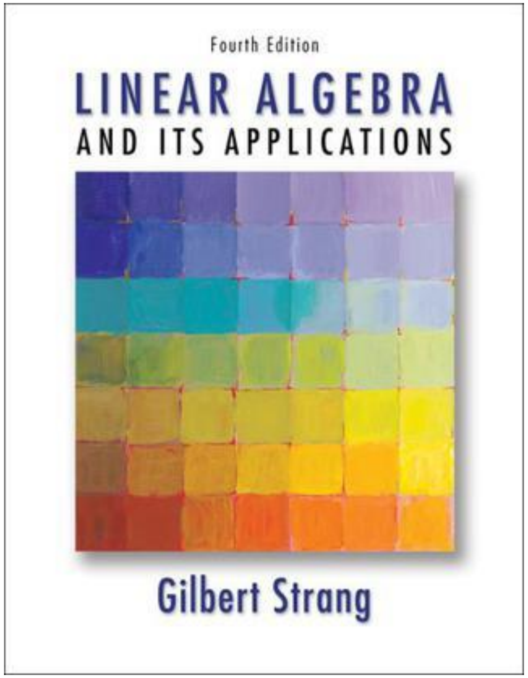
Linear Algebra



Instructor: Jing YAO

5 Eigenvalues and Eigenvectors （特征值与特征向量）

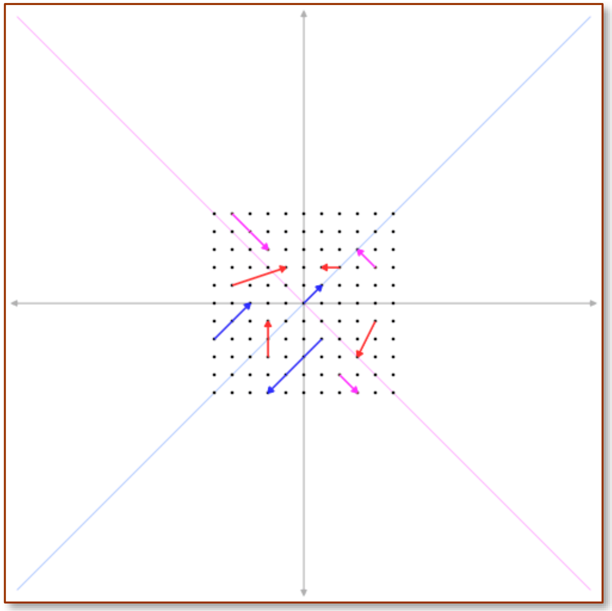
5.2

DIAGONALIZATION OF A MATRIX （矩阵的对角化）

Conditions

Examples

Powers and Products



I. Diagonalization – Conditions

A matrix $\boldsymbol{A}$ is called **diagonalizable (可对角化)** if there exists an invertible matrix $\boldsymbol{S}$ such that $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is a diagonal matrix. The matrix $\boldsymbol{S}$ is sometimes called a *diagonalizing matrix* for the matrix $\boldsymbol{A}$.

Example 1 Let $\boldsymbol{A} = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}$.

Then the eigenvalues of $\boldsymbol{A}$ are 2 and -1, and two eigenvectors are $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively.

Let $\boldsymbol{S} = \begin{bmatrix} 5 & 1 \\ 2 & 1 \end{bmatrix}$, then $\boldsymbol{S}^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix}$.

And

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \frac{1}{3}\begin{bmatrix} 1 & -1 \\ -2 & 5 \end{bmatrix}\begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix}\begin{bmatrix} 5 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}.$$

Theorem 1 Let $\boldsymbol{A}$ be a matrix of degree n, and have n linearly independent eigenvectors. Let $\boldsymbol{S}$ be a matrix with columns being the eigenvectors. Then $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ is a diagonal matrix $\boldsymbol{\Lambda}$. The eigenvalues of $\boldsymbol{A}$ are on the diagonal of $\boldsymbol{\Lambda}$:

