Matrix Computations
Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

Section 4.1 Eigendecomposition

Jie Lu
ShanghaiTech University

Eigenvalues and Eigenvectors

Definition: Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$). If there exists $\mathbf{v} \in \mathbb{C}^n, \mathbf{v} \neq 0$ s.t.

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

then we say $\mathbf{v}$ is a (right) eigenvector associated with eigenvalue λ of $\mathbf{A}$.

- In general, $\mathbf{Ax}$ differs from $\mathbf{x}$ in magnitude and direction. However, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$ and $\mathbf{A}, \mathbf{x}$ are real, then $\mathbf{Ax}$ and $\mathbf{x}$ are parallel
- $(*)$ is called an eigenvalue problem or eigen-equation
- Any solution $(\mathbf{v}, \lambda)$ to $(*)$ is called an eigen-pair of $\mathbf{A}$
- If $(\mathbf{v}, \lambda)$ is an eigen-pair of $\mathbf{A}$, $(\alpha\mathbf{v}, \lambda)$ for any $\alpha \in \mathbb{C}, \alpha \neq 0$ is also an eigen-pair of $\mathbf{A}$
- If there exists a **row** vector $\mathbf{w}$, $\mathbf{w} \neq 0$ s.t. $\mathbf{wA} = \lambda \mathbf{w}$ for some $\lambda \in \mathbb{C}$, we say $\mathbf{w}$ is a left eigenvector associated with eigenvalue λ of $\mathbf{A}$

Characteristic Polynomial

Every $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$) has n (possibly repeated) eigenvalues

- From the eigenvalue problem,

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

- $p(\lambda) := \det(\lambda\mathbf{I} - \mathbf{A})$ is called the characteristic polynomial of $\mathbf{A}$ (The characteristic polynomial can also be defined to be $\det(\mathbf{A} - \lambda\mathbf{I})$, which differs from $p(\lambda)$ by a sign $(-1)^n$)
- $p(\lambda) = 0 \Longleftrightarrow \lambda$ is an eigenvalue of $\mathbf{A}$
- It can be shown that $p(\lambda)$ is a polynomial of degree n, i.e., $p(\lambda) = \alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \ldots + \alpha_n\lambda^n$ where α_i's depend on $\mathbf{A}$ and in fact, $\alpha_n = 1$
- Therefore, $p(\lambda)$ has n roots, which are the n eigenvalues of $\mathbf{A}$
- $p(\lambda)$ can be factored as $p(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$, where $\lambda_1, \ldots, \lambda_n$ are the roots of $p(\lambda)$
- Given an eigenvalue λ of $\mathbf{A}$, $\mathrm{Null}(\lambda\mathbf{I} - \mathbf{A})$ is called the eigenspace of $\mathbf{A}$ associated with λ

Complex Eigenvalues and Eigenvectors

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

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

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

If **A** is real and there exists a real eigenvalue λ of **A**, the associated eigenvector **v** can be taken as real

- When $\lambda\mathbf{I} - \mathbf{A}$ is real, we can define $\mathcal{N}(\lambda\mathbf{I} - \mathbf{A})$ on $\mathbb{R}^n$
- If **v** is a complex eigenvector of a real **A** associated with a real λ, we can write $\mathbf{v} = \mathbf{v}_{\mathrm{R}} + \boldsymbol{j}\mathbf{v}_{\mathrm{I}}$, where $\mathbf{v}_{\mathrm{R}}, \mathbf{v}_{\mathrm{I}} \in \mathbb{R}^n$. We can verify that both of $\mathbf{v}_{\mathrm{R}}$ and $\mathbf{v}_{\mathrm{I}}$ are eigenvectors associated with λ

Some Properties

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$

$$\begin{aligned} \mathbf{A}\mathbf{v} &= \lambda\mathbf{v} \\ &\Updownarrow \quad \text{Hermitian transpose} \\ \mathbf{v}^H\mathbf{A}^H &= \lambda^*\mathbf{v}^H \\ &\Updownarrow \quad \text{transpose, } \mathbf{A}^H = \mathbf{A}^T \\ \mathbf{A}\mathbf{v}^* &= \lambda^*\mathbf{v}^* \end{aligned}$$

$$\begin{aligned} \mathbf{w}\mathbf{A} &= \lambda\mathbf{w} \\ &\Updownarrow \\ \mathbf{A}^H\mathbf{w}^H &= \lambda^*\mathbf{w}^H \\ &\Updownarrow \\ \mathbf{A}^\mathsf{T}\mathbf{w}^H &= \lambda^*\mathbf{w}^H \end{aligned}$$

- $\mathbf{v}^*$ is an eigenvector associated with eigenvalue λ^*
- Complex eigenvalues appear in conjugate pairs
- $\mathbf{w}^H$ is an eigenvector associated with eigenvalue λ^* of $\mathbf{A}^\mathsf{T}$
- $\mathbf{w}^\mathsf{T}$ is an eigenvector associated with eigenvalue λ of $\mathbf{A}^\mathsf{T}$
- $\mathbf{A}$ and $\mathbf{A}^\mathsf{T}$ have the same set of eigenvalues because $\det(\lambda\mathbf{I} - \mathbf{A}) = \det(\lambda\mathbf{I} - \mathbf{A})^\mathsf{T} = \det(\lambda\mathbf{I} - \mathbf{A}^\mathsf{T})$
- The set of eigenvalues corresponding to (right) eigenvectors is the set of eigenvalues corresponding to left eigenvectors

Some Properties (cont'd)

