线性代数课程

相似矩阵及二次型

2021 年 8 月 31 日

■ 武汉大学数学与统计学院 ■ 古月老师

第 5 章 相似矩阵及二次型

第 5 章 相似矩阵及二次型

本章要求：

(1) 理解特征值与特征向量的概念，熟练掌握计算特征值和特征向量的方法，熟练掌握特征值与特征向量的性质；

第 5 章　相似矩阵及二次型

本章要求：

(1) 理解特征值与特征向量的概念，熟练掌握计算特征值和特征向量的方法，熟练掌握特征值与特征向量的性质；

(2) 了解相似矩阵的概念、性质和矩阵相似对角化的条件，会判断一个矩阵在什么条件下可以相似对角化；

第 5 章 相似矩阵及二次型

本章要求：

(1) 理解特征值与特征向量的概念，熟练掌握计算特征值和特征向量的方法，熟练掌握特征值与特征向量的性质；

(2) 了解相似矩阵的概念、性质和矩阵相似对角化的条件，会判断一个矩阵在什么条件下可以相似对角化；

(3) 理解实对称矩阵的特征值和特征向量的性质，熟练掌握相似对角化的方法，和用正交变换化实对称矩阵为对角矩阵的方法；

第 5 章 相似矩阵及二次型

本章要求：

(1) 理解特征值与特征向量的概念，熟练掌握计算特征值和特征向量的方法，熟练掌握特征值与特征向量的性质；

(2) 了解相似矩阵的概念、性质和矩阵相似对角化的条件，会判断一个矩阵在什么条件下可以相似对角化；

(3) 理解实对称矩阵的特征值和特征向量的性质，熟练掌握相似对角化的方法，和用正交变换化实对称矩阵为对角矩阵的方法；

(4) 掌握二次型及其矩阵表示，了解合同变换与合同矩阵的概念，掌握化二次型为标准形的正交变换法，会用配方法化二次型为标准形；

第 5 章 相似矩阵及二次型

本章要求:

(1) 理解特征值与特征向量的概念, 熟练掌握计算特征值和特征向量的方法, 熟练掌握特征值与特征向量的性质;

(2) 了解相似矩阵的概念、性质和矩阵相似对角化的条件, 会判断一个矩阵在什么条件下可以相似对角化;

(3) 理解实对称矩阵的特征值和特征向量的性质, 熟练掌握相似对角化的方法, 和用正交变换化实对称矩阵为对角矩阵的方法;

(4) 掌握二次型及其矩阵表示, 了解合同变换与合同矩阵的概念, 掌握化二次型为标准形的正交变换法, 会用配方法化二次型为标准形;

(5) 了解二次型的秩、规范形的概念以及惯性定理, 理解正定二次型、正定矩阵的概念, 并掌握其判别法.

方阵的特征值与特征向量

定义 1　设 $\mathbf{A}$ 是 n 阶方阵, 若存在数 λ 和非零列向量 $\mathbf{x}$, 使得

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad (5-1)$$

则称数 λ 为 $\mathbf{A}$ 的特征值(eigenvalue 或 characteristic value), 称非零向量 $\mathbf{x}$ 为 $\mathbf{A}$ 的对应于 (或属于) 特征值 λ 的一个特征向量(eigenvector 或 characteristic vector), 同样也称 λ 为对应于特征向量 $\mathbf{x}$ 的特征值.

(5-1) 式也可写成

$$(\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}. \quad (5-2)$$

λ 是 n 阶矩阵 $\mathbf{A}$ 的特征值

$\Leftrightarrow (\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 有非零解

$\Leftrightarrow R(\lambda \mathbf{E} - \mathbf{A}) < n$

$\Leftrightarrow |\lambda \mathbf{E} - \mathbf{A}| = 0.$

(5-1) 式也可写成

$$(\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}. \quad (5-2)$$

λ 是 n 阶矩阵 $\mathbf{A}$ 的特征值

$\Leftrightarrow (\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 有非零解

$\Leftrightarrow R(\lambda \mathbf{E} - \mathbf{A}) < n$

$\Leftrightarrow |\lambda \mathbf{E} - \mathbf{A}| = 0.$

ξ 是 $\mathbf{A}$ 的属于特征值 λ 的一个特征向量

$\Leftrightarrow \xi$ 是 $(\lambda \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 的非零解.

设 $\mathbf{A}=(a_{ij})$, 对于固定的 λ, (5-2) 是含有 n 个方程 n 个未知数的关于 $\mathbf{x}$ 的齐次线性方程组, 它有非零解的充要条件是

$$|\lambda\mathbf{E}-\mathbf{A}|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \dots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \dots & -a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ -a_{n1} & -a_{n2} & \dots & \lambda-a_{nn}\end{vmatrix}=0. \quad (5-3)$$

(5-3) 是关于 λ 的一元 n 次方程, 称为方阵 **A** 的特征方程(characteristic equation), 而它左端的 n 次多项式

$$f(\lambda) = f_{\mathbf{A}}(\lambda) = |\lambda \mathbf{E} - \mathbf{A}|$$

称为**A** 的特征多项式(eigenpolynomial 或 characteristic polynomial).

有教材也称 $|\mathbf{A} - \lambda \mathbf{E}|$ 为 **A** 的特征多项式, 实际上两者之间至多只差一个符号.

特征方程的根就是方阵 $\mathbf{A}$ 的特征值, 方程组 (5-2) 的非零解向量 $\mathbf{x}$ 就是方阵 $\mathbf{A}$ 的对应于特征值 λ 的特征向量. 由此可得:

特征方程的根就是方阵 $\mathbf{A}$ 的特征值, 方程组 (5-2) 的非零解向量 $\mathbf{x}$ 就是方阵 $\mathbf{A}$ 的对应于特征值 λ 的特征向量. 由此可得:

①特征值有限, 但特征向量无穷多.

事实上, 若 η_1, η_2 均为 $\mathbf{A}$ 的对应于特征值 λ 的特征向量, 即 $\mathbf{A}\eta_1 = \lambda\eta_1$, $\mathbf{A}\eta_2 = \lambda\eta_2$, 则有

$$\mathbf{A}(c_1\eta_1 + c_2\eta_2) = c_1\mathbf{A}\eta_1 + c_2\mathbf{A}\eta_2 = \lambda(c_1\eta_1 + c_2\eta_2).$$

故在 $c_1\eta_1 + c_2\eta_2$ 不为零向量时, 均是 $\mathbf{A}$ 的对应于特征值 λ 的特征向量.

特征方程的根就是方阵 $\mathbf{A}$ 的特征值, 方程组 (5-2) 的非零解向量 $\mathbf{x}$ 就是方阵 $\mathbf{A}$ 的对应于特征值 λ 的特征向量. 由此可得:

①特征值有限, 但特征向量无穷多.

事实上, 若 η_1, η_2 均为 $\mathbf{A}$ 的对应于特征值 λ 的特征向量, 即 $\mathbf{A}\eta_1 = \lambda\eta_1, \mathbf{A}\eta_2 = \lambda\eta_2$, 则有

$$\mathbf{A}(c_1\eta_1 + c_2\eta_2) = c_1\mathbf{A}\eta_1 + c_2\mathbf{A}\eta_2 = \lambda(c_1\eta_1 + c_2\eta_2).$$

故在 $c_1\eta_1 + c_2\eta_2$ 不为零向量时, 均是 $\mathbf{A}$ 的对应于特征值 λ 的特征向量.

②每个特征向量有唯一特征值.

特征方程的根就是方阵 $\mathbf{A}$ 的特征值, 方程组 (5-2) 的非零解向量 $\mathbf{x}$ 就是方阵 $\mathbf{A}$ 的对应于特征值 λ 的特征向量. 由此可得:

①特征值有限, 但特征向量无穷多.

事实上, 若 η_1, η_2 均为 $\mathbf{A}$ 的对应于特征值 λ 的特征向量, 即 $\mathbf{A}\eta_1 = \lambda\eta_1$, $\mathbf{A}\eta_2 = \lambda\eta_2$, 则有

$$\mathbf{A}(c_1\eta_1 + c_2\eta_2) = c_1\mathbf{A}\eta_1 + c_2\mathbf{A}\eta_2 = \lambda(c_1\eta_1 + c_2\eta_2).$$

故在 $c_1\eta_1 + c_2\eta_2$ 不为零向量时, 均是 $\mathbf{A}$ 的对应于特征值 λ 的特征向量.

②每个特征向量有唯一特征值.

事实上, 若有 $\lambda\xi = \mathbf{A}\xi = \mu\xi$, 由 $(\lambda - \mu)\xi = \mathbf{0}$, 因为 $\xi \neq \mathbf{0}$, 得 $\lambda = \mu$.

由定义容易得到计算特征值与特征向量的步骤:

① 求出特征多项式 $|\lambda \mathbf{E} - \mathbf{A}|$.

由定义容易得到计算特征值与特征向量的步骤:

① 求出特征多项式 $|\lambda \mathbf{E} - \mathbf{A}|$.

② 求 $|\lambda \mathbf{E} - \mathbf{A}| = 0$ 的根, 得方阵 $\mathbf{A}$ 的特征值.

由定义容易得到计算特征值与特征向量的步骤:

① 求出特征多项式 $|\lambda \mathbf{E} - \mathbf{A}|$.

② 求 $|\lambda \mathbf{E} - \mathbf{A}| = 0$ 的根, 得方阵 $\mathbf{A}$ 的特征值.

③ 对每个特征值 λ_i, 求 $(\lambda_i \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 的基础解系 $\alpha_1, \alpha_2, \cdots, \alpha_r$, 则 A 的属于 λ_i 的全部特征向量为 $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r$, 其中 $k_1, k_2, \cdots, k_r$ 是不全为 0 的任意常数.

例 1 求 $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ 的特征值和特征向量.

解 $\mathbf{A}$ 的特征多项式为:

$$|\lambda \mathbf{E} - \mathbf{A}| = \begin{vmatrix} \lambda - 2 & -1 & -1 \\ -1 & \lambda - 2 & -1 \\ -1 & -1 & \lambda - 2 \end{vmatrix}$$

$$= (\lambda - 4) \begin{vmatrix} 1 & 1 & 1 \\ -1 & \lambda - 2 & -1 \\ -1 & -1 & \lambda - 2 \end{vmatrix} = (\lambda - 4)(\lambda - 1)^2 = 0,$$

故特征值为 $\lambda_1 = 4, \lambda_2 = \lambda_3 = 1.$

对 $\lambda_1 = 4$, 解 $(4\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 由

$$4E - A = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \xrightarrow{r} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

得基础解系 $\alpha_1 = (1, 1, 1)^{\mathrm{T}}$. 故属于 $\lambda_1 = 4$ 的全部特征向量为 $k(1, 1, 1)^{\mathrm{T}}(k \neq 0)$.

对 $\lambda_2 = \lambda_3 = 1$, 解 $(\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 由

$$E - A = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \xrightarrow{r} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得基础解系 $\alpha_2 = (-1, 1, 0)^\mathrm{T}$, $\alpha_3 = (-1, 0, 1)^\mathrm{T}$. 故属于 $\lambda_2 = \lambda_3 = 1$ 的全部特征向量为

$k_1(-1, 1, 0)^\mathrm{T} + k_2(-1, 0, 1)^\mathrm{T}$, k_1, k_2 不全为 0.

例 2　求方阵 $\mathbf{A} = \begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ 的特征值和特征向量.

解　由于 $\mathbf{A}$ 的特征多项式为

$$|\lambda \mathbf{E} - \mathbf{A}| = \begin{vmatrix} \lambda - 3 & 1 & -1 \\ -2 & \lambda & -1 \\ -1 & 1 & \lambda - 2 \end{vmatrix}$$

$$= (\lambda - 1)(\lambda - 2)^2 = 0,$$

故 $\mathbf{A}$ 的特征值为：$\lambda_1 = 1, \lambda_2 = \lambda_3 = 2.$

对于 $\lambda_1 = 1$, 解 $(\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$, 由

$$\mathbf{E}-\mathbf{A}=\begin{pmatrix}-2 & 1 & -1\\ -2 & 1 & -1\\ -1 & 1 & -1\end{pmatrix}\xrightarrow{r}\begin{pmatrix}0 & 1 & -1\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得 $\alpha=(0,1,1)^{\mathrm{T}}$. 故对应于 $\lambda_1 = 1$ 的特征向量全体为 $k\alpha(k\neq 0)$.

对于 $\lambda_2=\lambda_3=2$, 解 $(2\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$, 由

$$2\mathbf{E}-\mathbf{A}=\begin{pmatrix}-1&1&-1\\-2&2&-1\\-1&1&-1\end{pmatrix}\xrightarrow{r}\begin{pmatrix}-1&1&0\\0&0&1\\0&0&0\end{pmatrix},$$

得方程组的一个基础解系为 $\beta=(1,1,0)^{\mathrm{T}}$. 故对应于特征值 $\lambda_2=\lambda_3=2$ 的全部特征向量为 $k\beta(k\neq 0)$.

设 $\mathbf{A}$ 为 n 阶矩阵, 矩阵 $\mathbf{A}$ 的特征多项式 $f(\lambda) = |\lambda\mathbf{E} - \mathbf{A}|$ 是 n 次多项式, 所以矩阵 $\mathbf{A}$ 的特征多项式在复数范围内有 n 个根.

① n 个特征值中有可能有相同的, 称为重特征值, 即是 $f(\lambda) = 0$ 的重根. 如 n 阶单位矩阵有 n 个重根, 例 5.1 中, $\lambda = 1$ 为 2 重根.

① n 个特征值中有可能有相同的, 称为重特征值, 即是 $f(\lambda)=0$ 的重根. 如 n 阶单位矩阵有 n 个重根, 例 5.1 中, $\lambda=1$ 为 2 重根.

②即便 $\mathbf{A}$ 为实方阵, 其特征值也可能是复数.

① n 个特征值中有可能有相同的, 称为重特征值, 即是 $f(\lambda)=0$ 的重根. 如 n 阶单位矩阵有 n 个重根, 例 5.1 中, $\lambda=1$ 为 2 重根.

②即便 $\mathbf{A}$ 为实方阵, 其特征值也可能是复数.

例如 $\mathbf{A}=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 则 $|\lambda\mathbf{E}-\mathbf{A}|=\lambda^2+1$, $\mathbf{A}$ 的特征值为 $\lambda=\pm i$.

① n 个特征值中有可能有相同的, 称为重特征值, 即是 $f(\lambda)=0$ 的重根. 如 n 阶单位矩阵有 n 个重根, 例 5.1 中, $\lambda=1$ 为 2 重根.

②即便 $\mathbf{A}$ 为实方阵, 其特征值也可能是复数.

例如 $\mathbf{A}=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, 则 $|\lambda\mathbf{E}-\mathbf{A}|=\lambda^2+1$, $\mathbf{A}$ 的特征值为 $\lambda=\pm i$.

③根据多项式理论, 实矩阵的复特征值是成对出现的.

定理 1 设 $\xi_1, \xi_2, \cdots, \xi_m$ 分别是 $\mathbf{A}$ 的属于互不相等的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 的特征向量, 则 $\xi_1, \xi_2, \cdots, \xi_m$ 线性无关.

证　用数学归纳法.

当 $m=1$, 结论成立 (因 $\xi_1 \neq \mathbf{0}$) .

假设 $m=k$ 时结论成立, 当 $m=k+1$ 时, 设

$$a_1\xi_1 + a_2\xi_2 + \cdots + a_k\xi_k + a_{k+1}\xi_{k+1} = \mathbf{0}, \quad (5-4)$$

则

$$\mathbf{A}(a_1\xi_1 + a_2\xi_2 + \cdots a_k\xi_k + a_{k+1}\xi_{k+1}) = \mathbf{0},$$

即

$$a_1\lambda_1\xi_1 + a_2\lambda_2\xi_2 + \cdots + a_k\lambda_k\xi_k + a_{k+1}\lambda_{k+1}\xi_{k+1} = \mathbf{0} \quad (5-5)$$

将 (5-4) 式乘以 λ_{k+1}, 再减去 (5-5) 式得

$$a_1(\lambda_{k+1}-\lambda_1)\xi_1+a_2(\lambda_{k+1}-\lambda_2)\xi_2+\cdots+a_k(\lambda_{k+1}-\lambda_k)\xi_k=\mathbf{0}.$$

因为 $\xi_1,\xi_2,\cdots,\xi_k$ 线性无关, 故 $a_i(\lambda_{k+1}-\lambda_i)=0$, 而 $\lambda_{k+1}\neq\lambda_i$, 所以 $a_i=0$, $i=1,2,\cdots,k$. 代入 (5-4) 式, 得 $a_{k+1}\xi_{k+1}=\mathbf{0}$. 因为 $\xi_{k+1}\neq\mathbf{0}$, 所以 $a_{k+1}=0$, 故 $\xi_1,\xi_2,\cdots,\xi_{k+1}$ 线性无关.

