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 12 日

目录

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
    - 实对称矩阵的对角化

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
   - 实对称矩阵的对角化

定义 1 (特征值与特征向量) 设 $\boldsymbol{A}$ 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 $\boldsymbol{x}$ 使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

则称 λ 为矩阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 为 $\boldsymbol{A}$ 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 $\boldsymbol{A}$ 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 $\boldsymbol{x}$ 使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

则称 λ 为矩阵 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的对应于特征值 λ 的特征向量。

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

由定义，n 阶矩阵 $\boldsymbol{A}$ 的特征值，就是使齐次线性方程组

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$$

有非零解的 λ 值，即满足方程

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$$

的 λ 都是矩阵 $\boldsymbol{A}$ 的特征值。

由定义，n 阶矩阵 $\boldsymbol{A}$ 的特征值，就是使齐次线性方程组

$$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$$

有非零解的 λ 值，即满足方程

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=0$$

的 λ 都是矩阵 $\boldsymbol{A}$ 的特征值。

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

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

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

称为矩阵 $\boldsymbol{A}$ 的特征多项式，$\lambda \boldsymbol{I}-\boldsymbol{A}$ 称为 $\boldsymbol{A}$ 的特征矩阵，$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=0$ 称为 $\boldsymbol{A}$ 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 $\boldsymbol{A}=(a_{ij})$，则

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

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

(1) n 阶矩阵 $\boldsymbol{A}$ 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 $\boldsymbol{A}=(a_{ij})$，则

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

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

(1) n 阶矩阵 $\boldsymbol{A}$ 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

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

$$= (3-\lambda)\begin{vmatrix} 1 & -1 & -1 \\ 0 & 2-\lambda & 0 \\ 0 & -1 & 2-\lambda \end{vmatrix} = -(\lambda-3)(\lambda-2)^2 = 0$$

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

解 (续) 当 $\lambda_{1,2}=2$ 时，由 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，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 $\boldsymbol{x}_1=(1,1,2)^T$，因此 $k_1\boldsymbol{x}_1$（k_1 为非零任意常数）是 $\boldsymbol{A}$ 对应于 $\lambda_{1,2}=2$ 的全部特征向量。

解 (续)　当 $\lambda_{1,2}=2$ 时，由 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，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 $\boldsymbol{x}_1=(1,1,2)^T$，因此 $k_1\boldsymbol{x}_1$（k_1 为非零任意常数）是 $\boldsymbol{A}$ 对应于 $\lambda_{1,2}=2$ 的全部特征向量。

当 $\lambda_3=3$ 时，由 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\boldsymbol{0}$，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 $\boldsymbol{x}_2=(1,1,1)^T$，因此 $k_2\boldsymbol{x}_2$（k_2 为非零任意常数）是 $\boldsymbol{A}$ 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
    - 实对称矩阵的对角化

定理 1 若 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 都是 $\boldsymbol{A}$ 的对应于特征值 λ_0 的特征向量，则 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2$ 也是 $\boldsymbol{A}$ 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 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2\neq 0$）。

定理 1 若 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 都是 $\boldsymbol{A}$ 的对应于特征值 λ_0 的特征向量，则 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2$ 也是 $\boldsymbol{A}$ 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 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2\neq 0$）。

证明. 由于 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 是齐次线性方程组

$$(\lambda_0\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$$

的解，因此 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2$ 也是上式的解，故当 $k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2\neq 0$ 时，它是 $\boldsymbol{A}$ 属于 λ_0 的特征向量。 $\square$

