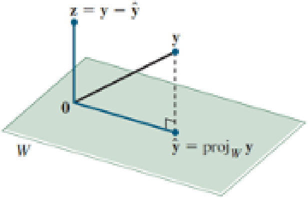
Orthogonal Projections

Linear Algebra

This work was initially funded by an Institutional OER Grant from the Colorado Department of Higher Education (CDHE). For similar OER materials in other courses funded by this project in the Department of Mathematical and Statistical Sciences at the University of Colorado Denver, visit
`https://github.com/CU-Denver-MathStats-OER`

Orthogonal Projection onto a Subspace

Given a subspace W of $\mathbb{R}^n$, in many instances it is useful to decompose a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$:

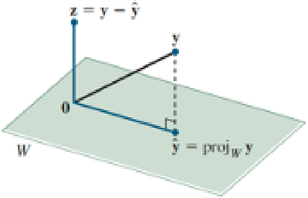


- The orthogonal projection of $\mathbf{y}$ onto W is $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$.
- The component of $\mathbf{y}$ orthogonal to W is $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

Orthogonal Projection onto a Subspace

Given a subspace W of $\mathbb{R}^n$, in many instances it is useful to decompose a vector $\mathbf{y}$ in $\mathbb{R}^n$ into the sum of two vectors $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$:

- The orthogonal projection of $\mathbf{y}$ onto W is $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$.
- The component of $\mathbf{y}$ orthogonal to W is $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.



Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof. Note that $\hat{\mathbf{y}}$ is in W, and that $\hat{\mathbf{y}}$ and $\mathbf{z}$ sum to $\mathbf{y}$. It remains to show that $\mathbf{z}$ is in $W^\perp$.

$$\begin{aligned}\mathbf{z}\cdot\mathbf{u}_i = (\mathbf{y}-\hat{\mathbf{y}})\cdot\mathbf{u}_i &= (\mathbf{y}\cdot\mathbf{u}_i) - \hat{\mathbf{y}}\cdot\mathbf{u}_i \\ &= (\mathbf{y}\cdot\mathbf{u}_i) - [c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p]\cdot\mathbf{u}_i\end{aligned}$$

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof. Note that $\hat{\mathbf{y}}$ is in W, and that $\hat{\mathbf{y}}$ and $\mathbf{z}$ sum to $\mathbf{y}$. It remains to show that $\mathbf{z}$ is in $W^\perp$.

$$\begin{aligned}
\mathbf{z}\cdot\mathbf{u}_i = (\mathbf{y} - \hat{\mathbf{y}})\cdot\mathbf{u}_i &= (\mathbf{y}\cdot\mathbf{u}_i) - \hat{\mathbf{y}}\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - [c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p]\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - c_1(\mathbf{u}_1\cdot\mathbf{u}_i) - \ldots - c_i(\mathbf{u}_i\cdot\mathbf{u}_i) - \ldots - c_p(\mathbf{u}_p\cdot\mathbf{u}_i)
\end{aligned}$$

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}, \qquad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof. Note that $\hat{\mathbf{y}}$ is in W, and that $\hat{\mathbf{y}}$ and $\mathbf{z}$ sum to $\mathbf{y}$. It remains to show that $\mathbf{z}$ is in $W^\perp$.

$$\begin{aligned}
\mathbf{z}\cdot\mathbf{u}_i = (\mathbf{y}-\hat{\mathbf{y}})\cdot\mathbf{u}_i &= (\mathbf{y}\cdot\mathbf{u}_i) - \hat{\mathbf{y}}\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - [c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p]\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - c_1(\mathbf{u}_1\cdot\mathbf{u}_i) - \ldots - c_i(\mathbf{u}_i\cdot\mathbf{u}_i) - \ldots - c_p(\mathbf{u}_p\cdot\mathbf{u}_i) \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - c_i(\mathbf{u}_i\cdot\mathbf{u}_i) = (\mathbf{y}\cdot\mathbf{u}_i) - \left(\frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}\right)(\mathbf{u}_i\cdot\mathbf{u}_i) = 0.
\end{aligned}$$

