§5.4 Eigenvectors and Linear Transformations

- Let V, W = n-dimensional and m-dimensional Vector Spaces.
- Let T = linear transformation from V to W.
- Let $\mathcal{B}$ and $\mathcal{C}$ be bases for V, W.

Given $\mathbf{x} \in V, \quad [T(\mathbf{x})]_{\mathcal{C}} = (?)\ [\mathbf{x}]_{\mathcal{B}}$

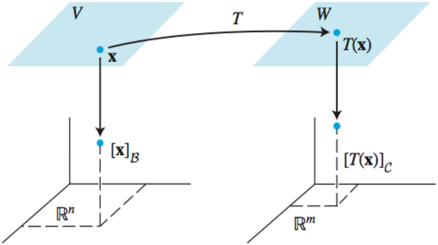


FIGURE 1 A linear transformation from V to W.

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

$$\text{Let}\quad \mathbf{x} = \alpha_1\,\mathbf{b}_1 + \cdots + \alpha_n\,\mathbf{b}_n \ \text{ so that }\ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

Let $\mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n$ so that $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$.

$$T(\mathbf{x}) = \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n),$$

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{C}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{C}} \end{aligned}$$

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{C}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{C}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{C}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{C}} \end{aligned}$$

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

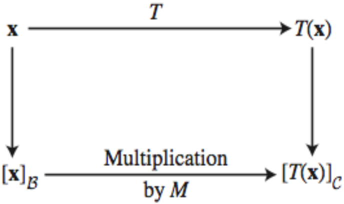
$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{C}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{C}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{C}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{C}} \\ &= [[T(\mathbf{b}_1)]_{\mathcal{C}}, [T(\mathbf{b}_2)]_{\mathcal{C}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{C}}] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \end{aligned}$$

Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \, \mathbf{b}_1 + \cdots + \alpha_n \, \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 \, T(\mathbf{b}_1) + \cdots + \alpha_n \, T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{C}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{C}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{C}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{C}} \\ &= [[T(\mathbf{b}_1)]_{\mathcal{C}}, [T(\mathbf{b}_2)]_{\mathcal{C}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{C}}] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \stackrel{def}{=} M \, [\mathbf{x}]_{\mathcal{B}} \end{aligned}$$



EX: Consider bases: $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for V, $\quad \mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ for W.
$T : V \longrightarrow W$ is linear transformation satisfying

$$T(\mathbf{b}_1) = 3\,\mathbf{c}_1 - 2\,\mathbf{c}_2 + 5\,\mathbf{c}_3 \quad \text{and} \quad T(\mathbf{b}_2) = 4\,\mathbf{c}_1 + 7\,\mathbf{c}_2 - \mathbf{c}_3.$$

Find M, Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$.

EX: Consider bases: $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ for V, $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$ for W.
$T : V \longrightarrow W$ is linear transformation satisfying

$$T(\mathbf{b}_1) = 3\,\mathbf{c}_1 - 2\,\mathbf{c}_2 + 5\,\mathbf{c}_3 \quad \text{and} \quad T(\mathbf{b}_2) = 4\,\mathbf{c}_1 + 7\,\mathbf{c}_2 - \mathbf{c}_3.$$

Find M, Matrix for T relative to $\mathcal{B}$ and $\mathcal{C}$.

SOLUTION: By definition,

$$[T(\mathbf{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix}, \quad \text{and} \quad [T(\mathbf{b}_2)]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 7 \\ -1 \end{bmatrix}.$$

$$\text{Hence} \quad M = \begin{bmatrix} 3 & 4 \\ -2 & 7 \\ 5 & -1 \end{bmatrix}.$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

$$\text{Let} \quad \mathbf{x} = \alpha_1\,\mathbf{b}_1 + \cdots + \alpha_n\,\mathbf{b}_n \;\text{ so that }\; [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \text{ so that } [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$T(\mathbf{x}) = \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n),$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{B}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{B}} \end{aligned}$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{B}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{B}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{B}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{B}} \end{aligned}$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

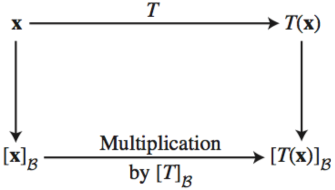
$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \ \text{ so that } \ [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{B}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{B}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{B}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{B}} \\ &= [[T(\mathbf{b}_1)]_{\mathcal{B}}, [T(\mathbf{b}_2)]_{\mathcal{B}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{B}}] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \end{aligned}$$

Matrix relative to $\mathcal{B}$ for $\boxed{T : V \to V.}$

$$\text{Let} \quad \mathbf{x} = \alpha_1 \mathbf{b}_1 + \cdots + \alpha_n \mathbf{b}_n \;\text{ so that }\; [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

$$\begin{aligned} T(\mathbf{x}) &= \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n), \\ [T(\mathbf{x})]_{\mathcal{B}} &= [\alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n)]_{\mathcal{B}} = \alpha_1 [T(\mathbf{b}_1)]_{\mathcal{B}} + \cdots + \alpha_n [T(\mathbf{b}_n)]_{\mathcal{B}} \\ &= [[T(\mathbf{b}_1)]_{\mathcal{B}}, [T(\mathbf{b}_2)]_{\mathcal{B}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{B}}] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \overset{def}{=} [T]_{\mathcal{B}} [\mathbf{x}]_{\mathcal{B}} \end{aligned}$$



EX: Consider linear transformation: $T : \mathcal{P}_2 \to \mathcal{P}_2$ defined by

$$T\left(a_0 + a_1\, t + a_2\, t^2\right) = a_1 + 2\, a_2\, t.$$

a. Find $[T]_{\mathcal{B}}$ for basis $\mathcal{B} = \left\{1,\ t,\ t^2\right\}$.

b. Verify $[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}\, [\mathbf{p}]_{\mathcal{B}}$ for each $\mathbf{p} \in \mathcal{P}_2$.

