§6.2 Orthogonal Sets

A set of vectors in $\mathbb{R}^n$ is an orthogonal set if each pair of distinct vectors from the set is orthogonal.

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

<u>Proof</u>: If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \dots, c_p$, then

$$\begin{aligned} 0 &= \mathbf{0} \cdot \mathbf{u}_1 \\ &= (c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p) \cdot \mathbf{u}_1 \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p \cdot \mathbf{u}_1) \\ &= c_1(\mathbf{u}_1 \cdot \mathbf{u}_1) \end{aligned}$$

since $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \dots, \mathbf{u}_p$. Because $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero and so $c_1 = 0$. It follows that $c_2, \dots, c_p$ must be zero as well. Hence S is linearly independent.■

An orthogonal basis for a subspace W of $\mathbb{R}^n$ is a basis for W that is also an orthogonal set.

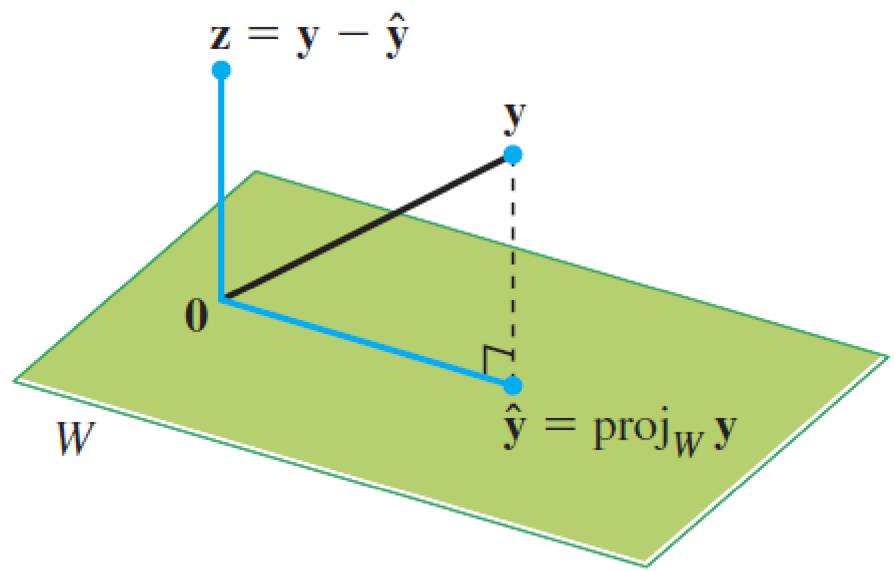
Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination $\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ are given by $c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \quad (j = 1, \ldots, p).$

<ins>Proof</ins>: The orthogonality of $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ implies that

$$\begin{aligned} \mathbf{y}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1). \end{aligned}$$

Since $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero, the previous equation can be solved for c_1. The argument can be extended to find c_j for $j = 2, \ldots, p$.$\blacksquare$

Let $\mathbf{u}$ be a nonzero vector in $\mathbb{R}^n$. We want to decompose a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors, one a multiple of $\mathbf{u}$ and the other orthogonal to $\mathbf{u}$. That is, $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$ where $\hat{\mathbf{y}} = \alpha\mathbf{u}$ for some scalar α and $\mathbf{z}$ is some vector orthogonal to $\mathbf{u}$. Now $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\mathbf{u}$ if and only if $0 = (\mathbf{y} - \alpha\mathbf{u}) \cdot \mathbf{u} = \mathbf{y} \cdot \mathbf{u} - \alpha(\mathbf{u} \cdot \mathbf{u})$. It follows that $\alpha = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}$ and $\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$. The vector $\hat{\mathbf{y}}$ is called the orthogonal projection of $\mathbf{y}$ onto $\mathbf{u}$, and the vector $\mathbf{z}$ is called the component of $\mathbf{y}$ orthogonal to $\mathbf{u}$.



This projection is determined by the subspace L spanned by $\mathbf{u}$ (the line through $\mathbf{u}$ and $\mathbf{0}$). Sometimes $\hat{\mathbf{y}}$ is called the orthogonal projection of $\mathbf{y}$ onto L and we write $\hat{\mathbf{y}} = \text{proj}_L \mathbf{y} = \frac{\mathbf{y}\cdot\mathbf{u}}{\mathbf{u}\cdot\mathbf{u}}\mathbf{u}$.

Note that the point identified with $\hat{\mathbf{y}}$ is the closest point of L to $\mathbf{y}$, and the distance from $\mathbf{y}$ to L is the length of the perpendicular line segment from $\mathbf{y}$ to the orthogonal projection $\hat{\mathbf{y}}$, that is, $\|\mathbf{y} - \hat{\mathbf{y}}\|$.

A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal set if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthonormal basis for W.