EE 635: Applied Linear Algebra
Assignment 5

Soumen Kumar Mondal
23M2157@iitb.ac.in

4. Suppose for a symmetric matrix $A \in \mathbb{R}^{n\times n}$, $\mathbb{U}$ is an A-invariant subspace. Show that $\mathbb{U}^{\perp}$ must also be A-invariant.

$\mathbb{U}$ is an A- invariant subspace.

$\Rightarrow \forall u \in \mathbb{U} \left(u \in \mathbb{U} \longrightarrow Au \in \mathbb{U}. \right)$

We have to prove that $\mathbb{U}^{\perp}$ is also A invariant.

$\mathbb{U}^{\perp} := \left\{ w \in \mathbb{V} \;\middle|\; \forall u \in \mathbb{U}, \langle w | u \rangle = 0_{\mathbb{F}} \right\}.$

Consider the standard inner product on $\mathbb{R}^n$ as,

$\langle w | u \rangle := w^T u$

So $\mathbb{U}^{\perp} := \left\{ w \in \mathbb{V} \;\middle|\; \forall u \in \mathbb{U}, \; w^T u = 0_{\mathbb{F}} \right\}.$

Take any arbitrary vector $w \in \mathbb{U}^{\perp}$.

We know that $\forall u \in \mathbb{U}, \quad w^T u = 0_{\mathbb{F}}.$

Since $u \in \mathbb{U} \rightarrow Au \in \mathbb{U}$ so it must be the case that,

$$w^T A u = 0_{\mathbb{F}}$$

$$\Rightarrow (w^T A u)^T = (0_{\mathbb{F}})^T = 0_{\mathbb{F}}$$

$$\Rightarrow u^T A^T w = 0_{\mathbb{F}}$$

$$\Rightarrow u^T A w = 0_{\mathbb{F}}. \qquad [\because A \text{ is symmetric}]$$

$$\Rightarrow (Aw)^T . u = 0_{\mathbb{F}}. \qquad (\text{taking transpose both side})$$

Note that $\forall u \in \mathbb{U}$, $(Aw)^T u = 0_{\mathbb{F}}$. This means, by definition of $\mathbb{U}^\perp$, $Aw \in \mathbb{U}^\perp$. Hence it is proved that $\forall w \in \mathbb{U}^\perp \ (w \in \mathbb{U}^\perp \longrightarrow Aw \in \mathbb{U}^\perp)$. That implies $\mathbb{U}^\perp$ is A invariant subspace.

6. For $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C \in \mathbb{R}^{p\times n}$, show that the following subspaces are A-invariant: (a) $\mathcal{C} = \operatorname{im}([B\ AB\ \ldots\ A^{n-1}B])$, and (b) $\bar{\mathcal{O}} = \ker([C^T\ A^TC^T\ \ldots\ (A^{n-1})^TC^T]^T)$.

(a) $C = \operatorname{im}\left(\left[B \ \ AB \ \ A^2B \ \ \cdots \ \ A^{n-1}B\right]\right)$

Suppose, $D = \left[B \ \ AB \ \ A^2B \ \ \cdots \ \ A^{n-1}B\right]$

$$= \left[[B]_{n\times m} \ \ [AB]_{n\times m} \ \ [A^2B]_{n\times m} \ \cdots \ [A^{n-1}B]_{n\times m} \right]$$

So $D \in \mathbb{R}^{n\times mn}$ because D has mn columns.

We have to prove that $\operatorname{im}(D)$ is invariant under A. Consider a vector $v \in \operatorname{im}(D)$. So we have to show that $Av \in \operatorname{im}(D)$.

let $P(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0$ be the characteristic polynomial of A.

Since $P(x) = 0$

$\Rightarrow x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0 = 0$

We also know that $p(A) = 0$ (Caley Hamilton thm)

$\Rightarrow A^n + c_{n-1}A^{n-1} + \dots + c_1 A + c_0 = 0$

$\Rightarrow A^n = -c_{n-1}A^{n-1} - \dots - c_1 A - c_0$

consider, $v \in \text{im}(D)$

$\Rightarrow \exists w$ such that $D.w = v$, $w \in \mathbb{R}^{mn}$, $v \in \mathbb{R}^n$

consider Av.

$$Av = ADw$$

$$= A\left[B \;\; AB \;\; A^2B \;\cdots\; A^{n-1}B\right] w$$

$$w = \begin{bmatrix} w_0 \\ \hdashline w_1 \\ \hdashline w_2 \\ \hdashline \vdots \\ \hdashline w_{n-1} \end{bmatrix} \begin{array}{l} \to \in \mathbb{R}^n \\ \to \in \mathbb{R}^n \\ \to \in \mathbb{R}^n \\ \\ \to \in \mathbb{R}^n \end{array}$$

$$\Rightarrow Av = A\left(Bw_0 + ABw_1 + A^2Bw_2 + \dots + A^{n-1}Bw_{n-1}\right)$$

$$= ABw_0 + A^2Bw_1 + A^3Bw_2 + \dots + A^nBw_{n-1}$$

$$= ABw_0 + A^2Bw_1 + \dots + \left(-c_{n-1}A^{n-1} - \dots - c_1A - c_0\right). B.w_{n-1}$$

$$= ABw_0 + A\tilde{B}w_1 + \cdots - c_{n-1}A^{n-1}Bw_{n-1} - \cdots$$

$$\cdots - c_1 ABw_{n-1} - c_0 Bw_{n-1}$$

$$= \begin{bmatrix} B & AB & \cdots & A^{n-1}B & A^{n-1}B \end{bmatrix} \begin{bmatrix} -c_0 w_{n-1} \\ w_0 - c_1 w_{n-1} \\ \vdots \\ w_{n-2} - c_{n-1} w_{n-1} \end{bmatrix}$$

$$= D \cdot w' \quad \text{where } w' \in \mathbb{R}^{mn}$$

Therefore $\exists w' \in \mathbb{R}^{mn}$. s.t $D \cdot w' = Av$

$\Rightarrow Av \in im(D)$.

Hence $im(D)$ is A invariant subspace. (proved)

(b) $\bar{O} = Ker\left(\left[C^T \; A^T C^T \; \cdots \; (A^{n-1})^T C^T \right]^T \right)$

let, $D = \left[C^T \; A^T C^T \cdots (A^{n-1})^T C^T \right]^T$.

$$= \left[C^T \; (CA)^T \cdots (CA^{n-1})^T \right]^T$$

$$= \begin{bmatrix} (C^T)^T \\ ((CA)^T)^T \\ \vdots \\ ((CA^{n-1})^T)^T \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Since $C \in \mathbb{R}^{p \times n}$ and $A \in \mathbb{R}^{n \times n}$ so $CA^i \in \mathbb{R}^{p \times n}$

Therefore, $D \in \mathbb{R}^{pn \times n}$

We have to prove that $\mathrm{Ker}(D)$ is A invariant.

To show that,

$$v \in \mathrm{Ker}(D) \longrightarrow Av \in \mathrm{Ker}(D).$$

Suppose $v \in \mathrm{Ker}(D)$

$\Rightarrow$ $Dv = 0$ where $D \in \mathbb{R}^{pn \times n}$ and $v \in \mathbb{R}^n$

$$\Rightarrow \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} v = 0 \Rightarrow \begin{bmatrix} Cv \\ CAv \\ CA^2v \\ \vdots \\ CA^{n-1}v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{—(i)}$$

Consider $D(Av)$

$$\Rightarrow \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} Av = \begin{bmatrix} CAv \\ CA^2v \\ CA^3v \\ \vdots \\ CA^n v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ CA^n v \end{bmatrix}$$

(from the equ (i) we see that $CA^i v = 0$ $(i \neq n)$

Let $P(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0$ be the characteristic polynomial of A.

Since $P(x) = 0$

$\Rightarrow x^n + c_{n-1}x^{n-1} + \cdots + c_1 x + c_0 = 0$

We also know that $P(A) = 0$ (Caley Hamilton thm)

$\Rightarrow A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 = 0$

$\Rightarrow A^n = -c_{n-1}A^{n-1} - \cdots - c_1 A - c_0$

$\Rightarrow CA^n = -c_{n-1}CA^{n-1} - \cdots - c_1 CA - c_0 C$

$\Rightarrow CA^n v = -c_{n-1}(CA^{n-1}v) - \cdots - c_1(CAv) - c_0(Cv)$

$= 0 - 0 - 0 \cdots - 0 - 0$ (from eqn 1)

$= 0$

Hence,

$$DAv = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} Av = \begin{bmatrix} CAv \\ CA^2 v \\ CA^3 v \\ \vdots \\ CA^n v \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = 0$$

$\Rightarrow Av \in \operatorname{Ker}(D)$. Hence $\operatorname{Ker}(D)$ is A invariant (proved)

12. For a matrix $A \in \mathbb{R}^{m\times n}$ whose SVD is given by $A = U \begin{bmatrix} \Sigma_{r\times r} & 0 \\ 0 & 0 \end{bmatrix} V^T$, the Moore-Penrose pseudoinverse is given by $A^\dagger = V \begin{bmatrix} \Sigma^{-1}_{r\times r} & 0 \\ 0 & 0 \end{bmatrix} U^T$. Prove the following:

(a) $A^{-1} = A^\dagger$ when A is invertible

(b) $(A^\dagger)^\dagger = A$

(c) $(A^\dagger)^T = (A^T)^\dagger$

(d) $A^\dagger = (A^T A)^{-1} A^T$ when A has full column rank and $A^\dagger = A^T (AA^T)^{-1}$ when A has full row rank.

