## 6.3 Orthogonal Projections

Suppose $\mathbf{u}, \mathbf{v} \in V$. We would like to write $\mathbf{u}$ as a scalar multiple of $\mathbf{v}$ plus a vector $\mathbf{w}$ orthogonal to $\mathbf{v}$. To discover how to write $\mathbf{u}$ as a scalar multiple of $\mathbf{v}$ plus a vector orthogonal to $\mathbf{v}$, let $a \in \mathbb{R}$ denote a scalar. Then

$$\mathbf{u} = a\mathbf{v} + (\mathbf{u} - a\mathbf{v}).$$

Thus we need to choose $a$ so that $\mathbf{v}$ is orthogonal to $(\mathbf{u} - a\mathbf{v})$. In other words, we want

$$0 = \langle \mathbf{u} - a\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - a\|\mathbf{v}\|^2.$$

The equation above shows that we should choose $a$ to be $\langle \mathbf{u}, \mathbf{v} \rangle / \|\mathbf{v}\|^2$, provided that $\mathbf{v} \neq \mathbf{0}$. Making this choice of $a$, we can write

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \left( \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right). \tag{1}$$

It is easy to verify that the equation above writes $\mathbf{u}$ as a scalar multiple of $\mathbf{v}$ plus a vector orthogonal to $\mathbf{v}$.

Suppose $U$ is a subspace of $V$. Each vector $\mathbf{v} \in V$ can be written uniquely in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{w},$$

where $\mathbf{u} \in U$ and $\mathbf{w} \in U^\perp$. We use this decomposition to define an operator on $V$, denoted $P_U$ (in the textbook, $\text{proj}_U$), called the *orthogonal projection* of $V$ onto $U$. For $\mathbf{v} \in V$, we define $P_U\mathbf{v}$ to be the vector $\mathbf{u}$ in the decomposition above.

It is easy to verify that $P_U$ is an operator that has the following properties:

- range $P_U = U$;
- null $P_U = U^\perp$;
- $\mathbf{v} - P_U\mathbf{v} \in U^\perp$ for every $\mathbf{v} \in V$;
- $P_U^2 = P_U$;
- $\|P_U\mathbf{v}\| \leq \|\mathbf{v}\|$ for every $\mathbf{v} \in V$.

**Theorem 8** (The Orthogonal Decomposition Theorem)**.** *Let $W$ be a subspace of $V$. Then each $\mathbf{y} \in V$ can be written uniquely in the form*

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{2}$$

*where $\hat{\mathbf{y}}$ is in $W$ and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is any orthogonal basis for $W$, then*

$$\hat{\mathbf{y}} = \frac{\langle \mathbf{y}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \cdots + \frac{\langle \mathbf{y}, \mathbf{u}_p \rangle}{\langle \mathbf{u}_p, \mathbf{u}_p \rangle}\mathbf{u}_p \tag{3}$$

*and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.*

If $\mathbf{y} \in W = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, then $P_W\mathbf{y} = \mathbf{y}$. Furthermore, if $\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ is an orthonormal basis of $W$, then

$$P_W\mathbf{y} = \langle \mathbf{y}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \cdots + \langle \mathbf{y}, \mathbf{e}_m \rangle \mathbf{e}_m \tag{4}$$

for every $\mathbf{y} \in V$.

**Theorem 9** (The Best Approximation Theorem)**.** *Suppose $U$ is a subspace of $V$ and $\mathbf{v} \in V$. Then*

$$\|\mathbf{v} - P_U\mathbf{v}\| \leq \|\mathbf{v} - \mathbf{u}\|$$

*for every $\mathbf{u} \in U$. Furthermore, if $\mathbf{u} \in U$ and the inequality above is an equality, then $\mathbf{u} = P_U\mathbf{v}$.*

In other words, $P_U\mathbf{v}$ is the closest point in $U$ to $\mathbf{v}$. The vector $P_U\mathbf{v}$ is called the best approximation to $\mathbf{v}$ by elements of $U$.

*Proof.* Suppose $\mathbf{u} \in U$. Then

$$\begin{aligned} \|\mathbf{v} - P_U\mathbf{v}\|^2 &\leq \|\mathbf{v} - P_U\mathbf{v}\|^2 + \|P_U\mathbf{v} - \mathbf{u}\|^2 && (5) \\ &= \|(\mathbf{v} - P_U\mathbf{v}) + (P_U\mathbf{v} - \mathbf{u})\|^2 && (6) \\ &= \|\mathbf{v} - \mathbf{u}\|^2, && (7) \end{aligned}$$

where (5) comes from the Pythagorean Theorem, which applies because $\mathbf{v} - P_U\mathbf{v} \in U^{\perp}$ and $P_U\mathbf{v} - \mathbf{u} \in U$. Taking square roots give the desired inequality.

Our inequality is an equality if and only if (4) is an equality, which happens if and only if $\|P_U\mathbf{v} - \mathbf{u}\| = 0$, which happens if and only if $\mathbf{u} = P_U\mathbf{v}$. ☐

The Best Approximation theorem is often combined with the formula (4) to compute explicit solutions to minimization problems.

**Theorem 10.** *If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace $W$ of $\mathbb{R}^n$, then If $U = [\mathbf{u}_1 \quad \cdots \mathbf{u}_p]$, then*

$$P_W\mathbf{y} = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \in \mathbb{R}^n. \tag{8}$$

**Example 1.** *Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ where $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -1 \end{bmatrix}$. Write $\mathbf{y} = \begin{bmatrix} 3 \\ 4 \\ 5 \\ 6 \end{bmatrix}$ as the sum of a vector in $W$ and a vector orthogonal to $W$.*

**Example 2.** *Find the closest point to $\mathbf{y} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ 13 \end{bmatrix}$ in the subspace $W$ spanned by $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 0 \\ 3 \end{bmatrix}$.*