# Existence of Eigendecomposition

Question: Not every $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) admits an eigendecomposition

Counter example: Consider

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

The characteristic polynomial is $p(\lambda) = -\lambda^3$, so $\lambda_1 = \lambda_2 = \lambda_3 = 0$

$$\mathcal{N}(\mathbf{A} - \lambda_1 \mathbf{I}) = \mathcal{N}(\mathbf{A}) = \mathcal{R}(\mathbf{A}^T)^\perp = \mathrm{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Any $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathcal{N}(\mathbf{A})$ are linearly dependent
Therefore, any $\mathbf{V}$ satisfying $\mathbf{AV} = \mathbf{V}\Lambda$ is singular

# Existence of Eigendecomposition (cont'd)

**Fact**: Eigenvectors associated with distinct eigenvalues are linearly independent

- If all the eigenvalues of **A** are distinct, i.e.,

$$\lambda_i \neq \lambda_j, \quad \text{for all } i, j \in \{1, \ldots, n\} \text{ with } i \neq j,$$

then **A** admits an eigendecomposition

$\sum_i \mu_i = n$

$\lambda_i$ distinct

> **Theorem**
> **A** *admits an eigendecomposition if and only if* $\mu_i = \gamma_i$ *for each eigenvalue* $\lambda_i$

Given $\lambda_i$, dim of eigenspace $= \gamma_i = \mu_i \Rightarrow \exists\ \mu_i$ linearly indep. eigenvectors

+ Fact $\Rightarrow$ $n$ linearly indep. eigenvectors associated with $\lambda_i$

# Proof of the Fact

Let $\lambda_1, \cdots, \lambda_k$ be all the distinct eigenvalues of $A$
$v_1, \cdots, v_k$ be the corresponding eigenvectors
Our goal is to show $v_1, \cdots, v_k$ are linearly indep.
For simplicity assume to the contrary ~~that~~

$$v_k = \alpha_1 v_1 + \cdots + \alpha_{k-1} v_{k-1} \quad (*) \qquad \alpha_1, \cdots, \alpha_{k-1} \text{ not all zero}$$

and $v_1, \cdots, v_{k-1}$ are linearly indep.

$$A v_k \overset{(*)}{=} \alpha_1 A v_1 + \cdots + \alpha_{k-1} A v_{k-1} = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_{k-1} \lambda_{k-1} v_{k-1} \quad (1)$$

$$A v_k = \lambda_k v_k \overset{(*)}{=} \alpha_1 \lambda_k v_1 + \cdots + \alpha_{k-1} \lambda_k v_{k-1} \quad (2)$$

$$(1) - (2) \Rightarrow 0 = \alpha_1 \underbrace{(\lambda_1 - \lambda_k)}_{\neq 0} v_1 + \cdots + \alpha_{k-1} \underbrace{(\lambda_{k-1} - \lambda_k)}_{\neq 0} v_{k-1}$$

Since $v_1, \cdots, v_{k-1}$ linearly indep., $\alpha_1 = \cdots \alpha_{k-1} = 0 \overset{(*)}{\Rightarrow} v_k = 0$

Contradiction!

# Matrix Computations
# Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

## Section 4.2 Schur Decomposition

Jie Lu
ShanghaiTech University

# Schur Decomposition

**Theorem**

*Let $\mathbf{A} \in \mathbb{C}^{n \times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. The matrix $\mathbf{A}$ admits a decomposition*

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$$

$U^H U = U U^H = I$

similarity transformation
A and T similar

*for some unitary $\mathbf{U} \in \mathbb{C}^{n \times n}$ and some upper triangular $\mathbf{T} \in \mathbb{C}^{n \times n}$ with $t_{ii} = \lambda_i$ for all $i$. If $\mathbf{A}$ is real and $\lambda_1, \ldots, \lambda_n$ are all real, $\mathbf{U}$ and $\mathbf{T}$ can be taken as real.*

- The above decomposition is called the Schur decomposition
- Suppose $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$ for some unitary $\mathbf{U}$ and upper triangular $\mathbf{T}$, but it's unknown whether $t_{ii} = \lambda_i$. Indeed, $t_{ii} = \lambda_i$ has to be true:

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \det(\lambda\mathbf{I} - \mathbf{T}) = \prod_{i=1}^{n}(\lambda - t_{ii})$$

- Any square matrix is similar to an upper triangular matrix whose diagonal entries are its eigenvalues and the "triangularizer" is unitary

# Proof of Schur Decomposition

## Lemma

special similarity transformation for block triangular matrices

*Let* $\mathbf{X} \in \mathbb{C}^{n \times n}$ *be block upper triangular in the form of*

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{X}_{12} \\ \mathbf{0} & \mathbf{X}_{22} \end{bmatrix}$$

*with* $\mathbf{X}_{11} \in \mathbb{C}^{k \times k}$, $\mathbf{X}_{22} \in \mathbb{C}^{(n-k)\times(n-k)}$, $0 \le k < n$. *There exists a unitary* $\mathbf{U} \in \mathbb{C}^{n \times n}$ *s.t.*

$$\mathbf{U}^H \mathbf{X} \mathbf{U} = \begin{bmatrix} \mathbf{X}_{11} & \mathbf{Y}_{12} \\ \mathbf{0} & \mathbf{Y}_{22} \end{bmatrix}, \quad \mathbf{Y}_{22} = \begin{bmatrix} \bar{\lambda} & \times \\ \mathbf{0} & \times \end{bmatrix} \in \mathbb{C}^{(n-k)\times(n-k)}, \ \bar{\lambda} \in \mathbb{C}$$

Proof of lemma: Let $\bar{\lambda}$ be any eigenvalue of $X_{22}$ and let $v \in \mathbb{C}^{n-k}$ be a corresponding eigenvector.

