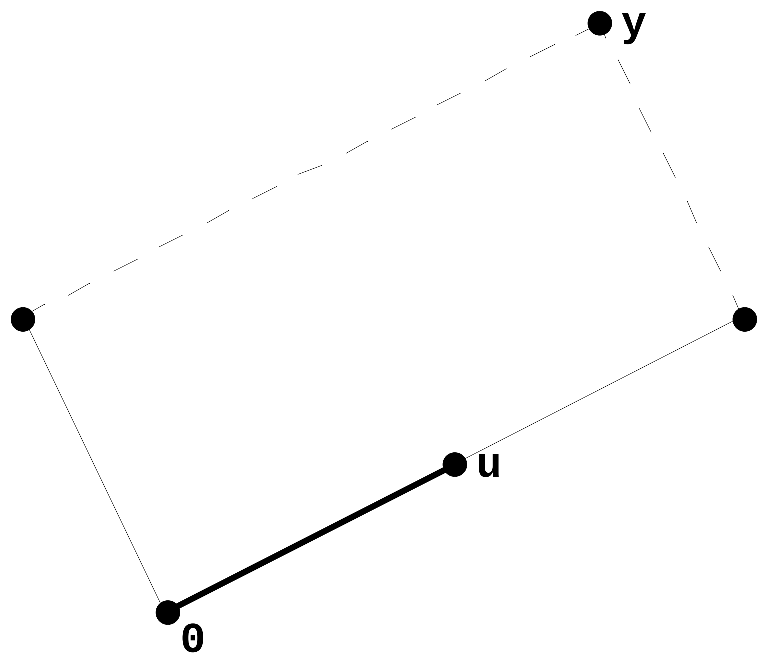
Section 6.3 Orthogonal Sets

Review

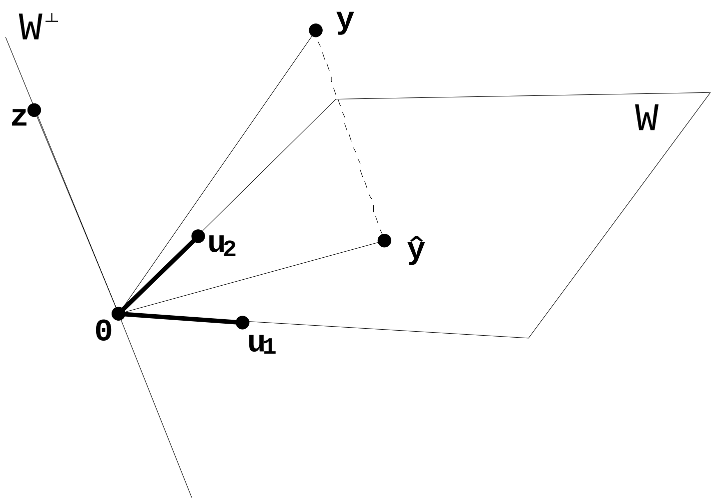
$\hat{\mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{u}}{\mathbf{u} \cdot \mathbf{u}} \mathbf{u}$ is the **orthogonal projection of ____ onto _____**.



Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W in $\mathbf{R}^n$. For each **y** in W,

$$\mathbf{y} = \left(\frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{y} \cdot \mathbf{u}_p}{\mathbf{u}_p \cdot \mathbf{u}_p}\right)\mathbf{u}_p$$

EXAMPLE: Suppose $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbf{R}^3$ and let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ in $\mathbf{R}^3$ as the sum of a vector $\hat{\mathbf{y}}$ in W and a vector $\mathbf{z}$ in $W^\perp$.



Solution: Write

$$\mathbf{y} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\right)\mathbf{u}_3$$

where

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2$$

$$\mathbf{z} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\right)\mathbf{u}_3.$$

To show that $\mathbf{z}$ is orthogonal to every vector in W, show that $\mathbf{z}$ is orthogonal to the vectors in $\{\mathbf{u}_1, \mathbf{u}_2\}$.

