Section 6.2: Orthogonal Sets

Definition. A set of vectors $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\}$ is said to be an orthogonal set if each vector is orthogonal to others, i.e., $\vec{u}_i \perp \vec{u}_j$ for any $i \neq j$.

Example 6.2.1. Show that (a) in R^n, the standard basis $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ is an orthogonal set, and (b) the following set is an orthogonal set:

$$\vec{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

Solution. (a) $\vec{e}_i \cdot \vec{e}j = 0$ if $i \neq j$. (b) we check by direct calculations:

$$\vec{u}_1 \cdot \vec{u}_2 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = 0,$$

$$\vec{u}_1 \cdot \vec{u}_3 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} = 0$$

$$\vec{u}_2 \cdot \vec{u}_3 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix} = 0.$$

Theorem. Any orthogonal set is linearly independent.

Proof. Suppose $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\}$ is an orthogonal set, and suppose

$$c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p = \vec{0}.$$

We dot-multiplying $\vec{u}_i$ on both sides of the equation and obtain

$$c_i\vec{u}_i \cdot \vec{u}_i = 0 \Longrightarrow c_i = 0.$$

■

Definition. A basis of a subspace is said to be an orthogonal basis if it is an orthogonal set.

Theorem. Let $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\}$ be an orthogonal basis for a subspace W. Then, for each $\vec{w} \in W$, its coordinate $[\vec{w}]_{\mathcal{B}}$ relative to this orthogonal basis can be expressed as

$$[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix}, \quad c_i = \frac{\vec{w} \cdot \vec{u}_i}{\|\vec{u}_i\|^2}, \ i = 1, 2, ..., p. \tag{1}$$

In other words,

$$\vec{w} = c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p, \tag{2}$$

Proof. Consider expression (2). All we need to do is to derive formula for c_i in (1). To this end, we dot-multiply (2) by $\vec{u}_i$:

$$\vec{w} \cdot \vec{u}_i = (c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p) \cdot \vec{u}_i = c_i\vec{u}_i \cdot \vec{u}_i$$

since $\vec{u}_j \cdot \vec{u}_i = 0$ unless $j = i$. It follows that

$$\vec{w} \cdot \vec{u}_i = c_i\vec{u}_i \cdot \vec{u}_i \Longrightarrow \ c_i = \frac{\vec{w} \cdot \vec{u}_i}{\|\vec{u}_i\|^2}.$$

■

Example 6.2.2. We know from Example 1 that

$$\vec{u}_1 = \begin{bmatrix}3\\1\\1\end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix}-1\\2\\1\end{bmatrix}, \ \vec{u}_3 = \begin{bmatrix}-1\\-4\\7\end{bmatrix}$$

form an orthogonal basis for R^3. Find the coordinate of

$$\vec{w} = \begin{bmatrix}6\\1\\-8\end{bmatrix} \quad \text{relative to this basis.}$$

Solution. Note that if the basis were not orthogonal, then we have to proceed as follows: solving linear system:

$$c_1\vec{u}_1 + c_2\ \vec{u}_2 + c_3\vec{u}_3 = \vec{w}$$

or

$$\begin{bmatrix}3 & -1 & -1\\1 & 2 & -4\\1 & 1 & 7\end{bmatrix}\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = \begin{bmatrix}6\\1\\-8\end{bmatrix}.$$

Since the basis is indeed orthogonal, we use formula (1):

$$c_1 = \frac{\vec{w} \cdot \vec{u}_1}{\|\vec{u}_1\|^2} = \begin{bmatrix}3\\1\\1\end{bmatrix} \cdot \begin{bmatrix}6\\1\\-8\end{bmatrix} \div \left\|\begin{bmatrix}3\\1\\1\end{bmatrix}\right\|^2 = 1$$

$$c_2 = \frac{\vec{w} \cdot \vec{u}_2}{\|\vec{u}_2\|^2} = \begin{bmatrix}-1\\2\\1\end{bmatrix} \cdot \begin{bmatrix}6\\1\\-8\end{bmatrix} \div \left\|\begin{bmatrix}-1\\2\\1\end{bmatrix}\right\|^2 = -2$$

$$c_3 = \frac{\vec{w} \cdot \vec{u}_3}{\|\vec{u}_2\|^3} = \begin{bmatrix}-1\\-4\\7\end{bmatrix} \cdot \begin{bmatrix}6\\1\\-8\end{bmatrix} \div \left\|\begin{bmatrix}-1\\-4\\7\end{bmatrix}\right\|^2 = -1$$

$$[\vec{w}]_{\{\vec{u}_1,\vec{u}_2,\vec{u}_3\}} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

Orthogonal Projections.