另证　由 $\mathbf{A}\xi = \lambda\xi$ 得 $\mathbf{A}^k\xi = \lambda^k\xi$, $k = 1, 2, \cdots$. 设

$$k_1\xi_1 + \cdots + k_m\xi_m = \mathbf{0},$$

只需证 $k_1 = \cdots = k_m = 0$. 等式两边同时左乘 $\mathbf{A}^s$, $s = 0, 1, \cdots, m-1$, 则有

$$\mathbf{A}^s(k_1\xi_1 + \cdots + k_m\xi_m) = \lambda_1^s k_1\xi_1 + \cdots + \lambda_m^s k_m\xi_m$$

$$= (k_1\xi_1, \cdots, k_m\xi_m)\begin{pmatrix} \lambda_1^s \\ \vdots \\ \lambda_m^s \end{pmatrix} = \mathbf{0}.$$

记

$$\mathbf{K}=(k_{ij})_{m\times m}=\begin{pmatrix}1 & \lambda_1 & \cdots & \lambda_1^{m-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1}\\ \vdots & \vdots & & \vdots\\ 1 & \lambda_m & \cdots & \lambda_m^{m-1}\end{pmatrix},$$

有 $(k_1\xi_1,\cdots,k_m\xi_m)(k_{ij})_{m\times m}=\mathbf{O}$. 因为 $\det(\mathbf{K})$ 是范德蒙德行列式, $\lambda_1,\cdots,\lambda_m$ 互不相等, 所以 $|\mathbf{K}|\neq 0$, 从而 $\mathbf{K}$ 可逆. 故 $(k_1\xi_1,\cdots,k_m\xi_m)=(\mathbf{0},\cdots,\mathbf{0})$, 得 $k_i\xi_i=\mathbf{0}$, $1\leqslant i\leqslant m$. 因 $\xi_i\neq\mathbf{0}$ $(1\leqslant i\leqslant m)$, 所以 $k_i=0$, $1\leqslant i\leqslant m$, 从而 $\xi_1,\xi_2,\cdots,\xi_m$ 线性无关.

推论 若 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 n 阶方阵 $\mathbf{A}$ 的不同特征值, 而 $\xi_{i1}, \xi_{i2}, \cdots, \xi_{ir_i}$ 是 $\mathbf{A}$ 的属于特征值 λ_i 的 r_i 个线性无关的特征向量 $(i = 1, 2, \cdots, m)$, 则向量组

$$\xi_{11}, \xi_{12}, \cdots, \xi_{1r_1}, \xi_{21}, \cdots, \xi_{2r_2}, \cdots, \xi_{m1}, \cdots, \xi_{mr_m}$$

也线性无关.

证　设

$$\sum_{j=1}^{r_1} c_{1j}\xi_{1j} + \sum_{j=1}^{r_2} c_{2j}\xi_{2j} + \cdots + \sum_{j=1}^{r_m} c_{mj}\xi_{mj} = \mathbf{0}, \quad (5-6)$$

记 $\mathbf{p}_i = \sum_{j=1}^{r_i} c_{ij}\xi_{ij}$, $i = 1, 2, \cdots, m$, 上式可写为

$$\mathbf{p}_1 + \mathbf{p}_2 + \cdots + \mathbf{p}_m = \mathbf{0},$$

其中 $\mathbf{p}_i$ 是对应于 λ_i 的特征向量或零向量, $i = 1, 2, \cdots, r$

证　设

$$\sum_{j=1}^{r_1} c_{1j}\xi_{1j} + \sum_{j=1}^{r_2} c_{2j}\xi_{2j} + \cdots + \sum_{j=1}^{r_m} c_{mj}\xi_{mj} = \mathbf{0}, \quad (5-6)$$

记 $\mathbf{p}_i = \sum_{j=1}^{r_i} c_{ij}\xi_{ij}$, $i = 1, 2, \cdots, m$, 上式可写为

$$\mathbf{p}_1 + \mathbf{p}_2 + \cdots + \mathbf{p}_m = \mathbf{0},$$

其中 $\mathbf{p}_i$ 是对应于 λ_i 的特征向量或零向量, $i = 1, 2, \cdots, m$. 由定理 5.1 知, $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_m$ 均不是特征向量 (否则若它们中有一个或几个是特征向量, 则由其线性无关性可知它们之和不等于 $\mathbf{0}$) , 故 $\mathbf{p}_i = \mathbf{0}$, $i = 1, 2, \cdots, m$, 由于 $\xi_{i1}, \xi_{i2}, \cdots, \xi_{ir_i}$ 线性无关, 因此 $c_{ij} = 0$, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, r_i$, 故结论成立.

设 $\mathbf{A} = (a_{ij})$ 是 n 阶矩阵, 称

$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}$$

为矩阵 A 的迹(trace).

定理 2　设 n 阶方阵 $\mathbf{A}$ 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$(重根按重数计算)，则

① $\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn} = \operatorname{tr}(\mathbf{A})$;

② $\lambda_1 \lambda_2 \cdots \lambda_n = |\mathbf{A}|$.

证 由条件，有

$$|\lambda \mathbf{E} - \mathbf{A}| = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

$$= \lambda^n - \left(\sum_{i=1}^{n} \lambda_i\right)\lambda^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i. \quad (5\text{-}7)$$

另一方面，由行列式定义，$|\lambda \mathbf{E} - \mathbf{A}|$ 中含有 λ^n, λ^{n-1} 的项只出现在主对角元素的乘积中，并且

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

故①成立；

在 (5-7) 式中令 $\lambda = 0$，得②成立.

推论　方阵 $\mathbf{A}$ 可逆当且仅当它的特征值全不为 0.

当 $\mathbf{A}$ 是上 (下) 三角矩阵或对角矩阵时, 特征值即对角线上的元素. 例如对 $\mathbf{A}=\begin{pmatrix}\lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3\end{pmatrix}$, 有

$$|\lambda\mathbf{E}-\mathbf{A}|=\begin{vmatrix}\lambda-\lambda_1 & -* & -* \\ 0 & \lambda-\lambda_2 & -* \\ 0 & 0 & \lambda-\lambda_3\end{vmatrix}=(\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3).$$

性质 1　设 λ 是 $\mathbf{A}=(a_{ij})_{n\times n}$ 的特征值, 则

①若 $\mathbf{A}$ 可逆, 则 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

性质 1 设 λ 是 $\mathbf{A}=(a_{ij})_{n\times n}$ 的特征值, 则

①若 $\mathbf{A}$ 可逆, 则 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

② λ^2 是 $\mathbf{A}^2$ 的特征值, 进一步地, 有

性质 1 设 λ 是 $\mathbf{A}=(a_{ij})_{n\times n}$ 的特征值, 则

①若 $\mathbf{A}$ 可逆, 则 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

② λ^2 是 $\mathbf{A}^2$ 的特征值, 进一步地, 有

若 $\varphi(x)=a_0+a_1x+\cdots+a_sx^s$, 则 $\varphi(\lambda)$ 是

$$\varphi(\mathbf{A})=a_0\mathbf{E}+a_1\mathbf{A}+\cdots+a_s\mathbf{A}^s$$

的特征值.

性质 1　设 λ 是 $\mathbf{A}=(a_{ij})_{n\times n}$ 的特征值, 则

①若 $\mathbf{A}$ 可逆, 则 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

② λ^2 是 $\mathbf{A}^2$ 的特征值, 进一步地, 有

若 $\varphi(x)=a_0+a_1x+\cdots+a_sx^s$, 则 $\varphi(\lambda)$ 是

$$\varphi(\mathbf{A})=a_0\mathbf{E}+a_1\mathbf{A}+\cdots+a_s\mathbf{A}^s$$

的特征值.

③若 $\lambda\neq 0$, 则 $\frac{1}{\lambda}|\mathbf{A}|$ 是 $\mathbf{A}^*$ 的特征值.

性质 1　设 λ 是 $\mathbf{A}=(a_{ij})_{n\times n}$ 的特征值, 则

①若 $\mathbf{A}$ 可逆, 则 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

② λ^2 是 $\mathbf{A}^2$ 的特征值, 进一步地, 有

若 $\varphi(x)=a_0+a_1x+\cdots+a_sx^s$, 则 $\varphi(\lambda)$ 是

$$\varphi(\mathbf{A})=a_0\mathbf{E}+a_1\mathbf{A}+\cdots+a_s\mathbf{A}^s$$

的特征值.

③若 $\lambda\neq 0$, 则 $\frac{1}{\lambda}|\mathbf{A}|$ 是 $\mathbf{A}^*$ 的特征值.

④ $\mathbf{A}^{\mathrm{T}}$ 的特征值与 $\mathbf{A}$ 的特征值相同.

证 ①若 $\mathbf{A}$ 可逆, 由定理 5.2 的推论可知 $\lambda \neq 0$. 设 ξ 是对应的特征向量, 由

$$\mathbf{A}\xi = \lambda\xi,$$

两边同时左乘 $\mathbf{A}^{-1}$ 可得

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

故

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

$\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值.

② λ 是方阵 $\mathbf{A}$ 的特征值, ξ 是对应的特征向量, 有 $\mathbf{A}\xi=\lambda\xi$, 两边同时左乘 $\mathbf{A}$, 得

$$\mathbf{A}^2\xi=\lambda\mathbf{A}\xi=\lambda^2\xi.$$

类似可得 $\mathbf{A}^k\xi=\lambda^k\xi$, 进一步有

$$\begin{aligned}\varphi(\mathbf{A})\xi&=(a_0\mathbf{E}+a_1\mathbf{A}+\cdots+a_s\mathbf{A}^s)\xi\\&=(a_0+a_1\lambda+\cdots+a_s\lambda^s)\xi=\varphi(\lambda)\xi.\end{aligned}$$

③设 $\lambda \neq 0$, 对应的特征向量为 ξ, 则 $A\xi = \lambda\xi$, 由于 $A^*A = |A|E$, 两边右乘 ξ 得

$$A^*A\xi = |A|\,\xi$$

$$\Rightarrow \quad \lambda A^*\xi = |A|\,\xi$$

$$\Rightarrow \quad A^*\xi = \frac{1}{\lambda}|A|\,\xi,$$

故 $\frac{1}{\lambda}|A|$ 是 A^* 的特征值.

③设 $\lambda \neq 0$, 对应的特征向量为 ξ, 则 $A\xi = \lambda\xi$, 由于 $A^*A = |A|E$, 两边右乘 ξ 得

$A^*A\xi = |A|\,\xi$

$$\Rightarrow \quad \lambda A^*\xi = |A|\,\xi$$

$$\Rightarrow \quad A^*\xi = \frac{1}{\lambda}|A|\,\xi,$$

故 $\frac{1}{\lambda}|A|$ 是 A^* 的特征值.

④因

$$\left|\lambda\mathbf{E} - \mathbf{A}^{\mathrm{T}}\right| = \left|(\lambda\mathbf{E} - \mathbf{A})^{\mathrm{T}}\right| = |\lambda\mathbf{E} - \mathbf{A}|,$$

故 $\mathbf{A}^{\mathrm{T}}$ 的特征值与 $\mathbf{A}$ 的特征值相同.

例 3　若 $\mathbf{A} = \mathbf{A}_{n\times n}$ 满足 $\mathbf{A}^2 = \mathbf{E}$, 证明：$\mathbf{A}$ 的特征值只能为 ± 1.

例 3 若 $\mathbf{A} = \mathbf{A}_{n\times n}$ 满足 $\mathbf{A}^2 = \mathbf{E}$, 证明：$\mathbf{A}$ 的特征值只能为 ± 1.

证 设 λ 是 $\mathbf{A}$ 的特征值, ξ 是对应的特征向量, 则

$$\mathbf{A}\xi = \lambda\xi,$$

有

$$\mathbf{A}^2\xi = \lambda\mathbf{A}\xi = \lambda^2\xi,$$

于是 $(1-\lambda^2)\xi = \mathbf{0}$, 故 $\lambda = \pm 1$.

例 4　已知 $\mathbf{A}=\mathbf{A}_{3\times 3}$, 且 $\mathbf{E}-\mathbf{A}$, $2\mathbf{E}-\mathbf{A}$ 和 $3\mathbf{E}-\mathbf{A}$ 均不可逆.

(1) 证明：$\mathbf{E}+2\mathbf{A}$ 可逆.

(2) 求 $|\mathbf{A}|$ 和 $\mathrm{tr}\mathbf{A}$.

证　(1) 由条件知

$$|\mathbf{E}-\mathbf{A}|=0,\ |2\mathbf{E}-\mathbf{A}|=0,\ |3\mathbf{E}-\mathbf{A}|=0,$$

故 1, 2, 3 均为 $\mathbf{A}$ 的特征值, 由于 $\mathbf{A}$ 为 3 阶矩阵, 只有 3 个特征值, 故 $-\frac{1}{2}$ 不是 $\mathbf{A}$ 的特征值, 从而

$$|\mathbf{E}+2\mathbf{A}|=\left|-2(-\frac{1}{2}\mathbf{E}-\mathbf{A})\right|=(-2)^3\left|-\frac{1}{2}\mathbf{E}-\mathbf{A}\right|\neq 0,$$

得 $\mathbf{E}+2\mathbf{A}$ 可逆.

证　(1) 由条件知

$$|\mathbf{E}-\mathbf{A}|=0,\ |2\mathbf{E}-\mathbf{A}|=0,\ |3\mathbf{E}-\mathbf{A}|=0,$$

故 1, 2, 3 均为 $\mathbf{A}$ 的特征值, 由于 $\mathbf{A}$ 为 3 阶矩阵, 只有 3 个特征值, 故 $-\frac{1}{2}$ 不是 $\mathbf{A}$ 的特征值, 从而

$$|\mathbf{E}+2\mathbf{A}|=\left|-2(-\frac{1}{2}\mathbf{E}-\mathbf{A})\right|=(-2)^3\left|-\frac{1}{2}\mathbf{E}-\mathbf{A}\right|\neq 0,$$

得 $\mathbf{E}+2\mathbf{A}$ 可逆.

(2) 由定理 5.2 知

$$|\mathbf{A}|=\prod_{i=1}^{3}\lambda_i=1\cdot2\cdot3=6,\mathrm{tr}\mathbf{A}=\sum_{i=1}^{n}a_{ii}=\sum_{i=1}^{3}\lambda_i=1+2+3=6.$$

例 5　设 $\mathbf{A}^*$ 为 3 阶方阵 $\mathbf{A}$ 的伴随矩阵, $\mathbf{A}^*$ 的特征值为 $1, -1, -4$, 求 $\mathbf{A}+2\mathbf{E}$ 的特征值.

例 5 设 $\mathbf{A}^*$ 为 3 阶方阵 $\mathbf{A}$ 的伴随矩阵, $\mathbf{A}^*$ 的特征值为 $1, -1, -4$, 求 $\mathbf{A} + 2\mathbf{E}$ 的特征值.

解 由于 $|A^*| = 1 \times (-1) \times (-4) = 4 \neq 0$, 故 $|A| \neq 0$, A^* 可逆, 由 $AA^* = |A|E$, 得 $A = |A|(A^*)^{-1}$, 且 $|A|^2 = |A^*| = 4$, 所以 $|A| = \pm 2$.

设矩阵 A 的特征值为 $\lambda_1, \lambda_2, \lambda_3$, 则 $|A| = \lambda_1\lambda_2\lambda_3 = \pm 2$. 得矩阵 A^* 的特征值为

$$\frac{1}{\lambda_1}|A| = 1, \qquad \frac{1}{\lambda_2}|A| = -1, \qquad \frac{1}{\lambda_3}|A| = -4,$$

(1) 当 $|A| = 2$ 时解得 A 的特征值为

$$\lambda_1 = 2,\ \lambda_2 = -2,\ \lambda_3 = -\frac{1}{2}.$$

记 $\varphi(x) = x + 2$, 得矩阵 $\varphi(A) = A + 2E$ 的特征值为

$$\varphi(\lambda_1) = 4,\ \varphi(\lambda_2) = 0,\ \varphi(\lambda_3) = \frac{3}{2}.$$

(1) 当 $|A|=2$ 时解得 A 的特征值为

$$\lambda_1=2,\ \lambda_2=-2,\ \lambda_3=-\frac{1}{2}.$$

记 $\varphi(x)=x+2$, 得矩阵 $\varphi(A)=A+2E$ 的特征值为

$$\varphi(\lambda_1)=4,\ \varphi(\lambda_2)=0,\ \varphi(\lambda_3)=\frac{3}{2}.$$

(2) 当 $|A|=-2$ 时解得 A 的特征值为

$$\lambda_1=-2,\ \lambda_2=2,\ \lambda_3=\frac{1}{2},$$

记 $\varphi(x)=x+2$, 得矩阵 $\varphi(A)=A+2E$ 的特征值为

$$\varphi(\lambda_1)=0,\ \varphi(\lambda_2)=4,\ \varphi(\lambda_3)=\frac{5}{2}.$$

定义 1 设 $\mathbf{A}, \mathbf{B}$ 都是 n 阶方阵, 若存在可逆矩阵 $\mathbf{P}$, 使得

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}, \quad (5-8)$$

则称**A 相似于 B**, 记作 $\mathbf{A} \sim \mathbf{B}$, 也则称 $\mathbf{B}$ 是 $\mathbf{A}$ 的相似矩阵(similar matrix).

对 $\mathbf{A}$ 进行运算 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$, 称为对 **A** 进行相似变换(similar transformation), 可逆矩阵 $\mathbf{P}$ 称为把 $\mathbf{A}$ 变成 $\mathbf{B}$ 的相似变换矩阵.