SOLUTION (**a**): $T(1) = 0, \qquad T(t) = 1, \qquad T\left(t^2\right) = 2\,t.$ Hence

$$[T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ [T(t)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},\ \left[T\left(t^2\right)\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}.$$

$$[T]_{\mathcal{B}} = \left[[T(1)]_{\mathcal{B}}, [T(t)]_{\mathcal{B}}, \left[T\left(t^2\right)\right]_{\mathcal{B}}\right] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

EX: Consider linear transformation: $T : \mathcal{P}_2 \to \mathcal{P}_2$ defined by

$$T\left(a_0 + a_1\, t + a_2\, t^2\right) = a_1 + 2\, a_2\, t.$$

b. Verify $[T(\mathbf{p})]_{\mathcal{B}} = [T]_{\mathcal{B}}\, [\mathbf{p}]_{\mathcal{B}}$ for each $\mathbf{p} \in \mathcal{P}_2$.

SOLUTION (**b**): For $\mathbf{p}(t) = a_0 + a_1\, t + a_2\, t^2, \quad T(\mathbf{p})(t) = a_1 + 2\, a_2\, t.$

,

$$[\mathbf{p}]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}, \quad [T(\mathbf{p})]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ 2\, a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = [T]_{\mathcal{B}}\, [\mathbf{p}]_{\mathcal{B}}.$$

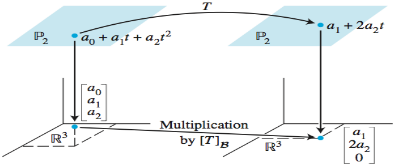


FIGURE 4 Matrix representation of a linear transformation.

Linear Transformation on $\mathcal{R}^n$

- For $A \in \mathcal{R}^{n \times n}$, define linear transformation on $\mathcal{R}^n$: $T(\mathbf{x}) = A\mathbf{x}$.

Linear Transformation on $\mathcal{R}^n$

- For $A \in \mathcal{R}^{n\times n}$, define linear transformation on $\mathcal{R}^n$: $T(\mathbf{x}) = A\,\mathbf{x}$.
- Given basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n\}$. Find $[T]_{\mathcal{B}} = [[T(\mathbf{b}_1)]_{\mathcal{B}}, [T(\mathbf{b}_2)]_{\mathcal{B}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{B}}]$?

By definition, $T(\mathbf{b}_1) = A\,\mathbf{b}_1$. Let $[T(\mathbf{b}_1)]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$ so that

$$A\,\mathbf{b}_1 = \alpha_1\,\mathbf{b}_1 + \alpha_2\,\mathbf{b}_2 + \cdots + \alpha_n\,\mathbf{b}_n = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n] \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix},$$

$$\Longrightarrow \quad [T(\mathbf{b}_1)]_{\mathcal{B}} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = P^{-1}(A\,\mathbf{b}_1), \quad \text{with} \quad P = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n].$$

Linear Transformation on $\mathcal{R}^n$

- Given basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n\}$.
  Find $[T]_{\mathcal{B}} = [[T(\mathbf{b}_1)]_{\mathcal{B}}, [T(\mathbf{b}_2)]_{\mathcal{B}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{B}}]$?

$$[T(\mathbf{b}_1)]_{\mathcal{B}} = P^{-1}(A\,\mathbf{b}_1), \quad \text{with} \quad P = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n].$$

$$\begin{aligned} [T]_{\mathcal{B}} &= [[T(\mathbf{b}_1)]_{\mathcal{B}}, [T(\mathbf{b}_2)]_{\mathcal{B}}, \cdots, [T(\mathbf{b}_n)]_{\mathcal{B}}] \\ &= \left[P^{-1}(A\,\mathbf{b}_1),\ P^{-1}(A\,\mathbf{b}_2),\ \cdots, P^{-1}(A\,\mathbf{b}_n)\right] \\ &= P^{-1}A\,[\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n] = P^{-1}A\,P \end{aligned}$$

Goal: Choose $\mathcal{B}$ to make $[T]_{\mathcal{B}}$ as simple as possible.

For any $\mathbf{x} \in \mathcal{R}^n$, $\quad [A\,\mathbf{x}]_{\mathcal{B}} = [T(\mathbf{x})]_{\mathcal{B}} = [T]_{\mathcal{B}}\,[\mathbf{x}]_{\mathcal{B}}.$

Linear Transformation on $\mathcal{R}^2$: **EX** 1

- For $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

Linear Transformation on $\mathcal{R}^2$: **EX 1**

- For $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

SOLUTION: From example in §5.3,

$$A = P\,D\,P^{-1}, \quad \text{where} \quad P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}, \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}.$$

Choose $\mathbf{b}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ to be columns of P. Then

$$[T]_{\mathcal{B}} = P^{-1}\,A\,P = D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}, \quad \text{(diagonal matrix)}$$

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

- Solution: Find eigenvalues of A

$$\mathbf{det}(A - \lambda I) = \mathbf{det}\begin{pmatrix} 4-\lambda & -9 \\ 4 & -8-\lambda \end{pmatrix} = (2+\lambda)^2.$$

So eigenvalues are $\lambda_1 = \lambda_2 = -2$.

- Find eigenvectors of A

$$(A - \lambda_1 I)\mathbf{v} = \begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix}\mathbf{v} = \mathbf{0}, \quad \Longrightarrow \quad \mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

- Choose $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $P = [\mathbf{v}_1, \ \mathbf{v}_2]$ and $\mathcal{B} = \{\mathbf{v}_1, \ \mathbf{v}_2\}$.

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, define linear transformation on $\mathcal{R}^2$: $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

- Solution: Find eigenvalues of A

$$\mathbf{det}(A - \lambda I) = \mathbf{det}\begin{pmatrix} 4-\lambda & -9 \\ 4 & -8-\lambda \end{pmatrix} = (2+\lambda)^2.$$

  So eigenvalues are $\lambda_1 = \lambda_2 = -2$.