在

$$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$$

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

在

$$(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$$

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

n 阶矩阵 $\boldsymbol{A}$ 的特征子空间是 n 维向量空间的子空间，其维数为

$$\dim\ V_\lambda = n - \mathrm{r}(\lambda \boldsymbol{I} - \boldsymbol{A}).$$

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

(1) $\displaystyle\sum_{i=1}^{n}\lambda_i=\mathrm{tr}(\boldsymbol{A})$

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

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

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

(1) $\displaystyle\sum_{i=1}^{n}\lambda_i=\operatorname{tr}(\boldsymbol{A})$

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

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

- 当 $\det(\boldsymbol{A})\neq 0$，即 $\boldsymbol{A}$ 为可逆矩阵时，其特征值全为非零数；
- 奇异矩阵 $\boldsymbol{A}$ 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 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于特征值 $\lambda_1, \lambda_2 (\lambda_1 \neq \lambda_2)$ 的特征向量，即有

$$\boldsymbol{A}\boldsymbol{x} = \lambda_1 \boldsymbol{x}, \quad \boldsymbol{A}\boldsymbol{x} = \lambda_2 \boldsymbol{x} \quad \Rightarrow \quad (\lambda_1 - \lambda_2)\boldsymbol{x} = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{0}$$

这与 $\boldsymbol{x} \neq \boldsymbol{0}$ 矛盾。 □

性质 2.1

表: 特征值与特征向量

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

性质 2.2 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^T$ 的特征值相同。

性质 2.2 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^T$ 的特征值相同。

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

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

因此，$\boldsymbol{A}$ 与 $\boldsymbol{A}^T$ 有完全相同的特征值。 □

定理 4 设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶矩阵，若

(1) $\sum\limits_{j=1}^{n}|a_{ij}|<1,\quad (i=1,2,\cdots,n)$

(2) $\sum\limits_{i=1}^{n}|a_{ij}|<1,\quad (j=1,2,\cdots,n)$

有一个成立，则 $\boldsymbol{A}$ 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) $\sum_{j=1}^{n} |a_{ij}| < 1, \quad (i=1,2,\cdots,n)$

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

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

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

记 $x_k = \max\limits_{1\le j\le n} |x_j|$，则有

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

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

定理 4 设 $A = (a_{ij})$ 是 n 阶矩阵，若

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

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

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

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

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

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

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

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。

$\square$

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

(i) 求 $\boldsymbol{A}$ 的特征值与特征向量

(ii) 求可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角阵。

解

(i) 由

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

知 $\boldsymbol{A}$ 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
   - 实对称矩阵的对角化

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

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

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：$\boldsymbol{A}\sim\boldsymbol{A}$
- 对称性：$\boldsymbol{A}\sim\boldsymbol{B} \ \Rightarrow\ \boldsymbol{B}\sim\boldsymbol{A}$
- 传递性：$\boldsymbol{A}\sim\boldsymbol{B},\ \boldsymbol{B}\sim\boldsymbol{C} \ \Rightarrow\ \boldsymbol{A}\sim\boldsymbol{C}$

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

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

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 $A^m \sim B^m$。 $\square$

相似矩阵的性质

性质 3.4

$$\boldsymbol{A} \sim \boldsymbol{B} \;\Rightarrow\; f(\boldsymbol{A}) \sim f(\boldsymbol{B}),$$

其中

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a0.$$

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 $\boldsymbol{A} \sim \boldsymbol{B}$，则存在可逆阵 $\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$$

于是

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

□

注 1 上述定理的逆命题不成立。例如,

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

都以 1 为二重特征值,但对于任何可逆阵 $\boldsymbol{P}$,都有

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

故 $\boldsymbol{A}$ 与 $\boldsymbol{I}$ 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
   - 实对称矩阵的对角化

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.

$\Rightarrow$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda} \implies \boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Lambda}$$

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.

$\Rightarrow$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} \implies \boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{\Lambda}$$

将 $\boldsymbol{P}$ 按列分块，即

$$\boldsymbol{P} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n),$$

则

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

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.

$\Rightarrow$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda} \implies \boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Lambda}$$

将 $\boldsymbol{P}$ 按列分块，即

$$\boldsymbol{P}=(\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n),$$

则

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

于是

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

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.

$\Rightarrow$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda} \implies \boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Lambda}$$

将 $\boldsymbol{P}$ 按列分块，即

$$\boldsymbol{P}=(\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n),$$

则

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

于是

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

故 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_n$ 是 $\boldsymbol{A}$ 分别对应于 λ_1，λ_2，$\cdots$，λ_n 的特征向量。

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

证明.