(e) $(A^T A)^\dagger = A^\dagger (A^T)^\dagger$ and $(AA^T)^\dagger = (A^T)^\dagger A^\dagger$

(a) When we use A^{-1} that means $A \in \mathbb{R}^{m\times m}$.

$A = U \Sigma V^T$ where $\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m \end{bmatrix}$

U and V are $\mathbb{R}^{m\times m}$ orthogonal matrices.

$A^+ = V \Sigma^{-1} U^T$ (from the given definition). —(i)

consider,

$A^{-1} = (U\Sigma V^T)^{-1} = (V^T)^{-1} \Sigma^{-1} . U^{-1}$

Since U and V are orthogonal matrix we know -

$UU^{-1} = UU^T = U^{-1}U = U^T U = I$

$\Rightarrow U^{-1} = U^T$

Similarly $V^T = V^{-1} \Rightarrow (V^T)^{-1} = (V^{-1})^{-1} = V$

Hence, $A^{-1} = (V^T)^{-1} \Sigma^{-1} U^{-1} = V \Sigma^{-1} U^T = A^+$ (proved)

(from (i))

(b) $A^+ = V\Sigma' U^T$ where, $\Sigma' = \begin{bmatrix} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{n\times m}$

$$(A^+)^+ = (V\Sigma' U^T)^+$$

$$= (U^T)^+ (\Sigma')^+ V^+ \quad [\because (AB)^+ = B^+A^+]$$

Since U & V are invertible matrices $(U^T)^+ = (U^T)^{-1}$ and $V^+ = V^{-1}$

$$\Rightarrow (A^+)^+ = (U^T)^{-1} (\Sigma')^+ V^{-1}$$

$$= (U^T)^T (\Sigma')^+ . V^T \quad [\because U \,\&\, V \text{ are orthogonal so } U^T = U^{-1}, V^T = V^{-1}]$$

$$= U(\Sigma')^+ . V^T$$

$$\Sigma'\Sigma = \begin{bmatrix} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{n\times m} . \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{m\times n}$$

$$= \begin{bmatrix} 1 & & & & & \\ & 1 & & & 0 & \\ & & \ddots & & & \\ & & & 1 & & \\ & 0 & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{n\times n}$$

Therefore, $(\Sigma')^+ = \Sigma$

Hence, $(A^+)^+ = U(\Sigma')^+ . V^T = U.\Sigma.V^T = A$

(Hence proved)

(c) $(A^+)^T = (A^T)^+$

$A = U\Sigma V^T$

$$\Rightarrow A^T = V\Sigma^T U^T = V.\begin{bmatrix} \sigma_1 & & & & & 0 \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}_{n\times m} U^T$$

$$\Rightarrow (A^T)^+ = U.\begin{bmatrix} 1/\sigma_1 & & & & & 0 \\ & 1/\sigma_2 & & & & \\ & & \ddots & & & \\ & & & 1/\sigma_r & & \\ & 0 & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}_{m\times n} V^T$$

$$= \left(\left(U\begin{bmatrix} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}_{m\times n} V^T\right)^T\right)^T$$

$$= \left(V.\begin{bmatrix} \sigma_1^{-1} & & & & & \\ & \sigma_2^{-1} & & & & \\ & & \ddots & & & \\ & & & \sigma_r^{-1} & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix}_{n\times m} .U^T\right)^T$$

$= (V.\Sigma'.U^T)^T = (A^+)^T$ (Hence proved)

(d) $A^+ = (A^TA)^{-1}A^T$ when A has full column rank.

A has full column rank $\longrightarrow$ $\text{rank}(A) = n$.

$$A = U \begin{bmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \vdots \\ & & \ddots & 0 \\ & & & \sigma_n \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{m\times n} V^T, \quad A^+ = V \underbrace{\begin{bmatrix} 1/\sigma_1 & & & 0 & 0 \\ & \ddots & & \vdots & \vdots \\ & & 1/\sigma_n & 0 & 0 \end{bmatrix}}_{\Sigma'} U^T \quad \text{——} (1)$$

$$A^T = V \begin{bmatrix} \sigma_1 & & & 0 & 0 \\ & \sigma_2 & & \vdots & \vdots \\ & & \ddots & & \\ 0 & \cdots & \sigma_n & 0 & 0 \end{bmatrix}_{n\times m} U^T$$

$$A^TA = V \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_n & 0 \cdots & 0 \end{bmatrix}_{n\times m} \underbrace{U^TU}_{I} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & & 0 \\ & & \vdots \\ & & 0 \end{bmatrix}_{m\times n} V^T$$

$$= V \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}_{n\times n} V^T$$

$$(A^TA)^{-1} = (V^T)^{-1} \begin{bmatrix} 1/\sigma_1^2 & & & \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ & & & 1/\sigma_n^2 \end{bmatrix} V^{-1}$$

$$= (V^T)^T \begin{bmatrix} 1/\sigma_1^2 & & & \\ & 1/\sigma_2^2 & & \\ & & \ddots & \\ & & & 1/\sigma_n^2 \end{bmatrix} V^T$$

$$= V \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_n^2} \end{bmatrix}_{n \times n} V^T$$

$$(A^TA)^{-1}A^T = V \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_n^2} \end{bmatrix}_{n \times n} \underbrace{V^T \cdot V}_{I} \begin{bmatrix} \sigma_1 & & & \vdots \\ & \ddots & & \vdots \\ & & \sigma_n \; 0 & 0 \end{bmatrix}_{n \times m} U^T$$

$$= V \cdot \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_n^2} \end{bmatrix}_{n \times n} \cdot \begin{bmatrix} \sigma_1 & & & 0 & 0 \\ & \ddots & & \vdots & \vdots \\ & & \sigma_n & 0 \cdots & 0 \end{bmatrix}_{n \times m} U^T$$

$$= V \cdot \begin{bmatrix} 1/\sigma_1 & & & & 0 & 0 \\ & 1/\sigma_2 & & & \vdots & \vdots \\ & & \ddots & & \vdots & \vdots \\ & & & 1/\sigma_n & 0 & 0 \end{bmatrix}_{n \times m} U^T$$

$= V\Sigma' U^T$ (from equation (i))

$= A^+$ (Hence proved)

$A^+ = A^T (AA^T)^{-1}$ when A has full row rank.

A has full row rank $\longrightarrow$ rank$(A) = m$

$$A = U \begin{bmatrix} \sigma_1 & & & & 0 & 0 \\ & \sigma_2 & & & \vdots & \vdots \\ & & \ddots & & & \\ & & & \sigma_m & 0 & 0 \end{bmatrix}_{m \times n} V^T \quad = U \Sigma V^T$$

$$A^+ = V \begin{bmatrix} 1/\sigma_1 & & & \\ & 1/\sigma_2 & & \\ & & \ddots & \\ & & & 1/\sigma_m \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix}_{n \times m} U^T \quad = V \Sigma' U^T \qquad \text{--- (i)}$$

Now consider,

$$AA^T = U \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_m & 0 & \cdots & 0 \end{bmatrix}_{m \times n} \underbrace{V^T . V}_{I} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & & 0 \\ & & \vdots \\ & & 0 \end{bmatrix}_{n \times m} U^T$$

$$= U \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_m^2 \end{bmatrix}_{m \times m} U^T$$

$$(AA^T)^{-1} = (U^T)^{-1} \cdot \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_m^2} \end{bmatrix}_{m\times m} U^{-1}$$

$$= U \cdot \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_m^2} \end{bmatrix}_{m\times m} U^T$$

$$A^T \cdot (AA^T)^{-1} = V \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_m \\ & \cdots & 0 \\ & & 0 \end{bmatrix}_{n\times m} \underbrace{U^T \cdot U}_{I} \begin{bmatrix} 1/\sigma_1^2 & & \\ & \ddots & \\ & & \frac{1}{\sigma_m^2} \end{bmatrix}_{m\times m} U^T$$

$$= V \begin{bmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & \\ & & \ddots & \\ & & & \sigma_m^{-1} \\ & & & 0 \\ & & & \vdots \\ & & & 0 \end{bmatrix}_{n\times m} U^T$$

$= V\Sigma' U^T$ (from equation (i))

$= A^+$ (Hence proved)

(e) $(A^TA)^+ = A^+(A^T)^+$

$$A = U \cdot \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{m\times n} V^T$$

$$A^TA = V \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{n\times m} \cdot \underbrace{U^T \cdot U}_{I} \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{m\times n} V^T$$

$$= V \begin{bmatrix} \sigma_1^2 & & & & & \\ & \sigma_2^2 & & & & \\ & & \ddots & & & \\ & & & \sigma_r^2 & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix}_{n\times n} \cdot V^T$$

$$(A^TA)^+ = V \cdot \begin{bmatrix} 1/\sigma_1^2 & & & & \\ & \ddots & & & \\ & & 1/\sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{n\times n} V^T$$

$$= V \begin{bmatrix} 1/\sigma_1 & & & & \\ & \ddots & & & \\ & & 1/\sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{n\times m} \underbrace{U^T U}_{I_{m\times m}} \begin{bmatrix} 1/\sigma_1 & & & & \\ & \ddots & & & \\ & & 1/\sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{m\times n} V^T$$

$$= \left(V \cdot \begin{bmatrix} 1/\sigma_1 & & & & \\ & \ddots & & & \\ & & 1/\sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{n\times m} U^T \right) \left(V \cdot \begin{bmatrix} 1/\sigma_1 & & & & \\ & \ddots & & & \\ & & 1/\sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}_{n\times m} U^T \right)^T$$

$$= A^+ \cdot (A^+)^T = A^+ (A^T)^+ \quad \text{(Hence } \underline{\text{proved}}\text{)}$$
(from c)

