Review

$S = \{\underline{u}_1, \underline{u}_2, \cdots, \underline{u}_k\}$.

- $S^\perp = \{\underline{v} \in \mathbb{R}^n \mid \underline{v} \cdot \underline{u}_i = 0\}$ $i = 1, \cdots, k$.

Thm 6.7. For any $\underline{v} \in \mathbb{R}^n$, $\underline{v} = \underline{w} + \underline{z}$ for a unique $\underline{w} \in W$ and unique $\underline{z} \in W^\perp$.

- $\underline{w}$ is the OP of $\underline{v}$ onto $W$.
- If $\{\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_k\}$ is an orthonormal basis for $W$, then

$$\underline{w} = (\underline{v} \cdot \underline{w}_1)\, \underline{w}_1 + \cdots + (\underline{v} \cdot \underline{w}_k)\, \underline{w}_k .$$

- Orthogonal projection matrix $P_W = C(C^TC)^{-1}C^T$.

let columns of $C$ be a basis for $W$.

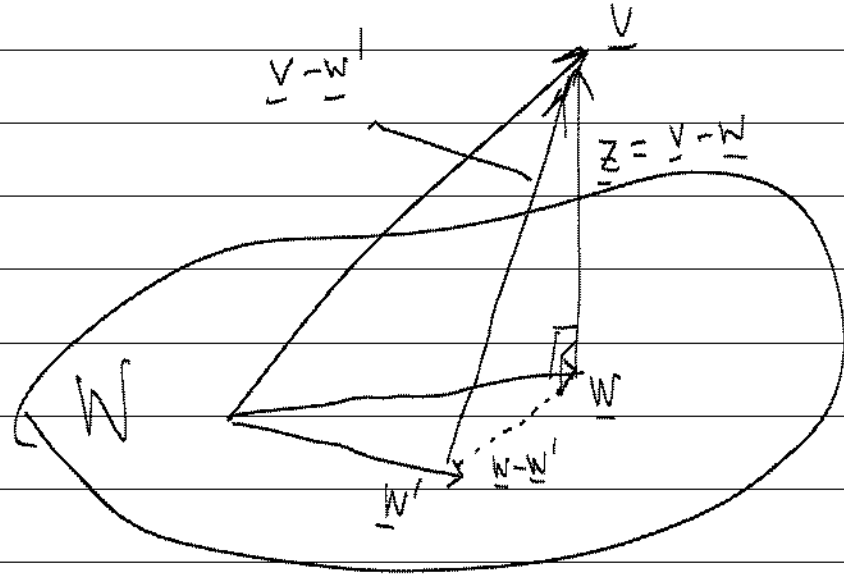
- $\underline{w} = P_W\, \underline{v}$.

---

Closest Vector Property.

- let $\underline{w}$ be the OP of $\underline{v}$ onto $W$.

Then $\underline{w}$ is the vector in $W$ that is closest to $\underline{v}$.




Pf let $\underline{w}'$ be another vector in $W$.

$$\|\underline{v} - \underline{w}'\|^2 = \|(\underline{v} - \underline{w}) + (\underline{w} - \underline{w}')\|^2 \underset{\text{thm}}{\overset{\text{Pythareon}}{=}} \|\underline{v} - \underline{w}\|^2 + \|\underline{w} - \underline{w}'\|^2$$

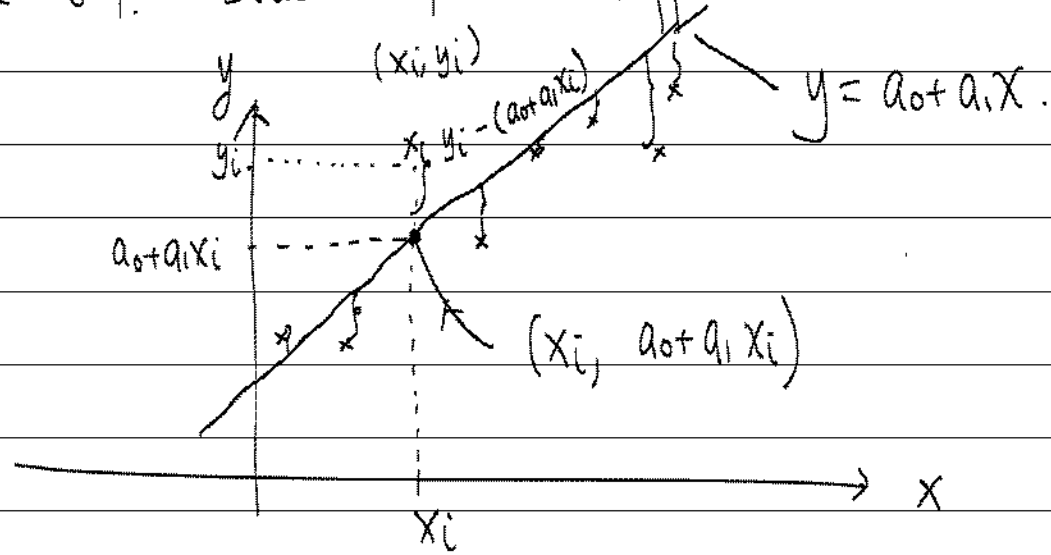
$\underline{v} - \underline{w} \in W^{\perp}$, $\underline{w} - \underline{w}' \in W$.