$\Rightarrow$

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda} \quad \Longrightarrow \quad \boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{\Lambda}$$

将 $\boldsymbol{P}$ 按列分块，即

$$\boldsymbol{P}=(\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n),$$

则

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

于是

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

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

□

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

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0} \qquad (1)$$

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0}$$

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad \boldsymbol{A}(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0}$$

$$\Longrightarrow \quad a_1\lambda_1\boldsymbol{x}_1 + a_2\lambda_2\boldsymbol{x}_2 + \cdots + a_k\lambda_k\boldsymbol{x}_k + a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0} \qquad (2)$$

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 $\lambda_1, \lambda_2, \cdots, \lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad A(a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \cdots + a_k\boldsymbol{x}_k + a_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0}$$

$$\Longrightarrow \quad a_1\lambda_1\boldsymbol{x}_1 + a_2\lambda_2\boldsymbol{x}_2 + \cdots + a_k\lambda_k\boldsymbol{x}_k + a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1} = \boldsymbol{0} \qquad (2)$$

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

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 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0}$$

$$\Longrightarrow \quad a_1\lambda_1\boldsymbol{x}_1+a_2\lambda_2\boldsymbol{x}_2+\cdots+a_k\lambda_k\boldsymbol{x}_k+a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (2)$$

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

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

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 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0}$$

$$\Longrightarrow \quad a_1\lambda_1\boldsymbol{x}_1+a_2\lambda_2\boldsymbol{x}_2+\cdots+a_k\lambda_k\boldsymbol{x}_k+a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (2)$$

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

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

$$\Longrightarrow \quad a_i=0, \quad i=1,2,\cdots,k$$

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$
\begin{aligned}
& a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (1)\\
\Longrightarrow\ & \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0}\\
\Longrightarrow\ & a_1\lambda_1\boldsymbol{x}_1+a_2\lambda_2\boldsymbol{x}_2+\cdots+a_k\lambda_k\boldsymbol{x}_k+a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (2)\\
\xRightarrow{(2)-\lambda_{k+1}(1)}\ & a_1(\lambda_{k+1}-\lambda_1)\boldsymbol{x}_1+a_2(\lambda_{k+1}-\lambda_2)\boldsymbol{x}_2+\cdots+a_k(\lambda_{k+1}-\lambda_k)\boldsymbol{x}_k=\boldsymbol{0}\\
\Longrightarrow\ & a_i(\lambda_{k+1}-\lambda_i)=0,\quad i=1,2,\cdots,k\\
\Longrightarrow\ & a_i=0,\quad i=1,2,\cdots,k\\
\Longrightarrow\ & a_{k+1}\boldsymbol{x}_{k+1}=0
\end{aligned}
$$

定理 7 $\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明. 设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (1)$$

$$\Longrightarrow \quad \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0}$$

$$\Longrightarrow \quad a_1\lambda_1\boldsymbol{x}_1+a_2\lambda_2\boldsymbol{x}_2+\cdots+a_k\lambda_k\boldsymbol{x}_k+a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1}=\boldsymbol{0} \qquad (2)$$

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

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

$$\Longrightarrow \quad a_i=0, \quad i=1,2,\cdots,k$$

$$\Longrightarrow \quad a_{k+1}\boldsymbol{x}_{k+1}=0$$

$$\Longrightarrow \quad a_{k+1}=0$$

定理 7　$\boldsymbol{A}$ 的属于不同特征值的特征向量是线性无关的。

证明.　设 $\boldsymbol{A}$ 的 m 个互不相同的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_m$，其相应的特征向量为 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_m$. 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 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 的特征向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，$\cdots$，$\boldsymbol{x}_k$ 线性无关。下面考虑 $k+1$ 个不同特征值的特征向量。设

