Eigenvalue Problems

Let $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$, $\underset{\sim}{x} \neq 0 \in \mathbb{C}^m$ ($\mathbb{C}$ is set of complex nos)

then $\underset{\sim}{x}$ is an eigenvector of $\underset{\sim}{A}$ and $\lambda \in \mathbb{C}$ is its corresponding eigenvalue

if

$$\boxed{\underset{\sim}{A}\underset{\sim}{x} = \lambda \underset{\sim}{x}}$$

* The set of all eigenvalues of a matrix $\underset{\sim}{A}$ is called the spectrum of $\underset{\sim}{A}$ denoted by $\Lambda(\underset{\sim}{A})$

Application areas :-

* Insights into evolution of system
  - → Vibration analysis
  - → Study of resonance
  - → Stability of structure
  - → fluid flows subjected to small perturbations

* Quantum mechanical modeling of materials (Solving Schrodinger equation)

* Principal stresses in solid mechanics

* PCA in data driven modeling

* Pagerank algorithm used in search engines is an eigenvalue problem

* Eigenvectors of graph Laplacian matrix actually help in construction of efficient filters Graph Convolution Neural Network!

Eigenvalue decomposition!

An eigenvalue decomposition of

$\underset{\sim}{A} \in \mathbb{R}^{m \times m}$ is a factorization of

$$\underset{\sim}{A} = \underset{\sim}{X} \underset{\sim}{\Lambda} \underset{\sim}{X}^{-1}$$ where

$\underset{\sim}{X}$ is non singular and $\underset{\sim}{\Lambda}$ is diagonal with $\underset{\sim}{X}$ comprising of eigenvectors of $\underset{\sim}{A}$ as columns!

<u>Note</u>:- Such decomposition may not always exist!

$$\underset{\sim}{A} \underset{\sim}{X} = \underset{\sim}{X} \underset{\sim}{\Lambda}$$

$$[A] \begin{bmatrix} | & | & | & & | \\ \underset{\sim}{x_1} & \underset{\sim}{x_2} & x_3 & \cdots & x_m \\ | & | & | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | & \\ x_1 & x_2 & x_3 & \cdots \\ | & | & | & \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$$

$$\underset{\sim}{A} \underset{\sim}{x_j} = \lambda_j x_j$$

i.e j^{th} column of $\underset{\sim}{X}$ is j^{th} eigen vector and (j,j) entry of Λ is corresponding eigenvalue!

→ <u>Geometric multiplicity</u> :- The geometric multiplicity of an eigenvalue λ is the number of linearly independent eigenvectors associated with that eigen value λ. If $\lambda \in \Lambda(A)$, eigenspace E_λ

E_λ is an invariant subspace of $\underset{\sim}{A}$

i.e $\underset{\sim}{A} E_\lambda \subseteq E_\lambda$

$\underset{\sim}{v_1}, \underset{\sim}{v_2}$ λ

$$\underset{\sim}{x} = \alpha_1 \underset{\sim}{v_1} + \alpha_2 \underset{\sim}{v_2}$$

$$\begin{aligned} \underset{\sim}{A}\underset{\sim}{x} &= \underset{\sim}{A}(\alpha_1 \underset{\sim}{v_1} + \alpha_2 \underset{\sim}{v_2}) \\ &= \alpha_1 \lambda \underset{\sim}{v_1} + \alpha_2 \lambda \underset{\sim}{v_2} \\ &= \lambda [\alpha_1 \underset{\sim}{v_1} + \alpha_2 \underset{\sim}{v_2}] \\ &= \lambda \underset{\sim}{x} \end{aligned}$$

The dimension of E_λ is the geometric multiplicity of λ i.e maximum number of linearly independent eigenvectors that can be found for a given λ

* Characteristic Polynomial :-

The characteristic polynomial p_A of $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$ is the m^{th} degree monic polynomial $\boxed{p_A(z) = \det(z\underset{\sim}{I} - \underset{\sim}{A})}$ ✓

(coefficient of z^m is 1 → monic polynomial)

