6.3 Orthogonal Projections

Theorem 8. The Orthogonal Decomposition Theorem
Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

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

$$\mathrm{proj}_W \vec{y} = \hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \tag{2}$$

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

The vector $\hat{\mathbf{y}}$ in (2) is called the **orthogonal projection of $\mathbf{y}$ onto W** and often is written as $\mathrm{proj}_W\,\mathbf{y}$.

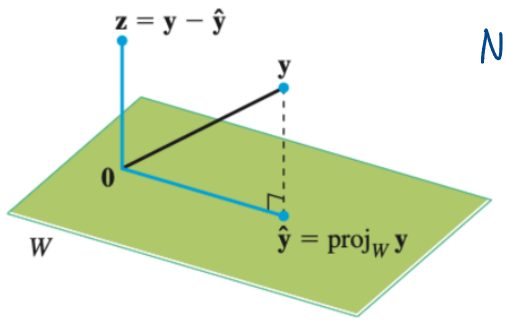

FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

Note: when W is one-dim'l. i.e. W is spanned by only one vector $\vec{u}$. Eq (2) matches the formula in § 6.2.

Example 1. Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = \mathrm{Span}\,\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

ANS: By Thm 8. we know $\vec{y} = \hat{\vec{y}} + \vec{z}$, where

$$\hat{\vec{y}} = \frac{\vec{y}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \frac{\vec{y}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\vec{u}_2 \in W$$

$$= \underbrace{\frac{2+10-3}{4+25+1}}_{\cancel{\frac{9}{30}}\ \frac{3}{10}}\vec{u}_1 + \underbrace{\frac{-2+2+3}{4+1+1}}_{\cancel{\frac{3}{6}}\ \frac{1}{2}}\vec{u}_2$$

$$= \begin{bmatrix} \frac{3}{5} - 1 \\ \frac{3}{2} + \frac{1}{2} \\ -\frac{3}{10} + \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ 2 \\ \frac{1}{5} \end{bmatrix}$$

and $\vec{z} = \vec{y} - \hat{\vec{y}} \in W^{\perp}$

$$= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{2}{5} \\ 2 \\ \frac{1}{5} \end{bmatrix} = \begin{bmatrix} \frac{7}{5} \\ 0 \\ \frac{14}{5} \end{bmatrix}$$

Thus the decomposition for $\vec{y}$ is

$$\vec{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \underbrace{\begin{bmatrix} -\frac{2}{5} \\ 2 \\ \frac{1}{5} \end{bmatrix}}_{\hat{\vec{y}} \in W} + \underbrace{\begin{bmatrix} \frac{7}{5} \\ 0 \\ \frac{14}{5} \end{bmatrix}}_{\vec{z} \in W^{\perp}}$$

A Geometric Interpretation of the Orthogonal Projection

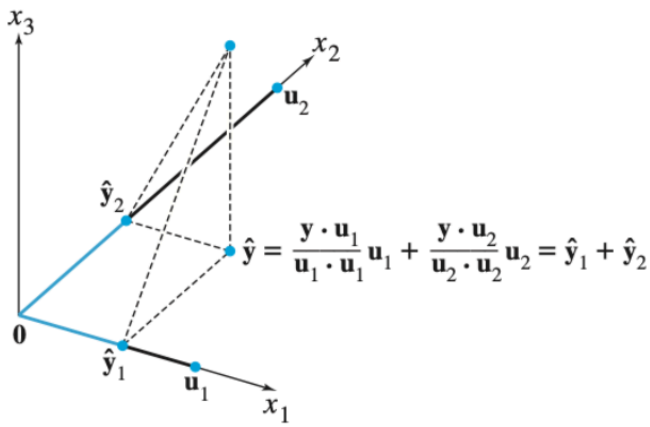

FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

Properties of Orthogonal Projections

Remark: If $\mathbf{y}$ is in $W = \text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$, then $\text{proj}_W\,\mathbf{y} = \mathbf{y}$

Theorem 9. The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

The vector $\hat{\mathbf{y}}$ in Theorem 9 is called **the best approximation to y by elements of W.**

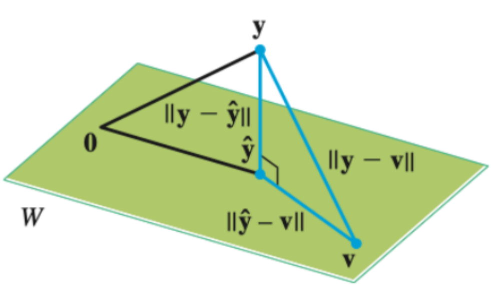

FIGURE 4 The orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$.

proof: Let $\vec{v} \in W$ distinct from $\hat{\vec{y}}$

then $\vec{y} - \vec{v} = (\vec{y} - \hat{\vec{y}}) + (\hat{\vec{y}} - \vec{v})$

By Thm 8, $\vec{y} - \hat{\vec{y}}$ is orthogonal to W. As $\hat{\vec{y}} - \vec{v}$ is in W, $\vec{y} - \hat{\vec{y}}$ is orthogonal to $\hat{\vec{y}} - \vec{v}$.

