

特征值与特征向量 矩阵的对角化



张晓平

武汉大学数学与统计学院



2017 年 12 月 8 日

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

定义 1 (特征值与特征向量) 设 A 为复数域 \mathbb{C} 上的 n 阶矩阵, 如果存在数 $\lambda \in \mathbb{C}$ 和非零的 n 维向量 x 使得

$$Ax = \lambda x$$

则称 λ 为矩阵 A 的特征值, x 为 A 的对应于特征值 λ 的特征向量。

定义 1 (特征值与特征向量) 设 A 为复数域 \mathbb{C} 上的 n 阶矩阵, 如果存在数 $\lambda \in \mathbb{C}$ 和非零的 n 维向量 x 使得

$$Ax = \lambda x$$

则称 λ 为矩阵 A 的特征值, x 为 A 的对应于特征值 λ 的特征向量。

- (1) 特征向量 $x \neq 0$;
- (2) 特征值问题是对方针而言的。

由定义, n 阶矩阵 A 的特征值, 就是使齐次线性方程组

$$(\lambda I - A)x = 0$$

有非零解的 λ 值, 即满足方程

$$\det(\lambda I - A) = 0$$

的 λ 都是矩阵 A 的特征值。

由定义, n 阶矩阵 A 的特征值, 就是使齐次线性方程组

$$(\lambda I - A)x = 0$$

有非零解的 λ 值, 即满足方程

$$\det(\lambda I - A) = 0$$

的 λ 都是矩阵 A 的特征值。

结论 1 特征值 λ 是关于 λ 的多项式 $\det(\lambda I - A)$ 的根。

定义 2 (特征多项式、特征矩阵、特征方程) 设 n 阶矩阵 $A = (a_{ij})$, 则

$$f(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为矩阵 A 的特征多项式, $\lambda I - A$ 称为 A 的特征矩阵, $\det(\lambda I - A) = 0$ 称为 A 的特征方程。

定义 2 (特征多项式、特征矩阵、特征方程) 设 n 阶矩阵 $A = (a_{ij})$, 则

$$f(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为矩阵 A 的特征多项式, $\lambda I - A$ 称为 A 的特征矩阵, $\det(\lambda I - A) = 0$ 称为 A 的特征方程。

(1) n 阶矩阵 A 的特征多项式是 λ 的 n 次多项式。

定义 2 (特征多项式、特征矩阵、特征方程) 设 n 阶矩阵 $A = (a_{ij})$, 则

$$f(\lambda) = \det(\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

称为矩阵 A 的特征多项式, $\lambda I - A$ 称为 A 的特征矩阵, $\det(\lambda I - A) = 0$ 称为 A 的特征方程。

- (1) n 阶矩阵 A 的特征多项式是 λ 的 n 次多项式。
- (2) 特征多项式的 k 重根称为 k 重特征值。

例 2.1 求矩阵

$$A = \begin{pmatrix} 5 & -1 & -1 \\ 3 & 1 & -1 \\ 4 & -2 & 1 \end{pmatrix}$$

的特征值与特征向量。

例 2.1 求矩阵

$$A = \begin{pmatrix} 5 & -1 & -1 \\ 3 & 1 & -1 \\ 4 & -2 & 1 \end{pmatrix}$$

的特征值与特征向量。

解

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 5-\lambda & -1 & -1 \\ 3 & 1-\lambda & -1 \\ 4 & -2 & 1-\lambda \end{vmatrix} = (3-\lambda) \begin{vmatrix} 1 & -1 & -1 \\ 1 & 1-\lambda & -1 \\ 1 & -2 & 1-\lambda \end{vmatrix} \\ &= (3-\lambda) \begin{vmatrix} 1 & -1 & -1 \\ 0 & 2-\lambda & 0 \\ 0 & -1 & 2-\lambda \end{vmatrix} = -(\lambda-3)(\lambda-2)^2 = 0 \end{aligned}$$

故 A 的特征值为 $\lambda_{1,2} = 2$, $\lambda_3 = 3$ 。

解 (续) 当 $\lambda_{1,2} = 2$ 时, 由 $(A - \lambda I)\mathbf{x} = \mathbf{0}$, 即

$$\begin{pmatrix} 3 & -1 & -1 \\ 3 & -1 & -1 \\ 4 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