- Find eigenvectors of A

$$(A - \lambda_1 I)\mathbf{v} = \begin{pmatrix} 6 & -9 \\ 4 & -6 \end{pmatrix}\mathbf{v} = \mathbf{0}, \quad \Longrightarrow \quad \mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

- Choose $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $P = [\mathbf{v}_1,\ \mathbf{v}_2]$ and $\mathcal{B} = \{\mathbf{v}_1,\ \mathbf{v}_2\}$.

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, let $T(\mathbf{x}) = A\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

Linear Transformation on $\mathcal{R}^2$: **EX** 2

- For $A = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix}$, let $T(\mathbf{x}) = A\,\mathbf{x}$. Find $\mathcal{B}$ to simplify $[T]_{\mathcal{B}}$

- SOLUTION: Eigenvalues $\lambda_1 = \lambda_2 = -2$. Only eigenvector $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$.

- Choose $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$, $P = [\mathbf{v}_1, \ \mathbf{v}_2]$ and $\mathcal{B} = \{\mathbf{v}_1, \ \mathbf{v}_2\}$.

$$\begin{aligned} A\,\mathbf{v}_1 &= -2\,\mathbf{v}_1, \quad A\,\mathbf{v}_2 = \begin{bmatrix} 4 & -9 \\ 4 & -8 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -35 \\ -32 \end{bmatrix} = -13\,\mathbf{v}_1 - 2\,\mathbf{v}_2 \\ A\,P &= A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} -2\mathbf{v}_1, -13\mathbf{v}_1 - 2\mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} \begin{bmatrix} -2 & -13 \\ 0 & -2 \end{bmatrix} \end{aligned}$$

Then $\quad [T]_{\mathcal{B}} = P^{-1}\,A\,P = \begin{bmatrix} -2 & -13 \\ 0 & -2 \end{bmatrix}, \quad$ (upper triangular matrix)

§6.1 Inner Product

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **inner product** of **u** and **v** $\stackrel{def}{=} u_1\, v_1 + \cdots + u_n\, v_n$

§6.1 Inner Product

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} \overset{def}{=} u_1 v_1 + \cdots + u_n v_n = \mathbf{u}^T \mathbf{v}$ ($\mathbf{u} \cdot \mathbf{v}$ in book)

§6.1 Inner Product

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} \overset{def}{=} u_1 v_1 + \cdots + u_n v_n = \mathbf{u}^T \mathbf{v}$ ($\mathbf{u} \cdot \mathbf{v}$ in book)
- **length** of $\mathbf{u}$ (denoted $\|\mathbf{u}\|$) $\overset{def}{=} \sqrt{u_1^2 + \cdots + u_n^2}$

§6.1 Inner Product

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} \overset{def}{=} u_1 v_1 + \cdots + u_n v_n = \mathbf{u}^T \mathbf{v}$ ($\mathbf{u} \cdot \mathbf{v}$ in book)
- **length** of $\mathbf{u}$ (denoted $\|\mathbf{u}\|$) $\overset{def}{=} \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\mathbf{u}^T \mathbf{u}}$

EX: Let $\mathbf{u} = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \in \mathcal{R}^3$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} = 3 \cdot 1 + (-5) \cdot 2 + 2 \cdot 1 = -5$

§6.1 Inner Product

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} \stackrel{def}{=} u_1\, v_1 + \cdots + u_n\, v_n = \mathbf{u}^T\, \mathbf{v}$ ($\mathbf{u} \cdot \mathbf{v}$ in book)
- **length** of $\mathbf{u}$ (denoted $\|\mathbf{u}\|$) $\stackrel{def}{=} \sqrt{u_1^2 + \cdots + u_n^2} = \sqrt{\mathbf{u}^T\, \mathbf{u}}$

EX: Let $\mathbf{u} = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \in \mathcal{R}^3$.

- **inner product** of $\mathbf{u}$ and $\mathbf{v} = 3 \cdot 1 + (-5) \cdot 2 + 2 \cdot 1 = -5$
- **length** of $\mathbf{u}$: $\|\mathbf{u}\| = \sqrt{3^2 + (-5)^2 + 2^2} = \sqrt{38}$

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **distance** between $\mathbf{u}$ and $\mathbf{v}$ (denoted $\mathbf{dist}(\mathbf{u}, \mathbf{v})) \stackrel{def}{=} \|\mathbf{u} - \mathbf{v}\|$

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **distance** between $\mathbf{u}$ and $\mathbf{v}$ (denoted $\mathbf{dist}\,(\mathbf{u}, \mathbf{v})$) $\stackrel{def}{=} \|\mathbf{u} - \mathbf{v}\|$
- $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if $\mathbf{u}^T \mathbf{v} = 0$

Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **distance** between **u** and **v** (denoted **dist** $(\mathbf{u}, \mathbf{v})$) $\overset{def}{=} \|\mathbf{u} - \mathbf{v}\|$
- **u** and **v** are **orthogonal** if $\mathbf{u}^T \mathbf{v} = 0$

EX: Let $\mathbf{u} = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathcal{R}^3$.