Thm :- λ is eigenvalue of $\underset{\sim}{A}$ if and only if $p_A(\lambda) = 0$

$$\underset{\sim}{A}\,\underset{\sim}{x} = \lambda \underset{\sim}{x}$$

$$(\underset{\sim}{A} - \lambda \underset{\sim}{I})\underset{\sim}{x} = 0$$

Note :- $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$

λ can be complex

any comple λ must appear in complex conjugate pairs

Eigenvector of $\underset{\sim}{A}$ lies in the null space of $(\underset{\sim}{A} - \lambda \underset{\sim}{I})$

i.e $\lambda = a + ib$ is an eigenvalue

$\lambda^* = a - ib$ is an eigenvalue

* Algebraic multiplicity :-

Since $p_A(z)$ is monic m-degree polynomial, it can be written as

$$p_A(z) = (z-\lambda_1)(z-\lambda_2) \cdots (z-\lambda_m)$$

for some $\lambda_j \in \mathbb{C}$ (roots of $p_A(z)$)

Each λ_j is an eigenvalue and in general may be repeated

"The multiplicity of λ as a root of $p_A(z)$ is the algebraic multiplicity of an eigenvalue λ"

Remark:- ⓐ If $A \in \mathbb{R}^{m \times m}$ then A has m eigenvalues counting algebraic multiplicity. In particular if roots of $p_A(z)$ are simple, then A has m distinct eigenvalues.

(b) The algebraic multiplicity of an eigenvalue λ is always atleast as large as its geometric multiplicity.

* Similarity transformation :-

If $\underset{\sim}{X} \in \mathbb{R}^{m \times m}$ is non singular, then

$$\underset{\sim}{A} \rightarrow \underset{\sim}{X}^{-1} \underset{\sim}{A} \underset{\sim}{X}$$ is called a similarity transformation of $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$

We say that two matrices $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are similar if there is a similarity transformation of one to another

i.e if there is a nonsingular $\underset{\sim}{X} \in \mathbb{R}^{m \times m}$ such that $\underset{\sim}{B} = \underset{\sim}{X}^{-1} \underset{\sim}{A} \underset{\sim}{X}$

then $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are said to similar!

Thm:- If $\underset{\sim}{X}$ is nonsingular, then $\underset{\sim}{A}$ and $\underset{\sim}{X}^{-1}\underset{\sim}{A}\underset{\sim}{X}$ have the same characteristic polynomial, eigenvalues and algebraic multiplicity and geometric multiplicity!

Pf:- $p(z) = \det(z\underset{\sim}{I} - \underset{\sim}{X}^{-1}\underset{\sim}{A}\underset{\sim}{X})$

$\underset{\sim}{X}^{-1}\underset{\sim}{A}\underset{\sim}{X} \quad = \det(z\underset{\sim}{X}^{-1}\underset{\sim}{X} - \underset{\sim}{X}^{-1}\underset{\sim}{A}\underset{\sim}{X})$

$= \det(\underset{\sim}{X}^{-1}(z\underset{\sim}{I} - \underset{\sim}{A})\underset{\sim}{X})$

$= \left[\det \underset{\sim}{X}^{-1}\right]\left[\det(z\underset{\sim}{I} - \underset{\sim}{A})\right](\det \underset{\sim}{X})$

$= (\det \underset{\sim}{X})^{-1}(\det(z\underset{\sim}{I} - \underset{\sim}{A}))(\det \underset{\sim}{X})$

$= \det(z\underset{\sim}{I} - \underset{\sim}{A}) = p_A(z)$

Hence $\underset{\sim}{A}$, $\underset{\sim}{X}^{-1}\underset{\sim}{A}\underset{\sim}{X}$ have the same eigenvalues (same roots of $p_A(z)$) and hence same algebraic multiplicity

Build a matrix E_λ whose column vectors span the eigenspace for the matrix $\underset{\sim}{A}$, corresponding to the eigen value λ

$$\left(\underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{X}\right)\underbrace{\left(\underset{\sim}{X}^{-1}E_\lambda\right)} = \underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{E}_\lambda$$

$$= \underbrace{\underset{\sim}{X}^{-1}E_\lambda}\,\Lambda$$

$\underset{\sim}{X}^{-1}E_\lambda$ is the eigenspace for $\underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{X}$ corresponding to the eigenvalue λ.

Hence geometric multiplicity of $\underset{\sim}{A}$ and $\underset{\sim}{X}^{-1}\underset{\sim}{A}\,\underset{\sim}{X}$ are the same.

($\because$ E_λ and $\underset{\sim}{X}^{-1}E_\lambda$ has the same rank)

<u>Defective eigenvalues and matrices</u>

* A generic matrix need not have distinct eigenvalues i.e algebraic multiplicity need not be 1 and geometric multiplicity need not be 1 as well and need not be equal to algebraic multiplicity as well)

$$\underset{\sim}{A} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} ; \quad B = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

Both $\underset{\sim}{A}$ and $\underset{\sim}{B}$ eigenvalue

$\lambda = 2$.