$\|\underline{v} - \underline{w}'\| \geq \|\underline{v} - \underline{w}\|$ with equality iff $\underline{w}' = \underline{w}$.

\#

Exercise 6.3, T&F, 8, 15, 28, 32, 58, 59, 62, 63, 66, 77, 84.

# Sec. 6.4. Least-Squares Approximations.




Find a line to fit $n$ points: $(x_1, y_1)$ $(x_2, y_2)$ $\cdots$ $(x_n, y_n)$

Error of Sum Squares.

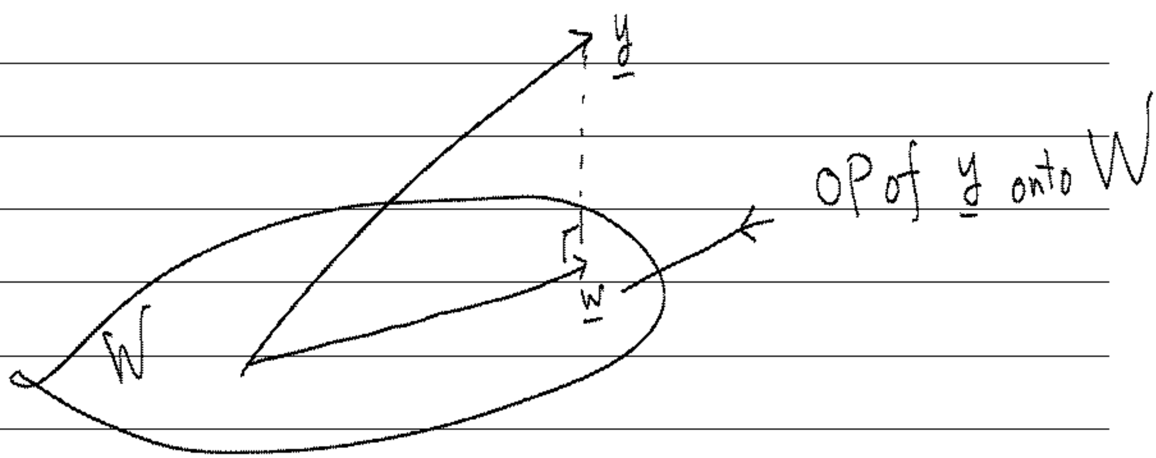
$$E = |y_1 - (a_0 + a_1 x_1)|^2 + |y_2 - (a_0 + a_1 x_2)|^2 + \cdots + |y_i - (a_0 + a_1 x_i)|^2 + \cdots + |y_n - (a_0 + a_1 x_n)|^2 . \quad \text{—— (1)}$$

- Find a line ($y = a_0 + a_1 x$) such that $E$ is a minimum.
- The solution is called the least-squares line.
- Define $\underline{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ $\underline{v_1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$, $\underline{v_2} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

$$\underline{e} = \underline{y} - (a_0 \underline{v}_1 + a_1 \underline{v}_2) = \begin{bmatrix} y_1 - (a_0 + a_1 x_1) \\ y_2 - (a_0 + a_1 x_2) \\ \vdots \\ y_n - (a_0 + a_1 x_n) \end{bmatrix}$$

From ①, we see that $E = \|\underline{e}\|^2 = \|\underline{y} - (a_0 \underline{v}_1 + a_1 \underline{v}_2)\|^2$

- Find $a_0$ and $a_1$ such that $E$ is minimized.
- Define $W = \mathrm{Span}\{\underline{v}_1, \underline{v}_2\}$. Then $a_0 \underline{v}_1 + a_1 \underline{v}_2$ is a vector in $W$.
- We want to find a vector in $W$ that is closest to $\underline{y}$.