Similar to the proof of $\mu_i \geq \gamma_i$ in Sec 4.1, $\exists\, q_2, \cdots, q_{n-k}$ s.t. $Q = [\, v \ \ q_2 \cdots q_{n-k} \,]$ is unitary. Then,

$$Q^H X_{22} Q = \begin{bmatrix} \bar{\lambda} & * \\ 0 & * \end{bmatrix} \qquad (*)$$

# Proof of Schur Decomposition (cont'd)

Let $U = \begin{bmatrix} I_k & 0 \\ 0 & Q \end{bmatrix}$ unitary. Then,

$$U^H X U = \begin{bmatrix} I_k & 0 \\ 0 & Q^H \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ 0 & X_{22} \end{bmatrix} \begin{bmatrix} I_k & 0 \\ 0 & Q \end{bmatrix}$$

$$= \begin{bmatrix} X_{11} & X_{12} \\ 0 & Q^H X_{22} \end{bmatrix} \begin{bmatrix} I_k & 0 \\ 0 & Q \end{bmatrix}$$

$$= \begin{bmatrix} X_{11} & \overset{Y_{12}}{\boxed{X_{12} Q}} \\ 0 & \underset{Y_{22}}{\boxed{Q^H X_{22} Q}} \end{bmatrix}$$

$\blacksquare$

Recursively apply the lemma to obtain Schur decomposition.

## Proof of Schur Decomposition (cont'd)

Let $A^{(0)} = A$. For each $i = 1, \cdots, n-1$, let

$$A^{(i)} = (U^{(i)})^H A^{(i-1)} U^{(i)}$$

where $U^{(i)}$ is unitary and obtained by applying the lemma with $X = A^{(i-1)}$ and $k = i-1$.

$(i = 1: \; X = X_{22} = A)$

Thus, $A^{(i)}$, $i = 1, \cdots, n-1$ takes the form $\begin{bmatrix} T_{ii} & * \\ 0 & * \end{bmatrix}$ with $T_{ii}$ upper triangular. and $A^{(n-1)}$ is upper triangular

$$T = A^{(n-1)} = \underbrace{(U^{(n-1)})^H \cdots (U^{(2)})^H (U^{(1)})^H}_{U^H} A \underbrace{U^{(1)} U^{(2)} \cdots U^{(n-1)}}_{U}$$

Let $U = U^{(1)} \cdots U^{(n-1)}$. $\Rightarrow T = U^H A U$.

# Computations of Schur Decomposition

- The proof of Schur Decomposition indicates how to compute the Schur factors $\mathbf{U}$ and $\mathbf{T}$

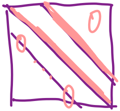
- From the lemma in the proof, we need two sub-algorithms to construct $\mathbf{U}$ and $\mathbf{T}$
  - An algorithm for computing an eigenvector of a given matrix (the power method, will be studied later)
  - An algorithm that finds a unitary matrix $\mathbf{Q}$ s.t. its first column is given (QR decomposition)

- There are other computationally more efficient methods for computing the Schur factors (key: QR decomposition)

# Discussion

- The Schur decomposition is a powerful tool
- For example, we can use it to show that for *any* square **A** (with or without eigendecomposition), $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$, $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$

  because A and T are similar
- We can also use it to prove the convergence of the power method (later) when eigendecomposition does not exist
- An enhancement of the Schur decomposition: Every square matrix **A** is also similar to a block diagonal (indeed upper triangular and tri-diagonal) matrix **J** called Jordan canonical form



$$\mathbf{A} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}, \quad \mathbf{S} \text{ is nonsingular}$$

  - We can apply the Schur decomposition to the proof of Jordan canonical form by showing that the Schur factor **T** is similar to **J** (non-trivial)

# A Consequence of Schur Decomposition

## Proposition

*Let $\mathbf{A} \in \mathbb{C}^{n\times n}$. For any $\varepsilon > 0$, there exists a matrix $\tilde{\mathbf{A}} \in \mathbb{C}^{n\times n}$ s.t. the $n$ eigenvalues of $\tilde{\mathbf{A}}$ are distinct and*

$$\|\mathbf{A} - \tilde{\mathbf{A}}\|_F^2 \leq \varepsilon.$$

**Implication**: For any square $\mathbf{A}$, we can always find $\tilde{\mathbf{A}}$ that is arbitrarily close to $\mathbf{A}$ and admits an eigendecomposition