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof. Note that $\hat{\mathbf{y}}$ is in W, and that $\hat{\mathbf{y}}$ and $\mathbf{z}$ sum to $\mathbf{y}$. It remains to show that $\mathbf{z}$ is in $W^\perp$.

$$\begin{aligned}
\mathbf{z}\cdot\mathbf{u}_i = (\mathbf{y}-\hat{\mathbf{y}})\cdot\mathbf{u}_i &= (\mathbf{y}\cdot\mathbf{u}_i) - \hat{\mathbf{y}}\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - [c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p]\cdot\mathbf{u}_i \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - c_1(\mathbf{u}_1\cdot\mathbf{u}_i) - \ldots - c_i(\mathbf{u}_i\cdot\mathbf{u}_i) - \ldots - c_p(\mathbf{u}_p\cdot\mathbf{u}_i) \\
&= (\mathbf{y}\cdot\mathbf{u}_i) - c_i(\mathbf{u}_i\cdot\mathbf{u}_i) = (\mathbf{y}\cdot\mathbf{u}_i) - \left(\frac{\mathbf{y}\cdot\mathbf{u}_i}{\mathbf{u}_i\cdot\mathbf{u}_i}\right)(\mathbf{u}_i\cdot\mathbf{u}_i) = 0.
\end{aligned}$$

Since $\mathbf{z}$ is orthogonal to each $\mathbf{u}_i$ in the basis, $\mathbf{z}$ is in $W^\perp$ by a previous theorem.

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^{\perp}$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof.

For uniqueness, suppose that $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$. Then $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$.

Orthogonal Decomposition Theorem

Theorem (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 $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthogonal basis of W, then

$$\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \ldots + c_p\mathbf{u}_p \text{ with weights } c_i = \frac{\mathbf{y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i}, \quad \text{and } \mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}.$$

Proof.

For uniqueness, suppose that $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$. Then $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$.

But $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 \in W$ and $\mathbf{z}_1 - \mathbf{z} \in W^\perp$, which implies that $\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z} = \mathbf{0}$.

Thus, $\hat{\mathbf{y}}_1 = \hat{\mathbf{y}}$ and $\mathbf{z}_1 = \mathbf{z}$. Therefore, $\hat{\mathbf{y}}$ and $\mathbf{z}$ are the only vectors satisfying the conditions. $\square$

Example

Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ where $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}$. Write $\mathbf{y} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix}$ as a sum of a vector in W and a vector orthogonal to W.

Example

Let $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ where $\mathbf{u}_1 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}$. Write $\mathbf{y} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix}$ as a sum of a vector in W and a vector orthogonal to W.

The orthogonal projection of $\mathbf{y}$ onto W is $\hat{\mathbf{y}} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ where we have weights

$$c_1 = \frac{\mathbf{y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} = \frac{9}{9} = 1 \text{ and } c_2 = \frac{\mathbf{y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} = \frac{26}{20} = \frac{13}{10}.$$

This gives

$$\hat{\mathbf{y}} = (1)\mathbf{u}_1 + \left(\frac{13}{10}\right)\mathbf{u}_2 = \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} + \frac{13}{10}\begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} -13/5 \\ 3 \\ 26/5 \end{bmatrix},$$

and thus

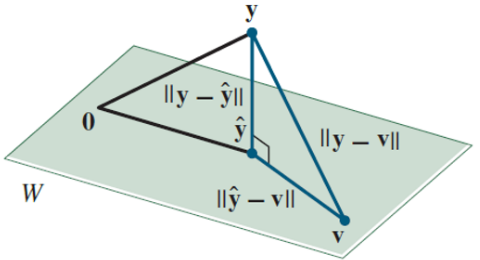
$$\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ 3 \\ 6 \end{bmatrix} - \begin{bmatrix} -13/5 \\ 3 \\ 26/5 \end{bmatrix} = \begin{bmatrix} 8/5 \\ 0 \\ 4/5 \end{bmatrix}.$$

