Learning Objectives:

1. Define orthogonal set, orthogonal projection, orthonormal basis, and orthogonal matrix.
2. Represent vectors as a linear combination of orthogonal basis vectors.
3. Explain properties of matrices with orthonormal columns.

1 Orthogonal sets

Definition: A set $\{\mathbf{v}_1, \dots \mathbf{v}_p\}$ is said to be an **orthogonal set** if each pair of distinct vectors is orthogonal: $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for $i \neq j$.

Definition: We say that a set $\{u_1, \dots u_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors (all have norm 1).

Example 1. *Show that the set* $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ *is orthogonal, where*

$$\mathbf{u}_1 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \ \mathbf{u}_2 = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \ \mathbf{u}_3 = \begin{pmatrix} -1/2 \\ -2 \\ 7/2 \end{pmatrix}.$$

Is the set orthonormal?

Solution. We can compute

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = (-3) + (2) + (1) = 0,$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = (1/2) + (-4) + (7/2) = 0,$$

and

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = (-3/2) + (-2) + (7/2) = 0.$$

Thus all the vectors are pairwise orthogonal and thus this forms an orthogonal set. This set is *not* orthonormal because the vectors are not all length 1. For example we see

$$\|\mathbf{u}_1\| = \sqrt{3^2 + 1^2 + 1^2} = \sqrt{11}.$$

We could make this set orthonormal by normalizing all of the vectors! Computing

$$\|\mathbf{u}_2\| = \sqrt{6}$$

and

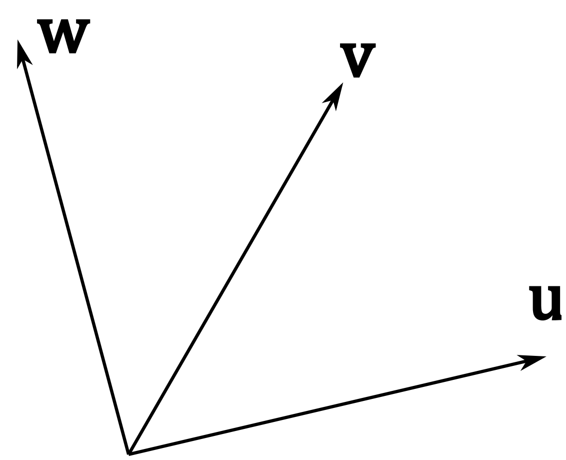
$$\|\mathbf{u}_3\| = \sqrt{33/2},$$

we thus see that the new set

$$\left\{ \begin{pmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \begin{pmatrix} -1/(2\sqrt{33/2}) \\ -2/\sqrt{33/2} \\ 7/(2\sqrt{33/2}) \end{pmatrix} \right\}$$

is orthonormal.

Intuition: Linearly independent sets are sets where no vectors are in the span of the remaining vectors. Intuitively, this means that all vectors "point in different directions." A set of orthogonal vectors is the **extreme** version of this: we not only want vectors to point in different directions, we want them to all be perpendicular. In some sense, they are pointing in *the most different directions* possible. For example: in the picture below we see that the set $\{\mathbf{u}, \mathbf{v}\}$ is linearly independent, but it is not orthogonal. On the other hand, the set $\{\mathbf{u}, \mathbf{w}\}$ is not just linearly independent, it is orthogonal! Thus, *orthogonality is an even stronger requirement!*



Theorem: If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and thus is a basis for the subspace spanned by S.

Proof: Suppose that the set is linearly dependent. Then, there is a non-trivial set of weights so that

$$0 = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p.$$

Without loss of generality assume that $c_1 \neq 0$. Then, take the inner product of both sides with $\mathbf{u}_1$:

$$0 = c_1 \|\mathbf{u}_1\|^2,$$

where all other terms vanish due to orthogonality. However this implies that $\mathbf{u}_1 = 0$, a contradiction.

Definition: An **orthogonal basis** for a subspace W is a basis for W that is also an orthogonal set. An **orthonormal basis** is a basis for W that is also an orthonormal set.

Example 2. *Why are orthogonal bases easier to work with? Consider the following bases of $\mathbb{R}^2$:*

$$\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\},$$

and

$$\left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Write $\mathbf{y} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}$ *as a linear combination of the basis vectors in each case.*

Solution. In the first case we must solve

$$c_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}$$

so we set up the augmented matrix and row reduce:

$$\begin{pmatrix} 1 & 2 & -5 \\ 2 & 1 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \end{pmatrix},$$

and so

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} - 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}.$$

On the other hand, it is pretty easy to see that

$$c_1 \begin{pmatrix} 2 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}$$

must mean

$$-5/2 \begin{pmatrix} 2 \\ 0 \end{pmatrix} - 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -5 \\ -1 \end{pmatrix}.$$

Remark: In general, if a set is orthogonal, then it is much easier to represent other vectors as linear combinations!

Theorem: Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for W. For each $y \in W$, we have

$$y = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p,$$

where

$$c_i = \frac{y \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}.$$

Example 3. *Take* $\mathbf{u}_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ *and* $\mathbf{u}_2 = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$. *Write* $\mathbf{x} = \begin{pmatrix} 9 \\ -7 \end{pmatrix}$ *as a linear combination of the* $\mathbf{u}_i$.

Solution. First, since $\mathbf{u}_1 \cdot \mathbf{u}_2 = 0$ we see that this forms an orthogonal set. Thus we can write

$$\mathbf{x} = \frac{u_1 \cdot x}{u_1 \cdot u_1}u_1 + \frac{u_2 \cdot x}{u_2 \cdot u_2}u_2.$$

We have

$$\frac{u_1 \cdot x}{u_1 \cdot u_1} = \frac{18+21}{4+9} = \frac{39}{13} = 3$$

and

$$\frac{u_2 \cdot x}{u_2 \cdot u_2} = \frac{54-28}{36+16} = \frac{26}{52} = \frac{1}{2}.$$

2 Orthogonal matrices

Example 4. *Suppose* $\{\mathbf{v}_1, \cdots, \mathbf{v}_3\}$ *is an orthonormal set of vectors from* $\mathbb{R}^3$. *Let* U *have columns formed by these vectors. What is*

$$U^TU?$$

Solution. The ij entry of U^TU is given by

$$\mathbf{v}_i \cdot \mathbf{v}_j,$$

which is either 0 or 1 (if $i = j$). Thus $U^TU = I_3$.

Definition: An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$. If U is square then we say U is **orthogonal**.

Theorem: Let U be an $m \times n$ matrix with orthonormal columns and $x, y \in \mathbb{R}^n$. Then

1. $\|Ux\| = \|x\|$
2. $(Ux) \cdot (Uy) = x \cdot y$
3. $(Ux) \cdot (Uy) = 0$ if and only if $x \cdot y = 0$.

Proof. Proving part 2 proves everything. Note that

$$(Ux) \cdot (Uy) = (Ux)^T(Uy) = x^T U^T U y = x^T y = x \cdot y.$$

Example 5. T/F: *An orthogonal matrix is necessarily invertible*

Solution. True! Since $U^T U = I$ then $U^T = U^{-1}$ and the matrix is always invertible.

Example 6. *Let $\theta = \pi/13$. Is*

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

orthogonal?

Solution. Yes! Write $x_1 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, and $x_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$. Then

$$x_1 \cdot x_2 = 0$$

and

$$x_1 \cdot x_1 = \cos^2\theta + \sin^2\theta = 1$$

and similarly for x_2. Thus the columns are orthonormal.