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Proof. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ be the n linearly independent eigenvectors of $\boldsymbol{A}$, corresponding to the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, respectively. Then $\boldsymbol{S} = [\boldsymbol{x}_1\ \ \boldsymbol{x}_2 \ \dots\ \ \boldsymbol{x}_n]$, and $\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$ with $1 \le i \le n$, so

$$\begin{aligned}\boldsymbol{A}\boldsymbol{S} &= \boldsymbol{A}[\boldsymbol{x}_1\ \ \boldsymbol{x}_2 \dots \boldsymbol{x}_n] = [\lambda_1\boldsymbol{x}_1\ \ \lambda_2\boldsymbol{x}_2\ \ \dots\ \ \lambda_n\boldsymbol{x}_n] \\ &= [\boldsymbol{x}_1\ \ \boldsymbol{x}_2 \dots\ \ \boldsymbol{x}_n]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix} = \boldsymbol{S}\boldsymbol{\Lambda}.\end{aligned}$$

***Crucial** to keep these matrices in the right order!*

$\boldsymbol{S}$ is invertible, because its columns (the eigenvectors) were assumed to be independent.

Therefore, $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$

Remarks:

(1)

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

Equivalently, $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\Lambda}$; $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{S}^{-1}$.

We also call $\boldsymbol{S}$ the "*eigenvector matrix*" ("*diagonalizing matrix*") and $\boldsymbol{\Lambda}$ the "*eigenvalue matrix*".

(2) The diagonalizing matrix $\boldsymbol{S}$ *is not unique*, since an eigenvector $\boldsymbol{x}$ can be multiplied by a constant, and remains an eigenvector.

For the trivial example $\boldsymbol{A} = \boldsymbol{I}$, any invertible $\boldsymbol{S}$ will do: $\boldsymbol{S}^{-1}\boldsymbol{I}\boldsymbol{S}$ is always diagonal ($\boldsymbol{\Lambda}$ is just $\boldsymbol{I}$). All vectors are eigenvectors of the identity matrix.

Remarks:

(3) Not all matrices possess n linearly independent eigenvectors, so ***not all matrices are diagonalizable***. (并非所有方阵都可以对角化)

An example: "defective matrix" $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Its eigenvalues are $\lambda_1 = \lambda_2 = 0$, since it is triangular with zeros on the diagonal.

All eigenvectors of this $\boldsymbol{A}$ are multiples of the vector $(1,0)^{\mathrm{T}}$.

$\lambda = 0$ is a double eigenvalue. But there is only one independent eigenvector. We can't construct $\boldsymbol{S}$.

(*A more direct proof*: Since $\lambda_1 = \lambda_2 = 0$, $\boldsymbol{\Lambda}$ would have to be the zero matrix. But if $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{\Lambda} = \boldsymbol{0}$, then $\boldsymbol{A} = \boldsymbol{0}$, which is not true.)

For $\lambda = 0$:

The *algebraic multiplicity* is 2. But the *geometric multiplicity* is 1.

(Explain in the next few slides.)

Lemma 1 If a matrix $\boldsymbol{A}$ has no repeated eigenvalues, i.e., $\lambda_1, \lambda_2, \dots, \lambda_n$ are distinct, then its n eigenvectors are linearly independent, and $\boldsymbol{A}$ is diagonalizable.

(In short, a matrix with n distinct eigenvalues can be diagonalized.

具有n个互不相同特征值的n阶方阵, 一定可以对角化.)

Proof Suppose first that $k = 2$, and that some combination of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ produces zero: $c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 = \mathbf{0}$. (*)

Multiplying (*) by $\boldsymbol{A}$, we find $c_1\lambda_1\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2 = \mathbf{0}$.

Multiplying (*) by λ_2, we find $c_1\lambda_2\boldsymbol{x}_1 + c_2\lambda_2\boldsymbol{x}_2 = \mathbf{0}$.

Subtraction makes the vector $\boldsymbol{x}_2$ disappears: $c_1(\lambda_1 - \lambda_2)\boldsymbol{x}_1 = \mathbf{0}$.

Since $\lambda_1 \neq \lambda_2$ and $\boldsymbol{x}_1 \neq \mathbf{0}$, we are forced into $c_1 = 0$.

Similarly $c_2 = 0$, and the two vectors are independent.