两个矩阵之间的相似关系具有三条性质：

(1) 自反性：$\mathbf{A} \sim \mathbf{A}$.

两个矩阵之间的相似关系具有三条性质:

(1) 自反性: $\mathbf{A} \sim \mathbf{A}$.

(2) 对称性: 若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{B} \sim \mathbf{A}$.

两个矩阵之间的相似关系具有三条性质：

(1) 自反性：$\mathbf{A} \sim \mathbf{A}$.

(2) 对称性：若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{B} \sim \mathbf{A}$.

(3) 传递性：若 $\mathbf{A} \sim \mathbf{B}$, $\mathbf{B} \sim \mathbf{C}$, 则 $\mathbf{A} \sim \mathbf{C}$.

两个矩阵之间的相似关系具有三条性质：

(1) 自反性：$\mathbf{A} \sim \mathbf{A}$.

(2) 对称性：若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{B} \sim \mathbf{A}$.

(3) 传递性：若 $\mathbf{A} \sim \mathbf{B}$, $\mathbf{B} \sim \mathbf{C}$, 则 $\mathbf{A} \sim \mathbf{C}$.

若存在对角矩阵 Λ, 使得 $\mathbf{A}$ 和对角矩阵 Λ 相似, 则称 $\mathbf{A}$可对角化(diagonalization).

两个矩阵之间的相似关系具有三条性质：

(1) 自反性：$\mathbf{A} \sim \mathbf{A}$.

(2) 对称性：若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{B} \sim \mathbf{A}$.

(3) 传递性：若 $\mathbf{A} \sim \mathbf{B}$, $\mathbf{B} \sim \mathbf{C}$, 则 $\mathbf{A} \sim \mathbf{C}$.

若存在对角矩阵 Λ, 使得 $\mathbf{A}$ 和对角矩阵 Λ 相似, 则称 $\mathbf{A}$可对角化(diagonalization).

若 $\mathbf{A}$ 相似于对角阵 $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 则 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda$, 即 $\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}^{-1}$, 于是 $\mathbf{A}^k = \mathbf{P}\Lambda^k\mathbf{P}^{-1}$, 因 $\Lambda^k = \mathrm{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k)$, 这样就可计算出 $\mathbf{A}^k$.

性质 1 若 $\mathbf{A} \sim \mathbf{B}$, 相似变换矩阵为 P, 即 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则有

① A 与 B 有相同的行列式与相同的秩, 即 $|\mathbf{A}| = |\mathbf{B}|$, $R(A) = R(B)$;

性质 1　若 $\mathbf{A} \sim \mathbf{B}$, 相似变换矩阵为 P, 即 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则有

① A 与 B 有相同的行列式与相同的秩, 即 $|\mathbf{A}| = |\mathbf{B}|$, $R(A) = R(B)$;

②当 A 可逆时,$A^{-1} \sim B^{-1}$, $A^* \sim B^*$, 相似变换矩阵皆为 P;

性质 1 若 $\mathbf{A} \sim \mathbf{B}$, 相似变换矩阵为 P, 即 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则有

① A 与 B 有相同的行列式与相同的秩, 即 $|\mathbf{A}| = |\mathbf{B}|$, $R(A) = R(B)$;

②当 A 可逆时,$A^{-1} \sim B^{-1}$, $A^* \sim B^*$, 相似变换矩阵皆为 P;

③ $\mathbf{A}^n \sim \mathbf{B}^n$, $k\mathbf{A} \sim k\mathbf{B}$, 其中 n 为正整数, k 为任意实数.

性质 1 若 $\mathbf{A} \sim \mathbf{B}$, 相似变换矩阵为 P, 即 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则有

① A 与 B 有相同的行列式与相同的秩, 即 $|\mathbf{A}| = |\mathbf{B}|$, $R(A) = R(B)$;

②当 A 可逆时,$A^{-1} \sim B^{-1}$, $A^* \sim B^*$, 相似变换矩阵皆为 P;

③ $\mathbf{A}^n \sim \mathbf{B}^n$, $k\mathbf{A} \sim k\mathbf{B}$, 其中 n 为正整数, k 为任意实数.

④ $f(\mathbf{A}) \sim f(\mathbf{B})$, 其中

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

为任意多项式.

证　①应用矩阵乘积的行列式性质, 对 $B=P^{-1}AP$ 两边取行列式, 得

$$|B|=\left|P^{-1}\right||A||P|=|P|^{-1}|A||P|=|A|,$$

由于矩阵 A 左乘可逆矩阵 P^{-1} 与右乘可逆矩阵 P, 其秩不变 (定理 2.9 推论) , 所以 A 与 B 有相同的秩.

证　①应用矩阵乘积的行列式性质, 对 $B=P^{-1}AP$ 两边取行列式, 得

$$|B|=\left|P^{-1}\right||A||P|=|P|^{-1}|A||P|=|A|,$$

由于矩阵 A 左乘可逆矩阵 P^{-1} 与右乘可逆矩阵 P, 其秩不变 (定理 2.9 推论) , 所以 A 与 B 有相同的秩.

②当 A 可逆时, B 可逆. 对 $P^{-1}AP=B$ 两边同时取逆

$$(P^{-1}AP)^{-1}=P^{-1}A^{-1}P=B^{-1},$$

表明 $A^{-1}\sim B^{-1}$, 相似变换矩阵为 P. 由于 $|A|=|B|$, 上式两边分别乘以 $|A|$, $|B|$ 得

$$|A|P^{-1}A^{-1}P=|B|B^{-1}\quad\Leftrightarrow\quad P^{-1}\left(|A|A^{-1}\right)P=|B|B^{-1},$$

因 $|A|A^{-1}=A^{*}$, $|B|B^{-1}=B^{*}$, 故 $\mathbf{P}^{-1}(\mathbf{A}^{*})P=B^{*}$, 即 $A^{*}\sim B^{*}$, 相似变换矩阵仍为 P.

③对任何正整数 n, 有

$$\mathbf{B}^n = (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdot(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdots\cdots(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathbf{P}^{-1}\mathbf{A}^n\mathbf{P},$$

故 $\mathbf{A}^n \sim \mathbf{B}^n$.

易证 $k\mathbf{A} \sim k\mathbf{B}$.

③对任何正整数 n, 有

$$\mathbf{B}^n = (\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdot(\mathbf{P}^{-1}\mathbf{A}\mathbf{P})\cdots\cdots(\mathbf{P}^{-1}\mathbf{A}\mathbf{P}) = \mathbf{P}^{-1}\mathbf{A}^n\mathbf{P},$$

故 $\mathbf{A}^n \sim \mathbf{B}^n$.

易证 $k\mathbf{A} \sim k\mathbf{B}$.

④ $f(\mathbf{B}) = a_n\mathbf{B}^n + a_{n-1}\mathbf{B}^{n-1} + \cdots + a_1\mathbf{B} + a_0\mathbf{E}$

$$= a_n\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P} + a_{n-1}\mathbf{P}^{-1}\mathbf{A}^{n-1}\mathbf{P} + \cdots + a_1\mathbf{P}^{-1}\mathbf{A}\mathbf{P} + a_0\mathbf{P}^{-1}\mathbf{E}\mathbf{P}$$

$$= \mathbf{P}^{-1}(a_n\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \cdots + a_1\mathbf{A} + a_0\mathbf{E})\mathbf{P}$$

$$= \mathbf{P}^{-1}f(\mathbf{A})\mathbf{P},$$

③对任何正整数 n, 有

$$\mathbf{B}^n=(\mathbf{P}^{-1}\mathbf{AP})\cdot(\mathbf{P}^{-1}\mathbf{AP})\cdots\cdots(\mathbf{P}^{-1}\mathbf{AP})=\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P},$$

故 $\mathbf{A}^n\sim\mathbf{B}^n$.

易证 $k\mathbf{A}\sim k\mathbf{B}$.

④ $f(\mathbf{B})=a_n\mathbf{B}^n+a_{n-1}\mathbf{B}^{n-1}+\cdots+a_1\mathbf{B}+a_0\mathbf{E}$

$=a_n\mathbf{P}^{-1}\mathbf{A}^n\mathbf{P}+a_{n-1}\mathbf{P}^{-1}\mathbf{A}^{n-1}\mathbf{P}+\cdots+a_1\mathbf{P}^{-1}\mathbf{AP}+a_0\mathbf{P}^{-1}\mathbf{EP}$

$=\mathbf{P}^{-1}(a_n\mathbf{A}^n+a_{n-1}\mathbf{A}^{n-1}+\cdots+a_1\mathbf{A}+a_0\mathbf{E})\mathbf{P}$

$=\mathbf{P}^{-1}f(\mathbf{A})\mathbf{P},$

得 $f(\mathbf{A})\sim f(\mathbf{B})$.

定理 1 若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征多项式相同, 从而 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征值也相同.

证 由 $\mathbf{A} \sim \mathbf{B}$, 存在可逆阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 故

$$|\lambda\mathbf{E} - \mathbf{B}|$$

$$= \left|\mathbf{P}^{-1}\lambda\mathbf{E}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right|$$

定理 1 若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征多项式相同, 从而 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征值也相同.

证 由 $\mathbf{A} \sim \mathbf{B}$, 存在可逆阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 故

$$
\begin{aligned}
&|\lambda \mathbf{E} - \mathbf{B}| \\
&\quad = \left|\mathbf{P}^{-1}\lambda \mathbf{E}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right| \\
&\quad = \left|\mathbf{P}^{-1}(\lambda \mathbf{E} - \mathbf{A})\mathbf{P}\right|
\end{aligned}
$$

定理 1　若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征多项式相同, 从而 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征值也相同.

证　　由 $\mathbf{A} \sim \mathbf{B}$, 存在可逆阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 故

$$
\begin{aligned}
&|\lambda \mathbf{E} - \mathbf{B}| \\
&\quad = \left|\mathbf{P}^{-1}\lambda \mathbf{E}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right| \\
&\quad = \left|\mathbf{P}^{-1}(\lambda \mathbf{E} - \mathbf{A})\mathbf{P}\right| \\
&\quad = \left|\mathbf{P}^{-1}\right| |\lambda \mathbf{E} - \mathbf{A}| \, |\mathbf{P}|
\end{aligned}
$$

定理 1 若 $\mathbf{A} \sim \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征多项式相同, 从而 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征值也相同.

证 由 $\mathbf{A} \sim \mathbf{B}$, 存在可逆阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 故

$$
\begin{aligned}
|\lambda \mathbf{E} - \mathbf{B}|
&= \left|\mathbf{P}^{-1}\lambda \mathbf{E}\mathbf{P} - \mathbf{P}^{-1}\mathbf{A}\mathbf{P}\right| \\
&= \left|\mathbf{P}^{-1}(\lambda \mathbf{E} - \mathbf{A})\mathbf{P}\right| \\
&= \left|\mathbf{P}^{-1}\right| |\lambda \mathbf{E} - \mathbf{A}| |\mathbf{P}| \\
&= |\lambda \mathbf{E} - \mathbf{A}|.
\end{aligned}
$$

注意 该结论反过来不对, 例如 $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ 特征多项式相同, 但不相似.

推论 若 n 阶方阵 $\mathbf{A}$ 与对角矩阵$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\mathbf{A}$ 的所有 n 个特征值.

推论 若 n 阶方阵 $\mathbf{A}$ 与对角矩阵$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\mathbf{A}$ 的所有 n 个特征值.

证 显然 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是对角矩阵 Λ 的特征值. 因为相似矩阵有相同的特征值, $\mathbf{A}$ 与对角阵 Λ 相似, 所以 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\mathbf{A}$ 的 n 个特征值.

若 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征向量之间的关系, 有:

η 是 $\mathbf{A}$ 的对应于 λ 的特征向量

$\Leftrightarrow P^{-1}\eta$ 是 $\mathbf{B}$ 的对应于 λ 的特征向量.

若 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 则 $\mathbf{A}$ 与 $\mathbf{B}$ 的特征向量之间的关系, 有:

η 是 $\mathbf{A}$ 的对应于 λ 的特征向量

$\Leftrightarrow P^{-1}\eta$ 是 $\mathbf{B}$ 的对应于 λ 的特征向量.

事实上, 有

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

$$\Leftrightarrow \mathbf{P}^{-1}\mathbf{A}\eta = \lambda\mathbf{P}^{-1}\eta$$

$$\Leftrightarrow \mathbf{P}^{-1}\mathbf{A}(\mathbf{P}\mathbf{P}^{-1})\eta = \lambda\left(\mathbf{P}^{-1}\eta\right)$$

$$\Leftrightarrow \mathbf{B}\left(\mathbf{P}^{-1}\eta\right) = \lambda\left(\mathbf{P}^{-1}\eta\right).\ (\eta \neq \mathbf{0} \text{ 时 } \mathbf{P}^{-1}\eta \neq \mathbf{0})$$

例 1　设 $\mathbf{A}$ 为 3 阶矩阵, ξ_1,ξ_2,ξ_3 是 3 个线性无关的向量, 且有

$$\mathbf{A}\xi_1=\xi_1,\ \mathbf{A}\xi_2=\xi_1+2\xi_2,\ \mathbf{A}\xi_3=2\xi_2+3\xi_3,$$

求矩阵 $\mathbf{A}$ 的特征值, 并用 ξ_1,ξ_2,ξ_3 来表示 $\mathbf{A}$ 的特征向量.

解　解法 1：设 $\mathbf{P}=(\xi_1,\xi_2,\xi_3)$, 则有

$$\mathbf{AP}=(\xi_1,\xi_2,\xi_3)\begin{pmatrix}1&1&0\\0&2&2\\0&0&3\end{pmatrix}=\mathbf{PB}.$$

因 ξ_1,ξ_2,ξ_3 线性无关, 故 $\mathbf{P}$ 可逆, 于是有 $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{B}$, 即 $\mathbf{A}\sim\mathbf{B}$.

下面先求 $\mathbf{B}$ 的特征值和特征向量. 由

$$|\lambda\mathbf{E}-\mathbf{B}|=\begin{vmatrix}\lambda-1&-1&0\\0&\lambda-2&-2\\0&0&\lambda-3\end{vmatrix}=(\lambda-1)(\lambda-2)(\lambda-3)=0,$$

可得 $\mathbf{B}$ 的特征值为 $\lambda_1=1,\lambda_2=2,\lambda_3=3$.

再求解齐次方程组 $(\lambda\mathbf{E}-\mathbf{B})\mathbf{x}=\mathbf{0}$, 可得矩阵 $\mathbf{B}$ 的对应于 $\lambda_1=1, \lambda_2=2, \lambda_3=3$ 的特征向量分别为

$$\eta_1=\begin{pmatrix}1\\0\\0\end{pmatrix}, \eta_2=\begin{pmatrix}1\\1\\0\end{pmatrix}, \eta_3=\begin{pmatrix}1\\2\\1\end{pmatrix}.$$

因为由 $\mathbf{B}\eta=\lambda\eta$, 可得 $\mathbf{AP}\eta=\lambda\mathbf{P}\eta$, 故 $\mathbf{A}$ 的特征值为 $\lambda_1=1, \lambda_2=2, \lambda_3=3$, 对应的特征向量分别为 $\mathbf{P}\eta_1, \mathbf{P}\eta_2, \mathbf{P}\eta_3$, 即 ξ_1, $\xi_1+\xi_2$, $\xi_1+2\xi_2+\xi_3$.

解法 2：将题给三个等式分别乘以常数 a,b,c 后相加得

$$a A\xi_1 + bA\xi_2 + cA\xi_3 = A(a\xi_1 + b\xi_2 + c\xi_3)$$
$$= a\xi_1 + b(\xi_1 + 2\xi_2) + c(2\xi_2 + 3\xi_3),$$

令

$$a\xi_1 + b(\xi_1 + 2\xi_2) + c(2\xi_2 + 3\xi_3) = \lambda(a\xi_1 + b\xi_2 + c\xi_3),$$

此式等价于

$$((1-\lambda)a + b)\xi_1 + ((2-\lambda)b + 2c)\xi_2 + (3-\lambda)c\xi_3 = \mathbf{0}.$$

由于 ξ_1, ξ_2, ξ_3 线性无关, 所以

$$\begin{cases} (1-\lambda)a + b = 0, \\ (2-\lambda)b + 2c = 0, \quad (*) \\ (3-\lambda)c = 0, \end{cases}$$

当 $\lambda = 1$ 时, 由 $(*)$ 式解得 $b = 0$, $c = 0$, 取 $a = 1$, 得 A 有特征值 $\lambda = 1$ 对应的特征向量为 ξ_1.

由于 ξ_1, ξ_2, ξ_3 线性无关, 所以

$$\begin{cases} (1-\lambda)a + b = 0, \\ (2-\lambda)b + 2c = 0, \quad (*) \\ (3-\lambda)c = 0, \end{cases}$$

当 $\lambda = 1$ 时, 由 $(*)$ 式解得 $b = 0$, $c = 0$, 取 $a = 1$, 得 A 有特征值 $\lambda = 1$ 对应的特征向量为 ξ_1.

当 $\lambda = 2$ 时, 由 $(*)$ 式解得 $c = 0$, $a = b$, 取 $a = 1$, 得 A 有特征值 $\lambda = 2$ 对应的特征向量为 $\xi_1 + \xi_2$.