- We should find $a_0$ and $a_1$ such that $a_0 \underline{v}_1 + a_1 \underline{v}_1$ is the OP of $\underline{y}$ onto $W$.
- Define $C = [\underline{v}_1 \ \underline{v}_2]$.

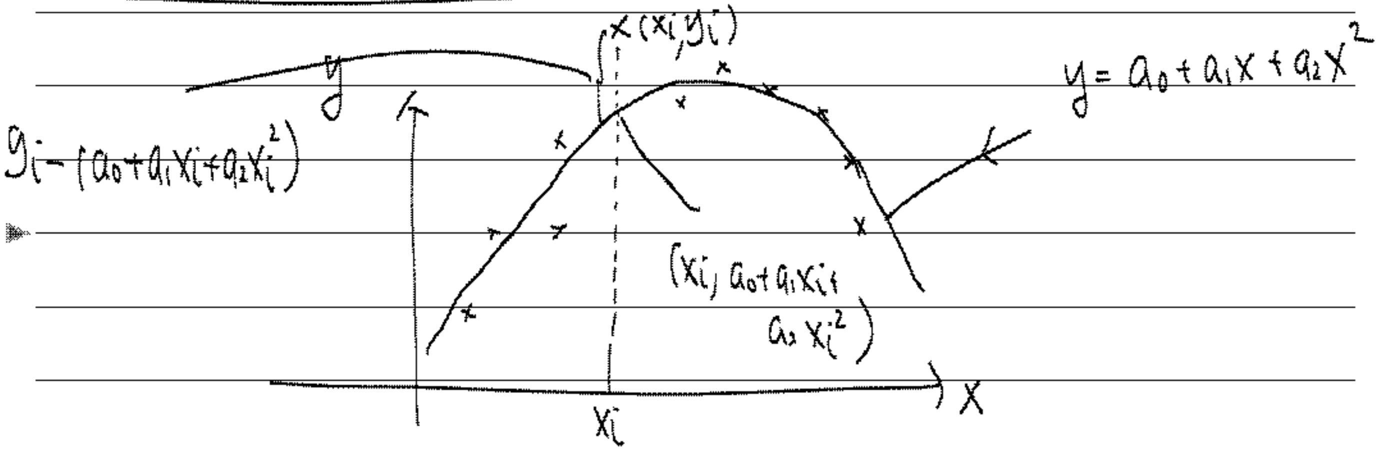
$$a_0 \underline{v}_1 + a_1 \underline{v}_2 = P_W \underline{y}.$$

$$C \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = C\left(C^T C\right)^{-1} C^T \underline{y}.$$

$$\left(C^T C\right) \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \underbrace{\left(C^T C\right)\left(C^T C\right)^{-1}}_{I} C^T \underline{y}$$

$$\begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \left(C^T C\right)^{-1} C^T \underline{y}.$$

---




Find a $2^{nd}$ degree polynomial to fit $n$ points $(x_1, y_1), \ldots, (x_n, y_n)$.

$$E = \left| y_1 - (a_0 + a_1 x_1 + a_2 x_1^2) \right|^2 + \left| y_2 - (a_0 + a_1 x_2 + a_2 x_2^2) \right|^2 + \cdots + \left| y_n - (a_0 + a_1 x_n + a_2 x_n^2) \right|^2.$$

Find $y = a_0 + a_1 x + a_2 x^2$ such that $E$ is a minimum.

Define $\underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, $\underline{v}_1 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, $\underline{v}_2 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, $\underline{v}_3 = \begin{bmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{bmatrix}$.

$$\underline{e} = \underline{y} - (a_0 \underline{v}_1 + a_1 \underline{v}_2 + a_2 \underline{v}_3) = \begin{bmatrix} y_1 - (a_0 + a_1 x_1 + a_2 x_1^2) \\ \vdots \\ y_n - (a_0 + a_1 x_n + a_2 x_n^2) \end{bmatrix}$$

$$E = \|\underline{e}\|^2 = \|\underline{y} - (a_0 \underline{v}_1 + a_1 \underline{v}_2 + a_2 \underline{v}_3)\|^2.$$

Define $W = \text{Span}\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$.