Now consider,

$$AA^T = U \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix}_{m\times n} \underbrace{V^T V}_{I_{n\times n}} \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots \; \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix}_{n\times m} \cdot U^T$$

$$= U \cdot \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & \\ & & \sigma_r^2 & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{m\times m} \cdot U^T$$

$$(AA^T)^+ = U \cdot \begin{bmatrix} 1/\sigma_1^2 & & & \\ & \ddots & & \\ & & 1/\sigma_r^2 & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{m\times m} \cdot U^T$$

$$= U \cdot \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{m\times n} \underbrace{V^T \cdot V}_{I_{n\times n}} \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{m\times n} U^T$$

$$= \left(V \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{n\times m} U^T \right)^T \cdot \left(V \cdot \begin{bmatrix} 1/\sigma_1 & & & \\ & \ddots & & \\ & & 1/\sigma_r & \\ & & & 0 \; \ddots \; 0 \end{bmatrix}_{m\times n} U^T \right)$$

$$= (A^+)^T \cdot A^+$$

$$= (A^T)^+ \cdot A^+ \quad \text{(Hence proved)} \quad \text{(from c)}$$

28. Evaluate $\cos(A)$, when $A = \begin{bmatrix} -\pi/2 & \pi/2 \\ \pi/2 & -\pi/2 \end{bmatrix}$.

$$A = \begin{bmatrix} -\pi/2 & \pi/2 \\ \pi/2 & -\pi/2 \end{bmatrix}$$

Solving for eigenvalue -

$$\det(A - \lambda I) = 0$$

$$\Rightarrow \lambda^2 + \pi\lambda = 0$$

$$\Rightarrow \lambda_1 = 0 \text{ and } \lambda_2 = -\pi$$

corresponding to $\lambda_1 = 0$, $v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $Av_1 = \lambda_1 v_1$

corresponding to $\lambda_2 = -\pi$, $v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $Av_2 = \lambda_2 v_2$

consider, $A = PDP^{-1}$

where, $P = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$

$$D = \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -\pi \end{bmatrix}$$

$$A^N = PDP^{-1}PDP^{-1}PDP^{-1} \cdots\cdots PDP^{-1}PDP^{-1}$$

(N times)

$$= PD^NP^{-1}$$

where, $D^N = \begin{bmatrix} 0 & 0 \\ 0 & (-\pi)^N \end{bmatrix}$

We also know that, $\cos(x) = \sum_{N=0}^{\infty} (-1)^N \frac{x^{2N}}{(2N)!}$

Therefore, $\cos(A) = \sum_{N=0}^{\infty} (-1)^N . \frac{A^{2N}}{(2N)!}$

$$\cos(A) = \sum_{N=0}^{\infty} (-1)^N . \frac{P.D^{2N}.P^{-1}}{(2N)!}$$

$$= P.\left(\sum_{N=0}^{\infty} (-1)^N . \frac{D^{2N}}{(2N)!}\right) P^{-1}$$

$$= P. \cos(D). P^{-1}$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \cos(0) & 0 \\ 0 & \cos(-\pi) \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

29. Show that e^A is an orthogonal matrix whenever $A = -A^T$.

We know that, $e^x = \sum_{N=0}^{\infty} \frac{x^N}{N!}$

Therefore, $e^A = \sum_{N=0}^{\infty} \frac{A^N}{N!}$

In order to prove e^A is orthogonal matrix we have to show $(e^A)^T \cdot e^A = I$.

$$(e^A)^T = \left(\sum_{N=0}^{\infty} \frac{A^N}{N!} \right)^T = \sum_{N=0}^{\infty} \frac{(A^N)^T}{N!}$$

We know that, $(A^2)^T = (A \cdot A)^T = A^T \cdot A^T = (A^T)^2$

So in general, $(A^N)^T = (A^T)^N$

$$\Rightarrow \left(e^{A}\right)^{T} = \sum_{N=0}^{\infty} \frac{(A^T)^N}{N!}$$

It is given that, $A = -A^T$

$$\Rightarrow A^T = -A$$

$$\Rightarrow \left(e^{A}\right)^{T} = \sum_{N=0}^{\infty} \frac{(-A)^N}{N!} = e^{-A} \quad \text{(By definition)}$$

Hence,

$$\left(e^{A}\right)^{T} \cdot \left(e^{A}\right) = e^{-A} \cdot e^{A}$$

We know that $e^{A} \cdot e^{B} = e^{A+B}$ if $AB = BA$

Here $-A \cdot A = A \cdot (-A)$ therefore,

$$\left(e^{A}\right)^{T}\left(e^{A}\right) = e^{-A+A} = e^{0} = I$$

Hence e^{A} is orthogonal (Proved)

32. Show that whenever $AB = BA$, we have $e^{A+B} = e^A e^B$ for $A, B \in \mathbb{R}^{n \times n}$. Provide a counterexample to this assertion if $AB \neq BA$.

$$e^{A} = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

$$e^{B} = I + B + \frac{B^2}{2!} + \frac{B^3}{3!} + \cdots$$

$$e^A \cdot e^B = \left(I + A + \frac{A^2}{2!} + \cdots\right) \cdot \left(I + B + \frac{B^2}{2!} + \cdots\right)$$

$$= I + B + \frac{B^2}{2!} + \cdots + A + AB + \frac{AB^2}{2!} + \cdots$$

$$+ \frac{A^2}{2!} + \frac{A^2 B}{2!} + \frac{A^2 B^2}{2!\,2!} + \cdots$$

$$= I + A + B + \frac{A^2 + 2AB + B^2}{2!} + \frac{A^3 + 3A^2B + 3AB^2 + B^3}{3!}$$

$$+ \cdots$$

if A, B commute so $AB = BA$ then,

$$(A+B)^2 = A^2 + AB + BA + B^2 = A^2 + 2AB + B^2$$

$$\Rightarrow e^A \cdot e^B = I + (A+B) + \frac{(A+B)^2}{2!} + \frac{(A+B)^3}{3!} + \cdots$$

$$= e^{(A+B)}$$

Therefore, if $AB = BA$ then $e^{A+B} = e^A \cdot e^B$.

consider, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, $B = \begin{bmatrix} r & s \\ t & u \end{bmatrix}$

$$AB = \begin{bmatrix} ar+bt & as+bu \\ cr+dt & cs+du \end{bmatrix}$$

$$BA = \begin{bmatrix} ar+cs & br+ds \\ at+cu & bt+du \end{bmatrix}$$

Suppose, $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Clearly, $AB = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $BA = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \neq AB.$

Write A, B in eigenspace,

$$A = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \qquad (A = P\Lambda P^{-1})$$

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad (B = P'\Lambda' P'^{-1})$$

$$e^A = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} e^0 & 0 \\ 0 & e^1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2.718 & 1.718 \\ 0 & 1 \end{bmatrix}$$

$$e^B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} e^0 & 0 \\ 0 & e^1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2.718 & 0 \\ 0 & 1 \end{bmatrix}$$

$$e^A \cdot e^B = \begin{bmatrix} 2.718 & 1.718 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2.718 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7.389 & 1.718 \\ 0 & 1 \end{bmatrix}$$

$$A+B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

Write $A+B$ in eigenspace.

$$A+B = \begin{bmatrix} -1/2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1/2 \end{bmatrix}$$

$$e^{A+B} = \begin{bmatrix} -1/2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} e^0 & 0 \\ 0 & e^2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1/2 \end{bmatrix} = \begin{bmatrix} 7.389 & 3.694 \\ 0 & 0 \end{bmatrix}$$

clearly $e^{A+B} \neq e^A \cdot e^B$. Since $AB \neq BA$.

27. Consider two diagonalizable matrices, $A, B \in \mathbb{R}^{n\times n}$. Show that the following statements are equivalent: 1. $AB = BA$, and 2. There exists an invertible matrix $S \in \mathbb{R}^{n\times n}$ such that $S^{-1}AS$ and $S^{-1}BS$ are both diagonal.

We have to prove that,

$AB = BA \longleftrightarrow S^{-1}AS$ and $S^{-1}BS$ are diagonal.

Suppose, $S^{-1}AS$ is diagonal
$S^{-1}BS$ is also diagonal.

let $D_1 = S^{-1}AS$ is diagonal
$D_2 = S^{-1}BS$ is diagonal

$$\Rightarrow \quad A = SD_1S^{-1}$$
$$B = SD_2S^{-1}$$

$$AB = (SD_1S^{-1})(SD_2S^{-1})$$
$$= SD_1D_2S^{-1}$$

We know that for any 2 diagonal matrix $D_1D_2 = D_2D_1$

So, $AB = SD_1D_2S^{-1} = SD_2D_1S^{-1}$

$$= \underbrace{SD_2S^{-1}}\,\underbrace{SD_1S^{-1}}$$

$$= B.A$$

Hence, $S^{-1}AS$, $S^{-1}BS$ are diagonal $\longrightarrow$ $AB = BA$

let's prove the other direction of the statement.

$AB = BA \longrightarrow S^{-1}AS$, $S^{-1}BS$ are diagonal.

Suppose, $AB = BA$.
Because A is diagonalizable, $\exists S \in \mathbb{R}^{n\times n}$ such that S is invertible and $D = S^{-1}AS$ is diagonal.

$$D = \begin{bmatrix} \lambda_1 I_{m_1} & & & 0 \\ & \lambda_2 I_{m_2} & & \\ & & \ddots & \\ 0 & & & \lambda_r I_{m_r} \end{bmatrix} \quad (\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_r)$$

where m_j is the mutiplicity of λ_j.