$$
\begin{aligned}
& a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\mathbf{0} \qquad (1)\\
\Longrightarrow\ & \boldsymbol{A}(a_1\boldsymbol{x}_1+a_2\boldsymbol{x}_2+\cdots+a_k\boldsymbol{x}_k+a_{k+1}\boldsymbol{x}_{k+1}=\mathbf{0}\\
\Longrightarrow\ & a_1\lambda_1\boldsymbol{x}_1+a_2\lambda_2\boldsymbol{x}_2+\cdots+a_k\lambda_k\boldsymbol{x}_k+a_{k+1}\lambda_{k+1}\boldsymbol{x}_{k+1}=\mathbf{0} \qquad (2)\\
\xRightarrow{(2)-\lambda_{k+1}(1)}\ & a_1(\lambda_{k+1}-\lambda_1)\boldsymbol{x}_1+a_2(\lambda_{k+1}-\lambda_2)\boldsymbol{x}_2+\cdots+a_k(\lambda_{k+1}-\lambda_k)\boldsymbol{x}_k=\mathbf{0}\\
\Longrightarrow\ & a_i(\lambda_{k+1}-\lambda_i)=0,\quad i=1,2,\cdots,k\\
\Longrightarrow\ & a_i=0,\quad i=1,2,\cdots,k\\
\Longrightarrow\ & a_{k+1}\boldsymbol{x}_{k+1}=0\\
\Longrightarrow\ & a_{k+1}=0\\
\Longrightarrow\ & \boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_k,\ \boldsymbol{x}_{k+1}\text{线性无关}
\end{aligned}
$$

□

推论 1 若 $\boldsymbol{A}$ 有 n 个互不相同的特征值，则 $\boldsymbol{A}$ 与对角阵相似。

例 3.1 设实对称矩阵

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

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

例 3.2　设 $\boldsymbol{A}=(a_{ij})_{n\times n}$ 是主对角元全为 2 的上三角矩阵，且存在 $a_{ij}\neq 0(i<j)$，问 $\boldsymbol{A}$ 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
    - 实对称矩阵的对角化

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

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

- $\overline{k\boldsymbol{A}} = \overline{k}\ \overline{\boldsymbol{A}}$
- $\overline{\boldsymbol{A}+\boldsymbol{B}} = \overline{\boldsymbol{A}} + \overline{\boldsymbol{B}}$
- $\overline{\boldsymbol{A}\boldsymbol{B}} = \overline{\boldsymbol{A}\boldsymbol{B}}$
- $\overline{(\boldsymbol{A}\boldsymbol{B})}^T = \overline{\boldsymbol{B}}^T\overline{\boldsymbol{A}}^T$
- $\overline{\boldsymbol{A}^{-1}} = (\overline{\boldsymbol{A}})^{-1}$
- $\det\bar{\boldsymbol{A}} = \overline{\det\boldsymbol{A}}$

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
   - 实对称矩阵的对角化

定理 8　实对称矩阵 $\boldsymbol{A}$ 的任一特征值都是实数。

定理 8 实对称矩阵 $\boldsymbol{A}$ 的任一特征值都是实数。

证明.

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \quad &\Longrightarrow \quad \overline{(\boldsymbol{A}\boldsymbol{x})}^T = \overline{\lambda\boldsymbol{x}}^T \\
&\Longrightarrow \quad \overline{\boldsymbol{x}}^T \ \overline{\boldsymbol{A}}^T \ \boldsymbol{x} = \overline{\lambda} \ \overline{\boldsymbol{x}}^T \ \boldsymbol{x} \\
&\Longrightarrow \quad \overline{\boldsymbol{x}}^T \ \boldsymbol{A}^T \ \boldsymbol{x} = \overline{\lambda} \ \overline{\boldsymbol{x}}^T \ \boldsymbol{x} \\
&\Longrightarrow \quad \lambda\overline{\boldsymbol{x}}^T \ \boldsymbol{x} = \overline{\lambda} \ \overline{\boldsymbol{x}}^T \ \boldsymbol{x} \\
&\Longrightarrow \quad \lambda = \overline{\lambda}
\end{aligned}
$$

$\square$

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 $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$， $\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$ $(\lambda_1 \neq \lambda_2)$， $\boldsymbol{A}^T = \boldsymbol{A}$，则

