Diagonalization

Let $A \in \mathbb{C}^n$ with **distinct** eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$.

For each of the eigenvalues we have an eigenvector $v_1, \ldots, v_n \in \mathbb{C}^n$. These eigenvectors are linearly independent. Since there are n of them, they form a basis.

Consider the matrix

$$V = \begin{pmatrix} | & \ldots & | \\ v_1 & \ldots & v_n \\ | & \ldots & | \end{pmatrix}$$

whose columns are the vectors $v_1, \ldots, v_n$.

Since the columns are basis, V is invertible.

Let $e_i \in \mathbb{C}^n$ with $1 \leq i \leq n$ be a standard coordinate vector. Then

$$\begin{aligned} V^{-1}AVe_i &= V^{-1}Av_i \\ &= V^{-1}\lambda v_i = \lambda V^{-1} v_i = \lambda e_i. \end{aligned}$$

Consequently, $V^{-1}AV$ is a diagonal matrix whose diagonal entries are the eigenvalues:

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Which matrices can be **diagonalized**?

Let $A \in \mathbb{C}^{n\times n}$. We define

$$\mathrm{Eig}(A,\lambda) := \{v \in \mathbb{C}^n \mid Av = \lambda v\}.$$

This is the **eigenspace** of A for the eigenvalue λ.

For every eigenvalue λ of A, the set $\mathrm{Eig}(A,\lambda)$ is a subspace of $\mathbb{C}^n$.

We call $\mu^{\mathrm{g}}(A,\lambda) = \dim \mathrm{Eig}(A,\lambda)$ the **geometric multiplicity** of A.

How does the geometric multiplicity $\mu^{\mathrm{g}}(A, \lambda)$ compare with the algebraic multiplicity $\mu^{\mathrm{a}}(A, \lambda)$?

Example

Consider

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad T - 1 \cdot \mathrm{Id} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

The characteristic polynomial of T is $p_T(\lambda) = (1 - \lambda)^2$. Hence 1 is an eigenvalue with algebraic multiplicity $\mu^{\mathrm{a}}(A, \lambda)$. But the kernel of $T - 1 \cdot \mathrm{Id}$ is the span of e_1. Hence

$$\mu^{\mathrm{g}}(A, \lambda) = 1 < 2 = \mu^{\mathrm{a}}(A, \lambda).$$

Can we generalize this?

Theorem

Let $A \in \mathbb{C}^{n\times n}$ with eigenvalue λ_i. Then $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$.

Proof.

Let $v_1, \ldots, v_r$ be a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $v_{r+1}, \ldots, v_n \in \mathbb{C}^n$ such that $v_1, \ldots, v_n$ is a basis of $\mathrm{Eig}(A, \lambda_i)$.

Theorem

Let $A \in \mathbb{C}^{n \times n}$ with eigenvalue λ_i. Then $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$.

Proof.

Let $v_1, \dots, v_r$ be a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $v_{r+1}, \dots, v_n \in \mathbb{C}^n$ such that $v_1, \dots, v_n$ is a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $V \in \mathbb{C}^{n \times n}$ with columns $v_1, \dots, v_n$. Set $B := V^{-1}AV$. We have

$$B = \left(\lambda_1 e_1 \;\middle|\; \lambda_1 e_2 \;\middle|\; \dots \;\middle|\; \lambda_1 e_r \;\middle|\; ??? \right)$$

Theorem

Let $A \in \mathbb{C}^{n \times n}$ with eigenvalue λ_i. Then $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$.

Proof.

Let $v_1, \dots, v_r$ be a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $v_{r+1}, \dots, v_n \in \mathbb{C}^n$ such that $v_1, \dots, v_n$ is a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $V \in \mathbb{C}^{n \times n}$ with columns $v_1, \dots, v_n$. Set $B := V^{-1}AV$. We have

$$B = \left(\lambda_1 e_1 \mid \lambda_1 e_2 \mid \dots \mid \lambda_1 e_r \mid ??? \right)$$

We have $p_A = p_B$.

