Orthogonal Sets

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an *orthogonal set* if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$ for all $\mathbf{u}_i$ and $\mathbf{u}_j$ in $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. That is, the set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is an orthogonal set if each pair of distinct $\mathbf{u}_i$ and $\mathbf{u}_j$ in $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ are orthogonal to each other.

Example. Consider $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$. Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set.

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

The three possible pairs of distinct vectors in the given set are $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$, where $\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1)+1(2)+1(1) = 0$, $\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\frac{1}{2}\right)+1(-2)+1\left(\frac{7}{2}\right) = 0$, and $\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$. Thus, since each pair of distinct vectors from $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal, it follows that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. ■

Suppose that $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors $\mathbf{u}_1, \ldots, \mathbf{u}_p \in \mathbb{R}^n$. Observe that if $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for suitable scalars $c_1, \ldots, c_p$, then the inner product $\mathbf{0} \cdot \mathbf{u}_i = 0$ can be expressed as $\mathbf{0} \cdot \mathbf{u}_i = (c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_i = 0$ or, equivalently, $\mathbf{0} \cdot \mathbf{u}_i = c_1(\mathbf{u}_1 \cdot \mathbf{u}_i) + \cdots + c_i(\mathbf{u}_i \cdot \mathbf{u}_i) + \cdots + c_p(\mathbf{u}_p \cdot \mathbf{u}_i) = 0$. Also observe that since $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$ for all $\mathbf{u}_i, \mathbf{u}_j \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, it follows that $\mathbf{0} \cdot \mathbf{u}_i = c_i(\mathbf{u}_i \cdot \mathbf{u}_i) = 0$ and, since every $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is nonzero, it also follows that $\mathbf{u}_i \cdot \mathbf{u}_i \neq 0$ and hence $c_i = 0$. Therefore, the set S is linearly independent.

> **Theorem** (Linear Independence of Orthogonal Set).
>
> If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors $\mathbf{u}_1, \ldots, \mathbf{u}_p \in \mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

The definition of an *orthogonal basis* follows directly from the preceding theorem.

> **Definition** (Orthogonal Basis).
>
> The **orthogonal basis** for a subspace $W \subseteq \mathbb{R}^n$ is a basis for W that is also an orthogonal set. That is, if $W = \operatorname{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, where $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W.

Suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for a subspace $W \subseteq \mathbb{R}^n$. Then $W = \operatorname{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ and thus each $\mathbf{y} \in W$ can be expressed as $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for suitable scalars $\{c_1, \ldots, c_p\}$. Observe that since $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set, $\mathbf{y} \cdot \mathbf{u}_i$ can be expressed as $\mathbf{y} \cdot \mathbf{u}_i = (c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_i = c_i(\mathbf{u}_i \cdot \mathbf{u}_i)$ and, since $\mathbf{u}_i \neq 0$ for all $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, it follows that $\mathbf{u}_i \cdot \mathbf{u}_i \neq 0$ and hence the equation $\mathbf{y} \cdot \mathbf{u}_i = c_i(\mathbf{u}_i \cdot \mathbf{u}_i)$ can be used to obtain each weight c_i in $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$.

$$\mathbf{y} \cdot \mathbf{u}_i = c_i(\mathbf{u}_i \cdot \mathbf{u}_i) \Longrightarrow c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \tag{1}$$

Theorem (Orthogonal Basis Linear Combination Weights).

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace $W \subseteq \mathbb{R}^n$. For each $\mathbf{y} \in W$, the weights $\{c_1, \ldots, c_p\}$ of $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ are given by $c_i = (\mathbf{y} \cdot \mathbf{u}_i) \,/\, (\mathbf{u}_i \cdot \mathbf{u}_i)$.

Example. Consider the vectors $\mathbf{u}_1 = \langle 3, 1, 1 \rangle$, $\mathbf{u}_2 = \langle -1, 2, 1 \rangle$, and $\mathbf{u}_3 = \langle -\frac{1}{2}, -2, \frac{7}{2} \rangle$. Recall from the preceding example that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbb{R}^3$. Express the vector $\mathbf{y} = \langle 6, 1, -8 \rangle$ as a linear combination of the vectors $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

Note that $(\mathbf{y} \cdot \mathbf{u}_1 / \mathbf{u}_1 \cdot \mathbf{u}_1) = \frac{11}{11}$, $(\mathbf{y} \cdot \mathbf{u}_2 / \mathbf{u}_2 \cdot \mathbf{u}_2) = -\frac{12}{6}$, and $(\mathbf{y} \cdot \mathbf{u}_3 / \mathbf{u}_3 \cdot \mathbf{u}_3) = -\frac{33}{33/2}$. Therefore, by the preceding theorem, $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3$ is defined by the following.

$$\begin{aligned}
\mathbf{y} &= c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3 \\
&= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3}\mathbf{u}_3 = \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\
&= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3
\end{aligned}$$