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

定理 9 实对称矩阵 $\boldsymbol{A}$ 对应于不同特征值的特征向量是正交的。

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

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

由于 $\lambda_1 \neq \lambda_2$，所以

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

$\square$

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
   - 实对称矩阵的对角化

定理 10 对于任何一个 n 阶实对称矩阵 $\boldsymbol{A}$，存在 n 阶正交矩阵 $\boldsymbol{T}$，使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

定理 10 对于任何一个 n 阶实对称矩阵 $\boldsymbol{A}$，存在 n 阶正交矩阵 $\boldsymbol{T}$，使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n).$$

证明. 用数学归纳法证明。

- 当 $n = 1$ 时，结论显然成立。

□

证明.

- 假设定理对任何一个 $n-1$ 阶实对称矩阵 $\boldsymbol{B}$ 成立，即存在 $n-1$ 阶正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{\Lambda}_1$。下证对 n 阶实对称矩阵 $\boldsymbol{A}$ 也成立。

证明.

- 假设定理对任何一个 $n-1$ 阶实对称矩阵 $\boldsymbol{B}$ 成立，即存在 $n-1$ 阶正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{\Lambda}_1$。下证对 n 阶实对称矩阵 $\boldsymbol{A}$ 也成立。

  设 $\boldsymbol{A}\boldsymbol{x}_1=\lambda_1\boldsymbol{x}_1$，其中 $\boldsymbol{x}_1$ 是长度为 1 的特征向量。现将 $\boldsymbol{x}_1$ 扩充为 $\mathbb{R}^n$ 的一组标准正交基

$$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$$

  其中 $\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 不一定 $\boldsymbol{A}$ 的特征向量，于是就有

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)&=(\boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{A}\boldsymbol{x}_2,\cdots,\boldsymbol{A}\boldsymbol{x}_n)\\&=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{bmatrix}\lambda_1 & b_{12} & \cdots & b_{1n}\\ & b_{22} & \cdots & b_{2n}\\ & \vdots & & \vdots\\ & b_{n2} & \cdots & b_{nn}\end{bmatrix}\end{aligned}$$

证明.

- 假设定理对任何一个 $n-1$ 阶实对称矩阵 $\boldsymbol{B}$ 成立，即存在 $n-1$ 阶正交矩阵 $\boldsymbol{Q}$，使得 $\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{Q}=\boldsymbol{\Lambda}_1$。下证对 n 阶实对称矩阵 $\boldsymbol{A}$ 也成立。

  设 $\boldsymbol{A}\boldsymbol{x}_1=\lambda_1\boldsymbol{x}_1$，其中 $\boldsymbol{x}_1$ 是长度为 1 的特征向量。现将 $\boldsymbol{x}_1$ 扩充为 $\mathbb{R}^n$ 的一组标准正交基

$$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$$

  其中 $\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$ 不一定 $\boldsymbol{A}$ 的特征向量，于是就有

$$\begin{aligned}\boldsymbol{A}(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)&=(\boldsymbol{A}\boldsymbol{x}_1,\boldsymbol{A}\boldsymbol{x}_2,\cdots,\boldsymbol{A}\boldsymbol{x}_n)\\&=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)\begin{bmatrix}\lambda_1 & b_{12} & \cdots & b_{1n}\\ & b_{22} & \cdots & b_{2n}\\ & \vdots & & \vdots\\ & b_{n2} & \cdots & b_{nn}\end{bmatrix}\end{aligned}$$

  记

$$\boldsymbol{P}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n)$$

  则上式可写为

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{bmatrix}\lambda_1 & \boldsymbol{b}\\ & \boldsymbol{B}\end{bmatrix}$$

□

证明. 由于 $\boldsymbol{P}^{-1}=\boldsymbol{P}^{T}, (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^{T}=\boldsymbol{P}^{T}\boldsymbol{A}(\boldsymbol{P}^{-1})^{T}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$，因此

$$\begin{bmatrix} \lambda_1 & \\ \boldsymbol{b}^T & \boldsymbol{B}^T \end{bmatrix} = \begin{bmatrix} \lambda_1 & \boldsymbol{b} \\ & \boldsymbol{B} \end{bmatrix}$$