Since

$$\mathbf{z}\cdot\mathbf{u}_1 = \qquad\qquad = \qquad\qquad = \mathbf{0}$$

$$\mathbf{z}\cdot\mathbf{u}_2 = \qquad\qquad = \qquad\qquad = \mathbf{0}$$

THEOREM 8 THE ORTHOGONAL DECOMPOSITION THEOREM

Let W be a subspace of $\mathbf{R}^n$. Then each $\mathbf{y}$ in $\mathbf{R}^n$ can be uniquely represented in the form

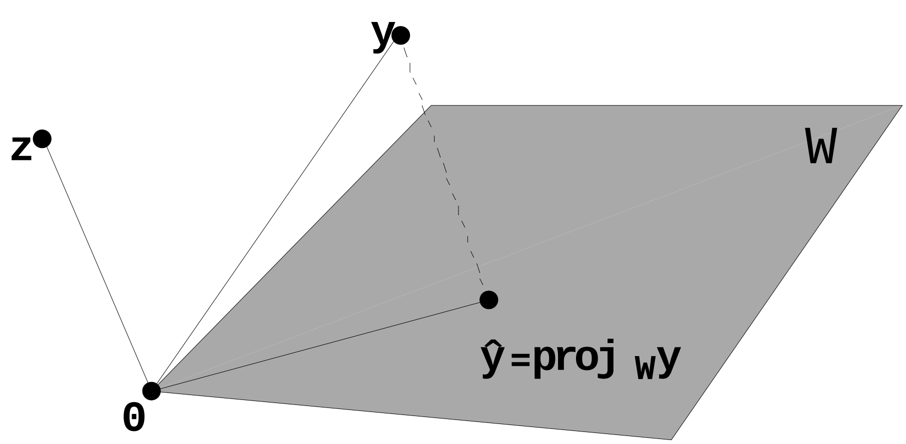
$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$$

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 of W, then

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\right)\mathbf{u}_p$$

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

The vector $\hat{\mathbf{y}}$ is called the **orthogonal projection of y onto** W.



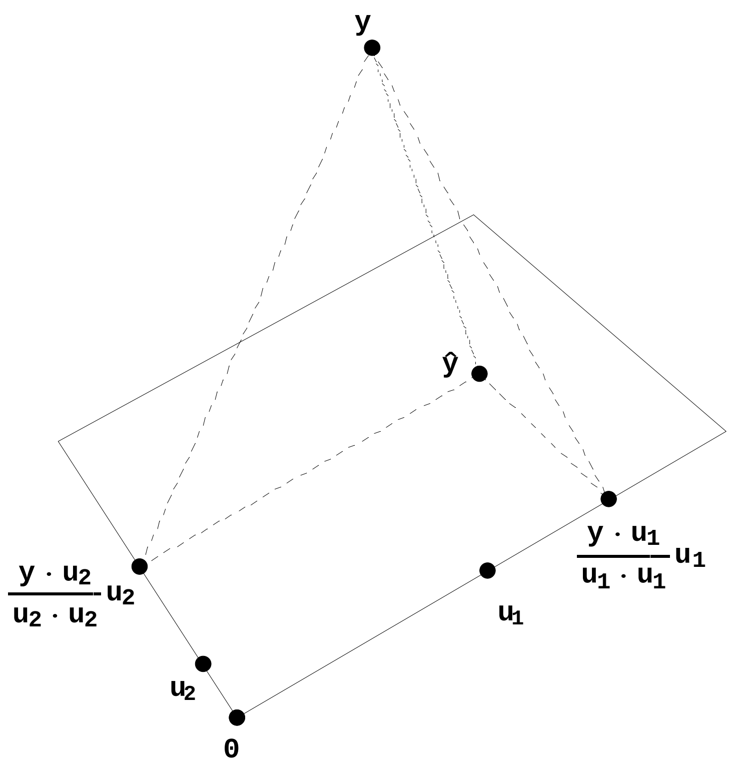
EXAMPLE: Let $\mathbf{u}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

Solution:

$$\text{proj}_W \mathbf{y} = \hat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2 = \left(\qquad\right)\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \left(\qquad\right)\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 0 \\ 3 \\ 10 \end{bmatrix} - \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 9 \end{bmatrix}$$