By mathematical induction, eigenvectors that come from *distinct* eigenvalues are automatically *independent*.

Example 2 Determine if the following matrix is diagonalizable.

$$\boldsymbol{A}=\begin{bmatrix}5 & -8 & 1\\ 0 & 0 & 7\\ 0 & 0 & -2\end{bmatrix}.$$

Solution Yes. The matrix is triangular, and its eigenvalues are obviously $5, 0$, and -2.

Since $\boldsymbol{A}$ is a 3×3 matrix with three distinct eigenvalues, $\boldsymbol{A}$ is diagonalizable.

Remark:

Diagonalization can fail only if there are repeated eigenvalues.

(只有当矩阵存在重复特征值时, 才有可能不能对角化)

Even then, it does not always fail.

Example: $\boldsymbol{A}=\boldsymbol{I}$ has repeated eigenvalues $1,1,\ldots,1$, but it is already diagonal! There is no shortage of eigenvectors in that case.

What if -- there are repeated eigenvalues?

What if -- there are repeated eigenvalues?

$$A \Rightarrow |A-\lambda I|=0 \Rightarrow (A-\lambda_i I)x=0$$

求特征值λ_i 求特征向量

The set of *all* solutions of $(\boldsymbol{A}-\lambda_i\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$ is just the nullspace of the matrix $\boldsymbol{A}-\lambda_i\boldsymbol{I}$.

So this set is a *subspace* of $\mathbf{R}^n$ and is called the **eigenspace (特征子空间)** of $\boldsymbol{A}$ corresponding to λ_i.

The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ_i.

algebraic multiplicity vs. *geometric multiplicity of an eigenvalue* λ_i

- **algebraic multiplicity (代数重数)** : multiplicity of λ_i as a root of the characteristic polynomial
- **geometric multiplicity (几何重数)** : dimension of the eigenspace for λ_i.

For $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, the failure of diagonalization was ***not*** a result of $\lambda = 0$. It came from $\lambda_1 = \lambda_2$:

$$\boldsymbol{A}_1 = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}, \qquad \boldsymbol{A}_2 = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$$

Eigenvalues: 3, 3 1, 1

Eigenvectors: $(1,0)^T$ $(1,1)^T$

For $\lambda = 3$ of $\boldsymbol{A}_1$ and $\lambda = 1$ of $\boldsymbol{A}_2$:

The *algebraic multiplicity* is 2. But the *geometric multiplicity* is 1.

$\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are not singular!

The problem is the shortage of independent eigenvectors — which are needed for constructing $\boldsymbol{S}$.

Attention :

- ***Diagonalizability of A depends on enough independent eigenvectors.***
- ***Invertibility of A depends on nonzero eigenvalues.***

Example 3 Diagonalize the following matrix, if possible:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

Solution The characteristic equation of $\boldsymbol{A}$:

$$0 = |\boldsymbol{A} - \lambda \boldsymbol{I}| = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2.$$

The eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -2$(*algebraic multiplicity =2*). However, it is easy to verify that each eigenspace is only one-dimensional:

Basis for the eigenspace of $\lambda_1 = 1$: $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$.

Basis for the eigenspace of $\lambda_2 = -2$: $\boldsymbol{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

(*geometric multiplicity =1*)

There are no other eigenvalues, and every eigenvector of $\boldsymbol{A}$ is a multiple of either $\boldsymbol{x}_1$ or $\boldsymbol{x}_2$. Thus $\boldsymbol{A}$ is *not* diagonalizable.

Theorem 2

Let $\boldsymbol{A}$ be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \lambda_2, \ldots, \lambda_p$.

(1) For $1 \le i \le p$, the dimension of the eigenspace for λ_i is less than or equal to the multiplicity of the eigenvalue λ_i as a root of characteristic polynomial.