由于 ξ_1,ξ_2,ξ_3 线性无关, 所以

$$\begin{cases} (1-\lambda)a+b=0, \\ (2-\lambda)b+2c=0, \quad (*) \\ (3-\lambda)c=0, \end{cases}$$

当 $\lambda=1$ 时, 由 $(*)$ 式解得 $b=0$, $c=0$, 取 $a=1$, 得 A 有特征值 $\lambda=1$ 对应的特征向量为 ξ_1.

当 $\lambda=2$ 时, 由 $(*)$ 式解得 $c=0$, $a=b$, 取 $a=1$, 得 A 有特征值 $\lambda=2$ 对应的特征向量为 $\xi_1+\xi_2$.

当 $\lambda=3$ 时, 由 $(*)$ 式解得 $2a=b=2c$, 取 $c=1$, 得 A 有特征值 $\lambda=3$ 对应的特征向量为 $\xi_1+2\xi_2+\xi_3$.

我们知道, 如果 $f(\lambda)$ 是矩阵 $\mathbf{A}$ 的特征多项式, 若 λ_0 是矩阵 $\mathbf{A}$ 的特征值, 则 $f(\lambda_0) = 0$. 如果把 λ_0 换成矩阵 $\mathbf{A}$ 的话, 该等式仍然成立, 即下面结论:

若 $f(\lambda)$ 为矩阵 $\mathbf{A}$ 的特征多项式, 则矩阵 $\mathbf{A}$ 的多项式 $f(\mathbf{A}) = \mathbf{O}$.

如果矩阵 $\mathbf{A}$ 和对角阵 Λ 相似时很容易证明, 这是因为:

$$f(\mathbf{A}) = \mathbf{P}f(\Lambda)\mathbf{P}^{-1} = \mathbf{P}\begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} P^{-1} = \mathbf{O}.$$

不是每个矩阵都相似于对角矩阵的, 例如 $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

事实上, 若存在可逆矩阵 $\mathbf{U}$, 使得 $\mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$,

从而 $\lambda_1 = \lambda_2 = 1$, 可得 $\mathbf{A} = \mathbf{E}$, 矛盾.

定理 2 n 阶方阵 $\mathbf{A}$ 相似于对角阵 (即 A 可对角化) 的充要条件是 $\mathbf{A}$ 有 n 个线性无关的特征向量.

定理 2 n 阶方阵 $\mathbf{A}$ 相似于对角阵 (即 A 可对角化) 的充要条件是 $\mathbf{A}$ 有 n 个线性无关的特征向量.

证 必要性. 若存在可逆矩阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, 该式可写成 $\mathbf{AP} = \mathbf{P}\Lambda$. 记 $\mathbf{P} = (\mathbf{p}_1, \cdots, \mathbf{p}_n)$, 则有

$$\mathbf{AP} = \mathbf{A}(\mathbf{p}_1, \cdots, \mathbf{p}_n) = (\mathbf{A}\mathbf{p}_1, \cdots, \mathbf{A}\mathbf{p}_n)$$

$$= (\mathbf{p}_1, \cdots, \mathbf{p}_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = (\lambda_1\mathbf{p}_1, \cdots, \lambda_n\mathbf{p}_n),$$

$$\Longleftrightarrow \quad Ap_j = \lambda_j p_j \quad (j = 1, 2, \cdots, n).$$

因为 $\mathbf{P}$ 可逆, 故 $\mathbf{p}_1, \cdots, \mathbf{p}_n$ 线性无关, 且 $\mathbf{p}_i$ 是 $\mathbf{A}$ 的对应于 λ_i 的特征向量 $(i = 1, 2, \cdots, n)$.

充分性. 若 $\mathbf{A}$ 有 n 个线性无关的特征向量 $\mathbf{p}_1, \cdots, \mathbf{p}_n$ 满足 $\mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i$, 记

$$\mathbf{P} = (\mathbf{p}_1, \cdots, \mathbf{p}_n), \ \Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n),$$

由必要性证明的推导过程倒推上去, 即可得 $\mathbf{A}$ 相似于对角阵.

推论　若 n 阶矩阵 $\mathbf{A}$ 有 n 个互不相等的特征值, 则 $\mathbf{A}$ 可对角化.

推论 若 n 阶矩阵 $\mathbf{A}$ 有 n 个互不相等的特征值, 则 $\mathbf{A}$ 可对角化.

注意本推论的逆不成立.

推论 若 n 阶矩阵 $\mathbf{A}$ 有 n 个互不相等的特征值, 则 $\mathbf{A}$ 可对角化.

注意本推论的逆不成立. 例如例 5.1 中的 $\mathbf{A}$ 有 3 个线性无关的特征向量, 故 $\mathbf{A}$ 相似于对角阵, 但 $\mathbf{A}$ 的 3 个特征值并不互异.

最后再给出一个矩阵可对角化的判别准则.

λ_0 作为 $|\lambda\mathbf{E}-\mathbf{A}|=0$ 的根出现的重数, 称为 λ_0 的**代数重数**(algebraic multiplicity).

若 λ_0 是矩阵 $\mathbf{A}$ 的特征值, 则 $(\lambda_0\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$ 的解空间的维数称为 λ_0 的**几何重数**(geometric multiplicity).

定理 3　n 阶方阵 $\mathbf{A}$ 相似于对角阵的充分必要条件是：$\mathbf{A}$ 的每个特征值的几何重数和代数重数相等，即 k_i 重特征值 λ_i 有 k_i 个线性无关的特征向量，也就是

$$R(\lambda_i \mathbf{E} - \mathbf{A}) = n - k_i.$$

关于特征值 λ 的代数重数 k 与几何重数 s 有重要性质：$1 \leqslant s \leqslant k$.

对于单特征值有 $s = k = 1$.

因此可得：一个矩阵能否对角化只需判别所有重特征值的几何重数是否等于其代数重数.

例 2　设 $A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & x \\ 1 & 0 & 0 \end{pmatrix}$. 问 x 为何值时, $\mathbf{A}$ 可对角化.

例 2　设 $A=\begin{pmatrix}0&0&1\\1&1&x\\1&0&0\end{pmatrix}$. 问 x 为何值时, $\mathbf{A}$ 可对角化.

解　$|\lambda\mathbf{E}-\mathbf{A}|=(\lambda-1)^2(\lambda+1)$. 所以 $\mathbf{A}$ 的所有不同特征值为 $\lambda_1=-1, \lambda_2=1$.

对于特征值 $\lambda_1=-1$, 其几何重数与代数重数皆为 1. 所以矩阵 A 可对角化当且仅当 2 重特征值 $\lambda_2=1$ 的几何重数等于 2, 此等价于 $R(\lambda_2 E-A)=1$. 由初等行变换

$$\lambda_2\mathbf{E}-\mathbf{A}=\mathbf{E}-\mathbf{A}=\begin{pmatrix}1 & 0 & -1\\ -1 & 0 & -x\\ -1 & 0 & 1\end{pmatrix}\xrightarrow{r}\begin{pmatrix}1 & 0 & -1\\ 0 & 0 & -x-1\\ 0 & 0 & 0\end{pmatrix}$$

得 $R(\mathbf{E}-\mathbf{A})=\begin{cases}1, & x=-1,\\ 2, & x\neq-1\end{cases}$,

所以当且仅当 $x=-1$ 时, 矩阵 A 可对角化.

若实矩阵 $\mathbf{A}$ 满足 $\mathbf{A}^{\mathsf{T}}=\mathbf{A}$, 则称 $\mathbf{A}$ 为实对称矩阵 (symmetric real matrix) .

若实矩阵 $\mathbf{A}$ 满足 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 则称 $\mathbf{A}$ 为实对称矩阵 (symmetric real matrix) .

性质 1　实对称阵的特征值必为实数.

证　设 λ 是对称矩阵 $\mathbf{A}$ 的特征值, 则存在 $\alpha \neq \mathbf{0}$, 使 $\mathbf{A}\alpha = \lambda\alpha$. 用 $\bar{\lambda}$ 表示 λ 的共轭复数, 设$\alpha = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$, 其共轭复向量 $\bar{\alpha} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n)^{\mathrm{T}}$.

对 $\mathbf{A}\alpha = \lambda\alpha$ 两边取共轭, 再取转置, 注意到 $\overline{\mathbf{A}} = \mathbf{A}$ 且 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 有 $\mathbf{A}\bar{\alpha} = \bar{\mathbf{A}}\bar{\alpha} = \overline{\mathbf{A}\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$,

得行向量的等式：$(\bar{\alpha})^{\mathrm{T}}\mathbf{A} = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}$,

证　设 λ 是对称矩阵 $\mathbf{A}$ 的特征值, 则存在 $\alpha \neq \mathbf{0}$, 使 $\mathbf{A}\alpha = \lambda\alpha$. 用 $\bar{\lambda}$ 表示 λ 的共轭复数, 设$\alpha = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$，其共轭复向量 $\bar{\alpha} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n)^{\mathrm{T}}$.

对 $\mathbf{A}\alpha = \lambda\alpha$ 两边取共轭, 再取转置, 注意到 $\overline{\mathbf{A}} = \mathbf{A}$ 且 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 有 $\mathbf{A}\bar{\alpha} = \bar{\mathbf{A}}\bar{\alpha} = \overline{\mathbf{A}\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$,

得行向量的等式：$(\bar{\alpha})^{\mathrm{T}}\mathbf{A} = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}$,

从而 $(\bar{\alpha})^{\mathrm{T}}A\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

证 设 λ 是对称矩阵 $\mathbf{A}$ 的特征值, 则存在 $\alpha \neq \mathbf{0}$, 使 $\mathbf{A}\alpha = \lambda\alpha$. 用 $\bar{\lambda}$ 表示 λ 的共轭复数, 设$\alpha = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$, 其共轭复向量 $\bar{\alpha} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n)^{\mathrm{T}}$.

对 $\mathbf{A}\alpha = \lambda\alpha$ 两边取共轭, 再取转置, 注意到 $\overline{\mathbf{A}} = \mathbf{A}$ 且 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 有 $\mathbf{A}\bar{\alpha} = \bar{\mathbf{A}}\bar{\alpha} = \overline{\mathbf{A}\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$,

得行向量的等式: $(\bar{\alpha})^{\mathrm{T}}\mathbf{A} = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}$,

从而 $(\bar{\alpha})^{\mathrm{T}}A\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

得 $\lambda(\bar{\alpha})^{\mathrm{T}}\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

证　设 λ 是对称矩阵 $\mathbf{A}$ 的特征值, 则存在 $\alpha \neq \mathbf{0}$, 使 $\mathbf{A}\alpha = \lambda\alpha$. 用 $\bar{\lambda}$ 表示 λ 的共轭复数, 设$\alpha = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$, 其共轭复向量 $\bar{\alpha} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n)^{\mathrm{T}}$.

对 $\mathbf{A}\alpha = \lambda\alpha$ 两边取共轭, 再取转置, 注意到 $\overline{\mathbf{A}} = \mathbf{A}$ 且 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 有 $\mathbf{A}\bar{\alpha} = \bar{\mathbf{A}}\bar{\alpha} = \overline{\mathbf{A}\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$,

得行向量的等式: $(\bar{\alpha})^{\mathrm{T}}\mathbf{A} = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}$,

从而 $(\bar{\alpha})^{\mathrm{T}}A\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

得 $\lambda(\bar{\alpha})^{\mathrm{T}}\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

故 $(\lambda - \bar{\lambda})(\bar{\alpha})^{\mathrm{T}}\alpha = 0$.

证　设 λ 是对称矩阵 $\mathbf{A}$ 的特征值, 则存在 $\alpha \neq \mathbf{0}$, 使 $\mathbf{A}\alpha = \lambda\alpha$. 用 $\bar{\lambda}$ 表示 λ 的共轭复数, 设 $\alpha = (a_1, a_2, \cdots, a_n)^{\mathrm{T}} \in \mathbb{C}^n$, 其共轭复向量 $\bar{\alpha} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_n)^{\mathrm{T}}$.

对 $\mathbf{A}\alpha = \lambda\alpha$ 两边取共轭, 再取转置, 注意到 $\overline{\mathbf{A}} = \mathbf{A}$ 且 $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$, 有 $\mathbf{A}\bar{\alpha} = \bar{\mathbf{A}}\bar{\alpha} = \overline{\mathbf{A}\alpha} = \overline{\lambda\alpha} = \bar{\lambda}\bar{\alpha}$,

得行向量的等式: $(\bar{\alpha})^{\mathrm{T}}\mathbf{A} = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}$,

从而 $(\bar{\alpha})^{\mathrm{T}}A\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

得 $\lambda(\bar{\alpha})^{\mathrm{T}}\alpha = \bar{\lambda}(\bar{\alpha})^{\mathrm{T}}\alpha$,

故 $(\lambda - \bar{\lambda})(\bar{\alpha})^{\mathrm{T}}\alpha = 0$.

因 $\alpha \neq \mathbf{0}$, 故 $(\bar{\alpha})^{\mathrm{T}}\alpha = \bar{a}_1 a_1 + \bar{a}_2 a_2 + \cdots + \bar{a}_n a_n > 0$, 于是 $\lambda - \bar{\lambda} = 0$, 即 λ 是实数.

因为特征向量是 $(\lambda\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$ 的非零解向量, 对于实对称阵 $\mathbf{A}$ 的任一特征值 λ, $(\lambda\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$ 的系数矩阵是实矩阵, 从而必存在它的实基础解系, 所以实对称阵 $\mathbf{A}$ 的特征向量都可以取为实向量.

性质 2　实对称阵 $\boldsymbol{A}$ 的属于不同特征值的特征向量相互正交.

性质 2 实对称阵 $\mathbf{A}$ 的属于不同特征值的特征向量相互正交.

证 设 λ_1, λ_2 是实对称阵 $\mathbf{A}$ 的不同特征值, α_i 是属于 λ_i 的特征向量, $i = 1, 2$, 则

$$\lambda_1 \alpha_1^{\mathsf{T}} = (\lambda_1 \alpha_1)^{\mathsf{T}} = (\mathbf{A}\alpha_1)^{\mathsf{T}} = \alpha_1^{\mathsf{T}} \mathbf{A}^{\mathsf{T}} = \alpha_1^{\mathsf{T}} \mathbf{A},$$

右乘向量 α_2 得 $\lambda_1 \alpha_1^{\mathsf{T}} \alpha_2 = \alpha_1^{\mathsf{T}} \mathbf{A} \alpha_2 = \lambda_2 \alpha_1^{\mathsf{T}} \alpha_2$, 即 $(\lambda_1 - \lambda_2)\alpha_1^{\mathsf{T}} \alpha_2 = 0$, 由于 $\lambda_1 - \lambda_2 \neq 0$, 从而 $\alpha_1^{\mathsf{T}} \alpha_2 = 0$, α_1 与 α_2 正交.

定理 1　设 $\mathbf{A}$ 为 n 阶实对称矩阵, 则必有正交矩阵 $\mathbf{P}$, 使

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n), \quad (5-9)$$

其中 $\lambda_1,\cdots,\lambda_n$ 是 $\mathbf{A}$ 的所有特征值. 此时称 $\mathbf{A}$正交相似(orthogonally similar) 于一个对角矩阵.

推论 设 $\mathbf{A}$ 为 n 阶实对称矩阵, λ_0 是 $\mathbf{A}$ 的特征多项式的 k 重根, 则 $R(\lambda_0\mathbf{E}-\mathbf{A})=n-k$, 从而对应于特征值 λ_0 恰好有 k 个线性无关的特征向量.

利用正交矩阵把实对称矩阵 **A** 对角化的一般步骤:

①写出 A 的特征多项式

$|\lambda\mathbf{E}-\mathbf{A}|=(\lambda-\lambda_1)^{k_1}(\lambda-\lambda_2)^{k_2}\cdots(\lambda-\lambda_s)^{k_s}$, $\lambda_i\neq\lambda_j(i\neq j)$,

求出 A 的互异特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_s$, 其中 λ_i 的代数重数为 $k_i(k_1+k_2+\cdots+k_s=n)$.

②对应于每个特征值 λ_i, 求出 $(\lambda_i\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$ 的基础解系: $\xi_{i1},\cdots,\xi_{i,k_i}$, 再把它们正交化并单位化, 得到 k_i 个两两正交的单位向量 $\mathbf{p}_{i1},\cdots,\mathbf{p}_{i,k_i}$.

③把②中求出的 n 个两两正交的单位向量合在一起, 以它们为列向量构成正交矩阵 **P**, 譬如令

$$\mathbf{P}=(\mathbf{p}_{11},\cdots,\mathbf{p}_{1,k_1},\cdots,\mathbf{p}_{s1},\cdots,\mathbf{p}_{s,k_s}),$$

由定理 4.5 知 **P** 是正交矩阵, 且

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \Lambda$$

$$= \mathrm{diag}\,(\underbrace{\lambda_1, \cdots \lambda_1}_{k_1}, \underbrace{\lambda_2, \cdots, \lambda_2}_{k_2}, \cdots, \underbrace{\lambda_s, \cdots, \lambda_s}_{k_s}).$$

由定理 4.5 知 $\mathbf{P}$ 是正交矩阵, 且

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \Lambda$$

$$= \operatorname{diag}(\underbrace{\lambda_1, \cdots \lambda_1}_{k_1}, \underbrace{\lambda_2, \cdots, \lambda_2}_{k_2}, \cdots, \underbrace{\lambda_s, \cdots, \lambda_s}_{k_s}).$$

注意 上面的对角矩阵中的主对角线上的特征值的排列次序和正交矩阵 $\mathbf{P}$ 中列向量的排列次序要对应一致.

