

线性代数

特征值与特征向量

武汉大学数学与统计学院

2017 年 8 月 30 日

1. 矩阵的特征值与特征向量 相似矩阵

2. 相似矩阵

3. 矩阵可对角化的条件

4. 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



1.1 特征值与特征向量



特征值与特征向量

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

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

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

特征值与特征向量

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

特征值与特征向量

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

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

有非零解的 λ 值, 即满足方程

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

的 λ 都是矩阵 \mathbf{A} 的特征值。

特征值与特征向量

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

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

有非零解的 λ 值, 即满足方程

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

的 λ 都是矩阵 \mathbf{A} 的特征值。

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

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

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

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

注

- (1) n 阶矩阵 \mathbf{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 1 求矩阵

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

的特征值与特征向量。

特征值与特征向量 I

解

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

故 \mathbf{A} 的特征值为 $\lambda_1 = 3$, $\lambda_2 = 2$ (二重特征值)。

当 $\lambda_1 = 3$ 时, 由 $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$, 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 $\mathbf{x}_1 = (1, 1, 1)^T$, 因此 $k_1 \mathbf{x}_1$ (k_1 为非零任意常数) 是 \mathbf{A} 对应于 $\lambda_1 = 3$ 的全部特征向量。

当 $\lambda_2 = 2$ 时, 由 $(\mathbf{A} - \lambda_2 \mathbf{I})\mathbf{x} = \mathbf{0}$, 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 $\mathbf{x}_2 = (1, 1, 2)^T$, 因此 $k_2 \mathbf{x}_2$ (k_2 为非零任意常数) 是 \mathbf{A} 对应于 $\lambda_2 = 2$ 的全部特征向量。

例 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。

1.2 特征值与特征值的性质



特征值与特征值的性质

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

特征值与特征值的性质

证明. 由于 \mathbf{x}_1 和 \mathbf{x}_2 是齐次线性方程组

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

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

特征值与特征值的性质

在 $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0$ 的解空间中, 除零向量以外的全体解向量就是 \mathbf{A} 的属于特征值 λ 的全体特征向量。因此, $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0$ 的解空间也称为 \mathbf{A} 关于特征值 λ 的特征子空间, 记作 V_λ 。 n 阶矩阵 \mathbf{A} 的特征子空间就是 n 维向量空间的子空间, 其维数为

$$\dim V_\lambda = n - \text{rank}(\mathbf{A} - \lambda \mathbf{I}).$$

特征值与特征值的性质

需要注意的是, n 维实矩阵的特征值可能是复数, 所以特征子空间一般是 n 维复向量空间 \mathbb{C}^n 的子空间。

上例中, 矩阵 $\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ 3 & 1 & -1 \\ 4 & -2 & 1 \end{pmatrix}$ 的两个特征子空间为

$$V_{\lambda_1} = \{k\mathbf{x} | \mathbf{x} = (1, 1, 1)^T, k \in \mathbb{C}\},$$

$$V_{\lambda_2} = \{k\mathbf{x} | \mathbf{x} = (1, 1, 2)^T, k \in \mathbb{C}\}.$$

特征值与特征值的性质

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

$$(1) \sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii};$$

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

其中 $\sum_{i=1}^n a_{ii}$ 是 \mathbf{A} 的主对角元之和, 称为 \mathbf{A} 的迹 (trace), 记为 $tr(\mathbf{A})$ 。

特征值与特征值的性质 I

证明. 设

$$\det(\mathbf{A} - \lambda \mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$
$$= \lambda^n + c_1 \lambda^{n-1} + c_2 \lambda^{n-2} + \cdots + c_{n-1} \lambda + c_n,$$

展开后含 λ^{n-1} 项的行列式有下面 n 个

$$\begin{vmatrix} a_{11} & & & \\ a_{21} & -\lambda & & \\ a_{31} & & -\lambda & \\ \vdots & & & \ddots \\ a_{n1} & & & & -\lambda \end{vmatrix}, \begin{vmatrix} -\lambda & a_{12} & & \\ a_{22} & & & \\ a_{32} & -\lambda & & \\ \vdots & & \ddots & \\ a_{n2} & & & -\lambda \end{vmatrix},$$
$$\dots, \begin{vmatrix} -\lambda & & & a_{1n} \\ & -\lambda & & a_{2n} \\ & & -\lambda & a_{3n} \\ & & & \ddots \\ & & & & a_{nn} \end{vmatrix}$$