■

An Orthogonal Projection

Given some nonzero $\mathbf{u} \in \mathbb{R}^n$, let $\mathbf{y} \in \mathbb{R}^n$ be defined by $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z} \in \mathbb{R}^n$ is orthogonal to $\mathbf{u}$ (that is, $\mathbf{u} \cdot \mathbf{z} = 0$). Observe that if α is any scalar, then $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \alpha\mathbf{u} + \mathbf{z}$ implies that $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \alpha\mathbf{u}$ is orthogonal to $\mathbf{u}$ if and only if $\mathbf{z} \cdot \mathbf{u} = (\mathbf{y} - \hat{\mathbf{y}}) \cdot \mathbf{u} = (\mathbf{y} - \alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - (\alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \alpha(\mathbf{u} \cdot \mathbf{u}) = 0$ and, therefore, the equation $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \alpha\mathbf{u} + \mathbf{z}$ is satisfied with $\mathbf{z}$ orthogonal to $\mathbf{u}$ if and only if $\alpha = (\mathbf{y} \cdot \mathbf{u}) \,/\, (\mathbf{u} \cdot \mathbf{u})$ and $\hat{\mathbf{y}} = [(\mathbf{y} \cdot \mathbf{u}) \,/\, (\mathbf{u} \cdot \mathbf{u})]\, \mathbf{u}$. The *orthogonal projection* of $\mathbf{y}$ onto $\mathbf{u}$ is the vector $\hat{\mathbf{y}}$ and *component* of $\mathbf{y}$ orthogonal $\mathbf{u}$ is the vector $\mathbf{z}$.

Observe that if c is any nonzero scalar, then by replacing $\mathbf{u}$ with $c\mathbf{u}$ in the definition of $\hat{\mathbf{y}}$, it follows that $\hat{\mathbf{y}} = [(\mathbf{y} \cdot c\mathbf{u}) \,/\, (c\mathbf{u} \cdot c\mathbf{u})]\, c\mathbf{u} = [c^2\,(\mathbf{y} \cdot \mathbf{u}) \,/ c^2\,(\mathbf{u} \cdot \mathbf{u})]\, \mathbf{u} = [(\mathbf{y} \cdot \mathbf{u}) \,/\, (\mathbf{u} \cdot \mathbf{u})]\, \mathbf{u} = \hat{\mathbf{y}}$ and hence $\hat{\mathbf{y}}$ is the *orthogonal projection* of $\mathbf{y}$ onto a subspace $L = \text{Span}\{\mathbf{u}\}$, rather than $\mathbf{u}$ itself.

$$\hat{\mathbf{y}} = \text{proj}_L(\mathbf{y}) = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} \tag{2}$$

Example. Let $\mathbf{y} = \langle 7, 6 \rangle$ and $\mathbf{u} = \langle 4, 2 \rangle$. Find $\hat{\mathbf{y}}$ and then express $\mathbf{y}$ as the sum of two orthogonal vectors such that one vector is an element of $L = \text{Span}\{\mathbf{u}\}$ and the other is orthogonal to $\mathbf{u}$. Finally, find the distance from $\mathbf{y}$ to L.

First note that $\mathbf{y} \cdot \mathbf{u} = 7(4) + 6(2) = 40$ and $\mathbf{u} \cdot \mathbf{u} = (4)^2 + (2)^2 = 20$ and thus $\hat{\mathbf{y}}$, the orthogonal projection of $\mathbf{y}$ onto L, and $\mathbf{z}$, the component of $\mathbf{y}$ orthogonal to $\mathbf{u}$, are defined by the following.

$$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}}\mathbf{u} = \frac{40}{20}\mathbf{u} = 2\begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} \qquad \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 7 \\ 6 \end{bmatrix} - \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

The sum of $\hat{\mathbf{y}}$ and $\mathbf{z}$ is $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, defined below.

$$\begin{bmatrix} 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

■

Orthonormal Sets

A set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is *orthonormal* if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of unit vectors. That is, the set $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is orthonormal if $\|\mathbf{u}_i\| = 1$ for all $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ and $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$ for all $\mathbf{u}_i, \mathbf{u}_j \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. Observe that if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is orthonormal, then $\|\mathbf{u}_i\| = 1$ for all $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ and hence $\|\mathbf{u}_i\|^2 = \mathbf{u}_i \cdot \mathbf{u}_i = 1$ for all $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, which further implies that $\mathbf{u}_i \neq \mathbf{0}$ for all $\mathbf{u}_i \in \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$. Therefore, if $W = \operatorname{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, then $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an *orthonormal basis* for W because $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors.

Example. Consider the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ defined below. Show that the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$.

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix} \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

First note that $\mathbf{v}_1 \cdot \mathbf{v}_2 = -\frac{3}{\sqrt{66}} + \frac{2}{\sqrt{66}} + \frac{1}{\sqrt{66}} = 0$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = -\frac{3}{\sqrt{726}} - \frac{4}{\sqrt{726}} + \frac{7}{\sqrt{726}} = 0$, and $\mathbf{v}_2 \cdot \mathbf{v}_3 = \frac{1}{\sqrt{396}} - \frac{8}{\sqrt{396}} + \frac{7}{\sqrt{396}} = 0$, so the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthogonal. To show that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors, it suffices to show that $\mathbf{v}_1 \cdot \mathbf{v}_1 = \frac{9}{11} + \frac{1}{11} + \frac{1}{11} = 1$, $\mathbf{v}_2 \cdot \mathbf{v}_2 = \frac{1}{6} + \frac{4}{6} + \frac{1}{6} = 1$, and $\mathbf{v}_3 \cdot \mathbf{v}_3 = \frac{1}{66} + \frac{16}{66} + \frac{49}{66} = 1$. Thus, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set and, since $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent, it follows that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$. ■