- **distance** between **u** and **v** : **dist** $(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(3-1)^2 + (-5-1)^2 + (2-1)^2} = \sqrt{41}$
- **u** and **v** are **orthogonal**: $\mathbf{u}^T \mathbf{v} = 3 \cdot 1 + (-5) \cdot 1 + 2 \cdot 1 = 0$

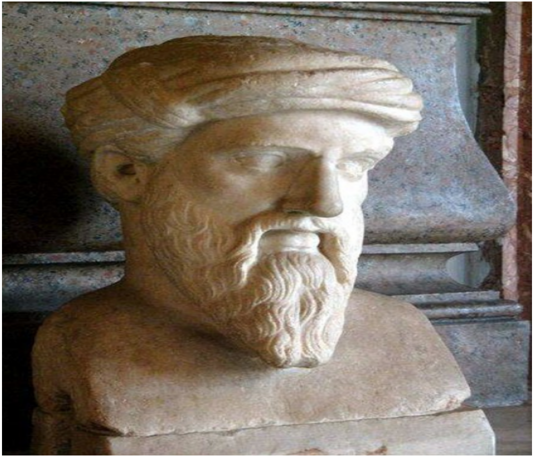
Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathcal{R}^n$.

- **distance** between $\mathbf{u}$ and $\mathbf{v}$ (denoted $\mathbf{dist}\,(\mathbf{u},\ \mathbf{v})$) $\stackrel{def}{=} \|\mathbf{u} - \mathbf{v}\|$
- $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal** if $\mathbf{u}^T\mathbf{v} = 0$

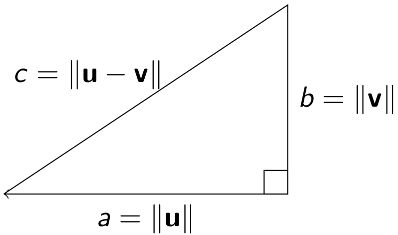
EX: Let $\mathbf{u} = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \mathcal{R}^3$.

- **distance** between $\mathbf{u}$ and $\mathbf{v}$: $\mathbf{dist}\,(\mathbf{u},\ \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(3-1)^2 + (-5-1)^2 + (2-1)^2} = \sqrt{41}$
- $\mathbf{u}$ and $\mathbf{v}$ are **orthogonal**: $\mathbf{u}^T\mathbf{v} = 3 \cdot 1 + (-5) \cdot 1 + 2 \cdot 1 = 0$
- **Pythagorean Thm:** $\|\mathbf{u} - \mathbf{v}\|^2 = (41 = 38 + 3 =)\ \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$

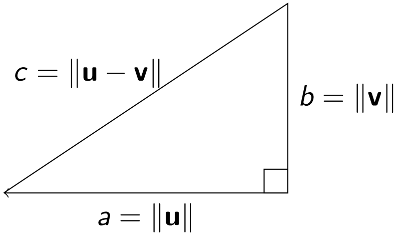
Pythagoras



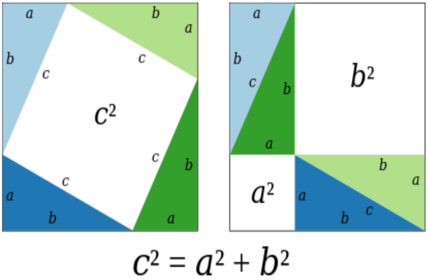
- **Pythagorean Thm**



- **Pythagorean Thm**



- **Pythagorean** Proof



Orthogonal Complement of subspace W of $\mathcal{R}^m$

- **orthogonal complement** of W (denoted $W^\perp$)
  $\stackrel{def}{=} \{\mathbf{u} \mid \mathbf{u}^T\mathbf{z} = 0, \quad \text{for all} \quad \mathbf{z} \in W\}$

Orthogonal Complement of subspace W of $\mathcal{R}^m$

- **orthogonal complement** of W (denoted $W^{\perp}$)
  $\stackrel{def}{=} \{\mathbf{u} \mid \mathbf{u}^T \mathbf{z} = 0, \quad \text{for all} \quad \mathbf{z} \in W\}$

EX: Let $W = \mathbf{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$ with $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

Find $W^{\perp}$.

Orthogonal Complement of subspace W of $\mathcal{R}^m$

- **orthogonal complement** of W (denoted $W^\perp$) $\stackrel{def}{=} \left\{\mathbf{u} \;\middle|\; \mathbf{u}^T\mathbf{z} = 0, \quad \text{for all} \quad \mathbf{z} \in W\right\}$

EX: Let $W = \textbf{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$ with $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

Find $W^\perp$.

- SOLUTION: First note $\mathbf{u} \in W^\perp \Longleftrightarrow \mathbf{u}^T[\mathbf{a}_1,\ \mathbf{a}_2] = [0,\ 0]$.

Orthogonal Complement of subspace W of $\mathcal{R}^m$

- **orthogonal complement** of W (denoted $W^\perp$)
  $\stackrel{def}{=} \left\{\mathbf{u} \;\middle|\; \mathbf{u}^T\mathbf{z} = 0, \quad \text{for all} \quad \mathbf{z} \in W\right\}$

EX: Let $W = \textbf{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$ with $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{a}_2 = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}.$

Find $W^\perp$.

- SOLUTION: First note $\mathbf{u} \in W^\perp \iff \mathbf{u}^T[\mathbf{a}_1, \ \mathbf{a}_2] = [0, \ 0]$.
- Hence $\mathbf{u} \in W^\perp \iff [\mathbf{a}_1, \ \mathbf{a}_2]^T \mathbf{u} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Orthogonal Complement of subspace W of $\mathcal{R}^m$

- **orthogonal complement** of W (denoted $W^\perp$) $\stackrel{def}{=} \{\mathbf{u} \mid \mathbf{u}^T\mathbf{z} = 0, \quad \text{for all} \quad \mathbf{z} \in W\}$

EX: Let $W = \mathbf{Span}\{\mathbf{a}_1, \mathbf{a}_2\}$ with $\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

Find $W^\perp$.