Because $AB = BA$ we can write,

$$(S^{-1}AS)(S^{-1}BS) = S^{-1}ABS = S^{-1}BAS = (S^{-1}BS)(S^{-1}AS)$$

Suppose, $C = S^{-1}BS$ then we have, $DC = CD$

$$C = \begin{bmatrix} C_1 & & & 0 \\ & C_2 & & \\ & & \ddots & \\ 0 & & & C_r \end{bmatrix} \quad C_j \in \mathbb{R}^{m_j \times m_j} \;\; \forall j$$

B is diagonalizable so $\exists R \in \mathbb{R}^{n \times n}$ s.t. R is invertible and $R^{-1} B R$ is diagonal.

$$R^{-1} S C S^{-1} R = R^{-1} S S^{-1} B S S^{-1} R = R^{-1} B R$$

So C is diagonalizable. We also know that in the block matrix each individual blocks are diagonalizable. so $\forall C_j$ are diagonalizable.

For each j, $\exists T_i \in \mathbb{R}^{m_j \times m_j}$ s.t. T_i^{-1} exists and $T_i^{-1} C_i T_i$ is diagonal,

let, $T = \begin{bmatrix} T_1 & & & 0 \\ & T_2 & & \\ & & \ddots & \\ 0 & & & T_r \end{bmatrix}$, $T^{-1} = \begin{bmatrix} T_1^{-1} & & & 0 \\ & T_2^{-1} & & \\ & & \ddots & \\ 0 & & & T_r^{-1} \end{bmatrix}$

$$T^{-1} S^{-1} B S T = T^{-1} C T = \begin{bmatrix} T_1^{-1} C_1 T_1 & & & 0 \\ & T_2^{-1} C_2 T_2 & & \\ & & \ddots & \\ 0 & & & T_r^{-1} C_r T_r \end{bmatrix}$$

which is also a diagonal matrix.

$$T^{-1} S^{-1} A S T = T^{-1} D T$$

$$T^{-1}DT = \begin{bmatrix} T_1^{-1}\lambda_1 I_{m_1} T_1 & & & 0 \\ & T_2^{-1}\lambda_2 I_{m_2} T_2 & & \\ & & \ddots & \\ 0 & & & T_r^{-1}\lambda_r I_{m_r} T_r \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 I_{m_1} & & & 0 \\ & \lambda_2 I_{m_2} & & \\ 0 & & \ddots & \\ & & & \lambda_r I_{m_r} \end{bmatrix}$$

$$= D$$

Thus both $T^{-1}S^{-1}AST$ and $T^{-1}S^{-1}BST$ are diagonal matrices. So D is diagonal and C is diagonal.

(Hence proved)

35. Obtain the minimal polynomial for the matrix $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 8 \end{bmatrix}$. Hence, evaluate its eigenvalues.

Consider the standard basis for $\mathbb{R}^3$ as.

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathrm{ann}_{v_i}(A) = \{ P(x) \in \mathbb{R}[x] \mid P(A)\cdot v_i = 0_V \}.$$

$\mu_{A_{v_i}}(x)$ = monic polynomial generator of smallest deg of the ideal $\mathrm{ann}_{v_i}(A)$.

construct the set $\{ Iv_1 , Av_1 \}$.

$$Av_1 = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} , \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$$ set is Linearly independent.

construct the set $\{ Iv_1, Av_1, A^2 v_1 \}$

$$A^2 v_1 = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 22 & 45 & 66 \\ 31 & 64 & 94 \\ 39 & 81 & 121 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 22 \\ 31 \\ 39 \end{bmatrix}$$

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} , \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} , \begin{bmatrix} 22 \\ 31 \\ 39 \end{bmatrix} \right\}$$ set is linearly independent

Therefore, clearly $\{ Iv_1, Av_1, A^2 v_1, A^3 v_1 \}$ is a linearly dependent set.

Hence, $A^3 v_1 + a_2 \cdot A^2 v_1 + a_1 \cdot A v_1 + a_0 I v_1 = O_N$.

$$A^3 = \begin{bmatrix} 22 & 45 & 66 \\ 31 & 64 & 94 \\ 39 & 81 & 121 \end{bmatrix} \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 7 \\ 3 & 6 & 8 \end{bmatrix} = \begin{bmatrix} 310 & 642 & 953 \\ 441 & 913 & 1355 \\ 564 & 1167 & 1730 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 310 \\ 441 \\ 564 \end{bmatrix} + a_2 \cdot \begin{bmatrix} 22 \\ 31 \\ 39 \end{bmatrix} + a_1 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + a_0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 22 & 1 & 1 \\ 31 & 2 & 0 \\ 39 & 3 & 0 \end{bmatrix} \begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} -310 \\ -441 \\ -564 \end{bmatrix}$$

Solving the equation, $a_2 = -13$, $a_1 = -19$, $a_0 = -5$

Therefore we get,

$$\mu_{A v_1}(x) = x^3 + a_2 x^2 + a_1 x + a_0$$

$$\boxed{\mu_{A v_1}(x) = x^3 - 13x^2 - 19x - 5}$$

construct the set $\{ I v_2, A v_2 \}$.

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix} \right\}$$ is linearly independent.

consider the set $\{ Iv_2, Av_2, A^2v_2 \}$.

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 45 \\ 64 \\ 81 \end{bmatrix} \right\}$$ is also linearly independent.

clearly, $\{ Iv_2, Av_2, A^2v_2, A^3v_2 \}$ is Linearly dependent set.

Hence, $A^3v_2 + a_2 \cdot A^2v_2 + a_1 Av_2 + a_0 Iv_2 = 0_V$.

$$\Rightarrow \begin{bmatrix} 642 \\ 913 \\ 1167 \end{bmatrix} + a_2 \begin{bmatrix} 45 \\ 64 \\ 81 \end{bmatrix} + a_1 \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system of equation we get -

$$\begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 45 & 3 & 0 \\ 64 & 4 & 1 \\ 81 & 6 & 0 \end{bmatrix}^{-1} \begin{bmatrix} -642 \\ -913 \\ -1167 \end{bmatrix} = \begin{bmatrix} -13 \\ -19 \\ -5 \end{bmatrix}$$

$$\Rightarrow \boxed{\mu_{Av_2}(x) = x^3 - 13x^2 - 19x - 5}$$

construct the set $\{ Iv_3, Av_3 \}$

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix} \right\}$$ is linearly independent.

consider the set $\{ Iv_3, Av_3, A^2v_3 \}$.

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 66 \\ 94 \\ 121 \end{bmatrix} \right\}$$ is also linearly independent.

clearly, $\{ Iv_3, Av_3, A^2v_3, A^3v_3 \}$ is Linearly dependent set.

Hence, $A^3v_3 + a_2 \cdot A^2v_3 + a_1 A v_3 + a_0 I v_3 = 0_V$.

$$\Rightarrow \begin{bmatrix} 953 \\ 1355 \\ 1730 \end{bmatrix} + a_2 \begin{bmatrix} 66 \\ 94 \\ 121 \end{bmatrix} + a_1 \begin{bmatrix} 5 \\ 7 \\ 8 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system of equation we get -

$$\begin{bmatrix} a_2 \\ a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 66 & 5 & 0 \\ 94 & 7 & 0 \\ 121 & 8 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -953 \\ -1355 \\ -1730 \end{bmatrix} = \begin{bmatrix} -13 \\ -19 \\ -5 \end{bmatrix}$$

$$\Rightarrow \boxed{\mu_{Av_3}(x) = x^3 - 13x^2 - 19x - 5}$$

The minimal polynomial of A is -

$$\mu_A(x) = LCM\left(\mu_{Av_1}(x), \mu_{Av_2}(x), \mu_{Av_3}(x)\right)$$

$$= LCM\left(x^3 - 13x^2 - 19x - 5,\ x^3 - 13x^2 - 19x - 5,\ x^3 - 13x^2 - 19x - 5\right)$$

$$= x^3 - 13x^2 - 19x - 5.$$

The characteristic polynomial of A is-

$$\chi_A(x) = \det(A - xI)$$

$$\chi_A(x) = \begin{vmatrix} 1-x & 3 & 5 \\ 2 & 4-x & 7 \\ 3 & 6 & 8-x \end{vmatrix}$$

$$\Rightarrow (1-x)\left[(4-x)(8-x) - 6\times 7\right] - 3\left(2\times(8-x) - 3\times 7\right) + 5\left[2\times 6 - 3\times(4-x)\right]$$

$$\Rightarrow -x^3 + 13x^2 + 19x + 5$$

To evaluate the eigenvalue, $\chi_A(x) = 0$

$$\Rightarrow x^3 - 13x^2 - 19x - 5 = 0$$

$$\Rightarrow (x+1)(x^2-14x-5) = 0$$

$$\Rightarrow x_1 = -1, \quad x_{2,3} = \frac{+14 \pm \sqrt{(14)^2 + 4\times 5}}{2}$$

$$= 7 \pm 3\sqrt{6}$$

Therefore the 3 eigenvalues will be -

$$\left.\begin{aligned} x_1 &= -1 \\ x_2 &= 7+3\sqrt{6} \\ x_3 &= 7-3\sqrt{6} \end{aligned}\right\}$$

Note: It could have been calculated by setting $\mu_A(x) = 0$ also as the roots of $\chi_A(x)$ and $\mu_A(x)$ are same without multiplicity.

13. (a) Argue why complex matrices A and A^H (for matrices over $\mathbb{C}$, $(\cdot)^H$ denotes entry-wise conjugation of the transpose of a matrix) must have eigenvalues that are conjugates of one another.