(2) The matrix $\boldsymbol{A}$ is diagonalizable if and only if the sum of the dimensions of the eigenspaces equals n, and this happens if and only if: the dimension of the eigenspace for each λ_i equals the multiplicity of λ_i.

algebraic multiplicity ≥ geometric multiplicity
(几何重数总是不超过代数重数)

The matrix $\boldsymbol{A}$ is diagonalizable if and only if *algebraic multiplicity = geometric multiplicity for each eigenvalue* λ_i.
(矩阵$\boldsymbol{A}$ 可以对角化 当且仅当 对于每一个特征值 λ_i都有:
其代数重数与几何重数相等)

Example 4 Diagonalize the following matrix, if possible:

$$\boldsymbol{A} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 1 & 4 & -3 & 0 \\ -1 & -2 & 0 & -3 \end{bmatrix}.$$

Solution The eigenvalues of $\boldsymbol{A}$ are 5 and -3, each with multiplicity 2.

For $\lambda_1 = 5$: $\boldsymbol{x}_1 = \begin{bmatrix} -8 \\ 4 \\ 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{x}_2 = \begin{bmatrix} -16 \\ 4 \\ 0 \\ 1 \end{bmatrix}$.

For $\lambda_2 = -3$: $\boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{x}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

So the matrix $\boldsymbol{S} = [\boldsymbol{x}_1 \;\; \boldsymbol{x}_2 \;\; \boldsymbol{x}_3 \;\; \boldsymbol{x}_4]$ is invertible, and $\boldsymbol{A} = \boldsymbol{S\Lambda S}^{-1}$, where $\boldsymbol{\Lambda} = \mathrm{diag}(5, 5, -3, -3)$.

The next theorem shows that *diagonalizing matrix* $\boldsymbol{S}$ must be formed by eigenvectors.

Theorem 3 Let $\boldsymbol{A}$ be a matrix of degree n, and assume that $\boldsymbol{S}$ is an invertible matrix such that

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \operatorname{diag}(d_1, d_2, \ldots, d_n).$$

Then $d_1, d_2, \ldots, d_n$ are the eigenvalues of $\boldsymbol{A}$, and column j of $\boldsymbol{S}$ is an eigenvector of $\boldsymbol{A}$ corresponding to d_j.

Proof. Let $\boldsymbol{S} = [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ldots \ \boldsymbol{v}_n]$, i.e., $\boldsymbol{v}_j$ is the j-th column of $\boldsymbol{S}$, and let $\boldsymbol{D} = \operatorname{diag}(d_1, d_2, \ldots, d_n)$. Then $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$, and so $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{D}$. Thus

$$[\boldsymbol{A}\boldsymbol{v}_1 \ \boldsymbol{A}\boldsymbol{v}_2 \ \ldots \boldsymbol{A}\boldsymbol{v}_n] = \boldsymbol{A}[\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ldots \boldsymbol{v}_n] = \boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{D}$$

$$= [\boldsymbol{v}_1 \ \boldsymbol{v}_2 \ldots \ \boldsymbol{v}_n]\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} = [d_1\boldsymbol{v}_1 \ d_2\boldsymbol{v}_2 \ldots \ d_n\boldsymbol{v}_n].$$

Therefore, $\boldsymbol{A}\boldsymbol{v}_j = d_j\boldsymbol{v}_j$ for $1 \le j \le n$, i.e., d_j is an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{v}_j$ is an eigenvector of $\boldsymbol{A}$ corresponding to d_j.

algebraic multiplicity ≥ geometric multiplicity

定理 设λ_0是 n 阶矩阵 $\boldsymbol{A}$的 k 重特征值, 属于λ_0的线性无关的特征向量的最大个数为 l, 则 $k \geq l$.

(代数重数 ≥ 几何重数)

证 由 $\boldsymbol{A}\boldsymbol{x}_i=\lambda_0\boldsymbol{x}_i$, $\boldsymbol{x}_i \neq \boldsymbol{0}$, $i=1,\cdots,l$ (1)