Fact: The eigenvalues of any triangular matrix are its diagonal entries

Fact: $\mathbf{A} \in \mathbb{R}^{n \times n}$ is nonsingular if and only if all its eigenvalues are nonzero

Fact: Suppose $(\mathbf{v}, \lambda)$ is an eigen-pair of $\mathbf{A}$, then $(\mathbf{v}, \lambda^k)$ is an eigen-pair of $\mathbf{A}^k$ for any positive integer k

Repeated Eigenvalues

- Let $\lambda_1, \dots, \lambda_n$ be the n eigenvalues of $\mathbf{A} \in \mathbb{R}^{n \times n}$
- WLOG, order $\lambda_1, \dots, \lambda_n$ so that $\{\lambda_1, \dots \lambda_k\}$, $k \leq n$ is the set of all **distinct** eigenvalues of $\mathbf{A}$: $\lambda_i \neq \lambda_j$ for all $i, j \in \{1, \dots, k\}$, $i \neq j$ and $\lambda_i \in \{\lambda_1, \dots \lambda_k\}$ for all $i \in \{1, \dots, n\}$
- Define the algebraic multiplicity of eigenvalue λ_i as the multiplicity of λ_i as root of $p(\lambda)$, denoted by μ_i
- Every λ_i may have multiple eigenvectors (scaling not counted)
- If $\dim \mathcal{N}(\lambda_i \mathbf{I} - \mathbf{A}) = r$, we can find r linearly independent $\mathbf{v}_i$'s
- Define the geometric multiplicity of eigenvalue λ_i as the maximum number of linearly independent eigenvectors associated with λ_i, denoted by γ_i
  - $\gamma_i = \dim \mathcal{N}(\lambda_i \mathbf{I} - \mathbf{A}) = n - \operatorname{rank}(\lambda_i \mathbf{I} - \mathbf{A})$

Repeated Eigenvalues (cont'd)

Fact: For every eigenvalue λ_i of $\mathbf{A}$, $\mu_i \geq \gamma_i$

Repeated Eigenvalues (cont'd)

Example: $\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$

Repeated Eigenvalues (cont'd)

Example: $\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$

Similarity Transformation

- Let $\mathbf{Q} = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_n] \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$) be a nonsingular matrix
    - The columns of Q form a basis of $\mathbb{R}^n$ (or $\mathbb{C}^n$)
- Let $A \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$). We call $\tilde{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ a similarity transformation
- A and $\tilde{A}$ are said to be similar
- Similar matrices represent the same linear map under two (possibly) different bases, with $\mathbf{Q}$ being the change of basis matrix
- **Interpretation**: Consider a linear system $\mathbf{A}\mathbf{x} = \mathbf{y}$ and let $\mathbf{x} = \mathbf{Q}\tilde{\mathbf{x}}$, $\mathbf{y} = \mathbf{Q}\tilde{\mathbf{y}}$

$$\mathbf{A}\mathbf{x} = \mathbf{y} \Leftrightarrow \mathbf{A}\mathbf{Q}\tilde{\mathbf{x}} = \mathbf{Q}\tilde{\mathbf{y}} \Leftrightarrow \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}\tilde{\mathbf{x}} = \tilde{\mathbf{y}} \Leftrightarrow \tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$$

Similarity Transformation (cont'd)

- Every **square** matrix is similar to itself
- If **A** is similar to **B** and **B** is similar to **C**, then **A** is similar to **C**

- If A, B are invertible and similar, then A^{-1} and B^{-1} are also similar

- Similar matrices have the same characteristic polynomial, determinant, rank, nullity, trace, eigenvalues, algebraic multiplicity, geometric multiplicity, etc.

Eigendecomposition

A matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ (or $\mathbb{C}^{n\times n}$) is said to be diagonalizable, or admit an eigendecomposition, if there exists a nonsingular $\mathbf{V} \in \mathbb{C}^{n\times n}$ s.t.

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

where $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, or, $\mathbf{A}$ is similar to a diagonal matrix

- In this definition, we didn't say that $(\mathbf{v}_i, \lambda_i)$ is an eigen-pair of $\mathbf{A}$, but it indeed has to be

$$\begin{aligned} \mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1} &\iff \mathbf{AV} = \mathbf{V}\Lambda,\ \mathbf{V} \text{ nonsingular} \\ &\iff \mathbf{Av}_i = \lambda_i\mathbf{v}_i,\ i = 1, \ldots, n, \\ &\qquad \mathbf{v}_1, \ldots, \mathbf{v}_n \text{ linearly independent} \end{aligned}$$

- The key lies in finding n linearly independent eigenvectors to form $\mathbf{V}$

Eigendecomposition (cont'd)

Facts: Suppose $\mathbf{A}$ admits an eigendecomposition

1. $\det(\mathbf{A}) = \prod_{i=1}^{n} \lambda_i$

2. $\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$

3. $\mathrm{rank}(\mathbf{A})$ = number of nonzero eigenvalues of $\mathbf{A}$

4. Suppose $\mathbf{A}$ is also nonsingular. Then, $\mathbf{A}^{-1} = \mathbf{V}\Lambda^{-1}\mathbf{V}^{-1}$

Note: Facts 1–2 are indeed true for any $\mathbf{A}$; Facts 3–4 may not hold when $\mathbf{A}$ does not admit an eigendecomposition

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 $\mathbf{A}$ are distinct, i.e.,

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

  then $\mathbf{A}$ admits an eigendecomposition

> **Theorem**
> $\mathbf{A}$ *admits an eigendecomposition if and only if* $\mu_i = \gamma_i$ *for each eigenvalue* λ_i

Proof of the Fact