Geometric Interpretation of Orthogonal Projections

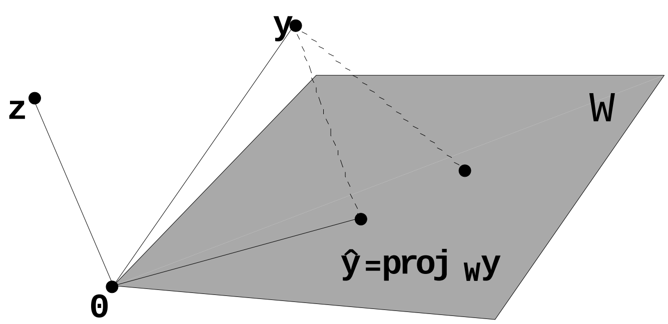


THEOREM 9 The Best Approximation Theorem

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

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

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



Outline of Proof: Let $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$. Then

$$\mathbf{v} - \hat{\mathbf{y}} \text{ is also in } W \quad \text{(why?)}$$

$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} \text{ is orthogonal to } W \quad \Rightarrow \quad \mathbf{y} - \hat{\mathbf{y}} \text{ is orthogonal to } \mathbf{v} - \hat{\mathbf{y}}$$

$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v}) \qquad \Rightarrow \qquad \|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2.$$

$$\|\mathbf{y} - \mathbf{v}\|^2 > \|\mathbf{y} - \hat{\mathbf{y}}\|^2$$

Hence, $\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|$. ■

EXAMPLE: Find the closest point to $\mathbf{y}$ in Span$\{\mathbf{u}_1, \mathbf{u}_2\}$ where $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -2 \end{bmatrix}$, $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, and $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$.

Solution: $\hat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2 = \left(\quad \right)\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \left(\quad \right)\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} =$

Part of Theorem 10 below is based upon another way to view matrix multiplication where A is $m \times p$ and B is $p \times n$

$$AB = \begin{bmatrix} \text{col}_1 A & \text{col}_2 A & \cdots & \text{col}_p A \end{bmatrix} \begin{bmatrix} \text{row}_1 B \\ \text{row}_2 B \\ \vdots \\ \text{row}_p B \end{bmatrix}$$

$$= (\text{col}_1 A)(\text{row}_1 B) + \cdots + (\text{col}_p A)(\text{row}_p B)$$

For example

$$\begin{bmatrix} 5 & 6 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 4 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 34 & 5 & 3 \\ 10 & 3 & 7 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 6 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 4 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \end{bmatrix} + \begin{bmatrix} 6 \\ 1 \end{bmatrix}\begin{bmatrix} 4 & 0 & -2 \end{bmatrix}$$

$$=$$

So if $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{bmatrix}$. Then $U^T = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \vdots \\ \mathbf{u}_p^T \end{bmatrix}$. So

$$UU^T = \mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T + \cdots + \mathbf{u}_p\mathbf{u}_p^T$$

$$(UU^T)\mathbf{y} = (\mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T + \cdots + \mathbf{u}_p\mathbf{u}_p^T)\mathbf{y}$$

$$=(\mathbf{u}_1\mathbf{u}_1^T)\mathbf{y} + (\mathbf{u}_2\mathbf{u}_2^T)\mathbf{y} + \cdots + (\mathbf{u}_p\mathbf{u}_p^T)\mathbf{y} = \mathbf{u}_1(\mathbf{u}_1^T\mathbf{y}) + \mathbf{u}_2(\mathbf{u}_2^T\mathbf{y}) + \cdots + \mathbf{u}_p(\mathbf{u}_p^T\mathbf{y})$$

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

$$\Rightarrow (UU^T)\mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p$$

THEOREM 10

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

$$\text{proj}_W\mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + \left(\mathbf{y} \cdot \mathbf{u}_p\right)\mathbf{u}_p$$

If $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_p \end{bmatrix}$, then

$$\text{proj}_W\mathbf{y} = UU^T\mathbf{y} \qquad \text{for all } \mathbf{y} \text{ in } \mathbf{R}^n.$$

Outline of Proof:

$$\text{proj}_W\mathbf{y} = \left(\tfrac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\tfrac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\right)\mathbf{u}_p$$

$$= (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + \left(\mathbf{y} \cdot \mathbf{u}_p\right)\mathbf{u}_p = UU^T\mathbf{y}.$$