LECTURE 22

SIMILARITY TRANSFORMATIONS

Let $A \in \mathbb{C}^{n\times n}$,
$S \in \mathbb{C}^{n\times n}$ be invertible.

The transformation

$$T: A \longrightarrow S^{-1}AS$$

is called a <u>similarity transformation</u>. The matrices A and $S^{-1}AS$ are said to be <u>similar</u>.

EXAMPLE

Consider the similarity transformation

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}$$

$$\longrightarrow$$

$$\begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}$$

The matrices

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}$$

are similar with the same eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 1$.

THM (Similarity Transformation & Eigenvalues)
Suppose $A, B \in \mathbb{C}^{n \times n}$ are similar matrices. Then A and B have exactly the same set of eigenvalues with the same algebraic and geometric multiplicities.

PROOF
There exists an S (invertible) such that

$$B = S^{-1} A S.$$

But then

$$\det(B - \lambda I) = \det(S^{-1} A S - \lambda I)$$
$$= \det(S^{-1} A S - \lambda S^{-1} S)$$

$$= \det(S^{-1}(A-\lambda I)S)$$

$$= \underbrace{\det(S^{-1})}_{\neq 0}\det(A-\lambda I)\underbrace{\det(S)}_{\neq 0}.$$

Therefore

$$\det(A-\lambda I)=0 \Longleftrightarrow \det(B-\lambda I)=0.$$

In other words A and B have the same characteristic polynomial meaning they share the same eigenvalues with same algebraic multiplicities.

Furthermore since S is invertible,

$$\operatorname{rank}(A-\lambda I) = \operatorname{rank}(S^{-1}(A-\lambda I)S) = \operatorname{rank}(B-\lambda I)$$

$$\Longrightarrow$$

$$\dim(\operatorname{Null}(A-\lambda I)) = \dim(\operatorname{Null}(B-\lambda I)).$$

Consequently eigenvalues of A and B have the same geometric multiplicities. $\square$

EXAMPLE (Algebraic & Geometric Multiplicities)

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

* $\lambda = 1$ is the only eigenvalue.
* $v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ are associated eigvecs.
* $E_\lambda = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2$ (eigenspace associated with $\lambda = 1$)
* algebraic multiplicity of $\lambda = 1$ is 2.
* geometric multiplicity of $\lambda = 1$ is 2.

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

* $\lambda = 1$ is the only eigenvalue.
* $E_\lambda = \text{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$
* algebraic multiplicity of $\lambda = 1$ is 2.
* geometric multiplicity of $\lambda = 1$ is 1.

THM (Algebraic & Geometric Multiplicities)

Let $A \in \mathbb{C}^{n\times n}$ and $\lambda \in \mathbb{C}$ be an eigenvalue of A. Then

algebraic multip. of λ

$\geqslant$

geometric multip of λ

PROOF

Let $\{q_1, q_2, \cdots, q_m\}$ be an orthonormal basis for $E_\lambda = Null(A-\lambda I)$.

Form an unitary matrix of the form

$$Q = [\underbrace{q_1 \ q_2 \ \cdots \ q_m}_{Q_m} \ \underbrace{q_{m+1} \ \cdots \ q_n}_{\hat{Q}}].$$

Now A is similar to

$$Q^* A Q = \begin{bmatrix} Q_m^* \\ \hat{Q}^* \end{bmatrix} A \, [Q_m \ \hat{Q}]$$

$$= \begin{bmatrix} \lambda Q_m^* Q_m & Q_m^* A \hat{Q} \\ \lambda \hat{Q}^* Q_m & \hat{Q}^* A \hat{Q} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda I_m & Q_m^* A \hat{Q} \\ 0 & \hat{Q}^* A \hat{Q} \end{bmatrix}.$$

Consequently the algebraic multiplicity of λ as an eigenvalue of Q^*AQ and A is at least $m = \dim(E_\lambda)$.

$\square$

<u>TERMINOLOGY</u>

An eigenvalue λ is called

* <u>defective</u> if its algebraic multip. is strictly greater than its geometric multip,
* <u>simple</u> if its algebraic multiplicity is one,
* <u>semi-simple</u> (or non-defective) if its algebraic multiplicity is equal to its geometric multiplicity.