特征值与特征值的性质 II

它们之和等于

$$(a_{11} + a_{22} + \cdots + a_{nn})(-\lambda)^{n-1} = (-1)^{n-1} \sum_{i=1}^n a_{ii} \lambda^{n-1},$$

即 $c_1 = (-1)^{n-1} \sum_{i=1}^n a_{ii}$ 。

展开后常数项为

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

即 $c_n = \det(\mathbf{A})$ 。

假设 \mathbf{A} 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 根据 n 次多项式的根与系数的关系, 得

$$(-1)^{n-1} \sum_{i=1}^n \lambda_i = c_1 = (-1)^{n-1} \sum_{i=1}^n a_{ii},$$

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

故

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



特征值与特征值的性质

注

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

特征值与特征值的性质

定理 3 一个特征向量不能属于不同的特征值。

特征值与特征值的性质

定理 3 一个特征向量不能属于不同的特征值。

证明.

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

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

这与 $\mathbf{x} \neq \mathbf{0}$ 矛盾。



特征值与特征值的性质

性质 1 若 λ 是矩阵 \mathbf{A} 的特征值, \mathbf{x} 是 \mathbf{A} 属于 λ 的特征向量, 则

- (i) $k\lambda$ 是 $k\mathbf{A}$ 的特征值;
- (ii) λ^m 是 \mathbf{A}^m 的特征值;
- (iii) 当 \mathbf{A} 可逆时, λ^{-1} 是 \mathbf{A}^{-1} 的特征值;

且 \mathbf{x} 仍是矩阵 $k\mathbf{A}, \mathbf{A}^m, \mathbf{A}^{-1}$ 分别对应于 $k\lambda, \lambda^m, \lambda^{-1}$ 的特征向量。

特征值与特征值的性质

证明.

- (i) 自行完成;
- (ii) 自行完成;
- (iii) 当 \mathbf{A} 可逆时, $\lambda \neq 0$, 由 $\mathbf{Ax} = \lambda\mathbf{x}$ 可得

$$\mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{A}^{-1}(\lambda\mathbf{x}) = \lambda\mathbf{A}^{-1}\mathbf{x},$$

因此

$$\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x},$$

故 λ^{-1} 是 \mathbf{A}^{-1} 的特征值, 且 \mathbf{x} 也是 \mathbf{A}^{-1} 对应于 λ^{-1} 的特征向量。



注 若 λ 是 \mathbf{A} 的特征值, 则 $\varphi(\lambda)$ 是 $\varphi(\mathbf{A})$ 的特征值, 其中

$$\begin{aligned}\varphi(\lambda) &= a_0 + a_1\lambda + \cdots + a_m\lambda^m, \\ \varphi(\mathbf{A}) &= a_0\mathbf{I} + a_1\mathbf{A} + \cdots + a_m\mathbf{A}^m.\end{aligned}$$

特征值与特征值的性质

例 3 设 3 阶矩阵 \mathbf{A} 的特征值为 $1, -1, 2$, 求 $|\mathbf{A}^* + 3\mathbf{A} - 2\mathbf{I}|$.

特征值与特征值的性质

例 3 设 3 阶矩阵 \mathbf{A} 的特征值为 $1, -1, 2$, 求 $|\mathbf{A}^* + 3\mathbf{A} - 2\mathbf{I}|$.

解 因 \mathbf{A} 的特征值全不为零, 故 \mathbf{A} 可逆, 从而 $\mathbf{A}^* = |\mathbf{A}|\mathbf{A}^{-1}$. 又因 $|\mathbf{A}| = \lambda_1\lambda_2\lambda_3 = -2$, 故

$$\mathbf{A}^* + 3\mathbf{A} - 2\mathbf{I} = -2\mathbf{A}^{-1} + 3\mathbf{A} - 2\mathbf{I}.$$