- SOLUTION: First note $\mathbf{u} \in W^\perp \iff \mathbf{u}^T[\mathbf{a}_1, \ \mathbf{a}_2] = [0, \ 0]$.
- Hence $\mathbf{u} \in W^\perp \iff [\mathbf{a}_1, \ \mathbf{a}_2]^T \mathbf{u} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
- $\boxed{W^\perp = \text{Nul } A^T}$, where $A^T \stackrel{def}{=} [\mathbf{a}_1, \ \mathbf{a}_2]^T = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -2 & 1 \end{bmatrix}$.
- $W^\perp = \mathbf{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ -4 \end{bmatrix}\right\} = (\text{Col } A)^\perp$

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

EX: Show the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}, \ \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

EX: Show the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

- SOLUTION: Must verify that $\mathbf{u}_i^T \mathbf{u}_j = 0$ for all $i < j$. Indeed

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

EX: Show the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4\}$ is orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}, \ \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

- SOLUTION: Must verify that $\mathbf{u}_i^T \mathbf{u}_j = 0$ for all $i < j$. Indeed

$$\begin{aligned} \mathbf{u}_1^T \mathbf{u}_2 &= 3 \cdot (-1) + 1 \cdot 2 + 1 \cdot 1 = 0, \\ \mathbf{u}_1^T \mathbf{u}_3 &= 3 \cdot (-1) + 1 \cdot (-4) + 1 \cdot (-7) = 0, \\ \mathbf{u}_2^T \mathbf{u}_3 &= (-1) \cdot (-1) + 2 \cdot (-4) + 1 \cdot (-7) = 0, \\ \mathbf{u}_i^T \mathbf{u}_4 &= 0, \quad i = 1, 2, 3 \end{aligned}$$

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

Thm: Assume $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ is orthogonal set of non-zero vectors. Then $\mathcal{S}$ is L.I.D set.

Orthogonal Set

- A set of vectors $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ in $\mathcal{R}^n$ is an **orthogonal set** if the vectors in the set are mutually orthogonal.

Thm: Assume $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ is orthogonal set of <u>non-zero</u> vectors. Then $\mathcal{S}$ is L.I.D set.

- PROOF: Let

$$\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p = \mathbf{0}. \qquad (\ell_1)$$

- Inner product with $\mathbf{u}_j$ on both sides of (ℓ_1), for $j = 1, \cdots, p$

$$\mathbf{u}_j^T (\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p) = 0. \qquad (\ell_2)$$

- All cross terms in (ℓ_2) die due to orthogonality:

$$\alpha_j \mathbf{u}_j^T \mathbf{u}_j = 0. \quad \Longrightarrow \quad \text{Must have } \alpha_j = 0 \text{ since } \mathbf{u}_j \neq \mathbf{0}.$$

- An **orthogonal basis** for a subspace $W \subseteq \mathcal{R}^n$ is a basis for W that is also an orthogonal set.

- An **orthogonal basis** for a subspace $W \subseteq \mathcal{R}^n$ is a basis for W that is also an <u>orthogonal set</u>.

Thm: Let $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ be an **orthogonal basis** for a subspace $W \subseteq \mathcal{R}^n$. Then for <u>each</u> $\mathbf{y} \in W$, we must have

$$\mathbf{y} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p, \quad \text{with} \quad \alpha_j = \frac{\mathbf{y}^T \mathbf{u}_j}{\mathbf{u}_j^T \mathbf{u}_j}, \quad j = 1, \cdots, p.$$

- An **orthogonal basis** for a subspace $W \subseteq \mathcal{R}^n$ is a basis for W that is also an orthogonal set.

Thm: Let $\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\}$ be an **orthogonal basis** for a subspace $W \subseteq \mathcal{R}^n$. Then for each $\mathbf{y} \in W$, we must have

$$\mathbf{y} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p, \quad \text{with} \quad \alpha_j = \frac{\mathbf{y}^T \mathbf{u}_j}{\mathbf{u}_j^T \mathbf{u}_j}, \quad j = 1, \cdots, p.$$

- PROOF: Given $\mathbf{y} \in W$, $\mathbf{y}$ is linear combination of basis vectors:

$$\mathbf{y} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p. \qquad (\ell_1)$$

- Inner product with $\mathbf{u}_j$ on both sides of (ℓ_1), for $j = 1, \cdots, p$

$$\mathbf{u}_j^T \mathbf{y} = \mathbf{u}_j^T (\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \cdots + \alpha_p \mathbf{u}_p). \qquad (\ell_2)$$

- All cross terms in (ℓ_2) die due to orthogonality:

$$\mathbf{u}_j^T \mathbf{y} = \alpha_j \mathbf{u}_j^T \mathbf{u}_j. \quad \Longrightarrow \quad \text{Must have } \alpha_j = \frac{\mathbf{y}^T \mathbf{u}_j}{\mathbf{u}_j^T \mathbf{u}_j}.$$

EX: Given the set $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal basis, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

Express $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as linear combination of vectors in $\mathcal{S}$.

EX: Given the set $\mathcal{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal basis, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

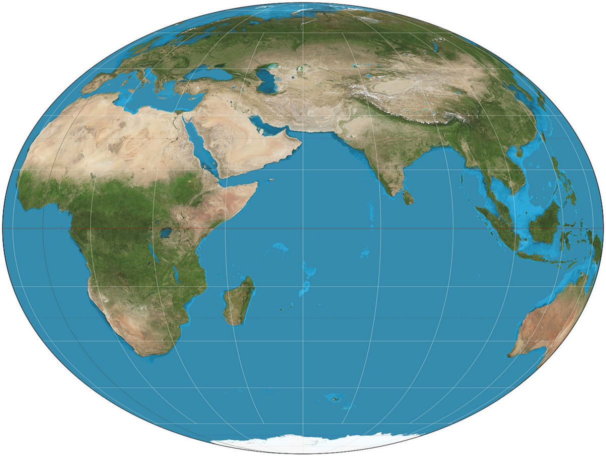
Express $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as linear combination of vectors in $\mathcal{S}$.