Find a vector in $W$ that is closest to $\underline{y}$.

define $C = [\underline{v}_1 \ \underline{v}_2 \ \underline{v}_3]$.

$$a_0 \underline{v}_1 + a_1 \underline{v}_2 + a_2 \underline{v}_3 = P_W \underline{y}.$$

$$C \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = C (C^T C)^{-1} C^T \underline{y}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = (C^T C)^{-1} C^T \underline{y}.$$

- It is not difficult to extend this to the approximation by a $k$-th degree polynomial $\underline{y} = a_0 + a_1 x + \cdots + a_k x^k$

$$\underline{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \underline{V_1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \underline{V_2} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad V_{k+1} = \begin{bmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{bmatrix}$$

$$C = \begin{bmatrix} \underline{V_1} & \underline{V_2} & \cdots & \underline{V_{k+1}} \end{bmatrix}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{bmatrix} = \left(C^T C\right)^{-1} C^T \underline{y}.$$

---

$$C = \begin{bmatrix} 1 & x_1 & x_1^2 & & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & & & \vdots \\ 1 & x_n & x_n^2 & & x_n^k \end{bmatrix} \qquad C \text{ has } (k+1) \text{ cols.}$$

- When $n = k+1$, $\overset{n\times n}{C}$ is called a Vandermonde matrix.
- It can be proven that when $x_1, x_2, \cdots, x_n$ are distinct, $C$ is invertible.
- That means, when $n \geq k+1$, then cols of $C$ are l.indep.

What if $n < k+1$? no soln because not enough point $(x_i, y_i)$.

Exercise Sec. 6.4, T & F, 13, 14, 34, 35.

## Sec. 6.5. Orthogonal Matrices.

Defn $T: \mathbb{R}^n \to \mathbb{R}^n$ linear. $T$ is said to be norm preserving if $\|T(\underline{v})\| = \|\underline{v}\|$ for every $\underline{v} \in \mathbb{R}^n$.

Let $A$ $(n \times n)$ be the standard matrix of $T$.

Then $\|A\underline{v}\| = \|\underline{v}\|$ for every $\underline{v} \in \mathbb{R}^n$. $A = [T(\underline{e}_1)\ T(\underline{e}_2) \cdots T(\underline{e}_n)]$

Claim $\|A\underline{v}\| = \|\underline{v}\|$ for all $\underline{v} \in \mathbb{R}^n$ iff its cols satisfy

(i) $\|\underline{a}_i\| = 1$ (ii) The cols of $A$ are orthogonal.

i.e. cols of $A$ are orthonormal.

Pf ($\Longrightarrow$) Suppose $\|A\underline{v}\| = \|\underline{v}\|$ for all $\underline{v} \in \mathbb{R}^n$.

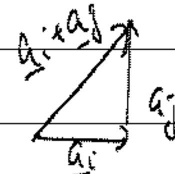
(i) Take $\underline{v} = \underline{e}_i$.

$\|\underline{a}_i\| = \|A\underline{e}_i\| = \|\underline{e}_i\| = 1$.

(ii) $\|\underline{a_i}+\underline{a_j}\|^2 = \|A\underline{e_i}+A\underline{e_j}\|^2 = \|A(\underline{e_i}+\underline{e_j})\|^2$
for $i \neq j$.

$$= \|\underline{e_i}+\underline{e_j}\|^2 = 2 = \|\underline{a_i}\|^2 + \|\underline{a_j}\|^2.$$

$$\|\underline{a_i}+\underline{a_j}\|^2 = \|\underline{a_i}\|^2 + \|\underline{a_j}\|^2$$




From Pythagorean thm, we conclude that $\underline{a_i}$ and $\underline{a_j}$ are orthogonal.

($\Leftarrow$) see Thm 6.9.

---

**Defn** An $n\times n$ matrix $Q$ is called an Orthogonal matrix if its cols are orthonormal.

**Defn** $T$ is called an orthogonal l. operator if its standard matrix is orthogonal.

---

Thm 6.9. Let $Q$ be $n\times n$. The following are equivalent.

(a) $Q$ is orthogonal

(b) $Q$ is invertible and $Q^{-1} = Q^T$.

(c) $Q\underline{u}\cdot Q\underline{v} = \underline{u}\cdot\underline{v}$ for any $\underline{u}, \underline{v} \in \mathbb{R}^n$.

(d) $\|Q\underline{v}\| = \|\underline{v}\|$ for all $\underline{v} \in \mathbb{R}^n$.

Pf (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (a) proved above.