令 $\varphi(\lambda) = -\frac{2}{\lambda} + 3\lambda - 2$, 则 $\varphi(\lambda)$ 为上述矩阵的特征值, 分别为 $\varphi(1) = -1, \varphi(-1) = -3, \varphi(2) = 3$, 于是

$$|\mathbf{A}^* + 3\mathbf{A} - 2\mathbf{I}| = (-1) \cdot (-3) \cdot 3 = 9.$$

特征值与特征值的性质

性质 2 矩阵 \mathbf{A} 与 \mathbf{A}^T 的特征值相同。

特征值与特征值的性质

性质 2 矩阵 \mathbf{A} 与 \mathbf{A}^T 的特征值相同。

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

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

故 \mathbf{A} 与 \mathbf{A}^T 有完全相同的特征值。



特征值与特征值的性质

例 4 设 $\mathbf{A} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 2 \\ -1 & 1 & -1 \end{pmatrix}$

- (i) 求 \mathbf{A} 的特征值与特征向量
- (ii) 求可逆矩阵 \mathbf{P} , 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ 为对角阵。

特征值与特征值的性质 I

解 由

$$\begin{aligned} |\mathbf{A} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & -1 & 1 \\ 2 & -2-\lambda & 2 \\ -1 & 1 & -1-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 2 & -\lambda & 2 \\ -1 & -\lambda & -1-\lambda \end{vmatrix} \\ &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 2 & -\lambda & 2 \\ -3 & 0 & -3-\lambda \end{vmatrix} = -\lambda[(\lambda-1)(\lambda+3)+3] = -\lambda^2(\lambda+2), \end{aligned}$$

知 \mathbf{A} 的特征值为 $\lambda_1 = \lambda_2 = 0$ 和 $\lambda_3 = -2$.

当 $\lambda_{1,2} = 0$ 时, 由 $(\mathbf{A} - 0\mathbf{I})\mathbf{x} = 0$, 即 $\mathbf{A}\mathbf{x} = 0$ 得基础解系

$$\mathbf{x}_1 = (1, 1, 0)^T, \quad \mathbf{x}_2 = (-1, 0, 1)^T,$$

故 \mathbf{A} 对应于 $\lambda_{1,2} = 0$ 的全体特征向量为

$$k_1\mathbf{x}_1 + k_2\mathbf{x}_2 = k_1(1, 1, 0)^T + k_2(-1, 0, 1)^T,$$

其中 k_1, k_2 为不全为零的任意常数。

特征值与特征值的性质 II

当 $\lambda_3 = -2$ 时, 由 $(\mathbf{A} - \lambda_3 \mathbf{I})\mathbf{x} = 0$, 即

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

得基础解系

$$\mathbf{x}_3 = (-1, -2, 1)^T,$$

故 \mathbf{A} 对应于 $\lambda_3 = -2$ 的全体特征向量为

$$k_3 \mathbf{x}_3 = k_3 (-1, -2, 1)^T,$$

其中 k_3 为非零的任意常数。

将 $\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i (i = 1, 2, 3)$ 表示成

$$\mathbf{A}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

取

$$\mathbf{P} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} 0 & & \\ & 0 & \\ & & -2 \end{pmatrix},$$

则 $\mathbf{AP} = \mathbf{P}\mathbf{\Lambda}$, 且 $|\mathbf{P}| = 2 \neq 0$, 故得

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

为对角阵。

特征值与特征值的性质

定理 4 设 $\lambda_1, \lambda_2, \dots, \lambda_m$ 是方阵 \mathbf{A} 的 m 个特征值, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ 依次是与之对应的特征向量, 若 $\lambda_1, \lambda_2, \dots, \lambda_m$ 互不相等, 则 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ 线性无关。

特征值与特征值的性质 I

证明. 设有常数 k_1, k_2, \dots, k_m 使得

$$k_1 \mathbf{x}_1 + k_2 \mathbf{x}_2 + \dots + k_m \mathbf{x}_m = 0,$$

则 $\mathbf{A}(k_1 \mathbf{x}_1 + k_2 \mathbf{x}_2 + \dots + k_m \mathbf{x}_m) = 0$, 即

$$\lambda_1 k_1 \mathbf{x}_1 + \lambda_2 k_2 \mathbf{x}_2 + \dots + \lambda_m k_m \mathbf{x}_m = 0,$$

以此类推, 有

$$\lambda_1^l k_1 \mathbf{x}_1 + \lambda_2^l k_2 \mathbf{x}_2 + \dots + \lambda_m^l k_m \mathbf{x}_m = 0, \quad l = 1, 2, \dots, m-1.$$