- SOLUTION: Write $\mathbf{y} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 + \alpha_3 \mathbf{u}_3$ with

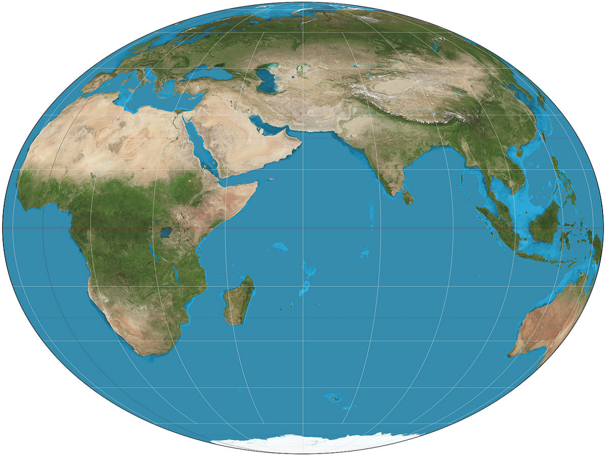
$$\begin{aligned} \alpha_1 &= \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1} = \frac{11}{11} = 1, \\ \alpha_2 &= \frac{\mathbf{y}^T \mathbf{u}_2}{\mathbf{u}_2^T \mathbf{u}_2} = \frac{-12}{6} = -2, \\ \alpha_3 &= \frac{\mathbf{y}^T \mathbf{u}_3}{\mathbf{u}_3^T \mathbf{u}_3} = \frac{-66}{66} = -1. \end{aligned}$$

- So $\mathbf{y} = \mathbf{u}_1 - 2\,\mathbf{u}_2 - \mathbf{u}_3$.

Eastern hemisphere?



Eastern hemisphere?



orthogonal projection: Humans evolutionary trained to see 3D in 2D views

Orthogonal Projection

- Given non-zero vector $\mathbf{u} \in \mathcal{R}^n$.
- For any vector $\mathbf{y} \in \mathcal{R}^n$, decompose

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}, \quad \text{where} \quad \widehat{\mathbf{y}} \in \textbf{Span}\{\mathbf{u}\}, \quad \mathbf{z} \in (\textbf{Span}\{\mathbf{u}\})^{\perp}. \qquad (\ell_1)$$

Orthogonal Projection

- Given non-zero vector $\mathbf{u} \in \mathcal{R}^n$.
- For any vector $\mathbf{y} \in \mathcal{R}^n$, decompose

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}, \quad \text{where} \quad \widehat{\mathbf{y}} \in \textbf{Span}\{\mathbf{u}\}, \quad \mathbf{z} \in (\textbf{Span}\{\mathbf{u}\})^{\perp}. \qquad (\ell_1)$$

- Let $\widehat{\mathbf{y}} = \alpha\,\mathbf{u}$. Re-write (ℓ_1) as

$$\mathbf{y} = \alpha\,\mathbf{u} + \mathbf{z}, \quad \text{where} \quad \mathbf{z} \in (\textbf{Span}\{\mathbf{u}\})^{\perp}. \qquad (\ell_2)$$

- Inner product with $\mathbf{u}$ on both sides of (ℓ_2),

$$\mathbf{u}^T\mathbf{y} = \mathbf{u}^T(\alpha\,\mathbf{u} + \mathbf{z}) = \alpha\,\mathbf{u}^T\mathbf{u}. \quad \Longrightarrow \quad \text{Must have } \alpha = \frac{\mathbf{y}^T\mathbf{u}}{\mathbf{u}^T\mathbf{u}}.$$

- **orthogonal projection** $\boxed{\widehat{\mathbf{y}} = \textbf{Proj}_L\,\mathbf{y} \stackrel{def}{=} \frac{\mathbf{y}^T\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\,\mathbf{u}}$,

  $(L = \textbf{Span}\{\mathbf{u}\})$

Orthonormal Set

- DEFINITION: Set $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is **orthonormal set** if $\mathcal{S}$ is orthogonal set of unit vectors.
- An **orthonormal basis** is basis that is orthonormal set.

- **EX:** Show the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

Orthonormal Set

- DEFINITION: Set $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is **orthonormal set** if $\mathcal{S}$ is orthogonal set of unit vectors.
- An **orthonormal basis** is basis that is orthonormal set.

- **EX:** Show the set $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \ \mathbf{u}_3 = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

- SOLUTION: Only need to verify mutual orthogonality:

$$\mathbf{u}_1^T \mathbf{u}_2 = 0, \ \ \mathbf{u}_1^T \mathbf{u}_3 = 0, \ \ \mathbf{u}_2^T \mathbf{u}_3 = 0.$$

Orthonormal Set

- DEFINITION: Set $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is **orthonormal set** if $\mathcal{S}$ is orthogonal set of unit vectors.
- An **orthonormal basis** is basis that is orthonormal set.

- EXAMPLE: Show the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthonormal, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

Orthonormal Set

- DEFINITION: Set $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is **orthonormal set** if $\mathcal{S}$ is orthogonal set of unit vectors.
- An **orthonormal basis** is basis that is orthonormal set.

- EXAMPLE: Show the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthonormal, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

- SOLUTION: Only need to verify vectors have unit length:

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \quad \mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \quad \mathbf{v}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}.$$

Orthonormal Set

- DEFINITION: Set $\mathcal{S} = \{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$ is **orthonormal set** if $\mathcal{S}$ is orthogonal set of unit vectors.
- An **orthonormal basis** is basis that is orthonormal set.

- EXAMPLE: Show the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthonormal, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

- SOLUTION: Only need to verify vectors have unit length:

$$\mathbf{v}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}, \quad \mathbf{v}_2 = \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|}, \quad \mathbf{v}_3 = \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|}.$$

- $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is orthonormal basis for $\mathcal{R}^3$

Thm: Matrix $U \in \mathcal{R}^{m \times n}$ has orthonormal columns $\iff U^T U = I$.

