SI231b: Matrix Computations

Lecture 13: Eigenvalue Problems

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Oct. 24, 2022

Outline

- Eigenvalues and Eigenvectors
- Characteristic Polynomials
- Eigenspaces
- Algebraic and Geometric Multiplicity
- Similarity Transformation
- Defective Eigenvalues and Matrices
- Eigenvalue Decomposition

Problem: given a $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$), find a vector $\mathbf{v} \in \mathbb{C}^n$ with $\mathbf{v} \neq \mathbf{0}$ such that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \qquad \text{for some } \lambda \in \mathbb{C} \tag{$*$}$$

- $(*)$ is called an eigenvalue problem or eigen-equation
- let $(\mathbf{v}, \lambda)$ be a solution to $(*)$. We call
  - $(\mathbf{v}, \lambda)$ an eigen-pair of $\mathbf{A}$
  - λ an eigenvalue of $\mathbf{A}$
  - $\mathbf{v}$ an eigenvector of $\mathbf{A}$ associated with λ
- if $(\mathbf{v}, \lambda)$ is an eigen-pair of $\mathbf{A}$, $(\alpha\mathbf{v}, \lambda)$ is also an eigen-pair for any $\alpha \in \mathbb{C}, \alpha \neq 0$
- unless specified, we will assume $\|\mathbf{v}\|_2 = 1$ in the sequel

Right Eigenvector

- $\mathbf{A}\mathbf{x} = \lambda \mathbf{x}$ for $\mathbf{x} \neq \mathbf{0}$

Left Eigenvector

- $\mathbf{x}^H \mathbf{A} = \lambda \mathbf{x}^H$ for $\mathbf{x} \neq \mathbf{0}$

Unless specified, eigenvectors in our lecture are right eigenvectors.

Spectral Radius

$$\rho(\mathbf{A}) = \max |\lambda(\mathbf{A})|$$

Numerical Range

$$W(\mathbf{A}) = \left\{ \mathbf{x}^H \mathbf{A} \mathbf{x} \,|\, \|\mathbf{x}\|_2 = 1 \right\}$$

Fact: Every $\mathbf{A} \in \mathbb{R}^{n \times n}$ (or $\mathbb{C}^{n \times n}$) has n eigenvalues.

- from the eigenvalue problem we see that

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{v} \text{ for some } \mathbf{v} \neq \mathbf{0} \iff (\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = \mathbf{0} \text{ for some } \mathbf{v} \neq \mathbf{0}$$
$$\iff \det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

- let $p(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$, called the characteristic polynomial of $\mathbf{A}$
- it can be shown that $p(\lambda)$ is a polynomial of degree n,

$$p(\lambda) = \alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \ldots + \alpha_n \lambda^n$$

  where $\{\alpha_i\}_{i=1}^{n+1}$ depend on $\mathbf{A}$

- as $p(\lambda)$ is a polynomial of degree n, it can be factored as $p(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda)$ where $\lambda_1, \ldots, \lambda_n$ are the roots of $p(\lambda)$
- we have $\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \iff \lambda \in \{\lambda_1, \ldots, \lambda_n\}$

Fact: an eigenvalue can be complex even if $\mathbf{A}$ is real.

- a polynomial $p(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_n\lambda^n$ with real coefficients α_i's can have complex roots
- example: consider

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

  - we have $p(\lambda) = \lambda^2 + 1$, so $\lambda_1 = \boldsymbol{j}$, $\lambda_2 = -\boldsymbol{j}$

Fact: if $\mathbf{A}$ is real and there exists a real eigenvalue λ of $\mathbf{A}$, the associated eigenvector $\mathbf{v}$ can be taken as real.

- obviously, when $\mathbf{A} - \lambda\mathbf{I}$ is real we can define $\mathcal{N}(\mathbf{A} - \lambda\mathbf{I})$ on $\mathbb{R}^n$
- or, if $\mathbf{v}$ is a complex eigenvector of a real $\mathbf{A}$ associated with a real λ, we can write $\mathbf{v} = \mathbf{v}_\mathrm{R} + \boldsymbol{j}\mathbf{v}_\mathrm{I}$, where $\mathbf{v}_\mathrm{R}, \mathbf{v}_\mathrm{I} \in \mathbb{R}^n$. It is easy to verify that $\mathbf{v}_\mathrm{R}$ and $\mathbf{v}_\mathrm{I}$ are eigenvectors associated with λ

For $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$), we should be careful

- the meaning of n eigenvalues: *they are defined as the n roots of the characteristic polynomial* $p(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$
- example: consider

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

  - from the original definition $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, one can verify that $\lambda = 1$ is the only eigenvalue of $\mathbf{A}$
  - from the characteristic polynomial, which is $p(\lambda) = (1 - \lambda)^2$, we see two roots $\lambda_1 = \lambda_2 = 1$ as two eigenvalues

Eigenspace

If $\mathbf{A}$ is an $n \times n$ square matrix and λ is an eigenvalue of $\mathbf{A}$, then the union of the zero vector $\mathbf{0}$ and the set of all eigenvectors corresponding to eigenvalues λ is a subspace of $\mathbb{R}^n$ known as the eigenspace of λ.

Subspace Interpretation

Denote $\mathcal{E}_\lambda$ the eigenspace of $\mathbf{A}$ associated with λ, then

- $\forall \mathbf{x} \in \mathcal{E}_\lambda$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ associated with eigenvalue λ.
- $\mathcal{E}_\lambda = \mathcal{N}(\mathbf{A} - \lambda \mathbf{I})$
- $\mathcal{E}_\lambda$ is an invariant subspace of $\mathbf{A}$, i.e., $\mathbf{A}\mathcal{E}_\lambda \subset \mathcal{E}_\lambda$

Repeated Eigenvalues

- order $\lambda_1, \ldots, \lambda_n$ such that $\{\lambda_1, \ldots \lambda_k\}$ ($k \leq n$) is the set of all distinct eigenvalues of $\mathbf{A}$, i.e., $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \ldots, k\}$ with $i \neq j$
- denote μ_i as the number of repeated eigenvalues of λ_i, $i = 1, \ldots, k$
  - μ_i is called the algebraic multiplicity of the eigenvalue λ_i

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

  with $\mu_1 + \mu_2 + \cdots + \mu_k = n$.
- every λ_i can have more than one eigenvector (scaling not counted)
  - if $\dim \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I}) = r$, we can find r linearly independent $\mathbf{v}_i$'s
  - denote $\gamma_i = \dim \mathcal{N}(\mathbf{A} - \lambda_i \mathbf{I})$, $i = 1, \ldots, k$
  - γ_i is the dimension of the eigenspace of λ_i
  - γ_i is called the geometric multiplicity of the eigenvalue λ_i

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 1 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.

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

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

Theorem 2: 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 the theorem 2, another term for nondefective is diagonalizable.

Properties of Eigenvalue Decomposition

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

- $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$
- $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$
- 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

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