例 1　已知 $\mathbf{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}$, 求一个正交矩阵 $\mathbf{P}$, 使 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda$ 为对角矩阵.

解 由 $\mathbf{A}$ 的特征多项式

$$|\lambda \mathbf{E}-\mathbf{A}| = \begin{vmatrix} \lambda-4 & 0 & 0 \\ 0 & \lambda-3 & -1 \\ 0 & -1 & \lambda-3 \end{vmatrix} = (\lambda-2)(\lambda-4)^2,$$

得 $\mathbf{A}$ 的特征值为 $\lambda_1 = 2, \lambda_2 = \lambda_3 = 4.$

解　由 $\mathbf{A}$ 的特征多项式

$$|\lambda\mathbf{E}-\mathbf{A}| = \begin{vmatrix} \lambda-4 & 0 & 0 \\ 0 & \lambda-3 & -1 \\ 0 & -1 & \lambda-3 \end{vmatrix} = (\lambda-2)(\lambda-4)^2,$$

得 $\mathbf{A}$ 的特征值为 $\lambda_1 = 2, \lambda_2 = \lambda_3 = 4.$

对 $\lambda_1 = 2$, 解 $(\lambda_1\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 得基础解系 $\alpha_1 = (0, 1, -1)^{\mathrm{T}}$, 单位化得 $\mathbf{p}_1 = (0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})^{\mathrm{T}}.$

对 $\lambda_2 = \lambda_3 = 4$, 解 $(\lambda_2 \mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 得基础解系 $\alpha_2 = (1, 0, 0)^{\mathsf{T}}$, $\alpha_2 = (0, 1, 1)^{\mathsf{T}}$. 两者已正交, 单位化可得 $\mathbf{p}_2 = (1, 0, 0)^{\mathsf{T}}$, $\mathbf{p}_3 = (0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^{\mathsf{T}}$, 令

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix},$$

则 $\mathbf{P}$ 是正交矩阵, 且 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda = \begin{pmatrix} 2 & & \\ & 4 & \\ & & 4 \end{pmatrix}$.

例 2　已知 3 阶实对称矩阵 $\mathbf{A}$ 的特征值为 $-6, 3, 3$, 且 $\xi = (2, -2, 1)^{\mathrm{T}}$ 是 $\mathbf{A}$ 的属于特征值 $\lambda_1 = -6$ 的特征向量, 求 $\mathbf{A}$.

例 2 已知 3 阶实对称矩阵 $\mathbf{A}$ 的特征值为 $-6, 3, 3$, 且 $\xi=(2,-2,1)^{\mathrm{T}}$ 是 $\mathbf{A}$ 的属于特征值 $\lambda_1=-6$ 的特征向量, 求 $\mathbf{A}$.

解法 1 因为 A 是实对称矩阵, 必可对角化, 所以特征值 $\lambda_2=\lambda_3=3$ 对应于两个线性无关的特征向量 α_1,α_2, 它们位于与 $\xi=(2,-2,1)^{\mathrm{T}}$ 垂直的平面上, 因此与 $\xi=(2,-2,1)^{\mathrm{T}}$ 垂直的任意非零向量 β 可表示为 $\beta=k_1\alpha_1+k_2\alpha_2(k_1,k_2$ 不全为 0) , 故 β 一定是属于 $\lambda_2=\lambda_3=3$ 的特征向量. 令 $\beta=(x_1,x_2,x_3)$, 得 $2x_1-2x_2+x_3=0$. 取 $x_2=1,x_3=0$ 得一特征向量 $\xi_1=(1,1,0)^{\mathrm{T}}$. 属于 3 的另一个与 ξ 和 ξ_1 都正交的特征向量 $\xi_2=(x_1,x_2,x_3)^{\mathrm{T}}$ 可由下式得到:

$$2x_1-2x_2+x_3=0,\ x_1+x_2=0,$$

联立得另一特征向量为 $\xi_2=(-1,1,4)^T$. 将正交的 ξ, ξ_1, ξ_2 单位化：

$$\eta_1=\frac{\xi}{\|\xi\|}=\begin{pmatrix}\frac{2}{3}\\-\frac{2}{3}\\\frac{1}{3}\end{pmatrix},\eta_2=\frac{\xi_1}{\|\xi_1\|}=\begin{pmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\end{pmatrix},\eta_3=\frac{\xi_2}{\|\xi_2\|}=\begin{pmatrix}-\frac{\sqrt{2}}{6}\\\frac{\sqrt{2}}{6}\\\frac{2\sqrt{2}}{3}\end{pmatrix}$$

记 $\mathbf{P}=(\eta_1,\eta_2,\eta_3)$, 则 $\mathbf{P}$ 为正交矩阵且 $\mathbf{P}^{-1}\mathbf{AP}=$ $\begin{pmatrix}-6&&\\&3&\\&&3\end{pmatrix}=\Lambda$, 于是

$$\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}^{-1} = \mathbf{P}\Lambda\mathbf{P}^{\mathsf{T}}$$

$$= \begin{pmatrix} \frac{2}{3} & \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{6} \\ -\frac{2}{3} & \frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{6} \\ \frac{1}{3} & 0 & \frac{2\sqrt{2}}{3} \end{pmatrix} \begin{pmatrix} -6 & & \\ & 3 & \\ & & 3 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{\sqrt{2}}{6} & \frac{\sqrt{2}}{6} & \frac{2\sqrt{2}}{3} \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 4 & -2 \\ 4 & -1 & 2 \\ -2 & 2 & 2 \end{pmatrix}.$$

解法 2 由于与 $\xi=(2,-2,1)^{\mathrm{T}}$ 垂直的任意非零向量 β 一定是属于 $\lambda_2=\lambda_3=3$ 的特征向量 (理由见解法 1)．记 $\beta=(x_1,x_2,x_3)$, 则 $2x_1-2x_2+x_3=0$, 易得此方程的两个解 $\beta_1=(1,1,0)^{\mathrm{T}}$, $\beta_2=(1,0,-2)^{\mathrm{T}}$, 此即属于 $\lambda_2=\lambda_3=3$ 的 2 个线性无关的特征向量. 记

$$P=(\xi,\beta_1,\beta_2)=\begin{pmatrix}2&1&1\\-2&1&0\\1&0&-2\end{pmatrix},\text{ 则}$$

$$P^{-1}AP=\Lambda=\begin{pmatrix}-6&0&0\\0&3&0\\0&0&3\end{pmatrix}.$$

下面先用初等行变换求 P^{-1}：

$$\begin{pmatrix} \mathbf{P} & \mathbf{E} \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ -2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -2 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r} \begin{pmatrix} 1 & 0 & 0 & \frac{2}{9} & -\frac{2}{9} & \frac{1}{9} \\ 0 & 1 & 0 & \frac{4}{9} & \frac{5}{9} & \frac{2}{9} \\ 0 & 0 & 1 & \frac{1}{9} & -\frac{1}{9} & -\frac{4}{9} \end{pmatrix},$$

所以 $P^{-1} = \frac{1}{9}\begin{pmatrix} 2 & -2 & 1 \\ 4 & 5 & 2 \\ 1 & -1 & -4 \end{pmatrix}$, 因此所求的矩阵为

$$A = P\Lambda P^{-1} = \frac{1}{9}\begin{pmatrix} 2 & 1 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & -2 \end{pmatrix}\begin{pmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 4 & 5 & 2 \\ 1 & -1 & -4 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 4 & -2 \\ 4 & -1 & 2 \\ -2 & 2 & 2 \end{pmatrix}.$$

注 1 可以通过直接凑出一组正交的属于 λ_2 的特征向量，从而避施密特正交化过程.

对方程 $2x_1 - 2x_2 + x_3 = 0$，下面构造一组正交的属于 λ_2 的特征向量：

要与 ξ 正交，可取 $\xi_1 = (1, 1, 0)^T$. 此时若要取 $\xi_2 = (x_1, x_2, x_3)^T$ 与 ξ_1 正交，则应有 $\xi_2 = (1, -1, \boxed{})^T$，还得与 ξ 正交，从而必须满足方程 $2x_1 - 2x_2 + x_3 = 0$，所以 $\xi_2 = (1, -1, -4)^T$，从而得到一组正交的属于 λ_2 的特征向量：

$$\xi_1 = (1, 1, 0)^T, \xi_2 = (1, -1, -4)^T.$$

注 1 可以通过直接凑出一组正交的属于 λ_2 的特征向量, 从而避免施密特正交化过程.

对方程 $2x_1 - 2x_2 + x_3 = 0$, 下面构造一组正交的属于 λ_2 的特征向量:

要与 ξ 正交, 可取 $\xi_1 = (1,1,0)^T$. 此时若要取 $\xi_2 = (x_1,x_2,x_3)^T$ 与 ξ_1 正交, 则应有 $\xi_2 = (1,-1,\sqcup)^T$, 还得与 ξ 正交, 从而必须满足方程 $2x_1 - 2x_2 + x_3 = 0$, 所以 $\xi_2 = (1,-1,-4)^T$, 从而得到一组正交的属于 λ_2 的特征向量:

$$\xi_1 = (1,1,0)^T, \xi_2 = (1,-1,-4)^T.$$

类似地, 可取 $\xi_1 = (1,0,-2)^T$, $\xi_2 = (4,5,2)^T$. 或者 $\xi_1 = (0,1,2)^T$, $\xi_2 = (-5,-4,2)^T$ 等.

注 2　通常情况下 P 不唯一, 譬如对于重特征值, 对应的特征向量可以排序不同, 所以 P 不唯一. 但是 $\mathbf{A}$ 是唯一的, 而且与对角矩阵中特征值的排序无关, 这由定理 5.4 可看出. 矩阵 $\mathbf{A}$ 可对角化时, $A = P\Lambda P^{-1}$, 即 $\mathbf{A}$ 由它的特征值与对应的特征向量唯一确定.

例 3　设实对称矩阵 $\mathbf{A}$ 和 $\mathbf{B}$ 相似, 证明：存在正交矩阵 $\mathbf{P}$, 使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$.

例 3 设实对称矩阵 $\mathbf{A}$ 和 $\mathbf{B}$ 相似，证明：存在正交矩阵 $\mathbf{P}$，使得 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$.

证 由于 $\mathbf{A}$ 和 $\mathbf{B}$ 相似，则 $\mathbf{A}$ 和 $\mathbf{B}$ 有相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$，故由定理 5.6，分别存在正交矩阵 $\mathbf{P}_1$，$\mathbf{P}_2$，使得

$$\mathbf{P}_1^{-1}A\mathbf{P}_1=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)=\mathbf{P}_2^{-1}\mathbf{B}\mathbf{P}_2,$$

故 $\mathbf{P}_2\mathbf{P}_1^{-1}A\mathbf{P}_1\mathbf{P}_2^{-1}=\mathbf{B}$，取 $\mathbf{P}=\mathbf{P}_1\mathbf{P}_2^{-1}$，则 $\mathbf{P}$ 是正交矩阵，且有 $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{B}$.

为看清二次曲线 $ax^2 + bxy + cy^2 = 1$ 的形状, 可以采用坐标变换

$$\begin{cases} x = x' \cos\theta - y' \sin\theta \\ y = x' \sin\theta + y' \cos\theta \end{cases}$$

化二次曲线为标准形 $a'x'^2 + b'y'^2 = 1$, 由此二次曲线的几何性质便一目了然.

第四节	二次型及其标准形
A	二次型及其标准形
B	正交变换法
C	配方法与初等变换法

定义 1 n 元二次齐次多项式

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n$$

$$+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots\cdots\cdots\cdots+a_{nn}x_n^2$$

$$=\sum_{i=1}^{n}a_{ii}x_i^2+2\sum_{1\leqslant i<j\leqslant n}a_{ij}x_ix_j \quad (5\text{-}10)$$

称为n 元二次型, 简称二次型(quadratic form). 如果 a_{ij} 都是复数, 则称 f 为复二次型. 如果系数 a_{ij} 和变量 x_i 都为实数, 则称 f 为实二次型. 我们只讨论实二次型.

如 $f(x_1,x_2,x_3)=3x_1^2-2x_2^2+x_3^2+4x_1x_2-6x_1x_3+5x_2x_3$ 是一个 3 元二次型.

令 $a_{ji}=a_{ij}$, 则 $2a_{ij}x_ix_j=a_{ij}x_ix_j+a_{ji}x_jx_i$, 从而

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

$$=a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n+\cdots$$

$$+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2$$

$$=x_1(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n)+\cdots$$

$$+x_n(a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n)$$

$$= (x_1, x_2, \cdots, x_n) \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{pmatrix}$$

$$= (x_1, x_2, \cdots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

令$\mathbf{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$, $\mathbf{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$，则二次型 f 可以表示为矩阵形式：

$$f(x_1,x_2,\cdots,x_n)$$

$$=(x_1,x_2,\cdots,x_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

$$=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x},\ (5\text{-}11)$$

其中 $\mathbf{A}$ 是对称矩阵.

二次型 (5-11) 也记为

$$f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}. \quad (5-12)$$

易见二次型 f 与对称矩阵 $\mathbf{A}$ 确立了一一对应关系, 即任给一个二次型 f 就唯一确定了一个对称矩阵 $\mathbf{A}$; 反之, 任给一个对称矩阵 $\mathbf{A}$, 也唯一确定了一个二次型 f. 称二次型 f 唯一确定的对称矩阵 $\mathbf{A}$ 为二次型 f 的矩阵, 也称二次型 f 是对称矩阵 $\mathbf{A}$ 的二次型. 矩阵 $\mathbf{A}$ 的秩称为二次型 f 的秩.

二次型 $f_1 = x_1^2 + 2x_2^2 - 3x_3^2 + 4x_1x_2 - 6x_2x_3$ 的矩阵

$\mathbf{A}_1 = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 2 & -3 \\ 0 & -3 & -3 \end{pmatrix}$； 二次型 $f_2 = 2x_1^2 + 2x_2^2 + 5x_3^2$

的矩阵 $\mathbf{A}_2 = \begin{pmatrix} 2 & & \\ & 2 & \\ & & 5 \end{pmatrix}$；

对称矩阵 $\mathbf{A}_3 = \begin{pmatrix} 4 & -1 & 3 \\ -1 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix}$ 确定的二次型为

$$f_3 = 4x_1^2 - 2x_2^2 + x_3^2 - 2x_1x_2 + 6x_1x_3.$$

对线性变换：

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n, \\ \cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases} \quad (5-13)$$

其矩阵形式为

$$\mathbf{x} = \mathbf{C}\mathbf{y}, \quad (5-14)$$

其中 $\mathbf{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, $\mathbf{y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}$,

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}.$$

设有一个 n 元二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 引进新的一组变量 $y_1, y_2, \cdots, y_n$, 并把 $x_1, x_2, \cdots, x_n$ 用线性变换 (5-13) 或 (5-14) 表示. 代入 $f(x_1, x_2, \cdots, x_n)$, 得到 $y_1, \cdots, y_n$ 的一个二次型 $g(y_1, \cdots, y_n)$, 即

$$f(x_1, \cdots, x_n) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{y}^{\mathrm{T}}\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}\,\mathbf{y} = g(y_1, \cdots, y_n),$$

于是二次型 $g(y_1, \cdots, y_n)$ 对应的矩阵为 $\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}$. 因

$$(\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}(\mathbf{C}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C},$$

故 $\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}$ 仍为对称阵.

称二次型经过可逆线性变换后得到的仅包含平方项的二次型为二次型的标准形(standard (或 canonical) form of a quadratic form).

形如：$x_1^2+x_2^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+q}^2$ 的二次型称为二次型的规范标准形(normal form of a quadratic form)(简称为规范形) .

标准二次型的矩阵是对角矩阵.

规范二次型的矩阵称为规范对角矩阵, 形式如下:

$$\begin{pmatrix} E_p & & \\ & -E_q & \\ & & O \end{pmatrix} = \mathrm{diag}(\underbrace{1,\cdots,1}_{p},\underbrace{-1,\cdots,-1}_{q},\underbrace{0,\cdots,0}_{n-p-q})$$

$(p+q=R(A))$.

规范二次型的矩阵称为规范对角矩阵, 形式如下:

$$\begin{pmatrix} E_p & & \\ & -E_q & \\ & & O \end{pmatrix} = \mathrm{diag}(\underbrace{1,\cdots,1}_{p},\underbrace{-1,\cdots,-1}_{q},\underbrace{0,\cdots,0}_{n-p-q})$$

$(p+q=R(A))$.

定义 2　两个 n 阶矩阵 $\mathbf{A}$ 和 $\mathbf{B}$, 若存在可逆矩阵 $\mathbf{C}$, 使得 $\mathbf{C}^{\mathrm{T}}\mathbf{A}\mathbf{C}=\mathbf{B}$, 则称 $\mathbf{A}$ 合同于 $\mathbf{B}$, 记为 $\mathbf{A}\simeq\mathbf{B}$, 称 C 为合同变换矩阵.