(b) If u and v are eigenvectors of complex matrices A and A^H, respectively, for two eigenvalues that are not complex conjugates of one another, show that u and v must be orthogonal to each other.

(c) For a matrix $A \in \mathbb{R}^{n\times n}$, prove that if λ is an eigenvalue with algebraic multiplicity of one (also called a *simple* eigenvalue), the left and right eigenvectors corresponding to λ cannot be orthogonal, irrespective of whether A is diagonalizable or not (left eigenvector is defined as $v \neq 0$ such that $v^*A = \lambda v^*$ for some eigenvalue λ).

(a) $A \in \mathbb{C}^{n\times n}$, $A^H \in \mathbb{C}^{n\times n}$

We have to prove that eigenvalues of A and eigenvalues of A^H are complex conjugate of one another.

Suppose A has eigenvalues α and eigenvector x

Suppose A^H has eigenvalues β and eigenvector y

We Know that $\quad Ax = \alpha x \quad \text{—— (i)}$

$$A^H y = \beta y \quad \text{—— (ii)}$$

Taking complex conjugali of equ (i) we get

$$\Rightarrow x^H A^H = \bar{\alpha} x^H \quad \text{—— (iii)}$$

Multiply (iii) by y and (ii) by x^H.

$$\Rightarrow \quad x^H A^H y = \bar{\alpha} x^H y \qquad (\text{Post multiply})$$

$$x^H A^H y = \beta x^H y \qquad (\text{Pre multiply})$$

Subtract : $\quad x^H A^H y - x^H A^H y = \bar{\alpha} x^H y - \beta x^H y$

$$\Rightarrow (\bar{\alpha} - \beta)\, x^H y = 0$$

Since x & y are eigenvectors so $x \neq 0$ and $y \neq 0$
hence $x^H y \neq 0$ as A and A^H need not be equal.

$$\Rightarrow \bar{\alpha} - \beta = 0$$

$$\Rightarrow \beta = \bar{\alpha}$$

Therefore it is proved that eigenvalues of A^H is the complex conjugali of eigenvalues of A.

(b) Eigenvector of A is u, corresponding eigenvalue $= \alpha$
Eigenvector of A^H is v, corresponding eigenvalue $= \beta$

We have, $Au = \alpha u$ —— (i)

$$A^H v = \beta v \quad \text{—— (ii)}$$

Perform complex conjugate on eqn (i)

$$u^H A^H = \bar{\alpha} u^H \quad \text{—— (iii)}$$

Premultiply by u^H to eqn (ii) & postmultiply by v to eqn (iii) we get -

$$u^H A^H v = \beta u^H v$$

$$u^H A^H v = \bar{\alpha} u^H v$$

Subtract : $u^H A^H v - u^H A^H v = (\beta - \bar{\alpha}) u^H v$

$$\Rightarrow (\beta - \bar{\alpha}) u^H v = 0$$

It is given that the 2 eigenvalues that are not complex conjugate of each other so $\beta \neq \bar{\alpha}$.
That means $\beta - \bar{\alpha} \neq 0$. Hence,

$$u^H v = 0$$

In $\mathbb{C}^n$, the definition of standard inner product is –

$$\langle u | v \rangle := u^H v$$

Therefore, $\langle u | v \rangle = 0$

$\Rightarrow$ u and v must be orthogonal to each other.

(Hence proved).

(c) $A \in \mathbb{R}^{n \times n}$

λ is an eigenvalue of A with algebraic multiplicity 1.
Suppose the right eigenvector of A is x. $(x \neq 0)$
Suppose the left eigenvector of A is y $(y \neq 0)$

We have $\quad Ax = \lambda x$

$$y^* A = \lambda y^*$$

$$\Rightarrow (A - \lambda I) x = 0$$

and $y^* (A - \lambda I) = 0$

$$\Rightarrow (A - \lambda I)^* y = 0$$

Since, algebraic multiplicity of $\lambda = 1$ so the geometric multiplicity of x, $\dim(A - \lambda I) \leq 1$.
Since $x \neq 0$ so, $\dim(A - \lambda I) = 1$.

Therefore, since x is an eigenvector we can say that, $\operatorname{Ker}(A-\lambda I) = \langle \{x\} \rangle$

Similarly, $\ker((A-\lambda I)^*) = \langle \{y\} \rangle$

To show that $y^* x \neq 0$ we assume the contrary that let $y^* x = 0$.

$\Rightarrow \langle y | x \rangle = 0$

$\Rightarrow x \in (\langle \{y\} \rangle)^\perp$ (orthogonal complement)

$$(\langle \{y\} \rangle)^\perp = (\ker((A-\lambda I)^*))^\perp$$

For any operator $T \in L(V, V)$ we know this already that,

$$(\operatorname{Ker}(T^*))^\perp = \operatorname{im}(T)$$

Using this fact we have -

$$(\langle \{y\} \rangle)^\perp = (\ker((A-\lambda I)^*))^\perp = \operatorname{im}(A-\lambda I)$$

Therefore, $x \in \operatorname{Im}(A-\lambda I)$.

We already have, $x \in \operatorname{Ker}(A-\lambda I)$.

Again, $(A-\lambda I)x = 0$

$\Rightarrow (A-\lambda I)^2 x = 0$

$\Rightarrow x \in \operatorname{Ker}((A-\lambda I)^2)$

Since $AM(\lambda) = 1$ that means,

$\dim(\operatorname{Ker}((A-\lambda I)^2)) = \dim(\operatorname{Ker}(A-\lambda I)) = 1.$

We have already obtained that,

$x \in \operatorname{Im}(A-\lambda I)$

$\Rightarrow (A-\lambda I)\cdot u = x \quad (\exists u \neq 0 \text{ and } u \in \mathbb{V}.)$

$\Rightarrow (A-\lambda I)^2 u = (A-\lambda I)x = 0$

$\Rightarrow u \in \operatorname{Ker}((A-\lambda I)^2).$

That means u can be written as LC of x. since the $\operatorname{Ker}((A-\lambda I)^2) = \langle \{x\} \rangle.$

So, $u = \alpha \cdot x$ where $\alpha \neq 0.$

$\Rightarrow (A-\lambda I)u = \alpha \cdot (A-\lambda I)x$

$\Rightarrow x = \alpha \cdot (A-\lambda I)x = \alpha \cdot 0 = 0.$

which is a contradiction of the fact that $x \neq 0$ since x is an eigenvector.

Therefore, the assumption was incorrect. Hence, $y^* x \neq 0$. So left & right eigenvector can't be orthogonal. (proved).

7. Let $\mathbb{W} \subseteq \mathbb{R}^n$ be an A-invariant subspace such that $\langle w_1, w_2, \ldots, w_k \rangle = \mathbb{W}$. Show that $\exists\ S \in \mathbb{R}^{k \times k}$ such that $AW = WS$. where $W = [w_1\ w_2\ \ldots\ w_k]$. Further, prove that A and S share at least r eigenvalues, where $r = \text{rank}(W)$.

$\mathbb{W}$ is a subspace of $\mathbb{R}^n$.
$\mathbb{W}$ is invariant under the operation A.

$$\mathbb{W} = \langle \{w_1, w_2, \ldots, w_K\} \rangle$$

$$\left.\begin{array}{lcl} w_1 \in \mathbb{W} & \longrightarrow & Aw_1 \in \mathbb{W} \\ w_2 \in \mathbb{W} & \longrightarrow & Aw_2 \in \mathbb{W} \\ \vdots & & \\ w_K \in \mathbb{W} & \longrightarrow & Aw_K \in \mathbb{W}. \end{array}\right\} \text{Because } \mathbb{W} \text{ is } A\text{-invariant.}$$

We can see that $Aw_1 \in \mathbb{W}$ so Aw_1 can be written as linear combination of $\{w_1, w_2, \ldots, w_K\}$.

$$Aw_1 = S_{11} w_1 + S_{12} w_2 + \cdots + S_{1K} \cdot w_K \qquad \forall S_{1K} \in \mathbb{R}.$$

Similarly we write $Aw_2, \ldots, Aw_K$ in terms of LC.

$$Aw_2 = s_{21} w_1 + s_{22} w_2 + \cdots + s_{2K} w_K$$

$$\vdots$$

$$Aw_K = s_{K1} w_1 + s_{K2} w_2 + \cdots + s_{KK} w_K$$

We can write in matrix form as:

$$s_{11} \begin{bmatrix} w_1 \\ \\ \\ \end{bmatrix}_{n\times 1} + s_{12} \begin{bmatrix} w_2 \\ \\ \\ \end{bmatrix}_{n\times 1} + \cdots + s_{1K} \begin{bmatrix} w_K \\ \\ \\ \end{bmatrix}_{n\times 1} = \begin{bmatrix} Aw_1 \\ \\ \\ \end{bmatrix}_{n\times 1}$$

$$\Rightarrow \begin{bmatrix} \\ w_1 & w_2 & \cdots & w_K \\ \\ \end{bmatrix}_{n\times K} \begin{bmatrix} s_{11} \\ s_{12} \\ \vdots \\ s_{1K} \end{bmatrix}_{K\times 1} = \begin{bmatrix} Aw_1 \\ \\ \\ \end{bmatrix}_{n\times 1}$$