A matrix $A \in \mathbb{C}^{n\times n}$ is called <u>non-defective</u> if

* it has n linearly independent eigenvectors, equivalently
* all eigenvalues of A are semi-simple (note that eigenvectors associated with different eigenvalues are linearly independent.)

EXAMPLE

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

is defective, because it has only one linearly independent eigenvector lying in span$\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$.

* This is due to the fact that the eigenvalue $\lambda = 1$ is defective.

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}$$

is non-defective, because it has two linearly independent eigenvectors $v_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ associated with eigenvalues $\lambda_1 = -3$ and $\lambda_2 = 1$, respectively.

* This is due to the fact that both eigenvalues are simple (therefore semi-simple).

In general given a non-defective matrix $A \in \mathbb{C}^{n\times n}$ with

* the set of linearly independent eigenvectors $\{v_1, v_2, \ldots, v_n\}$,
* and associated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

$$A v_1 = \lambda_1 v_1, \quad \ldots, \quad A v_n = \lambda_n v_n$$

$$\Longrightarrow$$

$$A \underbrace{[v_1 \cdots v_n]}_{V} = \underbrace{[v_1 \cdots v_n]}_{V} \underbrace{\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}}_{\Lambda}$$

$$\Longrightarrow$$

$$A = \underbrace{V}_{\substack{\text{matrix} \\ \text{of} \\ \text{eigenvecs}}} \underbrace{\Lambda}_{\substack{\text{matrix} \\ \text{of} \\ \text{eigenvals}}} V^{-1}$$

The decomposition

$$A = V \Lambda V^{-1}$$

where $V \in \mathbb{C}^{n\times n}$ is invertible and $\Lambda \in \mathbb{C}^{n\times n}$ is diagonal is called the <u>eigenvalue (or spectral) decomposition</u> of A.

XAMPLE

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \end{bmatrix} = (-3)\begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = (1)\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\Longrightarrow$$

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Longrightarrow$$

$\boxed{\text{igenvalue ecomposition}}$

$$\begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}}_{V}\underbrace{\begin{bmatrix} -3 & 0 \\ 0 & 1 \end{bmatrix}}_{\Lambda}\underbrace{\begin{bmatrix} -2 & 2 \\ 1 & 1 \end{bmatrix}^{-1}}_{V^{-1}}$$

DEFN (Diagonalizability)

A matrix $A \in \mathbb{C}^{n \times n}$ is called diagonalizable if it has an eigenvalue decomposition.

THM (Diagonalizable Matrices)

A matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if and only if

(i) it is non-defective, equivalently

(ii) all of its eigenvalues are semi-simple.

Computation of Eigenvalues (Overview)

Apply similarity transformations of form

$$(+)\quad A \longrightarrow Q_1^{-1} A Q_1 \longrightarrow Q_2^{-1} Q_1^{-1} A Q_1 Q_2$$
$$\cdots\cdots \longrightarrow \underbrace{Q_k^{-1} \cdots Q_1^{-1} A Q_1 \cdots Q_k}_{A_k}$$

so that

* $\lim\limits_{k\to\infty} A_k$ is an upper triangular matrix.

REMARKS

* Eigenvalues of a triangular matrix are given by its diagonal entries.

e.g. $\begin{bmatrix} -1 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

has the characteristic polynomial

$$p(\lambda) = (\lambda+1)(\lambda-3)(\lambda-2)$$

and eigenvalues

$$\lambda_1 = -1, \quad \lambda_2 = 3, \quad \lambda_3 = 2$$

* Reduction into a diagonal matrix by similarity transformations is too ambitious; not all matrices are diagonalizable.

* In (+) Q_k is indeed a unitary matrix so that

$$A_k = Q_k^* \cdots Q_1^* A Q_1 \cdots Q_k$$

The reduction into a triangular form by unitary similarity transformations is possible due to the existence of a Schur factorization for every matrix $A \in \mathbb{C}^{n \times n}$.