写成矩阵形式即为

$$(k_1 \mathbf{x}_1, k_2 \mathbf{x}_2, \dots, k_m \mathbf{x}_m) \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{m-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \dots & \lambda_m^{m-1} \end{pmatrix} = (0, 0, \dots, 0).$$

上式左边的第二个矩阵的行列式为范德蒙德行列式, 当 λ_i 互不相等时该行列式不为零, 从而该矩阵可逆. 于是有

$$(k_1 \mathbf{x}_1, k_2 \mathbf{x}_2, \dots, k_m \mathbf{x}_m) = (0, 0, \dots, 0),$$

特征值与特征值的性质 II

即 $k_j \mathbf{x}_j = 0$, $j = 1, 2, \dots, m$ 。但 $\mathbf{x}_j \neq 0$, 故 $k_j = 0$, $j = 1, 2, \dots, m$, 从而 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ 线性无关。 □

特征值与特征值的性质

例 5 设 λ_1 和 λ_2 是矩阵 \mathbf{A} 的两个不同特征值，对应的特征向量依次为 $\mathbf{x}_1, \mathbf{x}_2$ ，证明 $\mathbf{x}_1 + \mathbf{x}_2$ 不是 \mathbf{A} 的特征向量。

特征值与特征值的性质

证明. 按题设, 有 $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$, $A\mathbf{x}_2 = \lambda_2\mathbf{x}_2$, 故

$$A(\mathbf{x}_1 + \mathbf{x}_2) = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2.$$

(反证法) 假设 $\mathbf{x}_1 + \mathbf{x}_2$ 是 A 的特征向量, 则应存在 λ 使得 $A(\mathbf{x}_1 + \mathbf{x}_2) = \lambda(\mathbf{x}_1 + \mathbf{x}_2)$, 于是

$$\lambda(\mathbf{x}_1 + \mathbf{x}_2) = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2,$$

即

$$(\lambda_1 - \lambda)\mathbf{x}_1 + (\lambda_2 - \lambda)\mathbf{x}_2 = \mathbf{0}.$$

因 $\lambda_1 \neq \lambda_2$, 由上述定理知 $\mathbf{x}_1, \mathbf{x}_2$ 线性无关, 从而有

$$\lambda_1 - \lambda = \lambda_2 - \lambda,$$

即 $\lambda_1 = \lambda_2$, 与题设矛盾, 从而 $\mathbf{x}_1 + \mathbf{x}_2$ 不是 A 的特征向量。 □

2. 相似矩阵



定义 3 设 \mathbf{A}, \mathbf{B} 为 n 阶矩阵, 若存在可逆矩阵 \mathbf{P} , 使得

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B},$$

则称 \mathbf{B} 是 \mathbf{A} 的相似矩阵, 或 \mathbf{A} 与 \mathbf{B} 相似, 记为 $\mathbf{A} \sim \mathbf{B}$ 。对 \mathbf{A} 进行运算 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ 称为对 \mathbf{A} 进行相似变化, 可逆矩阵 \mathbf{P} 称为把 \mathbf{A} 变成 \mathbf{B} 的相似变换矩阵。

定理 5 若 $\mathbf{A} \sim \mathbf{B}$, 则 \mathbf{A} 与 \mathbf{B} 的 χ 特征多项式相同, 从而 \mathbf{A} 与 \mathbf{B} 的特征值相同。

证明. 因 $A \sim B$, 即有可逆阵 P 使得 $P^{-1}AP = B$ 。故

$$|B - \lambda I| = |P^{-1}AP - P^{-1}(\lambda I)P| = |P^{-1}(A - \lambda I)P| = |P^{-1}| |A - \lambda I| |P| = |A - \lambda I|.$$



推论 1 若 \mathbf{A} 与对角阵

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

相似, 则 $\lambda_1, \lambda_2, \dots, \lambda_n$ 为 \mathbf{A} 的 n 个特征值。

3. 矩阵可对角化的条件



矩阵可对角化的条件

矩阵可对角化，即矩阵与对角阵相似。

矩阵可对角化的条件

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

矩阵可对角化的条件

证明.