Similarly,

$$\begin{bmatrix} \\ w_1 & w_2 & \cdots & w_K \\ \\ \end{bmatrix}_{n\times K} \begin{bmatrix} s_{21} \\ s_{22} \\ \vdots \\ s_{2K} \end{bmatrix}_{K\times 1} = \begin{bmatrix} Aw_2 \\ \\ \\ \end{bmatrix}_{n\times 1}$$

$$\begin{bmatrix} & & \vdots & \\ w_1 & w_2 & \cdots & w_k \\ & & & \end{bmatrix}_{n\times k} \begin{bmatrix} s_{k1} \\ s_{k2} \\ \vdots \\ s_{kk} \end{bmatrix}_{k\times 1} = \begin{bmatrix} Aw_k \end{bmatrix}_{n\times 1}$$

Combining all the matrices equations together –

$$\underbrace{\begin{bmatrix} w_1 & w_2 & \cdots & w_k \\ & & & \end{bmatrix}_{n\times k}}_{\text{Given as } W.} \underbrace{\begin{bmatrix} s_{11} & s_{21} & & s_{k1} \\ s_{12} & s_{22} & & s_{k2} \\ \vdots & \vdots & \cdots & \vdots \\ s_{1k} & s_{2k} & & s_{kk} \end{bmatrix}_{k\times k}}_{\text{consider it as } S \in \mathbb{R}^{k\times k}.} = \begin{bmatrix} Aw_1 & Aw_2 & \cdots & Aw_k \\ & & & \end{bmatrix}_{n\times k}$$

$$\Rightarrow \quad W\cdot S = A\cdot \begin{bmatrix} w_1 & w_2 & \cdots & w_k \\ & & & \end{bmatrix} = A\cdot W$$

$\Rightarrow \quad A\cdot W = W\cdot S$ (Hence it is proved)

Given that $r = \text{rank}(W)$

$\Rightarrow$ W has r linearly independent columns.

$A \in \mathbb{R}^{n\times n}$, $S \in \mathbb{R}^{k\times k}$.

We have to prove that A and S shares at least r eigenvalues.

Consider an eigenvector u of S so $u \in \mathbb{R}^K$ $(u \neq 0)$

$Su = \lambda . u$ where λ is an eigenvalue of S.

$\Rightarrow WSu = \lambda . Wu$

$\Rightarrow AWu = \lambda . Wu.$ $(\because WS = AW)$

Suppose, $Wu = x$ where $x \in \mathbb{R}^n$.

$\Rightarrow Ax = \lambda x$

If we can prove that $x \neq 0$ then x will be an eigenvector of A with the same shared eigenvalue λ.

Clearly, $x = Wu$.

$$\begin{bmatrix} w_1 & w_2 & \cdots & w_K \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_K \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$\text{rank}(W) = r$

$\Rightarrow \dim(\text{Im}(W)) = r$

$\Rightarrow \dim(\text{Ker}(W)) = n - r$

That means $\exists$ $n-r$ basis vectors u_i in $\mathrm{Ker}(W)$ s.t. $Wu = 0$ so such x can't qualify for eigenvectors of A. ($u \neq 0$ but still $x = Wu = 0$)

Therefore, we have to pick x from $\mathrm{Im}(W)$ such that $\exists u \neq 0$ s.t. $Wu = x$ and $x \neq 0$. Therefore x can be a valid choice of eigenvector of A. As the $\mathrm{Im}(W)$ is spanned by r such basis vectors so we can get at least r such x_i's as eigenvectors so r such eigenvalues will be shared. As we know for r such $x \neq 0$, we have $u \neq 0$ in $\mathbb{R}^K$.

Therefore, A & S will share at least r eigenvalues.

(Hence proved)

40. Prove that for an operator $\varphi : \mathbb{V} \to \mathbb{V}$, with $\dim(\mathbb{V}) = n < \infty$, we have $\mathbb{V} = \mathrm{Ker}(\varphi^n) \oplus \mathrm{im}(\varphi^n)$.

In order to prove this we will proof some lemma and use it in the proof.

<u>lemma 1</u>: Suppose, $T \in L(V, V)$

$$\{0_V\} = \mathrm{Ker}(T^0) \subseteq \mathrm{Ker}(T) \subseteq \cdots \subseteq \mathrm{ker}(T^K) \subseteq \cdots$$

<u>proof</u>: Suppose, $v \in \mathrm{Ker}(T^K)$.

$\Rightarrow \quad T^k v = 0_V$

$\Rightarrow \quad T \cdot T^k \cdot v = 0_V$

$\Rightarrow \quad T^{k+1} \cdot v = 0_V$

So, $\quad v \in \ker(T^{k+1})$.

Therefore, $\quad \ker(T^k) \subseteq \ker(T^{k+1}) \qquad$ (Proved)

Lemma 2: Suppose $T \in L(V, V)$.

if $\ker(T^m) = \ker(T^{m+1}) \quad (m \geq 0)$ then

$\ker(T^m) = \ker(T^{m+1}) = \ker(T^{m+2}) = \cdots\cdots$

proof: Suppose $k \geq 0$

We know from lemma 1 that,

$$\ker(T^{m+k}) \subseteq \ker(T^{m+k+1}) \quad \text{———} (i)$$

Suppose $\quad v \in \ker(T^{m+k+1})$

$\Rightarrow \quad T^{m+k+1} \cdot v = 0_V$

$\Rightarrow \quad T^{m+1} \cdot (T^k v) = 0_V$

$\Rightarrow (T^K v) \in \ker(T^{m+1})$

$\Rightarrow (T^K v) \in \ker(T^m)$ [$\because$ Given that $\ker(T^m) = \ker(T^{m+1})$]

$\Rightarrow T^m (T^K v) = 0_V$

$\Rightarrow T^{m+K}(v) = 0_V$

$\Rightarrow v \in \ker(T^{m+K})$

Therefore, $\ker(T^{m+K+1}) \subseteq \ker(T^{m+K})$ — (ii)

Combining (i) & (ii), we get,

$\ker(T^{m+K}) = \ker(T^{m+K+1})$ (Hence proved)

Lemma 3: Suppose, $T \in L(V, V)$, $n = \dim(V)$

$\ker(T^n) = \ker(T^{n+1}) = \ker(T^{n+2}) = \cdots$

proof: From lemma 2, it is clear that when $m = n$ then the above equality holds.

Hence we just have to prove the hypothesis of lemma 2. which is to show that

$\ker(T^n) = \ker(T^{n+1})$.

Suppose it is not true.

From Lemma-1, we have,

$$\{0_V\} = \mathrm{Ker}(T^0) \subseteq \mathrm{Ker}(T) \subseteq \mathrm{Ker}(T^2) \subseteq \cdots$$

The subset equality will hold true only in the case when $\mathrm{Ker}(T^m) = \mathrm{Ker}(T^{m+1})$. But since, it is not the case we have -

$$\{0_V\} = \mathrm{Ker}(T^0) \subset \mathrm{Ker}(T) \subset \mathrm{Ker}(T^2) \subset \cdots$$
$$\cdots \subset \mathrm{Ker}(T^n) \subset \mathrm{Ker}(T^{n+1}) \subset \cdots$$

Since it is strictly inequality so the dimension of the subspaces must increase by at least 1.

$$\dim(\mathrm{Ker}(T^0)) = 0$$
$$\dim(\mathrm{Ker}(T)) \geqslant 1$$
$$\dim(\mathrm{Ker}(T^2)) \geqslant 2$$
$$\vdots$$
$$\dim(\mathrm{Ker}(T^n)) \geqslant n$$
$$\dim(\mathrm{Ker}(T^{n+1})) \geqslant n+1$$

Which is a contradiction of the fact that a subspace of $\mathbb{V}$ can't have larger dimension than n. So the assumption was false.

Hence $\operatorname{Ker}(T^n) = \operatorname{Ker}(T^{n+1})$ (proved)

<u>Theorem</u>: $\phi \in L(V, V)$, let $n = \dim(\mathbb{V})$

$$\mathbb{V} = \operatorname{Ker}(\varphi^n) \oplus \operatorname{Im}(\varphi^n)$$

<u>proof</u>: If we want to show that it is direct sum then we have to prove that,

$$\operatorname{Ker}(\varphi^n) \cap \operatorname{Im}(\varphi^n) = \{0_{\mathbb{V}}\}.$$

Suppose, $v \in \operatorname{Ker}(\varphi^n) \cap \operatorname{Im}(\varphi^n)$.

$\Rightarrow v \in \operatorname{Ker}(\varphi^n)$ and $v \in \operatorname{Im}(\varphi^n)$.

$\Rightarrow \phi^n v = 0_{\mathbb{V}}$. and $\phi^n . u = v \quad \exists u \in \mathbb{V}$.

$$\Rightarrow \phi^n . \phi^n . u = \phi^n . v$$

$$\Rightarrow \phi^{2n} . u = 0_{\mathbb{V}}.$$

$$\Rightarrow u \in \operatorname{Ker}(\varphi^{2n}).$$

From Lemma 3, we know that,

$$\mathrm{Ker}(\varphi^n) = \mathrm{Ker}(\varphi^{n+1}) = \cdots = \mathrm{Ker}(\varphi^{2n}) = \cdots$$

Therefore,

$$u \in \mathrm{Ker}(\varphi^{2n}) \longrightarrow u \in \mathrm{Ker}(\varphi^n).$$

$$\Rightarrow \phi^n . u = 0_{\mathbb{V}}.$$

We also have, $v = \phi^n . u = 0_{\mathbb{V}}$.

Therefore, $\mathrm{Ker}(T^n) \cap \mathrm{Im}(T^n) = \{0_{\mathbb{V}}\}$.

(proved)

We know that if $\mathbb{U}$ and $\mathbb{W}$ are the 2 subspaces of $\mathbb{V}$. Then $\mathbb{U} + \mathbb{W}$ is a direct sum if and only if $\mathbb{U} \cap \mathbb{W} = \{0_{\mathbb{V}}\}$.