Theorem

Let $A \in \mathbb{C}^{n \times n}$ with eigenvalue λ_i. Then $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$.

Proof.

Let $v_1, \ldots, v_r$ be a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $v_{r+1}, \ldots, v_n \in \mathbb{C}^n$ such that $v_1, \ldots, v_n$ is a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $V \in \mathbb{C}^{n \times n}$ with columns $v_1, \ldots, v_n$. Set $B := V^{-1}AV$. We have

$$B = \left(\; \lambda_1 e_1 \;\middle|\; \lambda_1 e_2 \;\middle|\; \ldots \;\middle|\; \lambda_1 e_r \;\middle|\; ??? \; \right)$$

We have $p_A = p_B$. The characteristic polynomial p_B contains the factor $(\lambda - \lambda_i)$ at most $\mu^{\mathrm{g}}(A, \lambda)$-times. The characteristic polynomial p_A contains the factor $(\lambda - \lambda_i)$ exactly $\mu^{\mathrm{a}}(A, \lambda)$-times.

Theorem

Let $A \in \mathbb{C}^{n \times n}$ with eigenvalue λ_i. Then $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$.

Proof.

Let $v_1, \dots, v_r$ be a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $v_{r+1}, \dots, v_n \in \mathbb{C}^n$ such that $v_1, \dots, v_n$ is a basis of $\mathrm{Eig}(A, \lambda_i)$. Let $V \in \mathbb{C}^{n \times n}$ with columns $v_1, \dots, v_n$. Set $B := V^{-1}AV$. We have

$$B = \left(\lambda_1 e_1 \mid \lambda_1 e_2 \mid \dots \mid \lambda_1 e_r \mid ??? \right)$$

We have $p_A = p_B$. The characteristic polynomial p_B contains the factor $(\lambda - \lambda_i)$ at most $\mu^{\mathrm{g}}(A, \lambda)$-times. The characteristic polynomial p_A contains the factor $(\lambda - \lambda_i)$ exactly $\mu^{\mathrm{a}}(A, \lambda)$-times. Since $p_A = p_B$, we see $\mu^{\mathrm{g}}(A, \lambda_i) \leq \mu^{\mathrm{a}}(A, \lambda_i)$ □

Let $\lambda_1, \ldots, \lambda_m$ be **distinct** eigenvalues of A. Then

$$\sum_{i=1}^{m} \mu^{\mathrm{g}}(A, \lambda_i) \leq \sum_{i=1}^{m} \mu^{\mathrm{a}}(A, \lambda_i).$$

More precisely,

$$\sum_{i=1}^{m} \mu^{\mathrm{g}}(A, \lambda_i) \leq n,$$

$$\sum_{i=1}^{m} \mu^{\mathrm{a}}(A, \lambda_i) = n$$

Let $A \in \mathbb{C}^{n \times n}$. We call A **diagonalizable** if it is similar to a diagonal matrix $D \in \mathbb{C}^{n \times n}$, i.e., there exists $V \in \mathbb{C}^{n \times n}$ invertible such that $A = VDV^{-1}$.

Let $\lambda_1, \ldots, \lambda_n$ be the diagonal entries of D. Then the i-th column of V is an eigenvector of A for the eigenvalue λ.

To see this, let v_i be the i-th column. Then $Ve_i = v_i$, and

$$Av_i = VDV^{-1}v_i = VDe_i = V\lambda_i e_i = \lambda_i Ve_i = \lambda v_i.$$

Since V is invertible, the columns of V are linearly independent.

Theorem

$A \in \mathbb{C}^{n\times n}$ is diagonalizable if and only if A has n linearly independent eigenvectors.

Proof.

If $A = VDV^{-1}$ with $D \in \mathbb{C}^{n\times n}$ diagonal and $V \in \mathbb{C}^{n\times n}$ invertible, then the columns of V are n linearly independent eigenvectors.

If $V \in \mathbb{C}^{n\times n}$ is a matrix of n linearly independent eigenvectors, then it is easily seen that $D = V^{-1}AV$ is a diagonal matrix. Hence $A = VDV^{-1}$. □

Questions?