PROOF: Let $U = (\mathbf{u}_1, \cdots, \mathbf{u}_n)$. Then

$$U^T U = \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} (\mathbf{u}_1, \cdots, \mathbf{u}_n) = \left(\mathbf{u}_i^T \mathbf{u}_j \right).$$

The (i, j) entry of $U^T U$ is $\mathbf{u}_i^T \mathbf{u}_j$.

- For $i = j : \mathbf{u}_i^T \mathbf{u}_i = 1$, each column of U has unit length.
- For $i \neq j : \mathbf{u}_i^T \mathbf{u}_j = 0$, each pair of columns of U is orthogonal.

Matrix of orthonormal columns, Example

- $U = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

Matrix of orthonormal columns, Example

- $U = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

Matrix of orthonormal columns, EXAMPLE

- $U = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

- $U^T U = I, \; U \in \mathcal{R}^{3\times 3}$.

Matrix of orthonormal columns, EXAMPLE

- $U = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$, where

$$\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \frac{1}{\sqrt{66}} \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}.$$

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

- $U^T U = I, \; U \in \mathcal{R}^{3\times 3}$.
- $\implies \quad U^{-1} = U^T$.

Thm: Let $U \in \mathcal{R}^{m \times n}$ have orthonormal columns. Then for $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$

a. $\|U\,\mathbf{x}\| = \|\mathbf{x}\|$. Preserves length.

b. $(U\,\mathbf{x})^T\,(U\,\mathbf{y}) = \mathbf{x}^T\,\mathbf{y}$. Preserves inner product.

Thm: Let $U \in \mathcal{R}^{m \times n}$ have orthonormal columns. Then for $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$

a. $\|U\,\mathbf{x}\| = \|\mathbf{x}\|$. Preserves length.
b. $(U\,\mathbf{x})^T\,(U\,\mathbf{y}) = \mathbf{x}^T\,\mathbf{y}$. Preserves inner product.

- Proof of (**b.**):

$$(U\,\mathbf{x})^T\,(U\,\mathbf{y}) = \mathbf{x}^T\,\left(U^T\,U\right)\,\mathbf{y} = \mathbf{x}^T\,\mathbf{y}.$$

Thm: Let $U \in \mathcal{R}^{m \times n}$ have orthonormal columns. Then for $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$. Preserves length.
b. $(U\mathbf{x})^T (U\mathbf{y}) = \mathbf{x}^T \mathbf{y}$. Preserves inner product.

- PROOF OF (**b.**):

$$(U\mathbf{x})^T (U\mathbf{y}) = \mathbf{x}^T \left(U^T U\right) \mathbf{y} = \mathbf{x}^T \mathbf{y}.$$

- PROOF OF (**a.**): Let $\mathbf{y} = \mathbf{x}$ in (**b.**).

Square matrix of orthonormal columns, Example

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

Square matrix of orthonormal columns, Example

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

- $U^T U = I, \; U \in \mathcal{R}^{3\times 3}. \implies \quad U^{-1} = U^T.$

Square matrix of orthonormal columns, Example

$$U = \begin{bmatrix} \frac{3}{\sqrt{11}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{2}{\sqrt{6}} & -\frac{4}{\sqrt{66}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{6}} & \frac{7}{\sqrt{66}} \end{bmatrix}.$$

- $U^T U = I, \ U \in \mathcal{R}^{3\times 3}. \implies \quad U^{-1} = U^T.$
- Definition: Square matrix of orthonormal columns is **orthogonal matrix**.

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W, \ \ \mathbf{z} \in W^{\perp}$, such that

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

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W, \ \mathbf{z} \in W^{\perp}$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^{\perp}$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

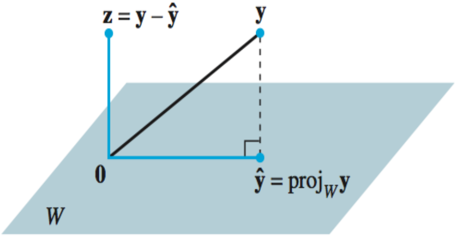


FIGURE 2 The orthogonal projection of **y** onto W.

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^{\perp}$, such that

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

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^\perp$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\,\mathbf{u}_1}{\mathbf{u}_1^T\,\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\,\mathbf{u}_p}{\mathbf{u}_p^T\,\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W, \;\; \mathbf{z} \in W^\perp$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

- PROOF: It is clear that $\widehat{\mathbf{y}} \in W$. Let

$$\mathbf{z} \stackrel{def}{=} \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - \left(\frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p\right).$$

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^\perp$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

- PROOF: It is clear that $\widehat{\mathbf{y}} \in W$. Let

$$\mathbf{z} \overset{def}{=} \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - \left(\frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p\right).$$

- Inner product of $\mathbf{u}_j$ and $\mathbf{z}$ for $j = 1, \cdots, p$:

$$\mathbf{u}_j^T\mathbf{z} = \mathbf{u}_j^T\mathbf{y} - \mathbf{u}_j^T\left(\frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p\right) = \mathbf{u}_j^T\mathbf{y} - \mathbf{u}_j^T\mathbf{u}_j\frac{\mathbf{y}^T\mathbf{u}_j}{\mathbf{u}_j^T\mathbf{u}_j} = 0.$$

§6.3 Orthogonal Projection

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- **Thm**: For any vector $\mathbf{y} \in \mathcal{R}^n$, there exists unique $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^{\perp}$, such that

$$\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}. \qquad \text{with} \quad \widehat{\mathbf{y}} = \frac{\mathbf{y}^T\,\mathbf{u}_1}{\mathbf{u}_1^T\,\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\,\mathbf{u}_p}{\mathbf{u}_p^T\,\mathbf{u}_p}\mathbf{u}_p. \qquad (\ell)$$