将$\{\boldsymbol{x}_1, \boldsymbol{x}_2,\cdots, \boldsymbol{x}_l\}$扩充为$\mathbf{R}^n$的基$\{\boldsymbol{x}_1,\cdots, \boldsymbol{x}_l, \boldsymbol{x}_{l+1},\cdots, \boldsymbol{x}_n\}$, $\boldsymbol{x}_{l+1},\cdots, \boldsymbol{x}_n$一般不是特征向量, 但 $\boldsymbol{A}\boldsymbol{x}_j\in\mathbf{R}^n$ $(j= l+1,\dots,n)$, 可用$\mathbf{R}^n$的这组基表示:

$$\boldsymbol{A}\boldsymbol{x}_j=b_{1j}\boldsymbol{x}_1+\cdots+b_{lj}\boldsymbol{x}_l+b_{l+1,j}\boldsymbol{x}_{l+1}+\cdots+b_{nj}\boldsymbol{x}_n ,$$

$$j= l+1,\dots,n \qquad (2)$$

将(1)、(2)式中的n个等式写成一个矩阵等式:

$$A[\boldsymbol{x}_1,\cdots \boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n]$$

$$=[\boldsymbol{x}_1,\cdots \boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n]\left[\begin{array}{ccc|ccc}\lambda_0 & \cdots & 0 & b_{1,l+1} & \cdots & b_{1n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & \cdots & \lambda_0 & b_{l,l+1} & \cdots & b_{\ln}\\ \hline 0 & \cdots & 0 & b_{l+1,l+1} & \cdots & b_{l+1,n}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & \cdots & 0 & b_{n,l+1} & \cdots & b_{nn}\end{array}\right] \quad (3)$$

其中λ_0有 l 个.

记 $\boldsymbol{P}$=[$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_l,\boldsymbol{x}_{l+1},\cdots,\boldsymbol{x}_n$], (3)式为:

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\lambda_0\boldsymbol{I}_l & \boldsymbol{B}_1\\ \boldsymbol{0} & \boldsymbol{B}_2\end{bmatrix}$$

$$P^{-1}AP=\begin{bmatrix}\lambda_0 I_l & B_1\\ 0 & B_2\end{bmatrix}$$

因为

$$|A-\lambda I|=|P^{-1}|\cdot|A-\lambda I|\cdot|P|=|P^{-1}(A-\lambda I)P|$$
$$=|P^{-1}AP-\lambda I|$$
$$=\begin{vmatrix}(\lambda_0-\lambda)I_l & -B_1\\ 0 & B_2-\lambda I_{n-l}\end{vmatrix}=(\lambda_0-\lambda)^l|B_2-\lambda I_{n-l}|.$$

由于$|B_2-\lambda I_{n-l}|$是 λ 的 $n-l$ 次多项式,

所以 , λ_0 是A的大于或等于l 重的特征值,

因此 $k\geq l$.

已知:λ_0是 n 阶矩阵 A的 k 重特征值

推论: n 阶方阵的线性无关的特征向量的个数不会超过n.

II. Diagonalization – Examples

Example 5 (Projection matrix)

Let $\boldsymbol{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$.

Then the eigenvalue matrix of $\boldsymbol{A}$ is $\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

The eigenvectors go into the columns of $\boldsymbol{S} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

And

$$\boldsymbol{AS} = \boldsymbol{S\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Therefore $\boldsymbol{S}^{-1}\boldsymbol{AS} = \boldsymbol{\Lambda}$.

Example 6 (Rotation matrix)

Let $\boldsymbol{K}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$. (90° rotation)

Then the characteristic polynomial is $|\boldsymbol{K}-\lambda\boldsymbol{I}|=\lambda^2+1$.

It has two roots—but those roots are *not real*.

The eigenvalues of $\boldsymbol{K}$ are *imaginary numbers*, $\lambda_1=i$ and $\lambda_2=-i$.

The eigenvectors are also not real.

$$(\boldsymbol{K}-\lambda_1\boldsymbol{I})\boldsymbol{x}_1=\begin{bmatrix}-i & -1\\ 1 & -i\end{bmatrix}\begin{bmatrix}y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix},\quad \text{and } \boldsymbol{x}_1=\begin{bmatrix}1\\ -i\end{bmatrix}.$$

$$(\boldsymbol{K}-\lambda_2\boldsymbol{I})\boldsymbol{x}_2=\begin{bmatrix}i & -1\\ 1 & i\end{bmatrix}\begin{bmatrix}y\\ z\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix},\quad \text{and } \boldsymbol{x}_2=\begin{bmatrix}1\\ i\end{bmatrix}.$$

The eigenvalues are *distinct*, even if imaginary, and the eigenvectors are *independent*. They go into the columns of $\boldsymbol{S}$:

$$\boldsymbol{S}=\begin{bmatrix}1 & 1\\ -i & i\end{bmatrix}\quad \text{and}\quad \boldsymbol{S}^{-1}\boldsymbol{K}\boldsymbol{S}=\begin{bmatrix}i & 0\\ 0 & -i\end{bmatrix}.$$

Remark: ***complex numbers may be needed even for real matrices***.