The Best Approximation Theorem

Theorem (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

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \quad \text{for all } \mathbf{v} \neq \mathbf{y} \text{ in } W.$$



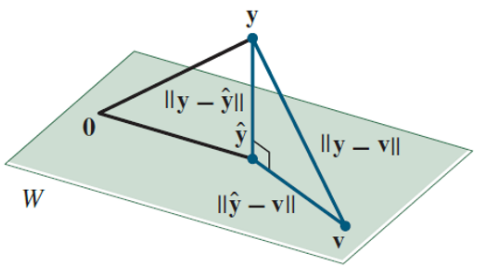
Theorem (The Best Approximation Theorem)

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

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \quad \text{for all } \mathbf{v} \neq \mathbf{y} \text{ in } W.$$

Proof.

For any $\mathbf{v} \in W$, $\hat{\mathbf{y}} - \mathbf{v}$ is orthogonal to $\mathbf{y} - \hat{\mathbf{y}}$.



The Best Approximation Theorem

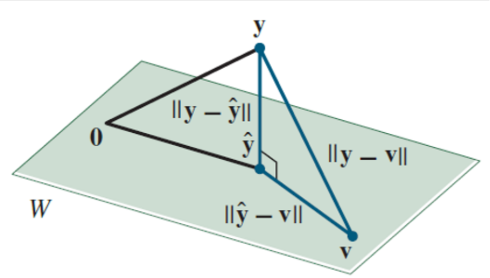
Theorem (The Best Approximation Theorem)

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

$$\|\mathbf{y}-\hat{\mathbf{y}}\| < \|\mathbf{y}-\mathbf{v}\| \quad \text{for all } \mathbf{v} \neq \mathbf{y} \text{ in } W.$$

Proof.

For any $\mathbf{v} \in W$, $\hat{\mathbf{y}}-\mathbf{v}$ is orthogonal to $\mathbf{y}-\hat{\mathbf{y}}$.
By Pyth., $\|\mathbf{y}-\mathbf{v}\|^2 = \|\mathbf{y}-\hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}}-\mathbf{v}\|^2$.
So $\|\mathbf{y}-\mathbf{v}\| > \|\mathbf{y}-\hat{\mathbf{y}}\|$ for $\mathbf{v} \neq \hat{\mathbf{y}}$. □

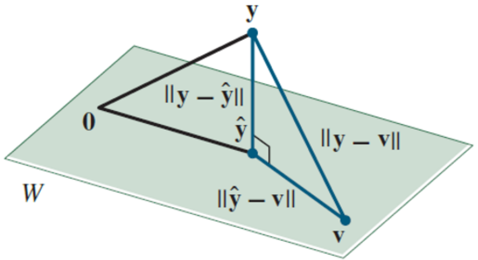


Theorem (The Best Approximation Theorem)

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

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \quad \text{for all } \mathbf{v} \neq \mathbf{y} \text{ in } W.$$

Proof.
For any $\mathbf{v} \in W$, $\hat{\mathbf{y}} - \mathbf{v}$ is orthogonal to $\mathbf{y} - \hat{\mathbf{y}}$.
By Pyth., $\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$.
So $\|\mathbf{y} - \mathbf{v}\| > \|\mathbf{y} - \hat{\mathbf{y}}\|$ for $\mathbf{v} \neq \hat{\mathbf{y}}$. □



- The vector $\hat{\mathbf{y}}$ is called the best approximation to $\mathbf{y}$ by elements of W.
- The distance $\|\mathbf{y} - \hat{\mathbf{y}}\| = \|\mathbf{z}\|$ is the error of the approximation.

Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$. Find the best approximation of $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$ in the subspace $W = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

Example

Let $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix}$. Find the best approximation of $\mathbf{y} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$ in the subspace $W = \operatorname{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

From the Best Approximation Theorem, we know that the orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$. The orthogonal projection of $\mathbf{y}$ onto W is $\hat{\mathbf{y}} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ where we have weights $c_1 = \dfrac{\mathbf{y}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1} = \dfrac{0}{14}$ and $c_2 = \dfrac{\mathbf{y}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2} = \dfrac{28}{42} = \dfrac{2}{3}$.