By the Pythagorean Thm

$$\|\vec{y} - \vec{v}\|^2 = \|\vec{y} - \hat{\vec{y}}\|^2 + \|\hat{\vec{y}} - \vec{v}\|^2.$$

As $\hat{\vec{y}}$ is distinct from $\vec{v}$, $\|\hat{\vec{y}} - \hat{\vec{v}}\|^2 > 0$,

$\|\vec{y} - \hat{\vec{y}}\| < \|y - \vec{v}\|$

Example 2. If $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, as in Example 1.

Determine the distance from $\mathbf{y}$ to the subspace W.

ANS: Note: the distance from a point in $\mathbb{R}^n$ to a subspace W is the distance from $\vec{y}$ to the nearest point $(\hat{\vec{y}})$ in W.

Thus by the Best Approximation Thm, it is

$$\underbrace{\|\vec{y} - \hat{\vec{y}}\|}_{\uparrow} \quad \text{proj}_W \vec{y}$$

Use the result in Example 1

$$\|\vec{z}\| = \sqrt{\left(\frac{7}{5}\right)^2 + \left(\frac{14}{5}\right)^2} = \sqrt{\frac{7^2 + 14^2}{5^2}}$$

$$= \sqrt{\frac{7^2 + 2^2 \cdot 7^2}{5^2}} = \sqrt{\frac{7^2 \cdot \cancel{5}}{5^{\cancel{2}}}} = \frac{7}{\sqrt{5}}$$

Example 3. Find the closest point to $\mathbf{y}$ ($\hat{y}$) in the subspace W spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$.

$$\mathbf{y} = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

Note $\vec{v}_1$ and $\vec{v}_2$ are orthogonal. The Best Approximation Thm states that $\hat{\vec{y}} = \text{proj}_W \vec{y}$ is the closest point to $\vec{y}$ in W.

$$\hat{\vec{y}} = \frac{\vec{y} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1} \vec{v}_1 + \frac{\vec{y} \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2} \vec{v}_2$$

$$= \overset{\frac{6}{\cancel{12}}\ \frac{1}{2}}{\cancel{\frac{9+1-5+1}{9+1+1+1}}} \vec{v}_1 + \overset{\frac{6}{4} = \frac{3}{2}}{\cancel{\frac{3-1+5-1}{1+1+1+1}}} \vec{v}_2$$

$$= \begin{bmatrix} \frac{3}{2} + \frac{3}{2} \\ \frac{1}{2} - \frac{3}{2} \\ -\frac{1}{2} + \frac{3}{2} \\ \frac{1}{2} - \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

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

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

$$\operatorname{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p$$

If $U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_p]$, then $\operatorname{proj}_W \mathbf{y} = UU^T\mathbf{y}$ for all $\mathbf{y}$ in $\mathbb{R}^n$.

Example 4. Let $\mathbf{y} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$, and $W = \operatorname{Span}\{\mathbf{u}_1\}$.

a. Let U be the 2×1 matrix whose only column is $\mathbf{u}_1$. Compute $\underset{1\times 2}{U^T}\underset{2\times 1}{U}$ and $\underset{2\times 1}{U}\underset{1\times 2}{U^T}$.

b. Compute $\operatorname{proj}_W \mathbf{y}$ and $(UU^T)\mathbf{y}$.

(a).

$$U = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

$$U^TU = \frac{1}{\sqrt{10}} \times \frac{1}{\sqrt{10}} \times \begin{bmatrix} 1 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \frac{1}{10}(1+9) = 1$$

$$UU^T = \frac{1}{\sqrt{10}} \times \frac{1}{\sqrt{10}} \cdot \begin{bmatrix} 1 \\ -3 \end{bmatrix} \begin{bmatrix} 1 & -3 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}$$

(b) By Thm 10.

$$\operatorname{proj}_W \vec{y} = (y \cdot \vec{u}_1)\, \vec{u}_1 = \frac{1}{\sqrt{10}} \times (7-27) \times \frac{1}{\sqrt{10}} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} -20 \\ -60 \end{bmatrix} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$$

$$UU^T\vec{y} = \frac{1}{10} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} \begin{bmatrix} 7 \\ 9 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} -20 \\ -21+81 \end{bmatrix} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$$

$//$

Exercise 5. Find the best approximation to $\mathbf{z}$ by vectors of the form $c_1\mathbf{v}_1 + c_2\mathbf{v}_2$.

$$\mathbf{z} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 2 \\ 0 \\ -1 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 5 \\ -2 \\ 4 \\ 2 \end{bmatrix}.$$

Solution. Note that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal. By the Best Approximation Theorem, the closest point in $\mathrm{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$ to $\mathbf{z}$ is $\hat{\mathbf{z}} = \dfrac{\mathbf{z}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\mathbf{v}_1 + \dfrac{\mathbf{z}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2}\mathbf{v}_2 = \dfrac{1}{2}\mathbf{v}_1 + 0\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1/2 \\ -3/2 \end{bmatrix}.$