得其基础解系为 $\mathbf{x}_1 = (1, 1, 2)^T$, 因此 $k_1 \mathbf{x}_1$ (k_1 为非零任意常数) 是 A 对应于 $\lambda_{1,2} = 2$ 的全部特征向量。

解(续) 当 $\lambda_{1,2}=2$ 时, 由 $(A-\lambda I)\mathbf{x}=\mathbf{0}$, 即

$$\begin{pmatrix} 3 & -1 & -1 \\ 3 & -1 & -1 \\ 4 & -2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

得其基础解系为 $\mathbf{x}_1 = (1, 1, 2)^T$, 因此 $k_1\mathbf{x}_1$ (k_1 为非零任意常数) 是 A 对应于 $\lambda_{1,2}=2$ 的全部特征向量。

当 $\lambda_3=3$ 时, 由 $(A-\lambda I)\mathbf{x}=\mathbf{0}$, 即

$$\begin{pmatrix} 2 & -1 & -1 \\ 3 & -2 & -1 \\ 4 & -2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

得其基础解系为 $\mathbf{x}_2 = (1, 1, 1)^T$, 因此 $k_2\mathbf{x}_2$ (k_2 为非零任意常数) 是 A 对应于 $\lambda_3=3$ 的全部特征向量。

例 2.2

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

的特征多项式为

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$$

故其 n 个特征值为 n 个对角元。

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

定理 1 若 x_1 和 x_2 都是 A 的对应于特征值 λ_0 的特征向量, 则 $k_1x_1 + k_2x_2$ 也是 A 的对应于特征值 λ_0 的特征向量 (其中 k_1, k_2 为任意常数, 但 $k_1x_1 + k_2x_2 \neq 0$)。

定理 1 若 x_1 和 x_2 都是 A 的对应于特征值 λ_0 的特征向量, 则 $k_1x_1 + k_2x_2$ 也是 A 的对应于特征值 λ_0 的特征向量 (其中 k_1, k_2 为任意常数, 但 $k_1x_1 + k_2x_2 \neq 0$)。

证明. 由于 x_1, x_2 是齐次线性方程组

$$(\lambda_0 I - A)x = 0$$

的解, 因此 $k_1x_1 + k_2x_2$ 也是上式的解, 故当 $k_1x_1 + k_2x_2 \neq 0$ 时, 它是 A 属于 λ_0 的特征向量。□

在

$$(\lambda I - A)x = 0$$

的解空间中，除零向量外的全体解向量就是 A 的属于特征值 λ 的全体特征向量。因此， $(\lambda I - A)x = 0$ 的解空间也称为矩阵 A 关于特征值 λ 的**特征子空间**，记作 V_λ 。

在

$$(\lambda I - A)x = 0$$

的解空间中，除零向量外的全体解向量就是 A 的属于特征值 λ 的全体特征向量。因此， $(\lambda I - A)x = 0$ 的解空间也称为矩阵 A 关于特征值 λ 的**特征子空间**，记作 V_λ 。

n 阶矩阵 A 的特征子空间是 n 维向量空间的子空间，其维数为

$$\dim V_\lambda = n - r(\lambda I - A).$$

定理 2 设 n 阶矩阵 $A = (a_{ij})$ 的 n 个特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则

$$(1) \sum_{i=1}^n \lambda_i = \text{tr}(A)$$

$$(2) \prod_{i=1}^n \lambda_i = \det(A)$$

其中 $\text{tr}(A) = \sum_{i=1}^n a_{ii}$ 为 A 的迹。

定理 2 设 n 阶矩阵 $A = (a_{ij})$ 的 n 个特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则

$$(1) \sum_{i=1}^n \lambda_i = \text{tr}(A)$$

$$(2) \prod_{i=1}^n \lambda_i = \det(A)$$

其中 $\text{tr}(A) = \sum_{i=1}^n a_{ii}$ 为 A 的迹。

- ▶ 当 $\det(A) \neq 0$, 即 A 为可逆矩阵时, 其特征值全为非零数;
- ▶ 奇异矩阵 A 至少有一个零特征值。

定理 3 一个特征向量不能属于不同的特征值。

定理 3 一个特征向量不能属于不同的特征值。

证明. 若 x 是 A 的属于特征值 $\lambda_1, \lambda_2 (\lambda_1 \neq \lambda_2)$ 的特征向量, 即有

$$Ax = \lambda_1 x, \quad Ax = \lambda_2 x \Rightarrow (\lambda_1 - \lambda_2)x = 0 \Rightarrow x = 0$$

