SI231b: Matrix Computations

Lecture 11: Eigenvalue Problems

Yue Qiu

qiuyue@shanghaitech.edu.cn

School of Information Science and Technology

ShanghaiTech University

Oct. 20, 2020

Outline

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

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

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

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

Right Eigenvector

- $Ax = \lambda x$ for $x \neq 0$

Left Eigenvector

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

Unless specified, eigenvectors in our lecture are right eigenvectors.

Spectral Radius

$$\rho(A) = \max |\lambda(A)|$$

Numerical Range

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

Characteristic Polynomial

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

- from the eigenvalue problem we see that

$$\begin{aligned} Av = \lambda v \text{ for some } v \neq 0 \quad &\Longleftrightarrow \quad (A - \lambda I)v = 0 \text{ for some } v \neq 0 \\ &\Longleftrightarrow \quad \det(A - \lambda I) = 0 \end{aligned}$$

- let $p(\lambda) = \det(A - \lambda I)$, called the characteristic polynomial of 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 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 $\det(A - \lambda I) = 0 \Longleftrightarrow \lambda \in \{\lambda_1, \ldots, \lambda_n\}$

Eigenvalues and Eigenvectors

Fact: an eigenvalue can be complex even if 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

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

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

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

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

For $\mathrm{A} \in \mathbb{R}^{n\times n}$ (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 $p(\lambda) = \det(\mathrm{A} - \lambda \mathrm{I})$*
- example: consider

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

  - from the original definition $\mathrm{Av} = \lambda \mathrm{v}$, one can verify that $\lambda = 1$ is the only eigenvalue of 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 A is an $n \times n$ square matrix and λ is an eigenvalue of A, then the union of the zero vector 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 A associated with λ, then

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

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 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 $\dim \mathcal{N}(\mathrm{A} - \lambda_i \mathrm{I}) = r$, we can find r linearly independent v_i's
  - denote $\gamma_i = \dim \mathcal{N}(\mathrm{A} - \lambda_i \mathrm{I})$, $i = 1, \ldots, k$
  - γ_i is the dimension of the eigenspace of λ_i
  - γ_i is called the geometric multiplicity of the eigenvalue λ_i

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

Theorem 1 If T is nonsingular, then A and $\mathrm{T}^{-1}\mathrm{AT}$ 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

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

Theorem 2: An $n \times n$ matrix A is nondefective if and only if it has an eigenvalue decomposition

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

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

Hint: you need the following lemma to prove the theorem

Lemma 2: Let $\mathrm{A} \in \mathbb{C}^{n\times n}$, 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 A. Also, let v_i be *any* eigenvector associated with λ_i. Then $\mathrm{v}_1, \ldots, \mathrm{v}_k$ must be linearly independent.

From the theorem 2, another term for nondefective is diagonalizable.

Properties of Eigenvalue Decomposition

If A admits an eigenvalue decomposition, the following properties can be shown (easily):

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

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