2 |\vec{v}_i|^2 = \sum_{i=1}^n \left( \frac{\langle \vec{v}, \vec{v}_i \rangle}{|\vec{v}_i|} \right)^2 = \left( \frac{\langle \vec{v}, \vec{v}_1 \rangle}{|\vec{v}_1|} \right)^2 + \left( \frac{\langle \vec{v}, \vec{v}_2 \rangle}{|\vec{v}_2|} \right)^2 = \left( \frac{7}{\sqrt{3}} \right)^2 + \left( \frac{-5}{\sqrt{15}} \right)^2 = 18.$$

Alternatively:  $|\vec{v}|^2 = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 18.$  ✓

d) Find an orthogonal basis  $\vec{u}_1, \vec{u}_2$  for this inner product.

$$\vec{u}_1 = \frac{\vec{v}_1}{|\vec{v}_1|} = \frac{(1,1)}{\sqrt{3}} = \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right). \quad \vec{u}_2 = \frac{\vec{v}_2}{|\vec{v}_2|} = \frac{(-2,1)}{\sqrt{15}} = \left( -\frac{2}{\sqrt{15}}, \frac{1}{\sqrt{15}} \right).$$

e) Write  $\vec{v}$  as a linear combination of the orthonormal basis, and verify

$$|\vec{v}| = \sqrt{c_1^2 + \dots + c_n^2} = \sqrt{\sum_{i=1}^n \langle \vec{v}, \vec{u}_i \rangle^2} \text{ from the first theorem in this section.}$$

$$\langle \vec{v}, \vec{u}_1 \rangle = \left\langle (3, 2), \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) \right\rangle = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} = \frac{7\sqrt{3}}{3}.$$

$$\langle \vec{v}, \vec{u}_2 \rangle = \left\langle (3, 2), \left( -\frac{2}{\sqrt{15}}, \frac{1}{\sqrt{15}} \right) \right\rangle = \begin{bmatrix} 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} -\frac{2}{\sqrt{15}} \\ \frac{1}{\sqrt{15}} \end{bmatrix} = -\frac{\sqrt{15}}{3}.$$

$$\vec{v} = c_1 \vec{u}_1 + c_2 \vec{u}_2 = \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}, \vec{u}_2 \rangle \vec{u}_2 = \frac{7\sqrt{3}}{3} \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right) - \frac{\sqrt{15}}{3} \left( -\frac{2}{\sqrt{15}}, \frac{1}{\sqrt{15}} \right) = \begin{bmatrix} 3 & 2 \end{bmatrix}. \quad \checkmark$$

$$|\vec{v}| = \sqrt{\sum_{i=1}^n \langle \vec{v}, \vec{u}_i \rangle^2} = \sqrt{\langle \vec{v}, \vec{u}_1 \rangle^2 + \langle \vec{v}, \vec{u}_2 \rangle^2}$$

So,  $|\vec{v}| = \sqrt{\left( \frac{7\sqrt{3}}{3} \right)^2 + \left( -\frac{\sqrt{15}}{3} \right)^2} = \sqrt{18}.$  ✓ (matches the result in part c)