矩阵间的合同关系满足自反性, 对称性和传递性:

矩阵间的合同关系满足自反性, 对称性和传递性:

(1) 自反性　　$\mathbf{A}$ 合同于 $\mathbf{A}$, 因为 $\mathbf{E}^\top \mathbf{A}\mathbf{E} = \mathbf{A}$;

矩阵间的合同关系满足自反性, 对称性和传递性:

(1) 自反性　　$\mathbf{A}$ 合同于 $\mathbf{A}$, 因为 $\mathbf{E}^{\mathsf{T}}\mathbf{A}\mathbf{E} = \mathbf{A}$;

(2) 对称性　　若 $\mathbf{A}$ 合同于 $\mathbf{B}$, 即存在可逆矩阵 $\mathbf{C}$, 使得 $\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C} = \mathbf{B}$, 从而 $(\mathbf{C}^{-1})^{\mathsf{T}}\mathbf{B}\mathbf{C}^{-1} = \mathbf{A}$, 即 $\mathbf{B}$ 合同于 $\mathbf{A}$;

矩阵间的合同关系满足自反性, 对称性和传递性:

(1) 自反性　　$\mathbf{A}$ 合同于 $\mathbf{A}$, 因为 $\mathbf{E}^{\top}\mathbf{A}\mathbf{E}=\mathbf{A}$;

(2) 对称性　　若 $\mathbf{A}$ 合同于 $\mathbf{B}$, 即存在可逆矩阵 $\mathbf{C}$, 使得 $\mathbf{C}^{\top}\mathbf{A}\mathbf{C}=\mathbf{B}$, 从而 $(\mathbf{C}^{-1})^{\top}\mathbf{B}\mathbf{C}^{-1}=\mathbf{A}$, 即 $\mathbf{B}$ 合同于 $\mathbf{A}$;

(3) 传递性　　若 $\mathbf{A}$ 合同于 $\mathbf{B}$, $\mathbf{B}$ 合同于 $\mathbf{C}$, 即存在可逆矩阵 $\mathbf{P},\mathbf{Q}$, 使得 $\mathbf{P}^{\top}\mathbf{A}\mathbf{P}=\mathbf{B}$, $\mathbf{Q}^{\top}\mathbf{B}\mathbf{Q}=\mathbf{C}$, 从而 $(\mathbf{P}\mathbf{Q})^{\top}\mathbf{A}(\mathbf{P}\mathbf{Q})=\mathbf{C}$, 得 $\mathbf{A}$ 合同于 $\mathbf{C}$.

矩阵间的合同关系满足自反性, 对称性和传递性:

(1) 自反性　　$\mathbf{A}$ 合同于 $\mathbf{A}$, 因为 $\mathbf{E}^{\mathsf{T}}\mathbf{A}\mathbf{E} = \mathbf{A}$;

(2) 对称性　　若 $\mathbf{A}$ 合同于 $\mathbf{B}$, 即存在可逆矩阵 $\mathbf{C}$, 使得 $\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C} = \mathbf{B}$, 从而 $(\mathbf{C}^{-1})^{\mathsf{T}}\mathbf{B}\mathbf{C}^{-1} = \mathbf{A}$, 即 $\mathbf{B}$ 合同于 $\mathbf{A}$;

(3) 传递性　　若 $\mathbf{A}$ 合同于 $\mathbf{B}$, $\mathbf{B}$ 合同于 $\mathbf{C}$, 即存在可逆矩阵 $\mathbf{P}, \mathbf{Q}$, 使得 $\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P} = \mathbf{B}$, $\mathbf{Q}^{\mathsf{T}}\mathbf{B}\mathbf{Q} = \mathbf{C}$, 从而 $(\mathbf{P}\mathbf{Q})^{\mathsf{T}}\mathbf{A}(\mathbf{P}\mathbf{Q}) = \mathbf{C}$, 得 $\mathbf{A}$ 合同于 $\mathbf{C}$.

因合同关系具有对称性, 所以 $\mathbf{A}$ 合同于 $\mathbf{B}$ 也说成 $\mathbf{A}$ 与 $\mathbf{B}$ 合同, 或称 $\mathbf{A}$, $\mathbf{B}$ 是合同矩阵(congruence of matrix).

性质 1 设 $A \simeq B$, 合同变换矩阵为 C, 即 $C^{\mathrm{T}}AC = B$, 则有:

① A 与 B 有相同的秩, 即 $R(A) = R(B)$;

性质 1　设 $A \simeq B$, 合同变换矩阵为 C, 即 $C^{\mathrm{T}}AC = B$, 则有:

① A 与 B 有相同的秩, 即 $R(A) = R(B)$;

② $A^{\mathrm{T}} \simeq B^{\mathrm{T}}$, 合同变换矩阵仍为 C;

性质 1　设 $A \simeq B$, 合同变换矩阵为 C, 即 $C^{\mathrm{T}}AC = B$, 则有:

① A 与 B 有相同的秩, 即 $R(A) = R(B)$;

② $A^{\mathrm{T}} \simeq B^{\mathrm{T}}$, 合同变换矩阵仍为 C;

③当 A 可逆时, $A^{-1} \simeq B^{-1}$, $A^{*} \simeq B^{*}$.

证　①由于矩阵 A 左乘可逆矩阵 C^{T} 与右乘可逆矩阵 C, 其秩不变 (定理 2.9 推论) , 故 A 与 B 有相同的秩.

证　①由于矩阵 A 左乘可逆矩阵 C^{T} 与右乘可逆矩阵 C, 其秩不变 (定理 2.9 推论) , 故 A 与 B 有相同的秩.

②对 $C^{\mathrm{T}}AC = B$ 两边同时取转置得

$$(C^{\mathrm{T}}AC)^{\mathrm{T}} = C^{\mathrm{T}}A^{\mathrm{T}}C = B^{\mathrm{T}},$$

此式表明 $A^{\mathrm{T}} \simeq B^{\mathrm{T}}$, 合同变换矩阵为 C.

③当 A 可逆时, 由①得 B 也可逆. 对 $C^{\mathrm{T}}AC = B$ 两边同时取逆得

$$(C^{\mathrm{T}}AC)^{-1} = C^{-1}A^{-1}(C^{\mathrm{T}})^{-1} = C^{-1}A^{-1}(C^{-1})^{\mathrm{T}} = B^{-1}.$$

取 $Q = (C^{-1})^{\mathrm{T}}$, 则 Q 为可逆矩阵, 上式化为 $Q^{\mathrm{T}}A^{-1}Q = B^{-1}$, 故 $A^{-1} \simeq B^{-1}$.

等式 $Q^{\mathrm{T}}A^{-1}Q = B^{-1}$ 两边分别乘以 $|C|^2\,|A|$ 与 $|B|$, 由于 $|C|^2\,|A| = |B|$, 所以

$$|C|^2\,|A|(Q^{\mathrm{T}}A^{-1}Q) = (|C|\,Q)^{\mathrm{T}}\left(|A|\,A^{-1}\right)(|C|\,Q) = |B|\,B^{-1}$$

因为 $|A|A^{-1} = A^*$, $|B|B^{-1} = B^*$, 令 $M = |C|Q$, 代入上式得 $M^{\mathrm{T}}A^*M = B^*$, 故 $A^* \simeq B^*$.

第四节	二次型及其标准形
A	二次型及其标准形
B	正交变换法
C	配方法与初等变换法

定理 1 对于 n 元二次型 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 存在正交变换 $\mathbf{x}=\mathbf{P}\mathbf{y}$, 将该二次型化为标准形:

$$
\begin{aligned}
f(\mathbf{x}) &= \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} \\
&= \mathbf{y}^{\mathrm{T}}(\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P})\mathbf{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2,
\end{aligned}
$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是实对称矩阵 $\mathbf{A}$ 的特征值, $\mathbf{P}$ 的列向量组 $\mathbf{p}_1,\mathbf{p}_2,\cdots,\mathbf{p}_n$ 是标准正交向量组, 且 $\mathbf{A}\mathbf{p}_i=\lambda_i\mathbf{p}_i(i=1,2,\cdots,n)$.

推论 任给二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 总存在可逆变换 $\mathbf{x} = \mathbf{C}\mathbf{z}$, 使得 $f(\mathbf{C}\mathbf{z})$ 为规范形.

证 按定理 5.7, 有

$$
\begin{aligned}
f(\mathbf{x}) &= \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = f(\mathbf{P}\mathbf{y}) \\
&= \mathbf{y}^{\mathrm{T}}(\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P})\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,
\end{aligned}
$$

设二次型的秩为 r, 则特征值中恰好 r 个不为零, 设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 不等于零, $\lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$, 令

$$
\mathbf{K} = \begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix}, \text{其中 } k_i = \begin{cases} \dfrac{1}{\sqrt{|\lambda_i|}}, & i \leqslant r \\ 1, & i > r \end{cases}
$$

则 $\mathbf{K}$ 可逆, 变换 $\mathbf{y}=\mathbf{K}\mathbf{z}$ 把 $f(\mathbf{P}\mathbf{y})$ 化为

$f(\mathbf{x})=f(\mathbf{P}\mathbf{y})=f(\mathbf{P}\mathbf{K}\mathbf{z})=\mathbf{z}^{\mathsf{T}}\mathbf{K}^{\mathsf{T}}\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P}\mathbf{K}\mathbf{z}=\mathbf{z}^{\mathsf{T}}\mathbf{K}^{\mathsf{T}}\Lambda\mathbf{K}\mathbf{z}$

而

$$\mathbf{K}^{\mathsf{T}}\Lambda\mathbf{K}=\operatorname{diag}\left(\frac{\lambda_1}{|\lambda_1|},\frac{\lambda_2}{|\lambda_2|},\cdots,\frac{\lambda_r}{|\lambda_r|},0,\cdots,0\right),$$

记 $\mathbf{C}=\mathbf{P}\mathbf{K}$, 即知可逆变换 $\mathbf{x}=\mathbf{C}\mathbf{z}$ 把原二次型 f 化成规范形

$$f(\mathbf{C}\mathbf{z})=\frac{\lambda_1}{|\lambda_1|}z_1^2+\frac{\lambda_2}{|\lambda_2|}z_2^2+\cdots+\frac{\lambda_r}{|\lambda_r|}z_r^2.$$

用正交变换将二次型化为标准形的步骤如下：

(1) 写出二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ 中的实对称矩阵 $\mathbf{A}$；

用正交变换将二次型化为标准形的步骤如下：

(1) 写出二次型 $f(\mathbf{x}) = \mathbf{x}^{\top}\mathbf{A}\mathbf{x}$ 中的实对称矩阵 $\mathbf{A}$；

(2) 由特征方程 $|\lambda\mathbf{E} - \mathbf{A}| = 0$, 求得 n 个特征值 (重根按重数计算)；

用正交变换将二次型化为标准形的步骤如下：

(1) 写出二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 中的实对称矩阵 $\mathbf{A}$；

(2) 由特征方程 $|\lambda\mathbf{E} - \mathbf{A}| = 0$, 求得 n 个特征值 (重根按重数计算)；

(3) 对每个不同的特征值 λ_i, 由齐次线性方程组 $(\lambda_i\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 求得 $\mathbf{A}$ 的对应于 λ_i 的线性无关的特征向量；

用正交变换将二次型化为标准形的步骤如下：

(1) 写出二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 中的实对称矩阵 $\mathbf{A}$；

(2) 由特征方程 $|\lambda\mathbf{E} - \mathbf{A}| = 0$, 求得 n 个特征值 (重根按重数计算)；

(3) 对每个不同的特征值 λ_i, 由齐次线性方程组 $(\lambda_i\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 求得 $\mathbf{A}$ 的对应于 λ_i 的线性无关的特征向量；

(4) 将每一个 λ_i 对应的特征向量先正交化再单位化, 构成正交矩阵 $\mathbf{P}$；

用正交变换将二次型化为标准形的步骤如下：

(1) 写出二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 中的实对称矩阵 $\mathbf{A}$；

(2) 由特征方程 $|\lambda\mathbf{E} - \mathbf{A}| = 0$, 求得 n 个特征值 (重根按重数计算)；

(3) 对每个不同的特征值 λ_i, 由齐次线性方程组 $(\lambda_i\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$, 求得 $\mathbf{A}$ 的对应于 λ_i 的线性无关的特征向量；

(4) 将每一个 λ_i 对应的特征向量先正交化再单位化, 构成正交矩阵 $\mathbf{P}$；

(5) 写出相应的正交变换 $\mathbf{x} = \mathbf{P}\mathbf{y}$ 及二次型的标准形.

例 1　(1) 用正交变换 $\mathbf{x}=\mathbf{P}\mathbf{y}$ 化二次型

$$f=4x_1^2+4x_2^2+4x_3^2+4x_1x_2+4x_1x_3+4x_2x_3$$

为标准形, 并给出正交变换矩阵 $\mathbf{P}$.

(2) 判别

$$4x_1^2+4x_2^2+4x_3^2+4x_1x_2+4x_1x_3+4x_2x_3=1$$

是什么曲面?

解　(1) f 的矩阵 $\mathbf{A}=\begin{pmatrix}4&2&2\\2&4&2\\2&2&4\end{pmatrix}$, 由

$$|\lambda\mathbf{E}-\mathbf{A}| \xlongequal{r_1+r_2+r_3} \begin{vmatrix}\lambda-8&\lambda-8&\lambda-8\\-2&\lambda-4&-2\\-2&-2&\lambda-4\end{vmatrix}$$

$$=(\lambda-8)\begin{vmatrix}1&1&1\\0&\lambda-2&0\\0&0&\lambda-2\end{vmatrix}=(\lambda-8)(\lambda-2)^2=0,$$

对于 $\lambda_1=\lambda_2=2$, 解 $(2\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$. 由 $2\mathbf{E}-\mathbf{A}\xrightarrow{r}\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}$, 可得 $x_1+x_2+x_3=0$, 它的一个基础解系为：$\alpha_1=(-1,1,0)^{\mathrm{T}},\alpha_2=(-1,0,1)^{\mathrm{T}}$, 正交化得：

$$\beta_1=\alpha_1=\begin{pmatrix}-1\\1\\0\end{pmatrix},\beta_2=\alpha_2-\frac{[\beta_1,\alpha_2]}{[\beta_1,\beta_1]}\beta_1=\begin{pmatrix}-\frac{1}{2}\\-\frac{1}{2}\\1\end{pmatrix}.$$

注意 对于 $\lambda_1 = \lambda_2 = 2$, 可直接观察出它的两个正交向量如: $\alpha_1 = (1,-1,0)^{\mathrm{T}}$, $\alpha_2 = (1,1,-2)^{\mathrm{T}}$ 或 $\alpha_1 = (1,0,-1)^{\mathrm{T}}$, $\alpha_2 = (1,-2,1)^{\mathrm{T}}$ 等.

对于 $\lambda_3 = 8$, 通过解 $(8\mathbf{E} - \mathbf{A})\mathbf{x} = \mathbf{0}$ 求得 $\lambda_3 = 8$ 的一个特征向量 $\alpha_3 = (1,1,1)^{\mathrm{T}}$.

再将 $\beta_1, \beta_2, \alpha_3$ 单位化得:

$$\mathbf{p}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \mathbf{p}_2 = \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}, \mathbf{p}_3 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}.$$

令

$$\mathbf{P}=(\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3)=\begin{pmatrix}-\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix},$$

则 $\mathbf{P}$ 即为所求正交变换矩阵, 它满足$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}=\begin{pmatrix}2 & & \\ & 2 & \\ & & 8\end{pmatrix}$.

于是正交变换 $\mathbf{x}=\mathbf{P}\mathbf{y}$ 可化二次型 f 为标准形:

$$f=2y_1^2+2y_2^2+8y_3^2.$$

(2) 因为正交变换不改变空间中的向量的长度和夹角，故正交变换后曲面的形状也不会改变，从而二次曲面

$$4x_1^2+4x_2^2+4x_3^2+4x_1x_2+4x_1x_3+4x_2x_3=1$$

与

$$2y_1^2+2y_2^2+8y_3^2=1$$

表示同一个曲面：是一个椭球面.

例 2 证明：当 $\|\mathbf{x}\| = 1$ 时，实二次型 $f(\mathbf{x}) = \mathbf{x}^{\top}\mathbf{A}\mathbf{x}$ 的最大值等于 $\mathbf{A}$ 的最大特征值.

证 因 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 为 n 元实二次型, 存在正交变换 $\mathbf{x}=\mathbf{P}\mathbf{y}$, 可将 $f(\mathbf{x})$ 化为标准形:

$$f=\mathbf{y}^{\mathrm{T}}(\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P})\mathbf{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2,$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\mathbf{A}$ 的特征值.