Therefore, the best approximation for $\mathbf{y}$ in W is

$$\hat{\mathbf{y}} = (0)\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} + \frac{2}{3}\begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix} \quad \text{with} \quad \|\mathbf{z}\| = \left\| \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} - \begin{bmatrix} 10/3 \\ 2/3 \\ 8/3 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -7/3 \\ 7/3 \\ 7/3 \end{bmatrix} \right\| \approx 4.0415.$$

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $U^T U$ and UU^T.

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $U^T U$ and UU^T.

We have

$$U^T U = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 \\ -2/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So U has orthonormal columns!

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $U^T U$ and UU^T.

We have

$$U^T U = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 \\ -2/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

So U has orthonormal columns!

$$UU^T = \begin{bmatrix} 1/3 & 2/3 \\ -2/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} = \begin{bmatrix} 5/9 & 2/9 & 4/9 \\ 2/9 & 8/9 & -2/9 \\ 4/9 & -2/9 & 5/9 \end{bmatrix}$$

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $(UU^T)\mathbf{y}$ and $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$ for $\mathbf{y} = (1, -2, 4)$.

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $(UU^T)\mathbf{y}$ and $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$ for $\mathbf{y} = (1, -2, 4)$.

We have

$$(UU^T)\mathbf{y} = \begin{bmatrix} 5/9 & 2/9 & 4/9 \\ 2/9 & 8/9 & -2/9 \\ 4/9 & -2/9 & 5/9 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17/9 \\ -22/9 \\ 28/9 \end{bmatrix},$$

Example

Let $\mathbf{u}_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$ and define the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$ of $\mathbb{R}^3$.

Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix}$. Compute $(UU^T)\mathbf{y}$ and $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$ for $\mathbf{y} = (1, -2, 4)$.

We have

$$(UU^T)\mathbf{y} = \begin{bmatrix} 5/9 & 2/9 & 4/9 \\ 2/9 & 8/9 & -2/9 \\ 4/9 & -2/9 & 5/9 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17/9 \\ -22/9 \\ 28/9 \end{bmatrix},$$

and

$$\hat{\mathbf{y}} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 = \frac{13}{3}\mathbf{u}_1 + \frac{2}{3}\mathbf{u}_2 = \frac{13}{3} \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 2/3 \\ 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 17/9 \\ -22/9 \\ 28/9 \end{bmatrix}.$$

Orthogonal projection onto a Subspace using an Orthonormal Basis

The formula for $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$ is simplified when the basis for W is an orthonormal basis.

Theorem

If $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \dots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n,$$

where the matrix $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_p \end{bmatrix}$.

The matrix UU^T is called the orthogonal projection matrix onto W.

Orthogonal projection onto a Subspace using an Orthonormal Basis

The formula for $\hat{\mathbf{y}} = \text{proj}_W \mathbf{y}$ is simplified when the basis for W is an orthonormal basis.

Theorem

If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W \mathbf{y} = (\mathbf{y} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y} \cdot \mathbf{u}_2)\mathbf{u}_2 + \ldots + (\mathbf{y} \cdot \mathbf{u}_p)\mathbf{u}_p = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n,$$

where the matrix $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_p \end{bmatrix}$.

The matrix UU^T is called the orthogonal projection matrix onto W.

Proof.

Note that $U^T\mathbf{y} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{y} \\ \mathbf{u}_2^T\mathbf{y} \\ \vdots \\ \mathbf{u}_p^T\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{y} \\ \mathbf{u}_2 \cdot \mathbf{y} \\ \vdots \\ \mathbf{u}_p \cdot \mathbf{y} \end{bmatrix}$. Thus, $UU^T\mathbf{y} = (\mathbf{u}_1 \cdot \mathbf{y})\mathbf{u}_1 + (\mathbf{u}_2 \cdot \mathbf{y})\mathbf{u}_2 + \ldots + (\mathbf{u}_p \cdot \mathbf{y})\mathbf{u}_p$.