Algebraic multiplicity of $\lambda = 2$ for A?

$\longrightarrow$ 3

Algebraic multiplicity of $\lambda = 2$ for B?

$\longrightarrow$ 3

For $\underset{\sim}{A}$, we can choose 3 linearly independent eigenvectors $\underset{\sim}{e_1}, \underset{\sim}{e_2}, \underset{\sim}{e_3}$ and geometric multiplicity is also 3.

For $\underset{\sim}{B}$, we can only have only linearly independent eigenvector $\underset{\sim}{e_1}$, the geometric multiplicity of $\underset{\sim}{B}$ is 1

* An eigenvalue whose algebraic multiplicity is greater than its geometric multiplicity is called a defective eigenvalue!

* A matrix that has atleast one defective eigenvalue is a defective matrix i.e it does not posses a full set of m linearly independent eigenvectors

* A diagonal matrix is non-defective

(algebraic multiplicity of eigenvalue λ = geometric multiplicity of λ)

Diagonalizability

Thm: $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$ is non-defective if and only if it has eigenvalue decomposition $\underset{\sim}{A} = \underset{\sim}{X} \underset{\sim}{\Lambda} \underset{\sim}{X}^{-1}$ where $\underset{\sim}{X} \in \mathbb{R}^{m \times m}$ non singular matrix.

$\boxed{\rightarrow \text{A non-defective matrix is diagonalizable}}$

Proof:- To show given $\underset{\sim}{A} = \underset{\sim}{X} \underset{\sim}{\Lambda} X^{-1}$, $\underset{\sim}{A}$ is non-defective!

$\Downarrow$

$\underset{\sim}{A} = \underset{\sim}{X} \underset{\sim}{\Lambda} \underset{\sim}{X}^{-1} \rightarrow \underset{\sim}{A}$ is similar to $\underset{\sim}{\Lambda}$

$\underset{\sim}{A}$ and $\underset{\sim}{\Lambda}$ has the same eigenvalues and same multiplicities.

Since $\underset{\sim}{\Lambda}$ is a diagonal matrix, it is non-defective and thus $\underset{\sim}{A}$ is non-defective.

→ To show if $\underset{\sim}{A}$ is non-defective

then $\underset{\sim}{A} = \underset{\sim}{X} \underset{\sim}{\Lambda} \underset{\sim}{X}^{-1}$ ↑

* Since $\underset{\sim}{A}$ is non-defective, it must have m linearly independent eigenvectors!

If $\underset{\sim}{X} = \left[\underset{\sim}{x}_1 \middle| \underset{\sim}{x}_2 \middle| \cdots \middle| \underset{\sim}{x}_m \right]$

comprises of these m linearly independent vectors, X is full rank

$$\underset{\sim}{A} \underset{\sim}{X} = \underset{\sim}{A} \left(\underset{\sim}{x}_1 \middle| \underset{\sim}{x}_2 \middle| \cdots \middle| \underset{\sim}{x}_m \right)$$

$$= \left[\underset{\sim}{x}_1 \middle| \underset{\sim}{x}_2 \middle| \cdots \middle| x_m \right] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}$$

$$= \underset{\sim}{X} \underset{\sim}{\Lambda}$$

$$\underset{\sim}{A}\,\underset{\sim}{X} = \underset{\sim}{X}\,\underset{\sim}{\Lambda}$$

$$\boxed{\underset{\sim}{A} = \underset{\sim}{X}\,\underset{\sim}{\Lambda}\,\underset{\sim}{X}^{-1}}$$

<u>Thm</u>:- Trace of $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$

$$tr(\underset{\sim}{A}) = \sum_{j=1}^{m} a_{jj} = \sum_{j=1}^{m} \lambda_j$$

Determinant of $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$

$$det(\underset{\sim}{A}) = \prod_{j=1}^{m} \lambda_j$$

<u>Unitary</u> / <u>orthogonal</u> <u>diagonalization</u>:-

* $\underset{\sim}{Q} \in \mathbb{C}^{m \times m}$ is a unitary matrix if $\underset{\sim}{Q}^{+}\underset{\sim}{Q} = \underset{\sim}{Q}\,\underset{\sim}{Q}^{+} = \underset{\sim}{I}$

when $\underset{\sim}{Q}^{+}$ is conjugate transpose of $\underset{\sim}{Q}$

* Unitary matrix reduces to orthogonal matrices if Q is real i.e $Q^T Q = Q Q^T = I$ is satisfied!