设 λ_{i_0} 是 $\mathbf{A}$ 的最大特征值, 因为

$$y_1^2 + y_2^2 + \cdots + y_n^2$$

$$= \mathbf{y}^\mathrm{T}\mathbf{y} = \mathbf{y}^\mathrm{T}\mathbf{P}^\mathrm{T}\mathbf{P}\mathbf{y} = (\mathbf{P}\mathbf{y})^\mathrm{T}\mathbf{P}\mathbf{y} = \mathbf{x}^\mathrm{T}\mathbf{x} = \|\mathbf{x}\|^2 = 1,$$

于是

$$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

$$\leqslant \lambda_{i_0} y_1^2 + \lambda_{i_0} y_2^2 + \cdots + \lambda_{i_0} y_n^2 = \lambda_{i_0}.$$

取 $\mathbf{y}$ 为第 i_0 个基本单位向量 $\mathbf{e}_{i_0}$, 则当 $\mathbf{x} = \mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{e}_{i_0}$ 时就有

$$f = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{y}^\mathsf{T}\mathbf{P}^\mathsf{T}\mathbf{A}\mathbf{P}\mathbf{y}$$

$$= \mathbf{e}_{i_0}^\mathsf{T} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \mathbf{e}_{i_0} = \lambda_{i_0},$$

即当 $\mathbf{x} = \mathbf{P}\mathbf{e}_{i_0}$ 时, $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ 确实可以取到最大值 λ_{i_0}.

第四节	二次型及其标准形
A	二次型及其标准形
B	正交变换法
C	配方法与初等变换法

拉格朗日配方法(Lagrange method of completing the square) 化二次型为标准形, 这种方法不用求解矩阵的特征值和特征向量.

例 3　将下列二次型化为标准形, 并求所用的可逆变换矩阵 **C**.

(1) $f = x_1^2 - 3x_2^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3$;

(2) $f = -4x_1x_2 + 2x_1x_3 + 2x_2x_3$.

解 (1) $f=(x_1-x_2+x_3)^2-(x_2-x_3)^2-3x_2^2-6x_2x_3$

$$=(x_1-x_2+x_3)^2-4x_2^2-4x_2x_3-x_3^2$$

$$=(x_1-x_2+x_3)^2-(2x_2+x_3)^2$$

令 $\begin{cases} y_1= x_1 -x_2 +x_3 \\ y_2= 2x_2 +x_3 \\ y_3= x_3 \end{cases}$ 则有 $\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix} 1 & \frac{1}{2} & -\frac{3}{2}\\ 0 & \frac{1}{2} & -\frac{1}{2}\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} y_1\\ y_2\\ y_3\end{pmatrix}$

得标准形为 $f=y_1^2-y_2^2$, 所用可逆变换矩阵

$$\mathbf{C}=\begin{pmatrix} 1 & \frac{1}{2} & -\frac{3}{2}\\ 0 & \frac{1}{2} & -\frac{1}{2}\\ 0 & 0 & 1\end{pmatrix}.$$

(2) 令$\begin{cases} x_1 = y_1 + y_2 \\ x_2 = y_1 - y_2 \\ x_3 = y_3 \end{cases}$，即$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$，

则

$$f = -4(y_1 + y_2)(y_1 - y_2) + 2(y_1 + y_2)y_3 + 2(y_1 - y_2)y_3$$

$$= -4y_1^2 + 4y_1y_3 + 4y_2^2$$

$$= -4(y_1 - \tfrac{1}{2}y_3)^2 + 4y_2^2 + y_3^2.$$

再令 $\begin{cases} z_1 = y_1 \qquad -\frac{1}{2}y_3 \\ z_2 = \qquad y_2 \\ z_3 = \qquad\qquad y_3 \end{cases}$, 则有 $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

从而 $f = -4z_1^2 + 4z_2^2 + z_3^2$ 为标准形. 且由

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & -1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

得变换矩阵 $C = \begin{pmatrix} 1 & 1 & \frac{1}{2} \\ 1 & -1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$, 显然可逆.

初等变换法(method of elementary transformation):

设 $\mathbf{A}$ 为某二次型对应的矩阵, 化二次型为标准形就是寻求可逆矩阵 $\mathbf{C}$, 使 $\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C}$ 成为对角矩阵. 由于任一可逆矩阵均可表示成有限个初等矩阵的乘积, 设 $\mathbf{C}=\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_s$, 其中 $\mathbf{P}_1$, $\mathbf{P}_2$, $\cdots$, $\mathbf{P}_s$ 为初等矩阵, 则 $\mathbf{C}^{\mathsf{T}}=\mathbf{P}_s^{\mathsf{T}}\cdots\mathbf{P}_2^{\mathsf{T}}\mathbf{P}_1^{\mathsf{T}}$, 从而

$$\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C}=\mathbf{P}_s^{\mathsf{T}}\cdots\mathbf{P}_2^{\mathsf{T}}\mathbf{P}_1^{\mathsf{T}}\mathbf{A}\mathbf{P}_1\mathbf{P}_2\cdots\mathbf{P}_s.$$

因用初等矩阵左 (右) 乘某矩阵等同于对该矩阵施行一次对应的初等行 (列) 变换, 故对任意 n 阶矩阵 $\mathbf{B}$ 及初等矩阵 $\mathbf{P}$, 易得 $\mathbf{P}^{\mathsf{T}}\mathbf{B}\mathbf{P}$ 等同于对矩阵 $\mathbf{B}$ 施行一次初等行变换再施行一次相同的初等列变换, 即

$$B\longrightarrow\mathbf{P}^{\mathsf{T}}\mathbf{B}\longrightarrow\mathbf{P}^{\mathsf{T}}\mathbf{B}\mathbf{P},$$

变换 $\mathbf{B} \to \mathbf{P}^{\mathsf{T}}\mathbf{B}\mathbf{P}$ 称为对 $\mathbf{B}$ 进行一次合同变换, 其中 $\mathbf{P}$ 为初等矩阵.

从而对矩阵 $\mathbf{A}$ 同时施行一系列合同变换化为对角矩阵, 而相应地将这一系列的初等列变换施加于单位矩阵 $\mathbf{E}$, 就得到变换矩阵 $\mathbf{C}$.

具体做法:

将 n 阶单位矩阵 $\mathbf{E}$ 放在二次型的矩阵 $\mathbf{A}$ 的下边 (或右边)，形成一个 $2n \times n$(或 $n \times 2n$) 矩阵, 对矩阵 $\mathbf{A}$ 作一系列合同变换化成对角矩阵的同时, 把单位矩阵 $\mathbf{E}$ 化成了可逆变换矩阵 $\mathbf{C}$(或 $\mathbf{C}^{\mathsf{T}}$)，这就是初等变换法, 即

$$\begin{pmatrix} \mathbf{A} \\ \mathbf{E} \end{pmatrix} \longrightarrow \begin{pmatrix} \mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C} \\ \mathbf{E}\mathbf{C} \end{pmatrix} \longrightarrow \begin{pmatrix} \Lambda \\ \mathbf{C} \end{pmatrix},$$

或

$$(\mathbf{A}, \mathbf{E}) \longrightarrow \left(\mathbf{C}^{\mathsf{T}}\mathbf{A}\mathbf{C}, \mathbf{C}^{\mathsf{T}}\mathbf{E}\right) \longrightarrow \left(\Lambda, \mathbf{C}^{\mathsf{T}}\right).$$

例 4　化二次型

$$f = x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$$

成标准形, 并求所用的变换矩阵.

例 4　化二次型

$$f = x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$$

成标准形, 并求所用的变换矩阵.

解　二次型 f 对应的矩阵为

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

进行如下合同变换, 有

$$\begin{pmatrix} A \\ E \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2-r_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 3 & 5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[c_2-c_1]{} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 5 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3-r_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[c_3-c_1]{} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3-2r_2}\begin{pmatrix}1&0&0\\0&1&2\\0&0&0\\1&-1&-1\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow[c_3-2c_2]{}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\\1&-1&1\\0&1&-2\\0&0&1\end{pmatrix}.$$

故令 $\mathbf{x}=\mathbf{C}\mathbf{y}$, 其中

$$C=\begin{pmatrix}1&-1&1\\0&1&-2\\0&0&1\end{pmatrix},$$

则二次型可化为标准形

$$f=y_1^2+y_2^2.$$

例 5　化二次型

$$f = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

成标准形, 并求所用的变换矩阵.

例 5　化二次型

$$f = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

成标准形, 并求所用的变换矩阵.

解　所给二次型 f 的矩阵

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix},$$

进行如下合同变换, 有

$$(\mathbf{A},\mathbf{E})=\begin{pmatrix}0 & 1 & 1 & 1 & 0 & 0\\ 1 & 0 & -3 & 0 & 1 & 0\\ 1 & -3 & 0 & 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{r_1+r_2}\begin{pmatrix}1 & 1 & -2 & 1 & 1 & 0\\ 1 & 0 & -3 & 0 & 1 & 0\\ 1 & -3 & 0 & 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow[c_1+c_2]{}\begin{pmatrix}2 & 1 & -2 & 1 & 1 & 0\\ 1 & 0 & -3 & 0 & 1 & 0\\ -2 & -3 & 0 & 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{r_2-\frac{1}{2}r_1}\begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -2 & -3 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[c_2-\frac{1}{2}c_1]{}\begin{pmatrix} 2 & 0 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -2 & -2 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3+r_1}\begin{pmatrix} 2 & 1 & -2 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -2 & -2 & 1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow[c_3+c_1]{}\begin{pmatrix} 2 & 0 & 0 & 1 & 1 & 0 \\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & -2 & -2 & 1 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{r_3-4r_2}\begin{pmatrix}2 & 0 & 0 & 1 & 1 & 0\\ 0 & -\frac{1}{2} & -2 & -\frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & 6 & 3 & -1 & 1\end{pmatrix}$$

$$\xrightarrow[c_3-4c_2]{}\begin{pmatrix}2 & 0 & 0 & 1 & 1 & 0\\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & \frac{1}{2} & 0\\ 0 & 0 & 6 & 3 & -1 & 1\end{pmatrix}.$$

故令 $\mathbf{x}=\mathbf{C}\mathbf{y}$, 其中

$$\mathbf{C}=\begin{pmatrix} 1 & -\frac{1}{2} & 3 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

则二次型可化为标准形

$$f=2y_1^2-\frac{1}{2}y_2^2+6y_3^2.$$

第一节	方阵的特征值与特征向量
第二节	相似矩阵
第三节	实对称矩阵的对角化
第四节	二次型及其标准形
第五节	**惯性定理与正定二次型**

第五节	惯性定理与正定二次型
A	惯性定理
B	正定二次型

定理 1 (惯性定理inertial Theorem) 对于一个 n 元二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 经任意一个可逆线性变换化为标准形 $f = k_1 y_1^2 + k_2 y_2^2 + \cdots + k_n y_n^2$ 后, 标准形中正平方项的个数 p 和负平方项的个数 q 都是唯一确定的, 且 $p + q = R(\mathbf{A})$.

证　设二次型的秩为 r. 因对任何二次型, 均存在可逆线性变换使其化为标准形, 再适当地交换变量的标号(也是可逆线性变换)，设有两个可逆线性变换

$$\mathbf{x}=\mathbf{P}\mathbf{y}, \mathbf{x}=\mathbf{Q}\mathbf{z},$$

使二次型分别化为:

$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_p y_p^2 - \lambda_{p+1} y_{p+1}^2 - \cdots - \lambda_r y_r^2$, $\lambda_i > 0, i = 1, 2, \cdots, r$

$f = \mu_1 z_1^2 + \mu_2 z_2^2 + \cdots + \mu_l z_l^2 - \mu_{l+1} z_{l+1}^2 - \cdots - \mu_r z_r^2$, $\mu_i > 0, i = 1, 2, \cdots, r$

下面证明 $p = l$.

用反证法. 设 $p > l$. 由假设, 有

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_p y_p^2 - \lambda_{p+1} y_{p+1}^2 - \cdots - \lambda_r y_r^2$$
$$= \mu_1 z_1^2 + \mu_2 z_2^2 + \cdots + \mu_l z_l^2 - \mu_{l+1} z_{l+1}^2 - \cdots - \mu_r z_r^2,$$
(5-15)

令 $\mathbf{C} = \mathbf{Q}^{-1}\mathbf{P} = (c_{ij})_{n\times n}$, 有 $\mathbf{z} = \mathbf{C}\mathbf{y}$. 考查齐次线性方程组

$$\begin{cases} c_{11}y_1 + c_{12}y_2 + \cdots c_{1n}y_n = 0, \\ \quad \vdots \\ c_{l1}y_1 + c_{l2}y_2 + \cdots c_{ln}y_n = 0, \\ y_{p+1} = 0, \\ \quad \vdots \\ y_n = 0 \end{cases}, \quad (5-16)$$

方程组的前 l 方程即为 $z_1=0, z_2=0, \cdots, z_l=0$. 此方程组有 n 个未知数, 有 $l+(n-p)=n-(p-l)<n$ 个方程, 所以有非零解, 设此非零解为 $\mathbf{y}^*=(y_1^*, y_2^*, \cdots, y_p^*, 0, \cdots, 0)^{\mathrm{T}}$, 由式 (5-16), 对应的有 $\mathbf{z}^*=\mathbf{C}\mathbf{y}^*=(0, \cdots, 0, z_{l+1}^*, \cdots, z_n^*)^{\mathrm{T}}$, 将此二式代入式 (5-15) 中, 左端为

$$\lambda_1 {y_1^*}^2+\lambda_2 {y_2^*}^2+\cdots+\lambda_p {y_p^*}^2>0,$$

而右端为

$$=-\mu_{l+1} {z_{l+1}^*}^2-\cdots-\mu_r {z_r^*}^2 \leqslant 0,$$

得到矛盾. 故 $p \leqslant l$. 同理可证 $p \geqslant l$, 故 $p=l$.

推论 1 任给二次型 $f(\mathbf{x}) = \mathbf{x}^{\top}\mathbf{A}\mathbf{x}$, 其规范形是唯一的.

推论 1 任给二次型 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 其规范形是唯一的.

二次型 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 的标准形中的正平方项的个数 p 称为二次型 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$(或 $\mathbf{A}$) 的正惯性指数(positive index of inertia), 负平方项的个数 q 称为二次型 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$(或 $\mathbf{A}$) 的负惯性指数(negative index of inertia). 若 $R(\mathbf{A})=r=p+q$ 时, 则 f 的规范形可确定为

$$f=y_1^2+y_2^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2.$$

其中 p 为正惯性指数, q 为负惯性指数.

(1) 两个 n 阶实对称矩阵合同的充要条件是它们有相同的正惯性指数和负惯性指数.

(1) 两个 n 阶实对称矩阵合同的充要条件是它们有相同的正惯性指数和负惯性指数.

(2) 全体 n 阶实对称矩阵, 按其合同规范形 (不考虑 $+1, -1, 0$ 的排列顺序) 分类, 共有 $\frac{(n+1)(n+2)}{2}$ 类.

例 1　将二次型 $f = -2x_1x_2 + 2x_1x_3 + 2x_2x_3$ 化为标准形, 并求相应的可逆变换和二次型的正惯性指数.

解　该二次型对应的矩阵为 $\mathbf{A} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

下面求 $\mathbf{A}$ 的特征值和特征向量. 由

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

解得特征值为：$\lambda = 1$(二重) , $\lambda = -2$.

对 $\lambda = 1$, 可求得相互正交的特征向量为

$\mathbf{p}_1 = \frac{1}{\sqrt{2}}(1,0,1)^{\mathrm{T}}$, $\mathbf{p}_2 = \frac{1}{\sqrt{6}}(-1,2,1)^{\mathrm{T}}$.

对 $\lambda = -2$, 可求得单位特征向量为：$\mathbf{p}_3 = \frac{1}{\sqrt{3}}(1,1,-1)^{\mathrm{T}}$

令

$$\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{pmatrix},$$

则有 $\mathbf{P}^{\mathrm{T}}\mathbf{AP} = \mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$

令 $\mathbf{x} = \mathbf{P}\mathbf{y}$, 则

$$f(\mathbf{P}\mathbf{y}) = \mathbf{y}^{\mathrm{T}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \mathbf{y} = y_1^2 + y_2^2 - 2y_3^2$$

为标准形. 二次型的正惯性指数为 2.

推论 2 若二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ 与二次型 $g(\mathbf{y}) = \mathbf{y}^{\mathsf{T}}\mathbf{B}\mathbf{y}$ 都是 n 元二次型，它们有相同的秩与正惯性指数，则必有可逆线性变换 $\mathbf{x} = \mathbf{P}\mathbf{y}$，使得

$$\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \mathbf{y}^{\mathsf{T}}\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P}\mathbf{y} = \mathbf{y}^{\mathsf{T}}\mathbf{B}\mathbf{y}.$$

证　设两二次型的秩为 r, 正惯性指数为 p, 则由定理 5.7 推论可知, 存在可逆线性变换

$$\mathbf{x} = \mathbf{P}_1\mathbf{z},\ \mathbf{y} = \mathbf{P}_2\mathbf{z},$$

使得

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{z}^{\mathrm{T}}\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1 z = z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2,$$

$$\mathbf{y}^{\mathrm{T}}\mathbf{B}\mathbf{y} = \mathbf{z}^{\mathrm{T}}\mathbf{P}_2^{\mathrm{T}}\mathbf{B}\mathbf{P}_2 z = z_1^2 + z_2^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2.$$

令 $\mathbf{P} = \mathbf{P}_1\mathbf{P}_2^{-1}$, 由上两式得：在线性变换 $\mathbf{x} = \mathbf{P}_1\mathbf{P}_2^{-1}\mathbf{y} = \mathbf{P}\mathbf{y}$ 下, 有

$$\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \mathbf{y}^{\mathrm{T}}\mathbf{B}\mathbf{y}.$$

第五节	惯性定理与正定二次型
A	惯性定理
B	正定二次型

定义 1 设有二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 若对 $\forall \mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n$, 有

(1) $f(\mathbf{x}) > 0$, 则称 f 为正定二次型(positive definite quadratic form), 并称对称矩阵 A 是正定矩阵(positive definite matrix);

(2) $f(\mathbf{x}) < 0$, 则称 f 为负定二次型(negative definite quadratic form), 并称对称矩阵 A 是负定矩阵(negative definite matrix);

(3) $f(\mathbf{x}) \geqslant 0$, 且至少存在一个 $\mathbf{x}_0 \neq \mathbf{0}$, 使 $f(\mathbf{x}_0) = 0$, 则称 f 为半正定二次型(positive semidefinite quadratic form), 并称对称矩阵 A 是半正定矩阵(positi semidefinite matrix);

(4) $f(\mathbf{x}) \leqslant 0$, 且至少存在一个 $\mathbf{x}_0 \neq \mathbf{0}$, 使 $f(\mathbf{x}_0) = 0$, 则称 f 为半负定二次型(negative semidefinite quadratic form), 并称对称矩阵 A 是半负定矩阵(negat semidefinite matrix).