Given a vector $\vec{u}$, the **orthogonal projection** of $\vec{y}$ onto $\vec{u}$, denoted by $\hat{y} = \text{Proj}_{\vec{u}}(\vec{y})$, is defined as the vector parallel to $\vec{u}$ such that

$$\vec{y} = \hat{y} + \vec{z}, \quad \vec{z} \perp \vec{u}, \ \hat{y}//\vec{u}.$$

Since $\hat{y}$ is parallel to $\vec{u}$, we have $\hat{y} = \alpha\vec{u}$. Hence

$$\vec{y} = \alpha\vec{u} + \vec{z}, \quad \vec{z} \perp \vec{u}.$$

Dot-multiplying by $\vec{u}$, we find

$$\vec{y}\cdot\vec{u} = (\alpha\vec{u} + \vec{z})\cdot\vec{u} = \alpha\vec{u}\cdot\vec{u} \Longrightarrow \alpha = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}.$$

$$\text{Proj}_{\vec{u}}(\vec{y}) = \alpha\vec{u} = \left(\frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\right)\vec{u} \tag{3}$$

In general, for any given subspace W, $\hat{y} = \text{Proj}_W(\vec{y})$ is defined as the vector in W such that

$$(\vec{y} - \hat{y}) \ \perp W.$$

In other words, any vector $\vec{y}$ can be decomposed into two components: one is the projection $\hat{y}$ on W (which is in W) and another component perpendicular to W. Suppose that W has an orthogonal basis $\mathcal{B}=\{\vec{u}_1,\vec{u}_2,...,\vec{u}_p\}$. Then we may write, since $\hat{y} \in W$,

$$\begin{aligned}\hat{y} &= Proj_W(\vec{y}) = c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p \\ \vec{y} &= \hat{y} + \vec{z} = c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p + \vec{z}, \quad \vec{z} \perp W.\end{aligned}$$

By dot-multiplying by $\vec{u}_i$, we find

$$\vec{y}\cdot\vec{u}_i = (c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p + \vec{z})\cdot\vec{u}_i = c_i\vec{u}_i\cdot\vec{u}_i \Longrightarrow c_i = \frac{\vec{y}\cdot\vec{u}_i}{\vec{u}_i\cdot\vec{u}_i}.$$

Therefore,

$$\hat{y} = Proj_W(\vec{y}) = c_1\vec{u}_1 + c_2\vec{u}_2 + ... + c_p\vec{u}_p, \quad c_i = \frac{\vec{y}\cdot\vec{u}_i}{\vec{u}_i\cdot\vec{u}_i}. \tag{4}$$

Example 6.2.3. Let

$$\vec{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \vec{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \vec{y} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

From Example 1 above, we know that $\vec{u}_1$ and $\vec{u}_2$ form an orthogonal basis for $W = Span\{\vec{u}_1, \vec{u}_2\}$. Find (a) $\text{Proj}_{\vec{u}_1}(\vec{y})$, (b) $\text{Proj}_{\vec{u}_2}(\vec{y})$, (c) $\text{Proj}_W(\vec{y})$.

Solution. (1) By (3),

$$Proj_{\vec{u}_1}(\vec{y}) = \left(\frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\right)\vec{u}_1 = \frac{6+3-1}{11}\begin{bmatrix}3\\1\\1\end{bmatrix} = \frac{8}{11}\begin{bmatrix}3\\1\\1\end{bmatrix}$$

(2) Analogously,

$$Proj_{\vec{u}_2}(\vec{y}) = \left(\frac{\vec{y}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\right)\vec{u}_2 = \frac{-2+6-1}{\sqrt{6}}\begin{bmatrix}-1\\2\\1\end{bmatrix} = \frac{3}{\sqrt{6}}\begin{bmatrix}-1\\2\\1\end{bmatrix}$$