III. Diagonalization – Powers and Products

The eigenvalue of $\boldsymbol{A}^2$ are exactly $\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2$, and every eigenvector of $\boldsymbol{A}$ is also an eigenvector of $\boldsymbol{A}^2$.

Corollary 1 If $\boldsymbol{A}$ is diagonalizable, then $\boldsymbol{A}^k$ is diagonalizable, and has same diagonalizing matrix.

This is true because when $\boldsymbol{S}$ diagonalizes $\boldsymbol{A}$, it also diagonalizes $\boldsymbol{A}^k$.

$$\boldsymbol{\Lambda}^k = (\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S})(\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}) \dots (\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}) = \boldsymbol{S}^{-1}\boldsymbol{A}^k\boldsymbol{S}.$$

Each $\boldsymbol{S}^{-1}$ cancels an $\boldsymbol{S}$, except for the first $\boldsymbol{S}^{-1}$ and the last $\boldsymbol{S}$.

If $\boldsymbol{A}$ is invertible this rule also applies to its inverse (the power $k = -1$).

Example 7 If $\boldsymbol{K}$ is rotation through 90°, then $\boldsymbol{K}^2$ is rotation through 180° (which means $-\boldsymbol{I}$) and $\boldsymbol{K}^{-1}$ is rotation through -90°:

$$\boldsymbol{K} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

$$\boldsymbol{K}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \text{ and } \boldsymbol{K}^{-1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

The eigenvalues of $\boldsymbol{K}$ are i and $-i$; their squares are -1 and -1; their reciprocals are $\frac{1}{i} = -i$ and $\frac{1}{-i} = i$.

Then $\boldsymbol{K}^4$ is a complete rotation through 360°:

$$\boldsymbol{K}^4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and also $\boldsymbol{\Lambda}^4 = \begin{bmatrix} i^4 & 0 \\ 0 & (-i)^4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$

Question: If λ is an eigenvalue of $\boldsymbol{A}$ and μ is an eigenvalue of $\boldsymbol{B}$, then $\boldsymbol{AB}$ has the eigenvalue $\lambda\mu$??
$\boldsymbol{A}+\boldsymbol{B}$ has the eigenvalue $\lambda+\mu$??

Usually not.

False proof $\quad \boldsymbol{ABx} = \boldsymbol{A}\mu\boldsymbol{x} = \mu\boldsymbol{Ax} = \mu\lambda\boldsymbol{x}.$

The mistake lies in assuming that $\boldsymbol{A}$ and $\boldsymbol{B}$ share the *same* eigenvector $\boldsymbol{x}$.

Counterexample (反例):

We could have two matrices with zero eigenvalues, while $\boldsymbol{AB}$ has $\lambda = 1$:

$$\boldsymbol{AB} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The eigenvectors of this $\boldsymbol{A}$ and $\boldsymbol{B}$ are completely different, which is typical.

For the same reason, the eigenvalues of $\boldsymbol{A}+\boldsymbol{B}$ generally have nothing to do with $\lambda+\mu$.

If the eigenvector is the same for $\boldsymbol{A}$ and $\boldsymbol{B}$, then $\boldsymbol{AB}$ has the eigenvalue $\lambda\mu$.

And finally, we have a nice result for product of matrices.

Theorem 4 Let $\boldsymbol{A}, \boldsymbol{B}$ be two diagonalizable matrices of degree n.

Then they have same eigenvectors if and only if $\boldsymbol{AB} = \boldsymbol{BA}$.