证明. 由于 $P^{-1}=P^T,(P^{-1}AP)^T=P^TA(P^{-1})^T=P^{-1}AP$，因此

$$\begin{bmatrix} \lambda_1 & \\ b^T & B^T \end{bmatrix} = \begin{bmatrix} \lambda_1 & b \\ & B \end{bmatrix}$$

因此 $b^T=b,B^T=B$，从而

$$P^{-1}AP=\begin{bmatrix} \lambda_1 & \\ & B \end{bmatrix}$$

由归纳假设，构造一个正交阵

$$S=\begin{bmatrix} 1 & \\ & Q \end{bmatrix}$$

便有

$$S^{-1}(P^{-1}AP)S=\begin{bmatrix} 1 & \\ & Q^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & \\ & B \end{bmatrix}\begin{bmatrix} 1 & \\ & Q \end{bmatrix}=\begin{bmatrix} \lambda_1 & \\ & Q^{-1}BQ \end{bmatrix}=\begin{bmatrix} \lambda_1 & \\ & \Lambda_1 \end{bmatrix}$$

证明. 由于 $P^{-1}=P^T,(P^{-1}AP)^T=P^TA(P^{-1})^T=P^{-1}AP$，因此

$$\begin{bmatrix} \lambda_1 & \\ b^T & B^T \end{bmatrix} = \begin{bmatrix} \lambda_1 & b \\ & B \end{bmatrix}$$

因此 $b^T=b,B^T=B$，从而

$$P^{-1}AP=\begin{bmatrix} \lambda_1 & \\ & B \end{bmatrix}$$

由归纳假设，构造一个正交阵

$$S=\begin{bmatrix} 1 & \\ & Q \end{bmatrix}$$

便有

$$S^{-1}(P^{-1}AP)S=\begin{bmatrix} 1 & \\ & Q^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & \\ & B \end{bmatrix}\begin{bmatrix} 1 & \\ & Q \end{bmatrix}=\begin{bmatrix} \lambda_1 & \\ & Q^{-1}BQ \end{bmatrix}=\begin{bmatrix} \lambda_1 & \\ & \Lambda_1 \end{bmatrix}$$

取 $T=PS,T^{-1}=S^{-1}P^{-1}$，则

$$T^{-1}AT=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

□

给定一个 n 阶实对称矩阵 $\boldsymbol{A}$，如何求正交矩阵 $\boldsymbol{A}$，使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{\Lambda}$ 呢？

给定一个 n 阶实对称矩阵 $\boldsymbol{A}$，如何求正交矩阵 $\boldsymbol{A}$，使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{\Lambda}$ 呢？

1. 有特征多项式 $|\boldsymbol{A}-\lambda\boldsymbol{I}|=\Pi_{i=1}^{m}(\lambda-\lambda_i)^{r_i}=0$ 求出全部互异的特征值 $\lambda_1,\cdots,\lambda_m$。
2. 因 $\boldsymbol{A}$ 可对角化，故 r_i 重特征值 λ_i 对应 r_i 个线性无关的特征向量 $\boldsymbol{x}_{i_1},\cdots,\boldsymbol{x}_{i_r}$。
3. 利用施密特正交化过程，得到 r_i 个相互正交的单位向量 $\boldsymbol{y}_{i_1},\cdots,\boldsymbol{y}_{i_r}$。
4. 因不同特征值对应的特征向量正交，故得到 $\{\boldsymbol{y}_{i_1},\cdots,\boldsymbol{y}_{i_{r_i}}\mid i=1,\cdots,m\}$ 为 n 个相互正交的单位特征向量，将其按列排成 n 阶矩阵，即得正交矩阵 $\boldsymbol{T}$.