For a non-defective matrix $A \in \mathbb{R}^{m \times m}$, it is possible not only to have m linearly independent eigenvectors, but these vectors can be orthogonal as well!

[If $A \in \mathbb{R}^{m \times m}$ eigenvalues can be complex and corresponding eigenvectors have to be complex as well]

Now in general for a non-defective $A \in \mathbb{R}^{m \times m}$, we define unitary diagonalizability i.e there exists a unitary matrix $\underset{\sim}{Q}$, such that

$$\underset{\sim}{A} = \underset{\sim}{Q}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Q}^{+} \quad \text{where}$$

$\underset{\sim}{Q}^{+}$ is conjugate transpose of $\underset{\sim}{Q}$ and columns of $\underset{\sim}{Q}$ are eigenvectors of $\underset{\sim}{A}$

Unitary diagonalizability reduces to orthogonal diagonalizability if $\underset{\sim}{Q}$ is real matrix and λ are real

i.e. $\underset{\sim}{A} = \underset{\sim}{Q}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Q}^{T}$

Eg:- A symmetric matrix $\underset{\sim}{S} \in \mathbb{R}^{m\times m}$ satisfies $\underset{\sim}{S} = \underset{\sim}{S}^{T}$ & has real eigen values

$$\longrightarrow \quad \underset{\sim}{S}\,\underset{\sim}{x} = \lambda\,\underset{\sim}{x}$$

Take complex conjugate both sids $\Rightarrow$ $\underbrace{\underset{\sim}{S}\,\underset{\sim}{x}^{*} = \lambda^{*}\,\underset{\sim}{x}^{*}}_{\text{where } \lambda^{*} \text{ is complex conjugate of } \lambda}$

$$\underset{\sim}{S}\underset{\sim}{x} = \lambda \underset{\sim}{x} \Rightarrow \underset{\sim}{x}^{*T}\underset{\sim}{S}\underset{\sim}{x} = \lambda \underset{\sim}{x}^{*T}\underset{\sim}{x}$$

$$\Rightarrow \underbrace{(\underset{\sim}{S}\underset{\sim}{x})^T\underset{\sim}{x}^*} = \lambda \underset{\sim}{x}^{*T}\underset{\sim}{x} \quad — ①$$

$$\underset{\sim}{S}\underset{\sim}{x}^* = \lambda^*\underset{\sim}{x}^* \Rightarrow \underset{\sim}{x}^T\underset{\sim}{S}\underset{\sim}{x}^* = \lambda^*\underset{\sim}{x}^T\underset{\sim}{x}^*$$

$$\Rightarrow \underset{\sim}{x}^T\underset{\sim}{S}^T\underset{\sim}{x}^* = \lambda^*\underset{\sim}{x}^T\underset{\sim}{x}^*$$

$$\Rightarrow \underbrace{(\underset{\sim}{S}\underset{\sim}{x})^T\underset{\sim}{x}^*} = \lambda^*\underset{\sim}{x}^T\underset{\sim}{x}^* \quad — ②$$

$$① - ② \Rightarrow 0 = \lambda\underset{\sim}{x}^{*T}\underset{\sim}{x} - \lambda^*\underset{\sim}{x}^T\underset{\sim}{x}^*$$

$$= (\lambda - \lambda^*)(x^Tx^*)$$

Since $\underset{\sim}{x}^T\underset{\sim}{x}^* \neq 0 \Rightarrow \lambda - \lambda^* = 0$

$$\Rightarrow \boxed{\lambda = \lambda^*}$$

Hence eigenvalues of a symmetric matrix real.