(\Rightarrow) 设

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

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。

(\Leftarrow) 上述步骤显然可逆, 故充分性也成立。



矩阵可对角化的条件

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

矩阵可对角化的条件

定理 7 \mathbf{A} 的属于不同特征值的特征向量是线性无关的。

矩阵可对角化的条件 I

证明. 设 \mathbf{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° 当 $m = 1$ 时, 结论显然成立。

2° 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

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

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

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

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

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

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

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

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

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



矩阵可对角化的条件

推论 2 若 \mathbf{A} 有 n 个互不相同的特征值, 则 \mathbf{A} 与对角阵相似。

矩阵可对角化的条件

例 6 设实对称矩阵

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

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

矩阵可对角化的条件 I

解 由

$$\begin{aligned} |\mathbf{A} - \lambda \mathbf{I}| &= \begin{vmatrix} 1-\lambda & -1 & -1 & -1 \\ -1 & 1-\lambda & -1 & -1 \\ -1 & -1 & 1-\lambda & -1 \\ -1 & -1 & -1 & 1-\lambda \end{vmatrix} \\ &= -(\lambda + 2) \begin{vmatrix} 1 & -1 & -1 & -1 \\ 1 & 1-\lambda & -1 & -1 \\ 1 & -1 & 1-\lambda & -1 \\ 1 & -1 & -1 & 1-\lambda \end{vmatrix} \\ &= -(\lambda + 2) \begin{vmatrix} 1 & -1 & -1 & -1 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 2-\lambda & 0 \\ 0 & 0 & 0 & 2-\lambda \end{vmatrix} = (\lambda + 2)(\lambda - 2)^3, \end{aligned}$$

故 \mathbf{A} 的特征值为 $\lambda_1 = -2$ (单根), $\lambda_2 = 2$ (三重根)。

矩阵可对角化的条件 II

由 $(\mathbf{A} - \lambda_1 \mathbf{I})\mathbf{x} = 0$, 即

$$\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

得 λ_1 对应的特征向量为 $\{k_1 \mathbf{x}_1 | \mathbf{x}_1 = (1, 1, 1, 1)^T, k_1 \neq 0\}$ 。

由 $(\mathbf{A} - \lambda_2 \mathbf{I})\mathbf{x} = 0$, 即

$$\begin{pmatrix} -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

得基础解系:

$$\mathbf{x}_{21} = (1, -1, 0, 0)^T, \quad \mathbf{x}_{22} = (1, 0, -1, 0)^T, \quad \mathbf{x}_{23} = (1, 0, 0, -1)^T.$$

因 \mathbf{A} 有 4 个线性无关的特征向量, 故 $\mathbf{A} \sim \mathbf{\Lambda}$ 。

矩阵可对角化的条件 III

取

$$\mathbf{P} = (\mathbf{x}_1, \mathbf{x}_{21}, \mathbf{x}_{22}, \mathbf{x}_{23}) = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix},$$

则

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} -2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix} = \mathbf{\Lambda}.$$

再由 $\mathbf{A} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$ 得

$$\mathbf{A}^k = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1})^k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1}$$

$$\begin{aligned} &= \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} (-2)^k & & & \\ & 2^k & & \\ & & 2^k & \\ & & & 2^k \end{pmatrix} \frac{1}{4} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \\ &= \begin{cases} 2^k \mathbf{I}_4, & k \text{ even,} \\ 2^{k-1} \mathbf{A}, & k \text{ odd.} \end{cases} \end{aligned}$$

矩阵可对角化的条件

例 7 设

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & x \\ 1 & 0 & 0 \end{pmatrix}$$

问 x 为何值时, 矩阵 \mathbf{A} 可对角化?

矩阵可对角化的条件

解 由

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} -\lambda & 0 & 1 \\ 1 & 1-\lambda & x \\ 1 & 0 & -\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = -(\lambda-1)^2(\lambda+1),$$

即 $\lambda_1 = -1, \lambda_2 = \lambda_3 = 1$ 。

对应于单根 $\lambda_1 = -1$ ，可求得线性无关的特征向量恰有 1 个，故 \mathbf{A} 可对角化的充分必要条件是对应重根 $\lambda_2 = \lambda_3 = 1$ ，有 2 个线性无关的特征向量，即 $(\mathbf{A} - \mathbf{I})\mathbf{x} = 0$ 有两个线性无关的解，亦即 $\mathbf{A} - \mathbf{I}$ 的秩 $R(\mathbf{A} - \mathbf{I}) = 1$ 。由

$$\mathbf{A} - \mathbf{I} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & x \\ 1 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & x+1 \\ 0 & 0 & 0 \end{pmatrix}$$

欲使 $R(\mathbf{A} - \mathbf{I}) = 1$ ，须有 $x+1=0$ ，即 $x=-1$ 。因此当 $x=-1$ 时，矩阵 \mathbf{A} 能对角化。

矩阵可对角化的条件

例 8 设 $\mathbf{A} = (a_{ij})_{n \times n}$ 是主对角元全为 2 的上三角矩阵, 且存在 $a_{ij} \neq 0 (i < j)$, 问 \mathbf{A} 是否可对角化?

