SI231b: Matrix Computations

Lecture 14: Eigenvalue Revealing Decomposition

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Oct. 26, 2022

Recap: Algebraic and Geometric Multiplicity

Algebraic Multiplicity

- Characteristic polynomial $p(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$
- denote μ_i as the number of repeated eigenvalues of λ_i $(i = 1, \ldots, k)$

$$p(\lambda) = (\lambda - \lambda_1)^{\mu_1}(\lambda - \lambda_2)^{\mu_2}\cdots(\lambda - \lambda_k)^{\mu_k},$$

  with $\mu_1 + \mu_2 + \cdots + \mu_k = n$ and λ_i is distinct with λ_j.
- μ_i is called the algebraic multiplicity of the eigenvalue λ_i

Geometric Multiplicity

- every λ_i can have more than one eigenvector (scaling not counted)
- eigenspace $\mathcal{E}_{\lambda_i}$ associated with λ_i, $\mathcal{E}_{\lambda_i} = \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$
- $\gamma_i = \dim(\mathcal{E}_{\lambda_i})$ is called the geometric multiplicity of the eigenvalue λ_i

Fact: $\mu_i \geq \gamma_i$ for each λ_i.

Outline

- Defective Eigenvalues and Matrices
- Diagonalization
- Similarity Transformation
- Schur Decomposition
- Eigenvalues of Hermitian Matrices

Lemma 1: the algebraic multiplicity of an eigenvalue λ_i is at least as great as its geometric multiplicity, i.e., $\mu_i \geq \gamma_i$.

You need to prove this.

Defective Eigenvalue

An eigenvalue whose algebraic multiplicity exceeds its geometric multiplicity is a defective eigenvalue.

Defective Matrix

A matrix that has one or more defective eigenvalues.

Examples: consider the following matrices (**A** nondefective, **B** defective)

$$\mathbf{A} = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}$$

Theorem 1: An $n \times n$ matrix $\mathbf{A}$ is nondefective if and only if it has an eigenvalue decomposition

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

with $\Lambda = \mathrm{diag}(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$ and the k-th column of $\mathbf{V}$ being the eigenvector $\mathbf{v}_k$ associated with λ_k.

Hint: you need the following lemma to prove the theorem

Lemma 2: Let $\mathbf{A} \in \mathbb{C}^{n\times n}$, and suppose that the eigenvalues $\lambda_1, \ldots, \lambda_n$ are ordered such that $\{\lambda_1, \ldots \lambda_k\}$, $k \leq n$, is the set of all distinct eigenvalues of $\mathbf{A}$. Also, let $\mathbf{v}_i$ be *any* eigenvector associated with λ_i. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ must be linearly independent.

From **Theorem 1**, another term for nondefective is diagonalizable.

If $\mathbf{A}$ admits an eigenvalue decomposition, the following properties can be shown (easily):

- $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$
- $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$
- the eigenvalues of $\mathbf{A}^k$ are $\lambda_1^k, \ldots, \lambda_n^k$
- $\mathbf{A}$ is nonsingular if and only if $\mathbf{A}$ does not have zero eigenvalues
- suppose that $\mathbf{A}$ is also nonsingular. Then, $\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^{-1}$

Note: the first three properties does not require the eigenvalue decomposition to prove.

For $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$), if $\mathbf{T} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) is nonsingular, the map $\mathbf{A} \mapsto \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ is called a similarity transformation of $\mathbf{A}$.

Theorem 2 If $\mathbf{T}$ is nonsingular, then $\mathbf{A}$ and $\mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ have the same

- characteristic polynomial
- eigenvalues
- algebraic multiplicity
- geometric multiplicity

Hint: using characteristic polynomial to show.

Let $\mathbf{A} \in \mathbb{C}^{n \times n}$, the Schur decomposition of $\mathbf{A}$ is given by

$$\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H,$$

where $\mathbf{Q}$ is unitary ($\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$), and $\mathbf{T}$ is upper-triangular.

Property: Since $\mathbf{A}$ and $\mathbf{T}$ are similar, the eigenvalues of $\mathbf{A}$ appear on the diagonal of $\mathbf{T}$.

Theorem 3: Every square matrix $\mathbf{A}$ has a Schur decomposition.

Hint: applying induction to prove.

Real Eigenvalues

Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be Hermitian ($\mathbf{A} = \mathbf{A}^H$), then

1. the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\mathbf{A}$ are real
2. suppose that λ_i's are ordered such that $\{\lambda_1, \ldots, \lambda_k\}$ is the set of all distinct eigenvalues of $\mathbf{A}$. Also, let $\mathbf{v}_i$ be *any* eigenvector associated with λ_i. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ must be orthonormal.

Remark:

- the above results apply to real symmetric matrices, recall that

$$\mathbf{A} = \mathbf{A}^T \Rightarrow \mathbf{A} = \mathbf{A}^H.$$

Corollary:

- for a real symmetric matrix, all eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ can be chosen as real

Diagonalization of Hermitian Matrices

Theorem 4: Every Hermitian matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$ has an eigenvalue decomposition given by

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H,$$

where $\mathbf{V} \in \mathbb{C}^{n\times n}$ is unitary ($\mathbf{V}^H\mathbf{V} = \mathbf{I}$), $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with $\lambda_i \in \mathbb{R}$ for all i. Also, if $\mathbf{A} \in \mathbb{R}^{n\times n}$ is symmetric, $\mathbf{V}$ is an orthogonal matrix.

Hint: can you use Schur decomposition to prove this?

Remark:

- does not require the assumption of $\mu_i = \gamma_i$ for all λ_i

Corollary:

- If $\mathbf{A}$ is Hermitian or real symmetric, $\mu_i = \gamma_i$ for all λ_i (no. of repeated eigenvalues = no. of linearly independent eigenvectors)

Factorize a matrix to a form in which eigenvalues are explicitly displayed

- Diagonalization, $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1}$, exists if and only if $\mathbf{A}$ is nondefective.
- Schur decompositon, $\mathbf{A} = \mathbf{Q}\mathbf{T}\mathbf{Q}^H$ always exists.
- Jordan canonical form, $\mathbf{A} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1}$ always exists (will not be introduced in our lecture), where

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & & \\ & \mathbf{J}_2 & & \\ & & \ddots & \\ & & & \mathbf{J}_k \end{bmatrix}$$

with

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & & & \\ & \lambda_i & & \\ & & \ddots & \\ & & & \lambda_i \end{bmatrix}, \quad \text{or} \quad \mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$$

Readings

You are supposed to read

- Gene H. Golub and Charles F. Van Loan. *Matrix Computations*, Johns Hopkins University Press, 2013.

  Chapter 7.1, 8.1