Consider the case where $\underset{\sim}{S}$ is symmetric $(\lambda_1, \underset{\sim}{x}_1)$ and $(\lambda_2, \underset{\sim}{x}_2)$ are two eigenpairs of $\underset{\sim}{S}$ where $\lambda_1 \neq \lambda_2$

$$\underset{\sim}{S}\underset{\sim}{x}_1 = \lambda_1\underset{\sim}{x}_1, \quad \underset{\sim}{S}\underset{\sim}{x}_2 = \lambda_2\underset{\sim}{x}_2$$

$$\Rightarrow \underline{\underset{\sim}{x}_2^T\underset{\sim}{S}\underset{\sim}{x}_1 = \lambda_1\underset{\sim}{x}_2^T\underset{\sim}{x}_1}, \quad \underline{\underset{\sim}{x}_1^T\underset{\sim}{S}\underset{\sim}{x}_2 = \lambda_2\underset{\sim}{x}_1^T\underset{\sim}{x}_2}$$

$$\underline{x}_2^T \underline{S}\, \underline{x}_1 = \lambda_1 \underline{x}_2^T \underline{x}_1$$
$$\Rightarrow \underline{x}_2^T \underline{S}^T \underline{x}_1 = \lambda_1 \underline{x}_2^T \underline{x}_1$$
$$\Rightarrow (\underline{S}\,\underline{x}_2)^T \underline{x}_1 = \lambda_1 \underline{x}_2^T \underline{x}_1 \quad - (1)$$

$$\underline{x}_1^T \underline{S}\, \underline{x}_2 = \lambda_2 \underline{x}_1^T \underline{x}_2$$
$$(\underline{S}\,\underline{x}_2)^T \underline{x}_1 = \lambda_2 \underline{x}_1^T \underline{x}_2 \quad - (2)$$

(1) - (2)

$$0 = (\lambda_1 - \lambda_2)\, \underline{x}_1^T \underline{x}_2$$

$$\lambda_1 - \lambda_2 \neq 0 \Rightarrow \underline{x}_1^T \underline{x}_2 = 0$$

i.e the eigenvectors $\underline{x}_1, \underline{x}_2$ are orthogonal to each other.

Now what happens if eigenvalues of $\underline{S}$ are degenerate, the above proof will fail. Let us consider one such case where the geometric multiplicity of $S \in \mathbb{R}^{m \times m}$ is $r < m$ corresponding to an eigenvalue λ. We need to show that algebraic multiplicity of S is also r. Since we assumed geometric multiplicity of $S \in \mathbb{R}^{m \times m}$ is r for an eigenvalue λ, we can always find r linearly independent vectors spanning the 'r' dimensional eigen subspace $\mathbb{V}^E$ corresponding to eigenvalue λ. This means I can choose r orthogonal eigenvectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots \underset{\sim}{v}_r$

$\mathbb{V}^E = \{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots \underset{\sim}{v}_r\}$. Let me now construct orthogonal vectors $\mathbb{V}^{\perp} = \{\underset{\sim}{v}_{r+1}, \underset{\sim}{v}_{r+2}, \ldots \underset{\sim}{v}_m\}$ spanning the subspace $\mathbb{V}^{\perp}$ orthogonal to $\mathbb{V}^E$.

Let $\underset{\sim}{V} = \left[\underset{\sim}{v}_1 \; \underset{\sim}{v}_2 \ldots \underset{\sim}{v}_r \; \underset{\sim}{v}_{r+1} \; \underset{\sim}{v}_{r+2} \ldots \underset{\sim}{v}_m \right]$

Now we have

$$\underset{\sim}{S}\,\underset{\sim}{v}_i = \lambda\,\underset{\sim}{v}_i \quad \text{for } i = 1 \text{ to } r$$

Also

$$\underset{\sim}{S}\,\underset{\sim}{v}_i = \sum_{j=1}^{m} c_{ij}\,\underset{\sim}{v}_j$$

$$c_{ij} = \lambda\,\delta_{ij} \quad \text{for } i = 1 \text{ to } r$$

Now we can write $\underset{\sim}{V}$

$$\underset{\sim}{S}\underbrace{\begin{bmatrix} | & | & | & & | & | & | & & | \\ \underset{\sim}{v}_1 & \underset{\sim}{v}_2 & \underset{\sim}{v}_3 & \cdots & \underset{\sim}{v}_r & \underset{\sim}{v}_{r+1} & \underset{\sim}{v}_{r+2} & \cdots & \underset{\sim}{v}_m \\ | & | & | & & | & | & | & & | \end{bmatrix}}_{\underset{\sim}{V}}$$