Using this fact, we have,

$$\mathrm{Ker}(T^n) + \mathrm{Im}(T^n) = \mathrm{Ker}(T^n) \oplus \mathrm{Im}(T^n).$$

Also we have,

$$\dim\left(\mathrm{Ker}(T^n) \oplus \mathrm{Im}(T^n)\right) = \dim\left(\mathrm{Ker}(T^n)\right) + \dim\left(\mathrm{Im}(T^n)\right)$$

Since $T^n: \mathbb{V} \to \mathbb{V}$ so applying rank-nullity theorem,

$$\dim(\operatorname{Ker}(T^n)) + \dim(\operatorname{Im}(T^n)) = \dim(\mathbb{V}).$$

Therefore,

$$\dim(\operatorname{Ker}(T^n) \oplus \operatorname{Im}(T^n)) = \dim(\mathbb{V}).$$

If the dimension of the subspace is exactly equal to the dimension of the vector space $\mathbb{V}$. then the subspace is equal to $\mathbb{V}$.

$$\Rightarrow \operatorname{Ker}(T^n) \oplus \operatorname{Im}(T^n) = \mathbb{V}. \quad \text{(proved)}$$

20. Which of the following are ideals and why (or why not)?

(a) {all polynomials in $\mathbb{C}[x]$ having the constant term equal to zero}

(b) {all polynomials in $\mathbb{C}[x]$ containing only even degree terms}

(c) $\{0, 2, 4\} \subseteq \mathbb{Z}_6$

(d) {all polynomials in $\mathbb{Z}[x]$ with even coefficients}

(e) $\mathbb{R} \subseteq \mathbb{R}[x]$

An ideal I of ring R is defined as: $(I \subseteq R)$

(i) $x, y \in I \longrightarrow x + y \in I$

(ii) $x \in I,\ y \in R \longrightarrow x \cdot y \in I$

We will check this 2 condition to see if it is ideal

(a) $I = \left\{ f(x) \in \mathbb{C}[x] \;\middle|\; f(x) = a_1x + a_2x^2 + \cdots + a_nx^n, \; \forall a_i \in \mathbb{C} \right\}$.

We have to cheek if I is ideal of $\mathbb{C}[x]$.

Consider $f(x), g(x) \in I$.

$$\left.\begin{aligned} f(x) &= a_1x + a_2x^2 + \cdots + a_nx^n \\ g(x) &= b_1x + b_2x^2 + \cdots + b_mx^m. \end{aligned}\right\} \forall a_i, b_i \in \mathbb{C}$$

$$f(x) + g(x) = (a_1+b_1)x + (a_2+b_2)x^2 + \cdots + (a_K+b_K)x^K$$

where $K = \max(m,n)$.

$$= c_1x + c_2x^2 + \cdots + c_Kx^K. \quad \forall c_i \in \mathbb{C}$$

$$\in \mathbb{C}[x].$$

$\Rightarrow f(x) + g(x) \in I$

Suppose, $f(x) \in I$ and $g(x) \in \mathbb{C}[x]$

$$\left.\begin{aligned} f(x) &= a_1x + a_2x^2 + \cdots + a_nx^n \\ g(x) &= b_0 + b_1x + b_2x^2 + \cdots + b_mx^m \end{aligned}\right\} \forall a_i, b_i \in \mathbb{C}.$$

$$f(x).g(x) = \sum_{i=1}^{n}\sum_{j=0}^{m} a_i b_j\, x^{i+j}$$

$$= \sum_{K=1}^{m+n} C_K \cdot x^k \qquad \forall C_K \in \mathbb{C}.$$

$$\in \mathbb{C}[x].$$

Moreover we can see that, $f(x).g(x) \in I$

That means I is an ideal of $\mathbb{C}[x]$.

(b) $I = \{ f(x) \in \mathbb{C}[x] \mid f(x) = a_0 + a_2x^2 + a_4x^4 + \cdots + a_{2n}x^{2n},$

$\forall a_{2i} \in \mathbb{C} \}.$

consider, $f(x), g(x) \in I$.

$$\left.\begin{aligned} f(x) &= a_0 + a_2x^2 + \cdots + a_{2n}x^{2n}. \\ g(x) &= b_0 + b_2x^2 + \cdots + b_{2m}x^{2m} \end{aligned}\right\} \forall a_{2i}, b_{2i} \in \mathbb{C}$$

$$f(x) + g(x) = (a_0+b_0) + (a_2+b_2)x^2 + \cdots + (a_{2K}+b_{2K})\, x^{2K}$$

$K = max(m, n)$.

$$= c_0 + c_2x^2 + \cdots + c_{2K}x^{2K} \qquad \forall c_{2i} \in \mathbb{C}$$

$$\in \mathbb{C}[x]$$

So $f(x) + g(x) \in I$.

Consider $f(x) \in I$ and $g(x) \in \mathbb{C}[x]$.

$$\left.\begin{aligned} f(x) &= a_0 + a_2 x^2 + \cdots + a_{2n} x^{2n} \\ g(x) &= b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m . \end{aligned}\right\} \begin{aligned} &\forall a_{2i} \in \mathbb{C} \\ &\forall b_i \in \mathbb{C}. \end{aligned}$$

$$f(x) \cdot g(x) = a_0 b_0 + a_0 b_1 x + \cdots$$
$$\in \mathbb{C}[x]$$

But note that $f(x) \cdot g(x)$ have x^1 term so,

$f(x) \cdot g(x) \notin I$.

Hence I is not ideal of $\mathbb{C}[x]$.

(c) $I = \{ z \in \mathbb{Z}_6 \mid z = 0 \text{ or } 2 \text{ or } 4 \}$

consider $z_1, z_2 \in I$

$$\begin{aligned} z_1 + z_2 &= 0 + 2 \pmod 6 = 2 \\ &= 0 + 4 \pmod 6 = 4 \\ &= 2 + 4 \pmod 6 = 0 \end{aligned}$$

Hence $z_1 + z_2 \in I$.

consider $z_1 \in I$ and $z_2 \in \mathbb{Z}_6$

$z_1 \in \{0,2,4\}$, $z_2 \in \{0,1,2,3,4,5\}$.

Since z_1 is even so,

$2 \cdot z_1$ is also even

$4 \cdot z_1$ is also even.

So when mod 6 will be applied, we will get either 0, 2, 4 as remainder So $z_1 \cdot z_2 \in I$.

Hence I is an ideal of $\mathbb{Z}_6$.

(d) $I = \{ f(x) \in \mathbb{Z}[x] \mid f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \ \forall a_i = 2 \cdot z_i, z_i \in \mathbb{Z} \}$

consider $f(x), g(x) \in I$.

$$\left.\begin{aligned} f(x) &= 2 \cdot z_0 + 2z_1 x + 2z_2 x^2 + \cdots + 2z_n x^n \\ g(x) &= 2w_0 + 2w_1 x + 2w_2 x^2 + \cdots + 2w_n x^n \end{aligned}\right\} \forall z_i, w_i \in \mathbb{Z}.$$

$f(x) + g(x) = 2(z_0 + w_0) + 2(z_1 + w_1)x + \cdots + 2(z_k + w_k)x^k$

$K = \max(m, n)$

$= 2c_0 + 2c_1 x + \cdots + 2c_k x^k$. $\forall c_k \in \mathbb{Z}$.

$\in \mathbb{Z}[x]$

clearly, $f(x) + g(x) \in I$.

Consider $f(x) \in I$ and $g(x) \in \mathbb{Z}[x]$.

$$\left.\begin{aligned} f(x) &= 2z_0 + 2z_1 x + 2z_2 x^2 + \dots + 2z_n x^n \\ g(x) &= a_0 + a_1 x + a_2 x^2 + \dots + a_m x^m \end{aligned}\right\} \begin{aligned} &\forall z_i \in \mathbb{Z} \\ &\forall a_i \in \mathbb{Z}. \end{aligned}$$

$$\begin{aligned} f(x) \cdot g(x) = \; & 2z_0 a_0 + 2z_0 a_1 x + \dots \dots + 2z_0 a_m x^n + \\ & 2z_1 a_0 x + 2z_1 a_1 x^2 + \dots \dots + 2z_1 a_m x^{n+1} + \\ & \dots + 2z_n a_0 x^n + \dots + 2z_n a_m x^{n+m} \end{aligned}$$

Note that $\forall i, j$, $2a_i z_j$ is an even number so.

$$f(x) \cdot g(x) = 2 \cdot P_0 + 2P_1 x^1 + 2P_2 x^2 + \dots + 2P_{n+m} \cdot x^{n+m}.$$

where $\forall P_i$ is even.

$\in \mathbb{Z}[x]$.

and clearly $f(x) \cdot g(x) \in I$.

So I is an ideal of $\mathbb{Z}[x]$.

(e) $I = \mathbb{R}$

$$= \{ f(x) \in \mathbb{R}[x] \mid f(x) = a, \; a \in \mathbb{R} \}.$$

Consider, $f(x), g(x) \in I$.

$$\left.\begin{array}{l} f(x) = a \\ g(x) = b. \end{array}\right\} \quad a, b \in \mathbb{R}.$$

$f(x) + g(x) = a + b = c \quad , c \in \mathbb{R}$.

$\Rightarrow f(x) + g(x) \in \mathbb{R}[x]$.

