LECTURE 22

SIMILARITY TRANSFORMATIONS

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

The transformation
\[
T : A \rightarrow S^{-1}AS
\]
is called a similarity transformation.
The matrices \( A \) and \( S^{-1}AS \) are said to be similar.

EXAMPLE

Consider the similarity transformation
\[
\begin{bmatrix}
-1 & 4 \\
1 & -1
\end{bmatrix}
\]
\[
\begin{bmatrix}
-2 & 2 \\
1 & 1
\end{bmatrix}^{-1} \cdot
\begin{bmatrix}
-1 & 4 \\
1 & -1
\end{bmatrix} \cdot
\begin{bmatrix}
-2 & 2 \\
1 & 1
\end{bmatrix} =
\begin{bmatrix}
-3 & 0 \\
0 & 1
\end{bmatrix}
\]
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 multiplicity.

**Proof**

There exists an \( S \) (invertible) such that

\[
B = S^{-1}AS.
\]

But then

\[
\det(B - \lambda I) = \det(S^{-1}AS - \lambda I)
\]

\[
= \det(S^{-1}AS - \lambda S^{-1}S)
\]

(2)
\[ = \det (S^{-1}(A-\lambda I)S) \]
\[ = \frac{\det (S^{-1}) \det (A-\lambda I) \det (S)}{\neq 0} \]

Therefore
\[ \det (A-\lambda I) = 0 \iff \det (B-\lambda I) = 0. \]

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,
\[ \text{rank} (A-\lambda I) = \text{rank} (S^{-1}(A-\lambda IS)) = \text{rank} (B-\lambda I) \]
\[ \implies \dim (\text{Null} (A-\lambda I)) = \dim (\text{Null} (B-\lambda I)). \]

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 eigenvectors.
* \( E_{\lambda} = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2 \) (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 $\lambda$ ≥
geometric multip. of $\lambda$

PROOF

Let $\{q_1, q_2, \ldots, q_m\}$ be an orthonormal basis for $E_\lambda = \text{Null}(A-\lambda I)$. Form an unitary matrix of the form

$$Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_m & q_{m+1} & \cdots & q_n \end{bmatrix}.$$

Now $A$ is similar to

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

(5)
\[
\begin{bmatrix}
\lambda I_m & Q^* A \hat{Q} \\
0 & \hat{Q}^* A \hat{Q}
\end{bmatrix}
\]

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

\[\square\]

**TERMINOLOGY**

An eigenvalue \( \lambda \) is called

* **defective** if its algebraic multiplicity is strictly greater than its geometric multiplicity,

* **simple** if its algebraic multiplicity is one,

* **semi-simple** if its algebraic multiplicity is equal to its geometric multiplicity.

A matrix \( A \in \mathbb{C}^{n \times n} \) is called **non-defective** 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 \( \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \).

* 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 $\lambda_1, \lambda_2, \ldots, \lambda_n$.

\[ A v_i = \lambda_i v_i, \quad \ldots, \quad A v_n = \lambda_n v_n \]

\[ A \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 \\
\vdots \\
\lambda_n \end{bmatrix} \]

\[ A = V \Lambda V^{-1} \]

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 eigenvalue (or spectral) decomposition of $A$. 

\[ \text{\textcircled{8}} \]
Example

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

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

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

**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

\[ (+) \ A \rightarrow Q_1^{-1} A Q_1 \rightarrow Q_2^{-1} Q_1^{-1} A Q_1 Q_2 \rightarrow \ldots \rightarrow Q_k^{-1} \ldots Q_1^{-1} A Q_1 \ldots Q_k \rightarrow \underbrace{A_k} \]

so that

\[ \lim_{k \to \infty} A_k \text{ is an upper triangular matrix.} \]

REMARKS

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

\[ \begin{bmatrix} -1 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} \]

has the characteristic polynomial

\[ p(\lambda) = (\lambda + 1)(\lambda - 3)(\lambda - 2) \]

and eigenvalues

\[ \lambda_1 = -1, \ \lambda_2 = 3, \ \lambda_3 = 2 \]
* 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

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

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

\[\text{THM (Schur Factorization)}\]

Every matrix $A \in \mathbb{C}^{n \times n}$ has a factorization of the form

$$(++) A = QTQ^*$$

where $Q \in \mathbb{C}^{n \times n}$ is unitary and $T \in \mathbb{C}^{n \times n}$ is upper triangular.
e.g.
\[
\begin{bmatrix}
4 & 1 \\
-2 & 7 \\
\end{bmatrix} = \begin{pmatrix}
\frac{1}{\sqrt{2}} & 1 \\
1 & -1 \\
\end{pmatrix} \begin{bmatrix}
5 & 3 \\
0 & 6 \\
\end{bmatrix} \begin{pmatrix}
\frac{1}{\sqrt{2}} & -1 \\
-1 & 1 \\
\end{pmatrix}
\]

PROOF OF THM - SCHUR FACTORIZATION

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

**Base case** \((n = 1)\)

Any scalar has trivially a Schur factorization.

**Inductive case** \((n = k > 1)\)

As the inductive hypothesis suppose \((k-1) \times (k-1)\) matrices have factorizations of the form \((++)\).

Let \(\lambda\) be an eigenvalue of \(A\) and \(\mathbf{q}\) be a unit eigenvector associated with \(\lambda\).
Consider a unitary matrix $Q \in \mathbb{C}^{n \times n}$ of form

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

satisfying

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

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

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

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

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.
Therefore
\[ Q^* A Q = \begin{bmatrix} \lambda & B \\ 0 & \bar{Q}^* \end{bmatrix} \]
\[ = \begin{bmatrix} 1 & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} \lambda & B \bar{Q} \\ 0 & \bar{T} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \bar{Q}^* \end{bmatrix} \]
\[ \Rightarrow A = Q \begin{bmatrix} 1 & 0 \\ 0 & \bar{Q} \end{bmatrix} \begin{bmatrix} \lambda & B \bar{Q} \\ 0 & \bar{T} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \bar{Q}^* \end{bmatrix} Q^* \]

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

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

\[ A^* = A \implies QTQ^* = QT^*Q^* \]
\[ \implies T = T^* \]
\[ \implies T = \Lambda \text{ is diagonal with real entries.} \]

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

\[ A = Q \Lambda Q^* \quad (\text{equivalently} \quad AQ = Q\Lambda) \]
EXAMPLE

\[
\begin{bmatrix}
1 & 2 \\
2 & 1
\end{bmatrix}
\]

* has real eigenvalues \( \lambda_1 = 3 \), \( \lambda_2 = -1 \)
* and the associated eigenvectors
\[
\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}
\]
are orthogonal.

THM (Symmetric Eigenvalue Problem)
Let \( A \in \mathbb{C}^{n \times n} \) be Hermitian. Then

(i) the eigenvalues \( \lambda_1, \ldots, \lambda_n \) of \( A \) are real, and
(ii) there exists a set of associated orthonormal eigenvectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \).