Orthogonal Bases and the Gram-Schmidt Process

DEFINITION: A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0$$

for any $i \neq j$. A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthonormal set** if it is an orthogonal set of unit vectors.

EXAMPLE: Let

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0 \\ 8 \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

Then $\{\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3\}$ is an orthogonal set.

EXAMPLE: Let

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

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

Solution: We have

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

$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\frac{1}{2}\right) + 1(-2) + 1\left(\frac{7}{2}\right) = 0$$

$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\frac{1}{2}\right) + 2(-2) + 1\left(\frac{7}{2}\right) = 0$$

THEOREM 6.1: Let $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal set of nonzero vectors in $\mathbb{R}^n$. Then S is a linearly independent set.

COROLLARY 6.2: If S is an orthogonal set of n nonzero vectors in $\mathbb{R}^n$, then S is an orthogonal basis for $\mathbb{R}^n$. Similarly, if S is an orthonormal set of n vectors in $\mathbb{R}^n$, then S is an orthonormal basis for $\mathbb{R}^n$.

EXAMPLE: Let

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Then $\{\mathbf{e}_1,\ \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is the orthonormal basis for $\mathbb{R}^n$.

EXAMPLE: Let $\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3$ be the same as above. Then $S = \{\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3\}$ is an orthogonal basis for $\mathbb{R}^3$ and

$$\mathbf{w}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{11}}\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$$

$$\mathbf{w}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = \sqrt{\frac{2}{33}}\begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix} = \frac{1}{\sqrt{66}}\begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

is the orthonormal basis for $\mathbb{R}^3$.

EXAMPLE: Let $\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3$ be the same as above. Find coordinates of $\mathbf{y} = (6, 1, -8)$ in $S = \{\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3\}$.

Solution: We have

$$\begin{bmatrix} 3 & -1 & -1/2 & 6 \\ 1 & 2 & -2 & 1 \\ 1 & 1 & 7/2 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -2 & 1 \\ 3 & -1 & -1/2 & 6 \\ 1 & 1 & 7/2 & -8 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & -7 & 5.5 & 3 \\ 0 & -1 & 5.5 & -9 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & -1 & 5.5 & -9 \\ 0 & -7 & 5.5 & 3 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & -1 & 5.5 & -9 \\ 0 & 0 & -33 & 66 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & 1 & -5.5 & 9 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

THEOREM 6.3: Let $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W the weights in the linear combination

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

are given by

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad (j = 1, \dots, p)$$

Proof: Let $c_1, \dots, c_p$ be such numbers that

$$\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \tag{$*$}$$

If we multiply both sides of $(*)$ by $\mathbf{u}_1$, we get

$$\begin{aligned}
\mathbf{y} \cdot \mathbf{u}_1 &= c_1\mathbf{u}_1 \cdot \mathbf{u}_1 + c_2\mathbf{u}_2 \cdot \mathbf{u}_1 + \ldots + c_p\mathbf{u}_p \cdot \mathbf{u}_1 \\
&= c_1\mathbf{u}_1 \cdot \mathbf{u}_1 + 0 + \ldots + 0 \\
&= c_1\mathbf{u}_1 \cdot \mathbf{u}_1
\end{aligned}$$

because of orthogonality of $\mathbf{u}_1, \dots, \mathbf{u}_p$. So, $\mathbf{y} \cdot \mathbf{u}_1 = c_1\mathbf{u}_1 \cdot \mathbf{u}_1$ therefore

$$c_1 = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}$$

Similarly, if we multiply both sides of $(*)$ by $\mathbf{u}_j$, we deduce

$$c_j = \frac{\mathbf{y} \cdot \mathbf{u}_j}{\mathbf{u}_j \cdot \mathbf{u}_j} \quad (j = 1, \dots, p).$$

EXAMPLE: Let $S = \{\mathbf{u}_1,\ \mathbf{u}_2,\ \mathbf{u}_3\}$, where

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

Find coordinates of $\mathbf{y} = (6, 1, -8)$ in S.

Solution: We have

$$\mathbf{y} \cdot \mathbf{u}_1 = 11, \quad \mathbf{y} \cdot \mathbf{u}_2 = -12, \quad \mathbf{y} \cdot \mathbf{u}_3 = -33$$

and

$$\mathbf{u}_1 \cdot \mathbf{u}_1 = 11, \quad \mathbf{u}_2 \cdot \mathbf{u}_2 = 6, \quad \mathbf{u}_3 \cdot \mathbf{u}_3 = 33/2$$

so

$$\begin{aligned}
c_1 &= \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \frac{11}{11} = 1 \\
c_2 &= \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \frac{-12}{6} = -2 \\
c_3 &= \frac{\mathbf{y} \cdot \mathbf{u}_3}{\mathbf{u}_3 \cdot \mathbf{u}_3} = \frac{-33}{33/2} = -2
\end{aligned}$$

therefore

$$[\mathbf{x}]_S = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}$$

THEOREM 6.4(The Gram-Schmidt Process): Given an arbitrary basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\ldots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \ldots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W.

EXAMPLE: Let

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

is the basis for a subspace W of $\mathbb{R}^4$. Find an orthogonal basis for W.

Solution:

Step 1: Put

$$\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Step 2: Put

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

Step 3: Put

$$\begin{aligned}
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}
\end{aligned}$$

Theorem 6.5 Let $\mathcal{W}$ be a subspace of $\mathbb{R}^n$. Then any orthogonal set of nonzero vectors in $\mathcal{W}$ is contained in (can be enlarged to) an orthogonal basis for $\mathcal{W}$. Similarly, any orthonormal set of vectors in $\mathcal{W}$ is contained in an orthonormal basis for $\mathcal{W}$.

Orthogonal Matrices

Definition A nonsingular (square) matrix $\mathbf{A}$ is **orthogonal** if and only if $\mathbf{A}^T = \mathbf{A}^{-1}$.

Theorem 6.6 If $\mathbf{A}$ and $\mathbf{B}$ are orthogonal matrices of the same size, then

(1) $|\mathbf{A}| = \pm 1$,

(2) $\mathbf{A}^T = \mathbf{A}^{-1}$ is orthogonal, and

(3) $\mathbf{AB}$ is orthogonal.

Proof:

(1) We have

$$|A| = |A^T| = |A^{-1}| = \frac{1}{|A|} \quad \Longrightarrow \quad |A|^2 = 1 \quad \Longrightarrow \quad |A| = \pm 1$$

(2) We have

$$(A^T)^T = (A^{-1})^T = (A^T)^{-1}$$

therefore A^T is orthogonal. Since $A^T = A^{-1}$, the matrix A^{-1} is also orthogonal.

(3) We have

$$\begin{aligned} AB(AB)^T = AB(B^T A^T) &= A(BB^T)A^T \\ &= A(BB^{-1})A^T = AIA^T = AA^T = AA^{-1} = I \end{aligned}$$

hence

$$(AB)^T = (AB)^{-1}$$

therefore AB is orthogonal.

Theorem 6.7 Let $\mathbf{A}$ be an $n \times n$ matrix. Then $\mathbf{A}$ is orthogonal

(1) if and only if the rows of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$

(2) if and only if the columns of $\mathbf{A}$ form an orthonormal basis for $\mathbb{R}^n$.

Theorem 6.8 Let B and C be ordered orthonormal bases for $\mathbb{R}^n$. Then the transition matrix from B to C is an orthogonal matrix.