(a) $\Rightarrow$ (b). let $Q = [\underline{q}_1 \ \underline{q}_2 \cdots \underline{q}_n]$ $\underline{q}_i$ : orthonormal.

$$Q^T Q = \begin{bmatrix} \underline{q}_1^T \\ \underline{q}_2^T \\ \vdots \\ \underline{q}_n^T \end{bmatrix} [\underline{q}_1 \ \underline{q}_2 \cdots \underline{q}_n] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I.$$

From Thm 2.6, $Q^{-1} = Q^T$.

(b) $\Rightarrow$ (c)

$$Q\underline{u} \cdot Q\underline{v} = (Q\underline{u})^T Q\underline{v} = \underline{u}^T \underbrace{Q^T Q}_{= I \text{ from (b)}} \underline{v} = \underline{u}^T \underline{v} = \underline{u} \cdot \underline{v}.$$

(c) $\Rightarrow$ (d) Take $\underline{u} = \underline{v}$. #

---

$$Q Q^T = \underline{I}_n \qquad Q = \begin{bmatrix} \underline{r}_1 \\ \underline{r}_2 \\ \vdots \\ \underline{r}_n \end{bmatrix}$$

$$\begin{bmatrix} \underline{r}_1 \\ \underline{r}_2 \\ \vdots \\ \underline{r}_n \end{bmatrix} [\underline{r}_1^T \ \underline{r}_2^T \cdots \underline{r}_n^T] = \begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & & 1 \end{bmatrix}.$$

- The row vectors of an orthogonal matrix are also orthonormal.

Thm 6.10 Let $P, Q$ be orthogonal. Then the following are true.

(a) $\det Q = \pm 1$.

(b) $PQ$ is orthogonal

(c) $Q^{-1}$ is orthogonal

(d) $Q^T$ is orthogonal.

Pf (a) $Q^TQ = I$ $\det(Q^TQ) = 1$.

$(\det Q)^2 = 1$.

(b) $(PQ)^T(PQ) = Q^TP^TPQ = I$

(c) $(Q^{-1})^T Q^{-1} = (Q^T)^T Q^T = QQ^T = I$

(d) $Q^T = Q^{-1}$.

---

Ex. Find an orthogonal l. operator $T$ such that

$$T\left(\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

More general case.

let $\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_k}\}$ and $\{\underline{w_1}, \underline{w_2}, \cdots, \underline{w_k}\}$ be orthonormal subsets of $\mathbb{R}^n$

Find an orthogonal l. operator $T$ such that $T(\underline{v_1}) = \underline{w_1}, \cdots, T(\underline{v_k}) = \underline{w_k}$.

$\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_k}\}$ ① extend it to a basis for $\mathbb{R}^n$ ② Use Gram-Schmidt process to make it an orthogonal basis for $\mathbb{R}^n$. ③ divide norm. $\{\underline{v_1}, \underline{v_2}, \cdots, \underline{v_k}, \underline{v_{k+1}}, \cdots, \underline{v_n}\}$ orthonormal basis for $\mathbb{R}^n$.

$\{\underline{w}_1, \cdots, \underline{w}_k\} \xrightarrow{①②③} \{\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_k, \underline{w}_{k+1}, \cdots, \underline{w}_n\}$

orthonormal basis for $\mathbb{R}^n$.

let $A$ be the standard matrix of $T$.

$$A \underbrace{[\underline{v}_1 \ \underline{v}_2 \ \cdots \ \underline{v}_n]}_{n \times n \quad V} = \underbrace{[\underline{w}_1 \ \underline{w}_2 \ \cdots \ \underline{w}_n]}_{n \times n \quad W}.$$

$V$ and $W$ are orthogonal matrices

$$A = WV^T \quad (A \text{ is orthogonal}).$$

$T_A$ is an orthogonal l. operator that satisfies $T(\underline{v}_1) = \underline{w}_1$

$\cdots \ T(\underline{v}_k) = \underline{w}_k$.

back to Ex.

$\{\underline{v}_1\} = \left\{ \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \right\}$ $\{\underline{w}_1\} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$.

$\underline{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \longrightarrow \{ \underline{v}_1, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\underline{u}_2}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\underline{u}_3} \}$.