**Proof** (construction of $\tilde{\mathbf{A}}$):

- Let $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$ be the Schur decomposition of $\mathbf{A}$. Let $\mathbf{D} = \mathrm{Diag}(d_1, \ldots, d_n)$ where $d_1, \ldots, d_n$ are chosen such that (1) $|d_i| \leq \left(\frac{\varepsilon}{n}\right)^{1/2}$ for all $i$ and (2) $t_{11} + d_1, \ldots, t_{nn} + d_n$ are distinct
- Let $\tilde{\mathbf{A}} = \mathbf{U}(\mathbf{T} + \mathbf{D})\mathbf{U}^H$ Schur decomposition of $\tilde{A}$
- We have $\|\mathbf{A} - \tilde{\mathbf{A}}\|_F^2 = \|\mathbf{U}\mathbf{D}\mathbf{U}^H\|_F^2 = \|\mathbf{D}\|_F^2 \leq \varepsilon$

# Matrix Computations
# Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

## Section 4.3 Hermitian matrices and the Variational Characterizations of Eigenvalues

Jie Lu
ShanghaiTech University

# Hermitian Matrices

Recall that

- A square matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ is said to be Hermitian if $a_{ij} = a_{ji}^*$ for all $i, j$ with $i \neq j$, or equivalently, if $\mathbf{A}^H = \mathbf{A}$

- We denote the set of all $n \times n$ complex Hermitian matrices by $\mathbb{H}^n$

- By definition, a real symmetric matrix is also Hermitian, i.e., $\mathbb{S}^n \subset \mathbb{H}^n$

- When we say that a matrix is Hermitian, we often imply that the matrix is complex–a real Hermitian matrix is simply real symmetric

# Eigenvalues and Eigenvectors of Hermitian Matrices

## Property

*The following properties hold for* $\mathbf{A} \in \mathbb{H}^n$:

1. *The* $n$ *eigenvalues of* $\mathbf{A}$ *are real*
2. *Suppose* $\{\lambda_1, \ldots, \lambda_k\}$ *is the set of all distinct eigenvalues of* $\mathbf{A}$, *and let* $\mathbf{v}_i$ *be* any *eigenvector associated with* $\lambda_i$. *Then* $\mathbf{v}_1, \ldots, \mathbf{v}_k$ *are orthogonal*
3. *There exists an orthonormal basis of* $\mathbb{C}^n$ *consisting of the eigenvectors of* $\mathbf{A}$ $\Longleftrightarrow$ A has n orthonormal eigenvectors

Any n linearly independent eigenvectors of A may NOT be orthogonal

**Corollary**: For any $\mathbf{A} \in \mathbb{S}^n$, there exist $n$ real orthonormal eigenvectors

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

# Proof of the Property

**Idea**: Use invariant subspace

Given $\mathbf{A} \in \mathbb{C}^{n \times n}$, a subspace $\mathcal{S} \subseteq \mathbb{C}^n$ with

$$\mathbf{x} \in \mathcal{S} \Longrightarrow \mathbf{Ax} \in \mathcal{S}$$

is said to be an invariant subspace for $\mathbf{A}$

E.g., any eigenvector of $\mathbf{A}$ spans a 1-dimensional invariant subspace

Let $v$ be an eigenvector of $A$ and $S = \text{span}\{v\}$.

$x \in S \Leftrightarrow x = \alpha v$ for some scalar $\alpha$

$Ax = \alpha A v = \alpha \lambda v \in S$

E.g., any $k$ eigenvectors of $\mathbf{A}$ spans an invariant subspace for $\mathbf{A}$

Let $v_1, \cdots, v_k$ be eigenvectors of $A$ and $S = \text{span}\{v_1, \cdots, v_k\}$.

$x \in S \Leftrightarrow x = \alpha_1 v_1 + \cdots + \alpha_k v_k$

$Ax = \alpha_1 A v_1 + \cdots + \alpha_k A v_k = \alpha_1 \lambda_1 v_1 + \cdots + \alpha_k \lambda_k v_k \in S$

**Fact**: If $\mathcal{Z}$ is a nonzero invariant subspace for $\mathbf{A}$, then $\mathbf{A}$ has an eigenvector in $\mathcal{Z}$

- A Consequence of the Fundamental Theorem of Algebra

# Proof of the Property (cont'd)

Property 1: Let $\lambda_1$ be an eigenvalue of $A$ with eigenvector $x \in \mathbb{C}^n$, $\|x\|_2 = 1$.

$$\lambda_1 = \lambda_1 \cdot \underbrace{x^H x}_{\|x\|_2^2 = 1} = x^H (\lambda_1 x) = x^H A x$$

$$\lambda_1^* = (x^H A x)^* = (x^H A x)^H = x^H A^H x \overset{A \in \mathcal{H}^n}{=} \underset{\overset{\|}{\lambda_1}}{x^H A x}$$

$\Rightarrow \lambda_1 \in \mathbb{R}$