$$= \begin{bmatrix} | & | & | & & | & | & | & & | \\ \underset{\sim}{v}_1 & \underset{\sim}{v}_2 & \underset{\sim}{v}_3 & \cdots & \underset{\sim}{v}_r & \underset{\sim}{v}_{r+1} & \underset{\sim}{v}_{r+2} & \cdots & \underset{\sim}{v}_m \\ | & | & | & & | & | & | & & | \end{bmatrix} \underbrace{\left[\begin{array}{cccc|c} \lambda & & & & \\ & \lambda & & & \\ & & \lambda & & \\ & & & \ddots\ \lambda & \\ \hline c_{r+1,1} & c_{r+1,2} & \cdots & & c_{r+1,m} \\ c_{r+2,1} & c_{r+2,2} & \cdots & & c_{r+2,m} \\ \vdots & & & & \\ c_{m,1} & c_{m,2} & \cdots & & c_{m,m} \end{array}\right]}_{\underset{\sim}{M}}$$

$$\underset{\sim}{S}\,\underset{\sim}{V} = \underset{\sim}{V}\,\underset{\sim}{M}$$

$$\Rightarrow \underset{\sim}{V}^{T}\,\underset{\sim}{S}\,\underset{\sim}{V} = \underset{\sim}{M}$$

Observe $V^T S V$ is symmetric because S is symmetric. This means M has to be symmetric

i.e. $$V^T S V = \overbrace{\left[\begin{array}{ccc|c} \lambda & & & \\ & \lambda & & 0 \\ & & \lambda \ddots & \\ \hline & 0 & & C \end{array}\right]}^{M} \quad \text{--- } (1)$$

where C is a symmetric matrix of size $(m-r) \times (m-r)$

Let us now write characteristic polynomial of LHS and RHS of (1), this gives $p_{V^TSV}(z) = p_M(z)$

$p_{V^TSV}(z) = p_S(z)$ since $V^T S V$ is similar to S. Hence

$$p_S(z) = p_M(z)$$

$$\Rightarrow \det(S - zI_m) = (\lambda - z)^r \det(C - zI_{m-r}) \quad -(2)$$

I_m is identity matrix of size m

I_{m-r} is identity matrix of size $m-r$.

Now our objective is to show algebraic multiplicity of S is also r. This means we need to show λ is not an eigenvalue of C.

Let us use indirect method of proof now.

Assume λ can be an eigenvalue of $\underset{\sim}{C}$, this means I have

$$\underset{\sim}{C}\,\underset{\sim}{u} = \lambda \underset{\sim}{u} \quad \text{where } \underset{\sim}{u} \in \mathbb{R}^{m-r} \text{ is}$$

eigenvector of $\underset{\sim}{C}$. Construct

$\hat{u} \in \mathbb{R}^m$ where $\hat{u} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \underset{\sim}{u} \end{bmatrix} \begin{array}{l} \to 1^{st} \text{ entry} \\ \\ \\ \\ \to r^{th} \text{ entry} \\ \\ \end{array}$

then we have

$$\underset{\sim}{V}^T \underset{\sim}{S} \underset{\sim}{V} \hat{u} = \underset{\sim}{M} \hat{u} = \left[\begin{array}{ccc|c} \lambda & & & \\ & \lambda \ddots & & \underset{\sim}{0} \\ & & \lambda & \\ \hline & \underset{\sim}{0} & & \underset{\sim}{C} \end{array}\right] \begin{bmatrix} \underset{\sim}{0} \\ \\ \underset{\sim}{u} \end{bmatrix}$$

$$= \begin{bmatrix} \underset{\sim}{0} \\ \\ \lambda \underset{\sim}{u} \end{bmatrix} = \lambda \begin{bmatrix} \underset{\sim}{0} \\ \\ \underset{\sim}{u} \end{bmatrix}$$

$$= \lambda \hat{u}$$

This means $\hat{u}$ is a eigenvector

of $\underset{\sim}{V}^T \underset{\sim}{S} \underset{\sim}{V}$ corresponding to eigenvalue λ.

i.e $\underset{\sim}{V}^T \underset{\sim}{S} \underset{\sim}{V} \hat{u} = \lambda \hat{u}$

$\Rightarrow \underset{\sim}{S} \underset{\sim}{V} \hat{\underset{\sim}{u}} = \lambda \underset{\sim}{V} \hat{\underset{\sim}{u}}$

$\Rightarrow \underset{\sim}{V} \hat{\underset{\sim}{u}}$ is an eigenvector of $\underset{\sim}{S}$ corresponding to eigenvalue λ.