<u>THM</u> (Schur Factorization)
Every matrix $A \in \mathbb{C}^{n\times n}$ has a factorization of the form

$$(++) \quad A = Q T Q^*$$

where $Q \in \mathbb{C}^{n\times n}$ is unitary and $T \in \mathbb{C}^{n \times n}$ is upper triangular.

e.g.

$$\underbrace{\begin{bmatrix} 4 & 1 \\ -2 & 7 \end{bmatrix}}_{A} = \underbrace{\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\right)}_{Q} \underbrace{\begin{bmatrix} 5 & 3 \\ 0 & 6 \end{bmatrix}}_{T} \underbrace{\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\right)}_{Q^T}$$

PROOF OF THM – SCHUR FACTORIZATION

The proof is by induction on the size of the matrix.

$\boxed{\text{Base case}}$ $(n=1)$
Any scalar has trivially a Schur factorization.

$\boxed{\text{Inductive case}}$ $(n=k>1)$
As the inductive hypothesis suppose $(k-1)\times(k-1)$ matrices have factorizations of the form $(++)$.

Let λ be an eigenvalue of A and q be a unit eigenvector associated with λ.

Consider a unitary matrix $Q \in \mathbb{C}^{n \times n}$ of form

$$Q = \begin{bmatrix} q & \hat{Q} \end{bmatrix}$$

satisfying

$$Q^* A Q = \begin{bmatrix} q^* \\ \hat{Q}^* \end{bmatrix} \begin{bmatrix} Aq & A\hat{Q} \end{bmatrix}$$

$$= \begin{bmatrix} q^* \\ \hat{Q}^* \end{bmatrix} \begin{bmatrix} \lambda q & A\hat{Q} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda q^* q & \overbrace{q^* A \hat{Q}}^{B} \\ \lambda \hat{Q}^* q & \underbrace{\hat{Q}^* A \hat{Q}}_{C} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda & B \\ 0 & C \end{bmatrix}.$$

By inductive hypothesis $C \in \mathbb{C}^{(k-1) \times (k-1)}$ has a factorization

$$C = \tilde{Q} \tilde{T} \tilde{Q}^*$$

where $\tilde{Q} \in \mathbb{C}^{(k-1) \times (k-1)}$ is unitary, $\tilde{T}$ is upper triangular.

refore

$$Q^* A Q = \begin{bmatrix} \lambda & B \\ 0 & \tilde{Q}\tilde{T}\tilde{Q}^* \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{bmatrix} \begin{bmatrix} \lambda & B\tilde{Q} \\ 0 & \tilde{T} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q}^* \end{bmatrix}$$

$$\Longrightarrow$$

$$A = \underbrace{Q \begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q} \end{bmatrix}}_{\hat{Q}} \underbrace{\begin{bmatrix} \lambda & B\tilde{Q} \\ 0 & \tilde{T} \end{bmatrix}}_{\hat{T}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \tilde{Q}^* \end{bmatrix} Q^*}_{\hat{Q}^*}$$

where $\hat{Q}$ is unitary, $\hat{T}$ is upper triangular implying the existence of a factorization of form (++). $\square$

Now suppose $A \in \mathbb{C}^{n\times n}$ is Hermitian.

$$A^* = A \Longrightarrow QTQ^* = QT^*Q^*$$
$$\Longrightarrow T = T^*$$
$$\Longrightarrow T = \Lambda \text{ is diagonal with real entries.}$$

For an Hermitian matrix A the Schur factorization becomes an orthogonal eigenvalue decomposition

$$A = Q\Lambda Q^* \quad (\text{equivalently } AQ = Q\Lambda)$$

EXAMPLE

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

* has real eigenvalues $\lambda_1 = 3$, $\lambda_2 = -1$
* and the associated eigenvectors

$v_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $v_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ are orthogonal

THM (Symmetric Eigenvalue Problem)

Let $A \in \mathbb{C}^{n \times n}$ be Hermitian. Then

(i) the eigenvalues $\lambda_1, \dots, \lambda_n$ of A are real, and

(ii) there exists a set of associated orthonormal eigenvectors $\{q_1, \dots, q_n\}$.