Proof. Suppose first that a matrix $\boldsymbol{S}$ diagonalizes both $\boldsymbol{A}, \boldsymbol{B}$, i.e., $\boldsymbol{S}^{-1}\boldsymbol{AS} = \boldsymbol{D}_1$ and $\boldsymbol{S}^{-1}\boldsymbol{BS} = \boldsymbol{D}_2$ are two diagonal matrices. Then

$$\boldsymbol{AB} = (\boldsymbol{SD}_1\boldsymbol{S}^{-1})(\boldsymbol{SD}_2\boldsymbol{S}^{-1}) = \boldsymbol{SD}_1\boldsymbol{D}_2\boldsymbol{S}^{-1},$$
$$\boldsymbol{BA} = (\boldsymbol{SD}_2\boldsymbol{S}^{-1})(\boldsymbol{SD}_1\boldsymbol{S}^{-1}) = \boldsymbol{SD}_2\boldsymbol{D}_1\boldsymbol{S}^{-1}.$$

Since $\boldsymbol{D}_1\boldsymbol{D}_2 = \boldsymbol{D}_2\boldsymbol{D}_1$ (diagonal matrices always commute), we have $\boldsymbol{AB} = \boldsymbol{BA}$. (*see next slide*)

Theorem 4 Let $\boldsymbol{A}, \boldsymbol{B}$ be two diagonalizable matrices of degree n. Then they have same eigenvectors if and only if $\boldsymbol{AB} = \boldsymbol{BA}$.

Proof. (*continued*)

Conversely, assume that $\boldsymbol{AB} = \boldsymbol{BA}$. Suppose that $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. Then

$$\boldsymbol{ABx} = \boldsymbol{BAx} = \boldsymbol{B}\lambda\boldsymbol{x} = \lambda\boldsymbol{Bx}.$$

Thus $\boldsymbol{Bx}$ is also an eigenvector of $\boldsymbol{A}$ corresponding to the same eigenvalue λ.

We only complete the proof for the simpler case where all eigenvalues of $\boldsymbol{A}$ are distinct.

Then the eigenspaces are all of dimension 1, so $\boldsymbol{Bx}$ must be a multiple of $\boldsymbol{x}$, i.e., $\boldsymbol{Bx} = \mu\boldsymbol{x}$, and $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{B}$, as claimed.

(The proof with repeated eigenvalues is a little longer.)

例 设三阶方阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1=1,\lambda_2=3,\lambda_3=4,$ 对应的特征向量是 $\xi_1=(1,1,0)^{\mathrm{T}},\xi_2=(-1,0,1)^{\mathrm{T}},$ $\xi_3=(1,1,2)^{\mathrm{T}}$，求 $\boldsymbol{A}$, $\boldsymbol{A}^{-1}$.

解 作矩阵

$$\boldsymbol{P}=[\xi_1\ \xi_2\ \xi_3]=\begin{bmatrix}1 & -1 & 1\\ 1 & 0 & 1\\ 0 & 1 & 2\end{bmatrix}.$$

因为 $\boldsymbol{P}$ 为可逆矩阵，且

$$\boldsymbol{P}^{-1}=\begin{bmatrix}-1/2 & 3/2 & -1/2\\ -1 & 1 & 0\\ 1/2 & -1/2 & 1/2\end{bmatrix}.$$

所以

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}1 & & \\ & 3 & \\ & & 4\end{bmatrix},$$

$$\boldsymbol{A}=\boldsymbol{P}\begin{bmatrix}1 & & \\ & 3 & \\ & & 4\end{bmatrix}\boldsymbol{P}^{-1}=\frac{1}{2}\begin{bmatrix}9 & -7 & 3\\ 3 & -1 & 3\\ 2 & -2 & 8\end{bmatrix},$$

$$\boldsymbol{A}^{-1}=\boldsymbol{P}\begin{bmatrix}1 & & \\ & 1/3 & \\ & & 1/4\end{bmatrix}\boldsymbol{P}^{-1}=\frac{1}{24}\begin{bmatrix}-1 & 25 & -9\\ -9 & 33 & -9\\ -2 & 2 & 6\end{bmatrix}.$$

Key words:

Conditions

Examples

Powers and Products

Homework

See Blackboard