(3) Using (4) and answer from part (1) & (2)

$$Proj_W(\vec{y}) = \left(\frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\right)\vec{u}_1 + \left(\frac{\vec{y}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\right)\vec{u}_2 = Proj_{\vec{u}_1}(\vec{y}) + Proj_{\vec{u}_2}(\vec{y}) = \frac{8}{11}\begin{bmatrix}3\\1\\1\end{bmatrix} + \frac{3}{\sqrt{6}}\begin{bmatrix}-1\\2\\1\end{bmatrix}$$

We conclude from this example that, in general, suppose that W has an orthogonal basis $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\}$. Then

$$Proj_W(\vec{y}) = Proj_{\vec{u}_1}(\vec{y}) + Proj_{\vec{u}_2}(\vec{y}) + ... + Proj_{\vec{u}_p}(\vec{y}).$$

Definition. A set $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_p\}$ is said to be an orthonormal set if it is an orthogonal set and if each vector is a unit vector, i.e.,

$$\vec{u}_i \cdot \vec{u}_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Example 6.2.4. Show that

$$\vec{u}_1 = \frac{1}{\sqrt{11}}\begin{bmatrix}3\\1\\1\end{bmatrix}, \quad \vec{u}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix}-1\\2\\1\end{bmatrix}, \quad \vec{u}_3 = \frac{1}{\sqrt{66}}\begin{bmatrix}-1\\-4\\7\end{bmatrix}$$

form an orthonormal set.

Solution. Direct computation show

$$\vec{u}_1\cdot\vec{u}_1 = \vec{u}_2\cdot\vec{u}_2 = \vec{u}_3\cdot\vec{u}_3 = 1$$
$$\vec{u}_1\cdot\vec{u}_2 = \vec{u}_2\cdot\vec{u}_3 = \vec{u}_1\cdot\vec{u}_3 = 0.$$

Theorem. Let $U = [\vec{u}_1, \vec{u}_2, ..., \vec{u}_n]$ be a $n \times n$ matrix with columns $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$. Suppose that the columns of U form an orthonormal set. Then

$$U^{-1} = U^T, \quad i.e., \; UU^T = U^T U = I.$$

We call it **orthonormal matrix.**

Proof. We observe that U may be written as

$$U = \begin{bmatrix} u_{11} & u_{12} & ... & u_{1n} \\ u_{21} & u_{22} & ... & u_{2n} \\ ... & ... & ... & ... \\ u_{n1} & u_{n2} & ... & u_{nn} \end{bmatrix} = [\vec{u}_1, \vec{u}_2, ..., \vec{u}_n], \quad \vec{u}_j = \begin{bmatrix} u_{1j} \\ u_{2j} \\ ... \\ u_{nj} \end{bmatrix}$$

and since $(\vec{u}_1)^T = [u_{11}, u_{21}, ..., u_{n1}]$ is a row-vector,

$$U^T = \begin{bmatrix} u_{11} & u_{21} & ... & u_{n1} \\ u_{12} & u_{22} & ... & u_{n2} \\ ... & ... & ... & ... \\ u_{1n} & u_{2n} & ... & u_{nn} \end{bmatrix} = \begin{bmatrix} (\vec{u}_1)^T \\ (\vec{u}_2)^T \\ ... \\ (\vec{u}_n)^T \end{bmatrix}.$$

Since, by orthonormality, $(\vec{u}_1)^T \vec{u}_1 = \vec{u}_1 \cdot \vec{u}_1 = 1$, $(\vec{u}_2)^T \vec{u}_1 = \vec{u}_2 \cdot \vec{u}_1 = 0$, ... we have

$$U^T U = \begin{bmatrix} (\vec{u}_1)^T \\ (\vec{u}_2)^T \\ ... \\ (\vec{u}_n)^T \end{bmatrix} [\vec{u}_1, \vec{u}_2, ..., \vec{u}_n] = \begin{bmatrix} (\vec{u}_1)^T \vec{u}_1 & (\vec{u}_1)^T \vec{u}_2 & ... & (\vec{u}_1)^T \vec{u}_n \\ (\vec{u}_2)^T \vec{u}_1 & (\vec{u}_2)^T \vec{u}_2 & ... & (\vec{u}_2)^T \vec{u}_n \\ ... & ... & ... & ... \\ (\vec{u}_n)^T \vec{u}_1 & (\vec{u}_n)^T \vec{u}_2 & ... & (\vec{u}_n)^T \vec{u}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