Clearly $f(x) + g(x) \in I$.

consider $f(x) \in I, \quad g(x) \in \mathbb{R}[x]$.

$$\left.\begin{array}{l} f(x) = a \\ g(x) = b_0 + b_1 x + b_2 x^2 + \dots + b_n x^n \end{array}\right\} \begin{array}{l} a \in \mathbb{R} \\ \forall b_i \in \mathbb{R}. \end{array}$$

$f(x) \cdot g(x) = a \cdot b_0 + a b_1 x + a b_2 x^2 + \dots + a b_n x^n. \in \mathbb{R}[x]$

clearly $f(x) \cdot g(x) \notin I$

Hence I is not an ideal of $\mathbb{R}[x]$.

24. Is $\mathbb{Z}[x]$ a PID? Justify your answer.

An integral domain R is a PID if all the ideals of R is a principal ideal.

$$\mathbb{Z}[x] := \left\{ f(x) \;\middle|\; f(x) = a_0 + a_1 x + \cdots + a_n x^n, \ \forall a_i \in \mathbb{Z} \right\}.$$

$\mathbb{Z}[x]$ is an integral domain.
But $\mathbb{Z}[x]$ is not an integral domain.

Consider an ideal,

$$I = \left\{ f_1(x) \cdot g_1(x) + f_2(x) \cdot g_2(x) \;\middle|\; g_1(x) = 2,\ g_2(x) = x,\ f_1(x), f_2(x) \in \mathbb{Z}[x] \right\}.$$

$$= \langle \{ g_1(x), g_2(x) \} \rangle$$

$$= \langle \{ 2, x \} \rangle$$

We can show that I is not a principal ideal.

<u>proof</u>:

$$\langle \{2, x\} \rangle = \left\{ 2 \cdot p(x) + x \cdot q(x) \;\middle|\; p, q \in \mathbb{Z}[x] \right\}$$

for the sake of contradiction assume that,

$$2 \cdot p(x) + x \cdot q(x) = a(x) \cdot s(x) \quad \text{where } a(x) \text{ is fixed and } s(x) \in \mathbb{Z}[x]$$

such that "I" can be generated by single element $a(x)$.

In order to generate 2 we must have,

$a(x).h(x) = 2 \quad \exists\, h(x) \in \mathbb{Z}[x]$

In order to generate x we must have,

$a(x).g(x) = x \quad \exists\, g(x) \in \mathbb{Z}[x]$.

$$\Rightarrow \frac{h(x)}{g(x)} = \frac{2}{x}$$

$$\Rightarrow 2.g(x) = x.h(x)$$

whatever is the coefficients of $g(x)$ since it is multiplied by 2 so the coefficients of $h(x)$ is always even as the coefficients belong to $\mathbb{Z}$.

$\Rightarrow h(x) = 2.r(x) \quad \exists\, r(x) \in \mathbb{Z}[x]$

Now, $a(x).g(x) = x$

$$\Rightarrow a(x).\frac{x.h(x)}{2} = x$$

$$\Rightarrow a(x).\frac{h(x)}{2} = 1$$

$$\Rightarrow a(x).\frac{2.r(x)}{2} = 1$$

$\Rightarrow a(x) \cdot r(x) = 1.$

$\Rightarrow 1 \in \langle \{a(x)\} \rangle$

We have assumed that

$$2 \cdot p(x) + x \cdot q(x) = a(x) \cdot s(x)$$

Since $1 \in \langle \{a(x)\} \rangle \longrightarrow 1 \in \langle \{2, x\} \rangle$

Therefore,

$$2 \cdot p(x) + x \cdot q(x) = 1 \quad \text{for some } p(x), q(x) \in \mathbb{Z}.$$

Let, $\left.\begin{aligned} p(x) &= a_0 + a_1 x + a_2 x^2 + \cdots \\ q(x) &= b_0 + b_1 x + b_2 x^2 + \cdots \end{aligned}\right\} a_i, b_i \in \mathbb{Z}.$

$$2 \cdot p(x) + x \cdot q(x) = 1$$

$\Rightarrow (2a_0 + 2a_1 x + \cdots) + (b_0 x + b_1 x^2 + b_2 x^3 + \cdots) = 1$

$\Rightarrow 2a_0 + (2a_1 + b_0)x + \cdots = 1$

So $2a_0 = 1 \qquad a_0 \in \mathbb{Z}.$

$\Rightarrow a_0 = \frac{1}{2}$ which is a contradiction of the fact that a_0 has to be $\mathbb{Z}$. (integer).

Therefore, we have assumed that $I = \langle \{2, x\} \rangle$ can be generated by a single element $a(x)$ is a false assumption.

Hence $I = \langle \{2, x\} \rangle$ can't be a principal ideal.
Therefore $\mathbb{Z}[x]$ can't be a PID. (Hence proved)

22. For a non-zero vector $v \in \mathbb{C}^n$, and a matrix $A \in \mathbb{C}^{n\times n}$, the sequence of vectors $\{v, Av, A^2v, \ldots, A^jv\}$ is called a Krylov sequence (named after the Russian mathematician Alexei Nikolaevich Krylov), and the subspace $\mathcal{K}_j$ spanned by vectors in the sequence is called a Krylov subspace. Suppose a Krylov subspace for vector v, say $\mathcal{K}_{n-1}$, spans $\mathbb{C}^n$, while the characteristic polynomial for A is given by $\chi_A(x) = \sum_{i=0}^{n} \alpha_i x^i$. Represent the matrix A in terms of the ordered basis given by the corresponding Krylov sequence. What is the minimal polynomial, $\mu_A(x)$, and why?

Given that $\langle \{v, Av, A^2v, \ldots, A^{n-1}v\} \rangle = \mathbb{C}^n$

$\Rightarrow B = \{v, Av, A^2v, \ldots, A^{n-1}v\}$ is a basis for $\mathbb{C}^n$.

Now,

$$[A]_B = \left[\, [Av]_B \; [AAv]_B \; \cdots \; [AA^{n-1}v]_B \,\right]$$

Since the basis is ordered basis,
$\{v_1, v_2, \ldots, v_n\}$ and $v_k = A^{k-1}v$ therefore,

$$Av_k = v_{k+1} = A^k v$$

$$\Rightarrow [Av_k]_B = [0 \; 0 \; \cdots 0 \; 1 \; 0 \cdots \; 0]^T$$

$k+1^{th}$ position.

Caley Hamilton theorem says, $\chi_A(A) = 0$

$$\Rightarrow A^n = -\sum_{K=0}^{n-1} \alpha_K A^K$$

$$\Rightarrow Av_n = A^n v = -\sum_{K=0}^{n-1} \alpha_K A^K v = -\sum_{K=0}^{n-1} \alpha_K v_{K+1}$$

$$\Rightarrow [Av_n]_B = [-\alpha_0 \ -\alpha_1 \ \cdots \ -\alpha_{n-1}]^T$$

$$[A]_B = \begin{bmatrix} 0 & 0 & & 0 & -\alpha_0 \\ 1 & 0 & & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & & 1 & -\alpha_{n-1} \end{bmatrix}$$

We know that $\{v, Av, A^2v, \dots, A^{n-1}v\}$ is linearly independent set.

$$\Rightarrow \sum_{K=0}^{n-1} \alpha_K A^K v = 0 \qquad \forall \alpha_K = 0$$

Minimal polynomial $= \mu_A(x)$. The deg of $\mu_A(x) = n$ because $\mu_{Av}(x)$ must be having degree at least n.

$$\mu_A(x) = \chi_A(x) \qquad [\because \mu_A(x) \mid \chi_A(x)].$$

34. For a symmetric matrix, $A \in \mathbb{R}^{n\times n}$ with zeros on its diagonal, suppose $a_{ij} = 0$ or 1 for $i \neq j$, show that the largest eigenvalue, λ_M is bounded by $\min_i \text{Rowsum}_i(A) \leq \lambda_M \leq \max_i \text{Rowsum}_i(A)$, where Rowsum_i denotes the sum of the entries of the i^{th} row.

Provided $a_{ij} \in \{0,1\}$.

Suppose the largest eigenvalue is λ_M and consider the eigenvector corresponding to λ_M is v_M.

$$\Rightarrow A v_M = \lambda_M v_M$$

$$v_M := [v_{M_1} \; v_{M_2} \; \ldots \; v_{M_n}]^T$$

Consider a vector $u = [u_1 \; u_2 \; \ldots \; u_n]^T$ where $u_K = |v_{M_K}| \; \forall K \in \{1, \ldots n\}$.

We have, $u^T u = v^T v$

$$\Rightarrow v_M^T A v = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} v_{M_i} v_{M_j}$$

$$\leq \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} |v_{M_i}| |v_{M_j}| = u^T u$$

We have $\dfrac{u^T A u}{u^T u} = \lambda_M$ atleast.

maximum possible $\dfrac{u^T A u}{u^T u}$ is λ_M

Suppose, $V_{M_i} = \max \{ v_{M_1}, v_{M_2}, \dots, v_{M_n} \}$.

$$\lambda_M = \frac{\sum_{j=1}^{n} a_{ij} v_{M_j}}{v_{M_j}} \leq \frac{\sum_{j=1}^{n} a_{ij} v_{M_i}}{v_{M_i}}$$

$$= \max_i \text{ rowsum}_i (A)$$

$V_{M_K} = \min \{ v_{M_1}, v_{M_2}, \dots, v_{M_n} \}$

$$\lambda_M = \frac{\sum_{j=1}^{n} a_{Kj} v_{M_j}}{v_{M_K}} \geqslant \frac{\sum_{j=1}^{n} a_{Kj} v_{M_K}}{v_{M_K}}$$

$$= \min_i \text{ rowsum}_i (A).$$

Therefore, $\min_i \text{ rowsum}_i (A) \leq \lambda_M \leq \max_i \text{ rowsum}_i (A)$