- PROOF: It is clear that $\widehat{\mathbf{y}} \in W$. Let

$$\mathbf{z} \overset{def}{=} \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - \left(\frac{\mathbf{y}^T\,\mathbf{u}_1}{\mathbf{u}_1^T\,\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\,\mathbf{u}_p}{\mathbf{u}_p^T\,\mathbf{u}_p}\mathbf{u}_p\right).$$

- Inner product of $\mathbf{u}_j$ and $\mathbf{z}$ for $j = 1, \cdots, p$:

$$\mathbf{u}_j^T\,\mathbf{z} = \mathbf{u}_j^T\,\mathbf{y} - \mathbf{u}_j^T\left(\frac{\mathbf{y}^T\,\mathbf{u}_1}{\mathbf{u}_1^T\,\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\,\mathbf{u}_p}{\mathbf{u}_p^T\,\mathbf{u}_p}\mathbf{u}_p\right) = \mathbf{u}_j^T\,\mathbf{y} - \mathbf{u}_j^T\,\mathbf{u}_j\,\frac{\mathbf{y}^T\,\mathbf{u}_j}{\mathbf{u}_j^T\,\mathbf{u}_j} = 0.$$

- $\Longrightarrow \mathbf{z} \in W^{\perp}$.

Orthogonal Projection, Example

Let $W \stackrel{def}{=} \textbf{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, with $\mathbf{v}_1 = \frac{1}{\sqrt{11}}\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$

and $\mathbf{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$. Write $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$, with $\widehat{\mathbf{y}} \in W$, $\mathbf{z} \in W^{\perp}$,

Orthogonal Projection, EXAMPLE

Let $W \stackrel{def}{=} \textbf{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$, with $\mathbf{v}_1 = \frac{1}{\sqrt{11}} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}$

and $\mathbf{y} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$. Write $\mathbf{y} = \widehat{\mathbf{y}} + \mathbf{z}$, with $\widehat{\mathbf{y}} \in W, \quad \mathbf{z} \in W^{\perp}$,

SOLUTION:

$$\widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{v}_1}{\mathbf{v}_1^T \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{y}^T \mathbf{v}_1}{\mathbf{v}_2^T \mathbf{v}_2}\mathbf{v}_2 = \frac{12}{11} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{33} \begin{bmatrix} 97 \\ 58 \\ 47 \end{bmatrix}.$$

and

$$\mathbf{z} = \mathbf{y} - \widehat{\mathbf{y}} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} - \frac{1}{33} \begin{bmatrix} 97 \\ 58 \\ 47 \end{bmatrix} = \frac{1}{33} \begin{bmatrix} 2 \\ 8 \\ -14 \end{bmatrix}$$

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- DEFINITION: For any vector $\mathbf{y} \in \mathcal{R}^n$, define **orthogonal projection** as

$$\mathbf{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T \mathbf{u}_p}{\mathbf{u}_p^T \mathbf{u}_p}\mathbf{u}_p$$

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- DEFINITION: For any vector $\mathbf{y} \in \mathcal{R}^n$, define **orthogonal projection** as

$$\mathbf{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T \mathbf{u}_p}{\mathbf{u}_p^T \mathbf{u}_p}\mathbf{u}_p$$

- **Thm**: For all $\mathbf{v} \in W, \mathbf{v} \neq \widehat{\mathbf{y}}$,

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

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- DEFINITION: For any vector $\mathbf{y} \in \mathcal{R}^n$, define **orthogonal projection** as

$$\mathbf{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T \mathbf{u}_p}{\mathbf{u}_p^T \mathbf{u}_p}\mathbf{u}_p$$

- **Thm**: For all $\mathbf{v} \in W, \mathbf{v} \neq \widehat{\mathbf{y}}$,

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

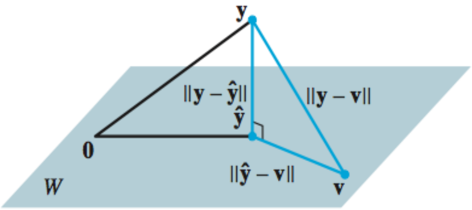


FIGURE 4 The orthogonal projection of **y** onto W is the closest point in W to **y**.

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- DEFINITION: For any vector $\mathbf{y} \in \mathcal{R}^n$, define **orthogonal projection** as

$$\mathbf{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1} \mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T \mathbf{u}_p}{\mathbf{u}_p^T \mathbf{u}_p} \mathbf{u}_p$$

- **Thm**: For all $\mathbf{v} \in W, \mathbf{v} \neq \widehat{\mathbf{y}}$,

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

Best Approximation

Let W be a subspace of $\mathcal{R}^n$ with orthogonal basis $\{\mathbf{u}_1, \cdots, \mathbf{u}_p\}$.

- DEFINITION: For any vector $\mathbf{y} \in \mathcal{R}^n$, define **orthogonal projection** as

$$\mathbf{proj}_W \mathbf{y} = \widehat{\mathbf{y}} = \frac{\mathbf{y}^T \mathbf{u}_1}{\mathbf{u}_1^T \mathbf{u}_1} \mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T \mathbf{u}_p}{\mathbf{u}_p^T \mathbf{u}_p} \mathbf{u}_p$$

- **Thm**: For all $\mathbf{v} \in W, \mathbf{v} \neq \widehat{\mathbf{y}}$,

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

PROOF: It is clear that $\mathbf{z} \stackrel{def}{=} \mathbf{y} - \widehat{\mathbf{y}} \in W^{\perp}$. Then

$$\mathbf{y} - \mathbf{v} = \mathbf{z} + (\widehat{\mathbf{y}} - \mathbf{v}) ,$$

with $\mathbf{z} \in W^{\perp}$, $\widehat{\mathbf{y}} - \mathbf{v} \in W$. By Pythagorean Thm,

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{z}\|^2 + \|\widehat{\mathbf{y}} - \mathbf{v}\|^2 > \|\mathbf{z}\|^2 . \quad \square$$