这与 $x \neq 0$ 矛盾。



性质 2.1

表: 特征值与特征向量

	特征值	特征向量
A	λ	x
kA	$k\lambda$	x
A^m	λ^m	x
A^{-1}	λ^{-1}	x

性质 2.2 矩阵 A 与 A^T 的特征值相同。

性质 2.2 矩阵 A 与 A^T 的特征值相同。

证明. 因 $(\lambda I - A)^T = \lambda I - A^T$, 故

$$\det(\lambda I - A) = \det(\lambda I - A^T)$$

因此, A 与 A^T 有完全相同的特征值。



定理 4 设 $A = (a_{ij})$ 是 n 阶矩阵, 若

$$(1) \sum_{j=1}^n |a_{ij}| < 1, \quad (i = 1, 2, \dots, n)$$

$$(2) \sum_{i=1}^n |a_{ij}| < 1, \quad (j = 1, 2, \dots, n)$$

有一个成立, 则 A 的所有特征值的模都小于 1。

定理 4 设 $A = (a_{ij})$ 是 n 阶矩阵, 若

$$(1) \sum_{j=1}^n |a_{ij}| < 1, \quad (i = 1, 2, \dots, n)$$

$$(2) \sum_{i=1}^n |a_{ij}| < 1, \quad (j = 1, 2, \dots, n)$$

有一个成立, 则 A 的所有特征值的模都小于 1。

证明. 设 λ 为 A 的任一特征值, x 为 λ 对应的特征向量, 由 $Ax = \lambda x$ 可得

$$\sum_{j=1}^n a_{ij} x_j = \lambda x_i, \quad i = 1, 2, \dots, n.$$

记 $x_k = \max_{1 \leq j \leq n} |x_j|$, 则有

$$|\lambda| = \left| \frac{\lambda x_k}{x_k} \right| = \left| \frac{\sum_{j=1}^n a_{kj} x_j}{x_k} \right| \leq \sum_{j=1}^n |a_{kj}| \left| \frac{x_j}{x_k} \right| \leq \sum_{j=1}^n |a_{kj}|.$$

由此可知, 若条件 (1) 成立, 则 $|\lambda| < 1$; 再由 λ 的任意性即得 $|\lambda_k| < 1, \quad k = 1, 2, \dots, n$ 。

定理 4 设 $A = (a_{ij})$ 是 n 阶矩阵, 若

$$(1) \sum_{j=1}^n |a_{ij}| < 1, \quad (i = 1, 2, \dots, n)$$

$$(2) \sum_{i=1}^n |a_{ij}| < 1, \quad (j = 1, 2, \dots, n)$$

有一个成立, 则 A 的所有特征值的模都小于 1。

证明. 设 λ 为 A 的任一特征值, x 为 λ 对应的特征向量, 由 $Ax = \lambda x$ 可得

$$\sum_{j=1}^n a_{ij} x_j = \lambda x_i, \quad i = 1, 2, \dots, n.$$

记 $x_k = \max_{1 \leq j \leq n} |x_j|$, 则有

$$|\lambda| = \left| \frac{\lambda x_k}{x_k} \right| = \left| \frac{\sum_{j=1}^n a_{kj} x_j}{x_k} \right| \leq \sum_{j=1}^n |a_{kj}| \left| \frac{x_j}{x_k} \right| \leq \sum_{j=1}^n |a_{kj}|.$$

由此可知, 若条件 (1) 成立, 则 $|\lambda| < 1$; 再由 λ 的任意性即得 $|\lambda_k| < 1, \quad k = 1, 2, \dots, n$ 。