Clearly by construction $\underset{\sim}{V} \hat{u}$ is a vector which lies in the space spanned by the $\{\underset{\sim}{v}_{r+1}, \underset{\sim}{v}_{r+2}, \dots \underset{\sim}{v}_m\}$ i.e $\mathbb{W}^{\perp}$ which is orthogonal to $\mathbb{W}^E$.

This means we are able to find another linearly independent eigenvector $\underset{\sim}{V} \hat{\underset{\sim}{u}}$ in addition to $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots \underset{\sim}{v}_r\}$.

This means that geometric multiplicity of $\underset{\sim}{S}$ is $r+1$. But this a contradiction to the assumption that geometric multiplicity of $\underset{\sim}{S}$ is r corresponding to the eigenvalue λ.

Hence λ is not an eigenvalue of $\underset{\sim}{C}$. This means algebraic multiplicity of $\underset{\sim}{S}$ is r which is same as the geometric multiplicity of $\underset{\sim}{S}$.

This means I can find 'r' orthogonal vectors which span this 'r' dimensional invariant subspace which are the eigenvectors for my $\underset{\sim}{S}$.

Thm :- A symmetric matrix $\underset{\sim}{S} \in \mathbb{R}^{m \times m}$ is always non-defective and is orthogonally diagonalizable with real eigenvalues i.e

$$\underset{\sim}{S} = \underset{\sim}{Q} \underset{\sim}{\Lambda} \underset{\sim}{Q}^T$$

where $\underset{\sim}{Q}$ is orthogonal matrix

$\underset{\sim}{\Lambda}$ is a diagonal matrix

Thm :- A skew symmetric matrix $\underset{\sim}{W} \in \mathbb{R}^{m \times m}$ which satisfies $\underset{\sim}{W} = -\underset{\sim}{W}^T$. This skew symmetric is non-defective and has purely imaginary eigenvalues and is unitarily diagonalizable

i.e $\underset{\sim}{W} = \underset{\sim}{Q} \underset{\sim}{\Lambda} \underset{\sim}{Q}^{+}$

It turns out that there is a class of unitarily diagonalizable matrices and these set of matrices satisfy the property $\underset{\sim}{A}^T \underset{\sim}{A} = \underset{\sim}{A} \underset{\sim}{A}^T$

* A matrix is normal if

$$\underset{\sim}{A}^T \underset{\sim}{A} = \underset{\sim}{A} \underset{\sim}{A}^T$$

<u>Thm</u>:- A matrix is unitarily diagonazable if and only if $\underset{\sim}{A}^T \underset{\sim}{A} = \underset{\sim}{A} \underset{\sim}{A}^T$ for $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$

Pf:- <u>Direction 1</u>

If a matrix is unitary diagonalizable then we need to show $\underset{\sim}{A}^T \underset{\sim}{A} = \underset{\sim}{A} \underset{\sim}{A}^T$

Since the given matrix is unitarily diagonalizable we know

$$A = Q \Lambda Q^\dagger$$

$$A^T A - A A^T = A^\dagger A - A A^\dagger$$

$$= (Q\Lambda Q^\dagger)^\dagger (Q \Lambda Q^\dagger) - (Q\Lambda Q^\dagger)(Q\Lambda Q^\dagger)^\dagger$$

$$= (Q^\dagger)^\dagger \Lambda^\dagger Q^\dagger Q \Lambda Q^\dagger - Q\Lambda Q^\dagger (Q^\dagger)^\dagger \Lambda^\dagger Q^\dagger$$

$$= Q\Lambda^\dagger I \Lambda Q^\dagger - Q\Lambda Q^\dagger Q \Lambda^\dagger Q^\dagger$$

$$= Q\Lambda^\dagger \Lambda Q^\dagger - Q\Lambda\Lambda^\dagger Q^\dagger$$

Since $\Lambda^\dagger \Lambda = \Lambda \Lambda^\dagger$ as Λ is a diagonal matrix

$$= 0$$

$$\Rightarrow \boxed{A^T A - A A^T = 0}$$

To show that $A^T A = A A^T$ implies unitary diagonalizability.