矩阵可对角化的条件

解 设

$$A = \begin{pmatrix} 2 & * & \cdots & * \\ & 2 & \cdots & * \\ & & \ddots & \vdots \\ & & & 2 \end{pmatrix}$$

其中 $*$ 为不全为零的任意常数, 则

$$|A - \lambda I| = (2 - \lambda)^n,$$

即 $\lambda = 2$ 为 A 的 n 重特征根, 而 $R(A - 2I) \geq 1$, 故 $(A - 2I)x = 0$ 的基础解系所含向量个数 $\leq n - 1$ 个, 即 A 的线性无关的特征向量的个数 $\leq n - 1$ 个, 因此 A 不与对角阵相似。

4. 实对称矩阵的对角化



例 9 实对称矩阵 \mathbf{A} 的任一特征值都是实数。

证明.

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



例 10 实对称矩阵 \mathbf{A} 对应于不同特征值的特征向量是正交的。

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

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

例 10 实对称矩阵 \mathbf{A} 对应于不同特征值的特征向量是正交的。

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

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

由于 $\lambda_1 \neq \lambda_2$, 所以

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



定理 8 设 \mathbf{A} 为 n 阶对称阵，则必有正交阵 \mathbf{Q} ，使得 $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^T\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ ，其中 $\mathbf{\Lambda}$ 是以 \mathbf{A} 的 n 个特征值为对角元的对角阵。

推论 3 设 \mathbf{A} 为 n 阶对称阵, λ 为 \mathbf{A} 的特征方程的 k 重根, 则矩阵 $\mathbf{A} - \lambda \mathbf{I}$ 的秩 $R(\mathbf{A} - \lambda \mathbf{I}) = n - k$, 从而对应特征值 λ 恰有 k 个线性无关的特征向量。

推论 3 设 \mathbf{A} 为 n 阶对称阵, λ 为 \mathbf{A} 的特征方程的 k 重根, 则矩阵 $\mathbf{A} - \lambda \mathbf{I}$ 的秩 $R(\mathbf{A} - \lambda \mathbf{I}) = n - k$, 从而对应特征值 λ 恰有 k 个线性无关的特征向量。

证明.

由上述定理知, 对称阵 \mathbf{A} 与对角阵 $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ 相似, 从而 $\mathbf{A} - \lambda \mathbf{I}$ 与 $\mathbf{\Lambda} - \lambda \mathbf{I} = \text{diag}(\lambda_1 - \lambda, \dots, \lambda_n - \lambda)$ 相似。当 λ 是 \mathbf{A} 的 k 重特征根时, 对角阵 $\mathbf{\Lambda} - \lambda \mathbf{I}$ 的对角元恰有 k 个等于 0, 于是 $R(\mathbf{\Lambda} - \lambda \mathbf{I}) = n - k$ 。而 $R(\mathbf{A} - \lambda \mathbf{I}) = R(\mathbf{\Lambda} - \lambda \mathbf{I})$, 故 $R(\mathbf{A} - \lambda \mathbf{I}) = n - k$ 。 □

将对称阵 \mathbf{A} 对角化的步骤：

1. 求出 \mathbf{A} 的全部互不相等的特征值 $\lambda_1, \dots, \lambda_s$ ，它们的重数依次为 $k_1, \dots, k_s (k_1 + \dots + k_s = n)$ ；
2. 对每个 k_i 重特征值 λ_i ，求 $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{x} = 0$ 的基础解系，得 k_i 个线性无关的特征向量；
3. 再把它们正交化、单位化，得 k_i 个两两正交的单位特征向量。因 $k_1 + \dots + k_s = n$ ，故总共可得 n 个两两正交的单位特征向量；
4. 将这 n 个两两正交的单位特征向量构成正交阵 \mathbf{Q} ，便有 $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{\Lambda}$ 。