同理, 若条件 (2) 成立, 则 A^T 的所有特征值, 亦即 A 的所有特征值的模小于 1。

□

例 2.3 设 $A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{pmatrix}$

(i) 求 A 的特征值与特征向量

(ii) 求可逆矩阵 P , 使得 $P^{-1}AP$ 为对角阵。

解

(i) 由

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -1 & 1 \\ 2 & -2-\lambda & 2 \\ -1 & 1 & -1-\lambda \end{vmatrix} = -\lambda^2(\lambda + 2)$$

知 A 的特征值为 $\lambda_1 = \lambda_2 = 0$ 和 $\lambda_3 = -2$ 。

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

定义 3 对于矩阵 A 和 B ，若存在可逆阵 P ，使得 $P^{-1}AP = B$ ，就称 A 相似于 B ，记为 $A \sim B$ 。

定义 3 对于矩阵 A 和 B ，若存在可逆阵 P ，使得 $P^{-1}AP = B$ ，就称 A 相似于 B ，记为 $A \sim B$ 。

矩阵的相似关系也是一种等价关系，即满足以下三条性质

- ▶ 反身性： $A \sim A$
- ▶ 对称性： $A \sim B \Rightarrow B \sim A$
- ▶ 传递性： $A \sim B, B \sim C \Rightarrow A \sim C$

相似矩阵的性质

性质 3.1

$$\boldsymbol{P}^{-1}(k_1 \boldsymbol{A}_1 + k_2 \boldsymbol{A}_2) \boldsymbol{P} = k_1 \boldsymbol{P}^{-1} \boldsymbol{A}_1 \boldsymbol{P} + k_2 \boldsymbol{P}^{-1} \boldsymbol{A}_2 \boldsymbol{P}.$$

性质 3.2

$$\boldsymbol{P}^{-1} \boldsymbol{A}_1 \boldsymbol{A}_2 \boldsymbol{P} = (\boldsymbol{P}^{-1} \boldsymbol{A}_1 \boldsymbol{P})(\boldsymbol{P}^{-1} \boldsymbol{A}_2 \boldsymbol{P})$$

相似矩阵的性质

性质 3.3

$$A \sim B \Rightarrow A^m \sim B^m (m \in \mathbb{Z}^+).$$

相似矩阵的性质

性质 3.3

$$A \sim B \Rightarrow A^m \sim B^m (m \in \mathbb{Z}^+).$$

证明. 因 $A \sim B$, 故存在可逆阵 P 使得

$$P^{-1}AP = B$$

于是

$$B^m = (P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP) = P^{-1}A^mP$$

故 $A^m \sim B^m$ 。



相似矩阵的性质

性质 3.4

$$A \sim B \Rightarrow f(A) \sim f(B),$$

其中

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0.$$

相似矩阵的性质

定理 5 相似矩阵的特征值相同。

相似矩阵的性质

定理 5 相似矩阵的特征值相同。

证明 只需证明相似矩阵有相同的特征多项式。设 $A \sim B$ ，则存在可逆阵 P ，使得

$$P^{-1}AP = B$$

于是

$$\begin{aligned} |\lambda I - B| &= |\lambda I - P^{-1}AP| \\ &= |P^{-1}(\lambda I - A)P| = |P^{-1}| |\lambda I - A| |P| \\ &= |\lambda I - A| \end{aligned}$$



注 1 上述定理的逆命题不成立。例如,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

都以 1 为二重特征值, 但对于任何可逆阵 P , 都有

$$P^{-1}IP = I \neq A,$$

故 A 与 I 不相似。

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

矩阵可对角化，即矩阵与对角阵相似。

定理 6 矩阵可对角化 \iff n 阶矩阵有 n 个线性无关的特征向量

定理 6 矩阵可对角化 $\iff n$ 阶矩阵有 n 个线性无关的特征向量
证明.

\Rightarrow

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

定理 6 矩阵可对角化 $\iff n$ 阶矩阵有 n 个线性无关的特征向量
证明.