Every $A \in \mathbb{R}^{m \times m}$ can be decomposed

as $A = A_s + A_{ss}$

$$A_s = \frac{1}{2}(A + A^T)$$

$$A_{ss} = \frac{1}{2}(A - A^T)$$

What does

$A^T A - A A^T = 0$ result in?

$$A^T A - A A^T = (A_s + A_{ss})^T (A_s + A_{ss}) - (A_s + A_{ss})(A_s + A_{ss})^T$$

$$= (A_s - A_{ss})(A_s + A_{ss}) - (A_s + A_{ss})(A_s - A_{ss})$$

$$= 2\underbrace{(A_s A_{ss} - A_{ss} A_s)}$$

We have $\underset{\sim}{A}^T \underset{\sim}{A} - \underset{\sim}{A}\underset{\sim}{A}^T = \underset{\sim}{0}$

$$(A_s A_{ss} - A_{ss} A_s) = 0$$

$$\Rightarrow \boxed{A_s A_{ss} = A_{ss} A_s}$$

(This means A_s and A_{ss} commute with each other)

Note :-

* Two matrices commute if and only if they have same eigenvectors!

Here Q is a unitary matrix

$$\underset{\sim}{A}_s = \underset{\sim}{Q}\,\underset{\sim}{\Lambda}_s \underset{\sim}{Q}^{+} \quad , \quad \underset{\sim}{A}_{ss} = \underset{\sim}{Q}\,\underset{\sim}{\Lambda}_{ss} \underset{\sim}{Q}^{+}$$

$$\Rightarrow \underset{\sim}{\Lambda}_s = \underbrace{\underset{\sim}{Q}^{+} \underset{\sim}{A}_s \underset{\sim}{Q}} \quad \Rightarrow \underset{\sim}{\Lambda}_{ss} = \underbrace{\underset{\sim}{Q}^{+} \underset{\sim}{A}_{ss} \underset{\sim}{Q}}$$

Consider

$$\underset{\sim}{Q}^{+} \underset{\sim}{A} \underset{\sim}{Q} = \underset{\sim}{Q}^{+} (\underset{\sim}{A}_s + \underset{\sim}{A}_{ss}) \underset{\sim}{Q}$$

$$= \underset{\sim}{\Lambda}_s + \underset{\sim}{\Lambda}_{ss} = \underset{\sim}{\Lambda}$$

Hence $\underset{\sim}{Q}^{+} \underset{\sim}{A} \underset{\sim}{Q} = \underset{\sim}{\Lambda}$

$\Rightarrow \boxed{\underset{\sim}{A} = \underset{\sim}{Q} \underset{\sim}{\Lambda} \underset{\sim}{Q}^{+}}$

<u>Schur factorization</u>

A factorization of $\underset{\sim}{A} \in \mathbb{C}^{m \times m}$ of the form $\underset{\sim}{A} = \underset{\sim}{Q} \underset{\sim}{T} \underset{\sim}{Q}^{+}$ where $\underset{\sim}{Q}$ is unitary $(Q^{+}Q = I)$ and $\underset{\sim}{T}$ is upper triangular is called Schur factorization

<u>Note</u>:- Since $\underset{\sim}{A}$ and $\underset{\sim}{T}$ are similar, the eigenvalues of $\underset{\sim}{A}$ and $\underset{\sim}{T}$ are same!

Thm:- Every square matrix $\underset{\sim}{A} \in \mathbb{C}^{m \times m}$ has a Schur factorization of the form

$\underset{\sim}{A} = \underset{\sim}{Q}\,\underset{\sim}{T}\,\underset{\sim}{Q}^{+}$ as discussed before.

If $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$, a factorization with real matrices exist $\boxed{\underset{\sim}{A} = \underset{\sim}{U}\,\underset{\sim}{T}\,\underset{\sim}{U}^{T}}$ where $\underset{\sim}{U}^{T}\underset{\sim}{U} = \underset{\sim}{U}\,\underset{\sim}{U}^{T} = I$ where $\underset{\sim}{T}$ is quasi upper triangular.

Quasi upper triangular matrices are those with diagonals with 1×1 block or 2×2 block

Summary:-

(i) A diagonalization $\underset{\sim}{A} = \underset{\sim}{X}\,\underset{\sim}{\Lambda}\,\underset{\sim}{X}^{-1}$ exists if and only if $\underset{\sim}{A}$ is non-defective

(ii) A unitary diagonalization $\underbrace{\underset{\sim}{A} = \underset{\sim}{Q}\,\Lambda\,\underset{\sim}{Q}^{+}}$

exists if $\underset{\sim}{A}$ is normal

(iii) A unitary triangularization (Schur Factorization)

$\underset{\sim}{A} = \underset{\sim}{Q}\,\underset{\sim}{T}\,\underset{\sim}{Q}^{\dagger}$ always exists