■

Note that the same technique may be used to calculate the inverse of a matrix $A = [u_1, u_2, ..., u_n]$, where the column vectors $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ form an orthogonal set, but not orthonormal set. In this case,

$$A^T A = \begin{bmatrix} (\vec{u}_1)^T \vec{u}_1 & (\vec{u}_1)^T \vec{u}_2 & ... & (\vec{u}_1)^T \vec{u}_n \\ (\vec{u}_2)^T \vec{u}_1 & (\vec{u}_2)^T \vec{u}_2 & ... & (\vec{u}_2)^T \vec{u}_n \\ ... & ... & ... & ... \\ (\vec{u}_n)^T \vec{u}_1 & (\vec{u}_n)^T \vec{u}_2 & ... & (\vec{u}_n)^T \vec{u}_n \end{bmatrix} = \begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & 0 & ... & 0 \\ 0 & \vec{u}_2 \cdot \vec{u}_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \vec{u}_n \cdot \vec{u}_n \end{bmatrix}.$$

So

$$\begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & 0 & ... & 0 \\ 0 & \vec{u}_2 \cdot \vec{u}_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \vec{u}_n \cdot \vec{u}_n \end{bmatrix}^{-1} A^T A = I,$$

i.e.,

$$A^{-1} = \begin{bmatrix} \vec{u}_1 \cdot \vec{u}_1 & 0 & ... & 0 \\ 0 & \vec{u}_2 \cdot \vec{u}_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \vec{u}_n \cdot \vec{u}_n \end{bmatrix}^{-1} A^T$$

$$= \begin{bmatrix} \frac{1}{\vec{u}_1 \cdot \vec{u}_1} & 0 & ... & 0 \\ 0 & \frac{1}{\vec{u}_2 \cdot \vec{u}_2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \frac{1}{\vec{u}_n \cdot \vec{u}_n} \end{bmatrix} \begin{bmatrix} (\vec{u}_1)^T \\ (\vec{u}_2)^T \\ ... \\ (\vec{u}_n)^T \end{bmatrix} = \begin{bmatrix} \frac{(\vec{u}_1)^T}{\vec{u}_1 \cdot \vec{u}_1} \\ \frac{(\vec{u}_2)^T}{\vec{u}_2 \cdot \vec{u}_2} \\ ... \\ \frac{(\vec{u}_n)^T}{\vec{u}_n \cdot \vec{u}_n} \end{bmatrix}.$$

Example 6.2.5. We know from previous examples that

$$\vec{u}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \vec{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{u}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

form an orthonormal basis, but

$$\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

form only an orthogonal basis. Set

$$U = [\vec{u}_1, \vec{u}_2, \vec{u}_3] = \begin{bmatrix} \frac{3}{\sqrt{11}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & \frac{-4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}$$

$$V = [\vec{v}_1, \vec{v}_2, \vec{v}_3] = \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & -4 \\ 1 & 1 & 7 \end{bmatrix}.$$

The first matrix U is an orthogonal matrix, and

$$U^{-1} = U^T = \begin{bmatrix} \frac{3}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} \\ -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{66}} & -\frac{4}{\sqrt{66}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

The matrix is not an orthonormal matrix. However,

$$V^T V = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & -4 & 7 \end{bmatrix} \begin{bmatrix} 3 & -1 & -1 \\ 1 & 2 & -4 \\ 1 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 11 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 66 \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} 11 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 66 \end{bmatrix}^{-1} V^T V = I,$$

or

$$\begin{aligned} V^{-1} &= \begin{bmatrix} 11 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 66 \end{bmatrix}^{-1} V^T \\ &= \begin{bmatrix} 11^{-1} & 0 & 0 \\ 0 & 6^{-1} & 0 \\ 0 & 0 & 66^{-1} \end{bmatrix} \begin{bmatrix} 3 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & -4 & 7 \end{bmatrix} \\ &= \begin{bmatrix} \frac{3}{11} & \frac{1}{11} & \frac{1}{11} \\ -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{66} & -\frac{2}{33} & \frac{7}{66} \end{bmatrix}. \end{aligned}$$