\Rightarrow

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

将 P 按列分块, 即

$$P = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

则

$$A(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

定理 6 矩阵可对角化 $\iff n$ 阶矩阵有 n 个线性无关的特征向量
证明.

\Rightarrow

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

将 P 按列分块, 即

$$P = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

则

$$A(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

于是

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (i = 1, 2, \dots, n).$$

定理 6 矩阵可对角化 $\iff n$ 阶矩阵有 n 个线性无关的特征向量
证明.

\Rightarrow

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

将 P 按列分块, 即

$$P = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

则

$$A(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

于是

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (i = 1, 2, \dots, n).$$

故 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 是 A 分别对应于 $\lambda_1, \lambda_2, \dots, \lambda_n$ 的特征向量。

定理 6 矩阵可对角化 $\iff n$ 阶矩阵有 n 个线性无关的特征向量
证明.

\Rightarrow

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

将 P 按列分块, 即

$$P = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n),$$

则

$$A(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

于是

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i \quad (i = 1, 2, \dots, n).$$

故 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 是 A 分别对应于 $\lambda_1, \lambda_2, \dots, \lambda_n$ 的特征向量。由于 P 可逆, 所以它们是线性无关的。

□

若 A 与 Λ 相似, 则 Λ 的主对角元都是 A 的特征值。若不计 λ_k 的排列次序, 则 Λ 是唯一的, 称 Λ 为 A 的相似标准型。

定理 7 A 的属于不同特征值的特征向量是线性无关的。

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^0 当 $m = 1$ 时, 结论显然成立。

2^0 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^o 当 $m = 1$ 时, 结论显然成立。

2^o 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^0 当 $m=1$ 时, 结论显然成立。

2^0 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$\begin{aligned} & a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1) \\ \Rightarrow & A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0} \end{aligned}$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1⁰ 当 $m=1$ 时, 结论显然成立。

2⁰ 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^0 当 $m=1$ 时, 结论显然成立。

2^0 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xRightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^o 当 $m=1$ 时, 结论显然成立。

2^o 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xrightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

$$\Rightarrow a_i (\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \dots, k$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1⁰ 当 $m=1$ 时, 结论显然成立。

2⁰ 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xRightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

$$\Rightarrow a_i (\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_i = 0, \quad i = 1, 2, \dots, k$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^0 当 $m=1$ 时, 结论显然成立。

2^0 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xRightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

$$\Rightarrow a_i (\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_i = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_{k+1} \mathbf{x}_{k+1} = \mathbf{0}$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^o 当 $m=1$ 时, 结论显然成立。

2^o 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ 。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xRightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

$$\Rightarrow a_i (\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_i = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_{k+1} \mathbf{x}_{k+1} = \mathbf{0}$$

$$\Rightarrow a_{k+1} = 0$$

定理 7 A 的属于不同特征值的特征向量是线性无关的。

证明. 设 A 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_m$, 其相应的特征向量为 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$. 对 m 做数学归纳法。

1^o 当 $m=1$ 时, 结论显然成立。

2^o 设 k 个不同特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 的特征向量 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$. 下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (1)$$

$$\Rightarrow A(a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 + \dots + a_k \mathbf{x}_k + a_{k+1} \mathbf{x}_{k+1}) = \mathbf{0}$$

$$\Rightarrow a_1 \lambda_1 \mathbf{x}_1 + a_2 \lambda_2 \mathbf{x}_2 + \dots + a_k \lambda_k \mathbf{x}_k + a_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0} \quad (2)$$

$$\xRightarrow{(2)-\lambda_{k+1}(1)} a_1 (\lambda_{k+1} - \lambda_1) \mathbf{x}_1 + a_2 (\lambda_{k+1} - \lambda_2) \mathbf{x}_2 + \dots + a_k (\lambda_{k+1} - \lambda_k) \mathbf{x}_k = \mathbf{0}$$

$$\Rightarrow a_i (\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_i = 0, \quad i = 1, 2, \dots, k$$

$$\Rightarrow a_{k+1} \mathbf{x}_{k+1} = \mathbf{0}$$

$$\Rightarrow a_{k+1} = 0$$

$$\Rightarrow \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k, \mathbf{x}_{k+1} \text{ 线性无关}$$



推论 1 若 A 有 n 个互不相同的特征值, 则 A 与对角阵相似。

例 3.1 设实对称矩阵

$$A = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$$

问 A 是否可对角化？若可对角化，求对角阵 Λ 及可逆矩阵 P 使得 $P^{-1}AP = \Lambda$ ，再求 A^k 。

例 3.2 设 $A = (a_{ij})_{n \times n}$ 是主对角元全为 2 的上三角矩阵, 且存在 $a_{ij} \neq 0 (i < j)$, 问 A 是否可对角化?

