Math 2331 – Linear Algebra

6.3 Orthogonal Projections

Jiwen He

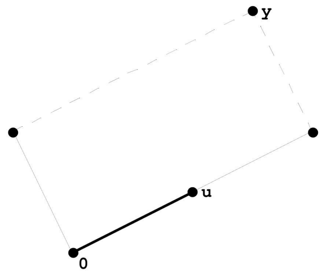
Department of Mathematics, University of Houston

jiwenhe@math.uh.edu
math.uh.edu/∼jiwenhe/math2331

6.3 Orthogonal Projections

Orthogonal Projection: Review

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



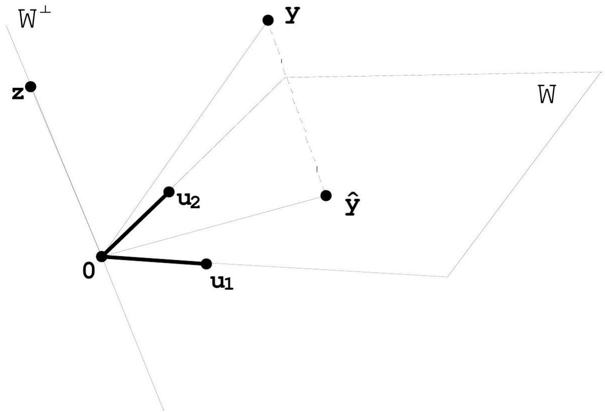
Suppose $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal basis for W in $\mathbf{R}^n$. For each $\mathbf{y}$ in W,

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

Orthogonal Projection: Example

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 $\widehat{\mathbf{y}}$ in W and a vector $\mathbf{z}$ in $W^\perp$.



Orthogonal Projection: Example (cont.)

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

$$\widehat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2, \qquad \mathbf{z} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\right)\mathbf{u}_3.$$

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

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

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

The Orthogonal Decomposition Theorem

Theorem (8)

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

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

where $\widehat{\mathbf{y}}$ 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

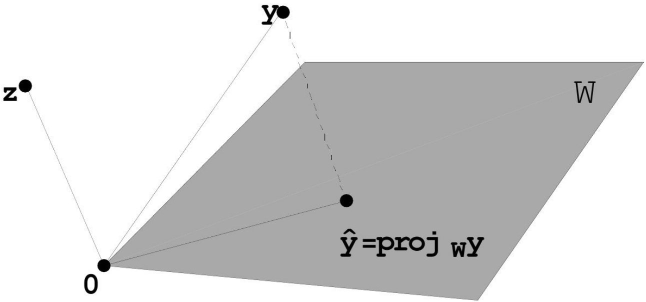
$$\widehat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\right)\mathbf{u}_p$$

and

$$\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}}.$$

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

The Orthogonal Decomposition Theorem



The Orthogonal Decomposition: Example

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 $W =$Span$\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

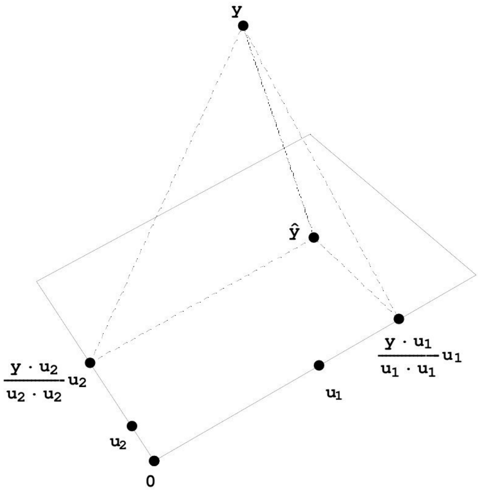
Solution:

$$\text{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \left(\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\right)\mathbf{u}_2$$

$$= (\qquad) \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + (\qquad) \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$$

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

Geometric Interpretation of Orthogonal Projections



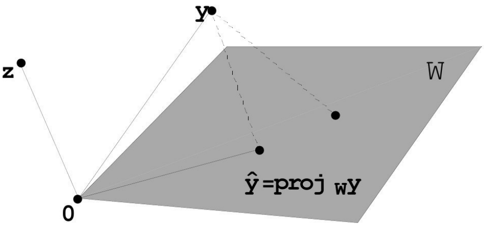
The Best Approximation Theorem

Theorem (9 The Best Approximation Theorem)

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

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

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



The Best Approximation Theorem (cont.)

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

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

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

$$\mathbf{y}-\mathbf{v} = (\mathbf{y}-\widehat{\mathbf{y}}) + (\widehat{\mathbf{y}}-\mathbf{v}) \quad \Longrightarrow \quad \|\mathbf{y}-\mathbf{v}\|^2 = \|\mathbf{y}-\widehat{\mathbf{y}}\|^2 + \|\widehat{\mathbf{y}}-\mathbf{v}\|^2.$$

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

Hence, $\|\mathbf{y}-\widehat{\mathbf{y}}\| < \|\mathbf{y}-\mathbf{v}\|$.$\blacksquare$

The Best Approximation Theorem: Example

Example

Find the closest point to $\mathbf{y}$ in Span$\{\mathbf{u}_1, \mathbf{u}_2\}$ where

$$\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 0 \\ -2 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } \mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Solution:

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

$$= (\qquad) \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + (\qquad) \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} =$$

New View of Matrix Multiplication

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

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

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

New View of Matrix Multiplication: Example

Example

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

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

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

$$=$$

Orthogonal Matrix

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

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

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

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

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

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

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

Orthogonal Projection: Theorem

Theorem (10)

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

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

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

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

Outline of Proof:

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

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