例 4.1　设

$$A=\begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{bmatrix}$$

求正交阵 T，使得 $T^{-1}AT$ 为对角阵。

例 4.1 设

$$A = \begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{bmatrix}$$

求正交阵 T，使得 $T^{-1}AT$ 为对角阵。

解 由

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & -2 & 2 \\ -2 & -2-\lambda & 4 \\ 2 & 4 & -2-\lambda \end{vmatrix} = -(\lambda-2)^2(\lambda+7)$$

得 $\lambda_{1,2} = 2$ 和 $\lambda_3 = -7$。

例 4.1 设

$$A=\begin{bmatrix} 1 & -2 & 2 \\ -2 & -2 & 4 \\ 2 & 4 & -2 \end{bmatrix}$$

求正交阵 T，使得 $T^{-1}AT$ 为对角阵。

解 由

$$|A-\lambda I|=\begin{vmatrix} 1-\lambda & -2 & 2 \\ -2 & -2-\lambda & 4 \\ 2 & 4 & -2-\lambda \end{vmatrix}=-(\lambda-2)^2(\lambda+7)$$

得 $\lambda_{1,2}=2$ 和 $\lambda_3=-7$。

解 当 $\lambda_{1,2}=2$ 时，由

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\begin{bmatrix} -1 & -2 & 2 \\ -2 & -4 & 4 \\ 2 & 4 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

得线性无关的特征向量 $\boldsymbol{x}_1=(2,-1,0)^T$ 和 $\boldsymbol{x}_2=(2,0,1)^T$。

解 当 $\lambda_{1,2}=2$ 时，由

$$(\boldsymbol{A}-\lambda \boldsymbol{I})\boldsymbol{x}=\begin{bmatrix} -1 & -2 & 2 \\ -2 & -4 & 4 \\ 2 & 4 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

得线性无关的特征向量 $\boldsymbol{x}_1=(2,-1,0)^T$ 和 $\boldsymbol{x}_2=(2,0,1)^T$。

用施密特正交化过程：先正交化

$$\boldsymbol{\beta}_1=\boldsymbol{x}_1,$$

$$\boldsymbol{\beta}_2=\boldsymbol{x}_2-\frac{(\boldsymbol{x}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\frac{1}{5}\begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix},$$

再单位化得

$$\boldsymbol{y}_1=\begin{bmatrix} \frac{2\sqrt{5}}{5} \\ -\frac{\sqrt{5}}{5} \\ 0 \end{bmatrix},\boldsymbol{y}_2=\begin{bmatrix} \frac{2\sqrt{5}}{15} \\ \frac{4\sqrt{5}}{15} \\ \frac{\sqrt{5}}{3} \end{bmatrix}$$

解　当 $\lambda_3=-7$ 时，由

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{x}=\begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

得特征向量 $\boldsymbol{x}_3=(1,2,-2)^T$，单位化得 $\boldsymbol{y}_3=(\frac{1}{3},\frac{2}{3},-\frac{2}{3})^T$。

解 取正交阵

$$T=(y_1,y_2,y_3)=\begin{bmatrix} \frac{2\sqrt{5}}{5} & \frac{2\sqrt{5}}{15} & \frac{1}{3} \\ -\frac{\sqrt{5}}{5} & \frac{4\sqrt{5}}{15} & \frac{2}{3} \\ 0 & \frac{\sqrt{5}}{3} & -\frac{2}{3} \end{bmatrix}$$

则 $T^{-1}AT=\mathrm{diag}(2,2,-7)$。

例 4.2　设实对称矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是相似矩阵，证明：存在正交矩阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{B}$.

例 4.3　设 n 阶实对称矩阵 $\boldsymbol{A},\boldsymbol{B}$ 有完全相同的 n 个特征值，证明：村早正交阵 $\boldsymbol{T}$ 和 n 阶矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{A}=\boldsymbol{QT}$ 和 $\boldsymbol{B}=\boldsymbol{TQ}$ 同时成立。

例 4.4　设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是 n 实对称矩阵，若存在正交矩阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$ 都是对角阵，则 $\boldsymbol{A}\boldsymbol{B}$ 是实对称阵。