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

定义 4 元素为复数的矩阵和向量，称为复矩阵和复向量。

定义 5 设 a_{ij} 为复数， $A = (a_{ij})_{m \times n}$ ， $\bar{A} = (\bar{a}_{ij})_{m \times n}$ ， \bar{a}_{ij} 是 a_{ij} 的共轭复数，则称 \bar{A} 是 A 的共轭矩阵。

- ▶ $\overline{\bar{A}} = A$
- ▶ $\bar{A}^T = \overline{A^T}$
- ▶ 当 A 为实对称矩阵时， $\bar{A}^T = A^T$

- ▶ $\overline{kA} = \overline{k} \overline{A}$
- ▶ $\overline{A+B} = \overline{A} + \overline{B}$
- ▶ $\overline{AB} = \overline{A} \overline{B}$
- ▶ $\overline{(AB)^T} = \overline{B}^T \overline{A}^T$
- ▶ $\overline{A^{-1}} = (\overline{A})^{-1}$
- ▶ $\det \overline{A} = \overline{\det A}$

1. 矩阵的特征值与特征向量 相似矩阵

- ▶ 特征值与特征向量
- ▶ 特征值与特征值的性质
- ▶ 相似矩阵及其性质

2. 矩阵可对角化的条件

3. 实对称矩阵的对角化

- ▶ 实对称矩阵的特征值和特征向量

定理 8 实对称矩阵 A 的任一特征值都是实数。

定理 8 实对称矩阵 A 的任一特征值都是实数。

证明.

$$\begin{aligned} Ax = \lambda x &\implies \overline{(Ax)}^T = \overline{\lambda x}^T \\ &\implies \bar{x}^T \bar{A}^T x = \bar{\lambda} \bar{x}^T x \\ &\implies \bar{x}^T A^T x = \bar{\lambda} \bar{x}^T x \\ &\implies \lambda \bar{x}^T x = \bar{\lambda} \bar{x}^T x \\ &\implies \lambda = \bar{\lambda} \end{aligned}$$



定理 9 实对称矩阵 A 对应于不同特征值的特征向量是正交的。

定理 9 实对称矩阵 A 对应于不同特征值的特征向量是正交的。

证明. 设 $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$ ($\lambda_1 \neq \lambda_2$), $A^T = A$, 则

$$\lambda_1 \mathbf{x}_2^T \mathbf{x}_1 = \mathbf{x}_2^T A \mathbf{x}_1 = \mathbf{x}_2^T A^T \mathbf{x}_1 = (A \mathbf{x}_2)^T \mathbf{x}_1 = (\lambda_2 \mathbf{x}_2)^T \mathbf{x}_1 = \lambda_2 \mathbf{x}_2^T \mathbf{x}_1$$

定理 9 实对称矩阵 A 对应于不同特征值的特征向量是正交的。

证明. 设 $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$ ($\lambda_1 \neq \lambda_2$), $A^T = A$, 则

$$\lambda_1 \mathbf{x}_2^T \mathbf{x}_1 = \mathbf{x}_2^T A \mathbf{x}_1 = \mathbf{x}_2^T A^T \mathbf{x}_1 = (A \mathbf{x}_2)^T \mathbf{x}_1 = (\lambda_2 \mathbf{x}_2)^T \mathbf{x}_1 = \lambda_2 \mathbf{x}_2^T \mathbf{x}_1$$

由于 $\lambda_1 \neq \lambda_2$, 所以

$$\mathbf{x}_2^T \mathbf{x}_1 = 0.$$