$v_1 = \underline{v}_1$

$$\underline{v}_2' = \underline{u}_2 - \frac{\underline{u}_2 \cdot \underline{v}_1}{\|\underline{v}_1\|^2}\underline{v}_1 = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}, \quad \|\underline{v}_2'\|^2 = 1/2$$

$$\underline{v}_3' = \underline{u}_3 - \frac{\underline{u}_3 \cdot \underline{v}_1}{\|\underline{v}_1\|^2}\underline{v}_1 - \frac{\underline{u}_3 \cdot \underline{v}_2'}{\|\underline{v}_2'\|^2}\underline{v}_2' = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\|\underline{v}_3'\|^2 = 1.$$

$$\left\{ \underbrace{\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}}_{\underline{v}_1}, \underbrace{\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}}_{\underline{v}_2}, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\underline{v}_3} \right\}.$$ orthonormal basis for $\mathbb{R}^3$

$$\underline{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \longrightarrow \left\{ \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\underline{w}_1}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\underline{w}_2}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\underline{w}_3} \right\}$$ orthonormal basis for $\mathbb{R}^3$.

$$A \begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \underline{v}_3 \end{bmatrix} = \begin{bmatrix} \underline{w}_1 & \underline{w}_2 & \underline{w}_3 \end{bmatrix}$$

$$A = \begin{bmatrix} \underline{w}_1 & \underline{w}_2 & \underline{w}_3 \end{bmatrix} \begin{bmatrix} \underline{v}_1^T \\ \underline{v}_2^T \\ \underline{v}_3^T \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}.$$

$T_A$ is the orthogonal l. operator that satisfies $T(\underline{v}_1) = \underline{w}_1$.

Exercise Sec. 6.5, T&F, 38, 40, 41, 44, 49,

53, 54, 55.

---

(1)(e) Find CP of $A^{-1}$.

46%

$A = PDP^{-1}$ $\quad A^{-1} = PD^{-1}P^{-1}$. (eigenvalues of $A^{-1}$ are the reciprocals of eigenvalues of $A$)

(a) CP of $A = -(t-3)^2(t-5)$.

4% CP of $A^{-1} = -(t-\frac{1}{3})^2(t-\frac{1}{5})$.

(2). (a) (b) 26%

(c) $\underline{v} = \begin{bmatrix} 4 \\ 2 \\ 1 \\ 5 \end{bmatrix} = c_1\underline{v_1} + c_2\underline{v_2} + c_3\underline{v_3}$.

$$c_i = \frac{\underline{v} \cdot \underline{v_i}}{\|\underline{v_i}\|^2}.$$

---

(3) Given CP of $A = -t^3 + P_2 t^2 + P_1 t + P_0$.

(a) $\det(A^T - tI) = \det(A - tI)^T = \det(A - tI)$ #

(b) CP of $-A$ $= \det(-A - tI) = \det[-(A + tI)]$.

$= \det[-\underbrace{(A - (-t)I)}_{B}] = (-1)^3 \det B$

3x3.

$= (-1)^3 \det(A - (-t)I)$.

$= (-1)^3 \left(-(-t)^3 + P_2(-t)^2 + P_1(-t) + P_0\right)$

$= -t^3 - P_2 t^2 + P_1 t - P_0$.

(c) CP of $A^2$

$\det[A^2 - z^2 I] = \det\left((A + zI)(A - zI)\right)$

$= \det(A - zI) \det(A + zI)$.

$= \det(A - zI) \det(A - (-z)I)$.

$$= \left(-z^3 + p_2 z^2 + p_1 z + p_0\right) \cdot \left(-(-z)^3 + p_2(-z)^2 + p_1(-z) + p_0\right)$$

$$= \left(\underbrace{-z^3 + p_1 z}_{a} \underset{+}{+} \underbrace{p_2 z^2 + p_0}_{b}\right)\left(\underbrace{z^3 - p_1 z}_{-a} \underset{+}{+} \underbrace{p_2 z^2 + p_0}_{b}\right)$$

$$= -\left(-z^3 + p_1 z\right)^2 + \left(p_2 z^2 + p_0\right)^2$$

$$= -\left(z^6 - 2p_1 z^4 + p_1^2 z^2\right) + \left(p_2 z^4 + 2p_0 p_2 z^2 + p_0^2\right)$$

$$\det\left(A^2 - z^2 I\right) = -z^6 + \left(2p_1 + p_2\right) z^4 + \left(-p_1^2 + 2p_0 p_2\right) z^2 + p_0^2.$$

$$\det\left(A^2 - tI\right) = -t^3 + \left(2p_1 + p_2\right) t^2 + \left(-p_1^2 + 2p_0 p_2\right) t + p_0^2$$