正定和半正定以及负定和半负定二次型, 统称为有定二次型.

例如:

$f_1(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_3^2$ 是正定二次型.

$f_2(x_1,x_2,x_3)=4x_1^2+x_2^2$ 是半正定二次型.

$f_3(x_1,x_2,x_3)=-2x_1^2-2x_2^2-3x_3^2$ 是负定二次型.

而 $f_4(x_1,x_2,x_3)=x_1^2-2x_2^2-x_3^2$ 既不是正定 (或负定) 二次型, 也不是半正定 (或半负定) 二次型, 称为不定二次型.

性质 1　对二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 有

①实对称阵 $\mathbf{A}$ 正定当且仅当 $-\mathbf{A}$ 负定.

②二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 经任何可逆线性变换 $\mathbf{x} = \mathbf{P}\mathbf{y}$ 后所得的二次型

$$f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = (\mathbf{P}\mathbf{y})^{\mathrm{T}}\mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{y}^{\mathrm{T}}(\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P})\mathbf{y} = \mathbf{y}^{\mathrm{T}}\mathbf{B}\mathbf{y} = g(\mathbf{y})$$

的正定性不变.

证 ①显然.

②设二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ 正定. $\forall \mathbf{y} \in \mathbb{R}^n$ 且 $\mathbf{y} \neq \mathbf{0}$, 则对任意可逆矩阵 $\mathbf{P}$, 有 $\mathbf{x} = \mathbf{P}\mathbf{y} \neq \mathbf{0}$, 经可逆线性变换 $\mathbf{x} = \mathbf{P}\mathbf{y}$ 后, 有

$$g(\mathbf{y}) = \mathbf{y}^{\mathsf{T}}(\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P})\mathbf{y} = (\mathbf{P}\mathbf{y})^{\mathsf{T}}\mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} > 0,$$

即 $g(\mathbf{y})$ 也正定.

同理, 若 $g(\mathbf{y}) = \mathbf{y}^{\mathsf{T}}\mathbf{B}\mathbf{y}$ 正定, 则经线性变换 $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ 后得到的二次型 $f(\mathbf{x}) = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x}$ 也正定.

定理 2　设 $\mathbf{A} = \mathbf{A}_{n\times n}$ 为实对称阵, $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 则以下几个命题等价:

① $\mathbf{A}$ 正定, 或 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型;

定理 2 设 $\mathbf{A} = \mathbf{A}_{n\times n}$ 为实对称阵, $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 则以下几个命题等价:

① $\mathbf{A}$ 正定, 或 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型;

② $\mathbf{A}$ 的特征值全大于零;

定理 2 设 $\mathbf{A} = \mathbf{A}_{n\times n}$ 为实对称阵，$f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$，则以下几个命题等价：

① $\mathbf{A}$ 正定，或 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型；

② $\mathbf{A}$ 的特征值全大于零；

③ $\mathbf{A}$ 的正惯性指数为 n；

定理 2　设 $\mathbf{A} = \mathbf{A}_{n\times n}$ 为实对称阵，$f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$，则以下几个命题等价：

① $\mathbf{A}$ 正定，或 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型；

② $\mathbf{A}$ 的特征值全大于零；

③ $\mathbf{A}$ 的正惯性指数为 n；

④ $\mathbf{A}$ 合同于单位阵 $\mathbf{E}$；

定理 2 设 $\mathbf{A}=\mathbf{A}_{n\times n}$ 为实对称阵, $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$, 则以下几个命题等价:

① $\mathbf{A}$ 正定, 或 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型;

② $\mathbf{A}$ 的特征值全大于零;

③ $\mathbf{A}$ 的正惯性指数为 n;

④ $\mathbf{A}$ 合同于单位阵 $\mathbf{E}$;

⑤存在可逆阵 $\mathbf{B}$, 使得 $\mathbf{A}=\mathbf{B}^{\mathrm{T}}\mathbf{B}$.

证 ① $\Rightarrow$ ②

设 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 经正交变换 $\mathbf{x}=\mathbf{Q}\mathbf{y}$ 化为标准形：

$$f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}=\mathbf{y}^{\mathrm{T}}(\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q})\mathbf{y}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2,$$

其中 λ_i 是 $\mathbf{A}$ 的特征值. 令 $\mathbf{y}_i=\mathbf{e}_i=(0,\cdots,0,1,0,\cdots,$
则 $\mathbf{x}_i=\mathbf{Q}\mathbf{y}_i\neq\mathbf{0}$. 由 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 是正定二次型得

$$0<f(\mathbf{x}_i)=\mathbf{x}_i^{\mathrm{T}}\mathbf{A}\mathbf{x}_i=\mathbf{y}_i^{\mathrm{T}}\left(\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q}\right)\mathbf{y}_i=\lambda_i(i=1,2,\cdots,n).$$

② ⇒ ③

若 $\mathbf{A}$ 的特征值全大于零, 因 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 经正交线性变换 $\mathbf{x}=\mathbf{Q}\mathbf{y}$ 化为标准形:

$$f=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2,$$

其中 λ_i 为 $\mathbf{A}$ 的特征值. 因为 λ_i 大于零, 故 $\mathbf{A}$ 的正惯性指数为 n.

③ ⇒ ④

若 $\mathbf{A}$ 的正惯性指数为 n, 则 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 可经适当可逆线性变换 $\mathbf{x}=\mathbf{P}\mathbf{y}$ 化为规范形 $f(\mathbf{x})=\mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}=y_1^2+y_2^2+\cdots+y_n^2=\mathbf{y}^{\mathrm{T}}\mathbf{E}\mathbf{y}$, 即存在可逆阵 $\mathbf{P}$ 使得 $\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P}=\mathbf{E}$, 故 $\mathbf{A}$ 合同于单位阵 $\mathbf{E}$.

④ $\Rightarrow$ ⑤

若 $\mathbf{A}$ 合同于单位阵 $\mathbf{E}$, 则存在可逆阵 $\mathbf{P}$ 使得 $\mathbf{P}^\mathsf{T}\mathbf{A}\mathbf{P} = \mathbf{E}$, 由此 $\mathbf{A} = (\mathbf{P}^\mathsf{T})^{-1}\mathbf{P}^{-1}$. 记 $\mathbf{B} = \mathbf{P}^{-1}$, 则 $(\mathbf{P}^\mathsf{T})^{-1} = (\mathbf{P}^{-1})^\mathsf{T} = \mathbf{B}^\mathsf{T}$. 即 $\mathbf{A} = \mathbf{B}^\mathsf{T}\mathbf{B}$.

⑤ $\Rightarrow$ ①

若存在可逆阵 $\mathbf{B}$ 使得 $\mathbf{A} = \mathbf{B}^\mathsf{T}\mathbf{B}$, 则 $\forall \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}$, 有 $\mathbf{B}\mathbf{x} \neq \mathbf{0}$, 故

$$f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{B}^\mathsf{T}\mathbf{B}\mathbf{x} = (\mathbf{B}\mathbf{x})^\mathsf{T}(\mathbf{B}\mathbf{x}) = \|\mathbf{B}\mathbf{x}\|^2 > 0,$$

即 $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$ 是正定二次型 (或 $\mathbf{A}$ 正定) .

定理 3　若 $\mathbf{A}$ 为实对称矩阵, 则下列条件等价:

(1) $\mathbf{A}$ 为半正定矩阵;

定理 3　若 $\boldsymbol{A}$ 为实对称矩阵, 则下列条件等价:

(1) $\boldsymbol{A}$ 为半正定矩阵;

(2) $\boldsymbol{A}$ 的特征值均大于等于零, 且至少有一个等于零;

定理 3 若 $\mathbf{A}$ 为实对称矩阵, 则下列条件等价:

(1) $\mathbf{A}$ 为半正定矩阵;

(2) $\mathbf{A}$ 的特征值均大于等于零, 且至少有一个等于零;

(3) $\mathbf{A}$ 的正惯性指数为 $R(\mathbf{A})$ 且 $R(\mathbf{A}) < n$;

定理 3　若 $\mathbf{A}$ 为实对称矩阵, 则下列条件等价:

(1) $\mathbf{A}$ 为半正定矩阵;

(2) $\mathbf{A}$ 的特征值均大于等于零, 且至少有一个等于零;

(3) $\mathbf{A}$ 的正惯性指数为 $R(\mathbf{A})$ 且 $R(\mathbf{A}) < n$;

(4) $\mathbf{A} \simeq \mathrm{diag}(1, 1, \cdots, 1, 0, \cdots, 0)$, 其中 1 有 $R(\mathbf{A})$ 个, $R(\mathbf{A}) < n$;

定理 3 若 $\mathbf{A}$ 为实对称矩阵, 则下列条件等价:

(1) $\mathbf{A}$ 为半正定矩阵;

(2) $\mathbf{A}$ 的特征值均大于等于零, 且至少有一个等于零;

(3) $\mathbf{A}$ 的正惯性指数为 $R(\mathbf{A})$ 且 $R(\mathbf{A}) < n$;

(4) $\mathbf{A} \simeq \mathrm{diag}(1, 1, \cdots, 1, 0, \cdots, 0)$, 其中 1 有 $R(\mathbf{A})$ 个, $R(\mathbf{A}) < n$;

(5) 存在非满秩矩阵 $\mathbf{B}$, 使得 $\mathbf{A} = \mathbf{B}^{\mathsf{T}}\mathbf{B}$.

例 2　判定下列二次型是否是有定二次型：

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_3^2+2x_1x_2-4x_2x_3.$$

例 2　判定下列二次型是否是有定二次型：

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_3^2+2x_1x_2-4x_2x_3.$$

解　因

$$f(x_1,x_2,x_3)=x_1^2+2x_2^2+4x_3^2+2x_1x_2-4x_2x_3$$

$$=(x_1+x_2)^2+(x_2-2x_3)^2\geqslant 0,$$

且易找到满足 $x_1+x_2=0, x_2-2x_3=0$ 的非零解，如 $(-2,2,1)^{\mathrm{T}}$ 等，故原二次型为半正定二次型.

也可以利用二次型矩阵 **A** 的特征值来判断. 因为

$$|\lambda \mathbf{E}-\mathbf{A}|=\begin{vmatrix}\lambda-1 & -1 & 0\\ -1 & \lambda-2 & 2\\ 0 & 2 & \lambda-4\end{vmatrix}=\lambda(\lambda^2-7\lambda+9),$$

故 **A** 的特征值为 $\lambda_1=0$, $\lambda_2=\frac{7+\sqrt{13}}{2}$, $\lambda_3=\frac{7-\sqrt{13}}{2}$, 从而原二次型为半正定二次型.

例 3　设 $\mathbf{A} = \mathbf{A}_{n\times n}, \mathbf{B} = \mathbf{B}_{n\times n}$ 都正定，证明 $\mathbf{A} + \mathbf{B}$, $\begin{pmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{pmatrix}_{2n\times 2n}$ 也都正定.

证 $\mathbf{A}+\mathbf{B}$ 显然对称. 对 $\forall \alpha \in \mathbb{R}^n$, $\alpha \neq \mathbf{0}$, 由 $\mathbf{A}=\mathbf{A}_{n\times n}, \mathbf{B}=\mathbf{B}_{n\times n}$ 都正定可得

$$\alpha^{\mathsf{T}}(\mathbf{A}+\mathbf{B})\alpha=\alpha^{\mathsf{T}}\mathbf{A}\alpha+\alpha^{\mathsf{T}}\mathbf{B}\alpha>0.$$

故 $\mathbf{A}+\mathbf{B}$ 正定.

记 $\mathbf{C}=\begin{pmatrix}\mathbf{A} & \mathbf{O}\\ \mathbf{O} & \mathbf{B}\end{pmatrix}$, 因 $\mathbf{A},\mathbf{B}$ 实对称,

$$C^{\mathsf{T}}=\begin{pmatrix}A & O\\ O & B\end{pmatrix}^{\mathsf{T}}=\begin{pmatrix}A^{\mathsf{T}} & O\\ O & B^{\mathsf{T}}\end{pmatrix}=\begin{pmatrix}A & O\\ O & B\end{pmatrix}=C,$$

故 $\mathbf{C}$ 是实对称矩阵. 下面用两种方法证明 C 是正定矩阵.

方法 1. 考虑矩阵 C 的特征多项式,

$$|\lambda E_{2n} - C| = \left| \begin{pmatrix} \lambda E_n & O \\ O & \lambda E_n \end{pmatrix} - \begin{pmatrix} A & O \\ O & B \end{pmatrix} \right|$$

$$= \begin{vmatrix} \lambda E_n - A & O \\ O & \lambda E_n - B \end{vmatrix} = |\lambda E_n - A|\,|\lambda E_n - B|,$$

所以 **C** 的 $2n$ 个特征值是由 **A** 的 n 个特征值和 **B** 的 n 个特征值合并组成, 由于 A, B 都是正定矩阵, 所以 A 的 n 个特征值与 B 的 n 个特征值都全大于 0, 故 **C** 的特征值也全大于零, 所以 **C** 也正定.

方法 2. $\forall x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{R}^{2n}, x \neq \mathbf{0}$ 则 $\alpha \neq \mathbf{0} (\alpha \in \mathbb{R}^n)$, 或 $\beta \neq \mathbf{0} (\beta \in \mathbb{R}^n)$, 譬如 $\alpha \neq \mathbf{0}$, 由于 A, B 都是正定矩阵, 所以 $\alpha^{\mathrm{T}} A \alpha > 0$, $\beta^{\mathrm{T}} B \beta \geqslant 0$, 因此

$$x^{\mathrm{T}} C x = (\alpha^{\mathrm{T}}, \beta^{\mathrm{T}}) \begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

$$= (\alpha^{\mathrm{T}} A, \beta^{\mathrm{T}} B) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha^{\mathrm{T}} A \alpha + \beta^{\mathrm{T}} B \beta > 0,$$

即得 C 是正定矩阵.

定理 4 实对称阵 $\mathbf{A}=(a_{ij})_{n\times n}$ 正定的充要条件是它的各阶顺序主子式全大于零, 即

$$A_1=a_{11}>0,\ A_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}>0,$$

$$A_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}>0,\ \cdots,\ A_n=|\mathbf{A}|>0.$$

证　必要性. 因为实对称矩阵 $\mathbf{A}$ 正定, 从而 $f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x}$ 为正定二次型. 取

$$\mathbf{x} = (x_1, x_2, \cdots, x_k, 0, \cdots, 0)^{\mathrm{T}} \neq \mathbf{0},$$

代入二次型可得

$$f(\mathbf{x}) = \sum_{j=1}^{k}\sum_{i=1}^{k} a_{ij}x_ix_j = \mathbf{x}_k^{\mathrm{T}}\mathbf{A}_k\mathbf{x}_k > 0,$$

其中 $\mathbf{x}_k = (x_1, x_2, \cdots, x_k)^{\mathrm{T}} \neq \mathbf{0}$, $\mathbf{A}_k = (a_{ij})_{k\times k}$. 故 $\mathbf{A}_k$ 为 k 阶正定矩阵, $k = 1, 2, \cdots, n$, 所以 $|A_k| > 0 (k = 1, 2, \cdots, n)$, 从而 $\mathbf{A}$ 的各阶顺序主子式均为正.

充分性. 对 n 用数学归纳法. 当 $n=1$ 时, $f=a_{11}x_1^2$, 因 $a_{11}>0$, 所以 f 正定. 假设充分性对 $n-1$ 元二次型成立, 下面证明对 n 元二次型也成立. 将 $\mathbf{A}$ 分块为

$$\mathbf{A}=\begin{pmatrix} \mathbf{A}_{n-1} & \alpha \\ \alpha^{\mathrm{T}} & a_{nn} \end{pmatrix},$$

其中 $\alpha=(a_{1n},a_{2n},\cdots,a_{n-1,n})^{\mathrm{T}}$, 由归纳假设 $\mathbf{A}_{n-1}$ 正定, 所以 $\mathbf{A}_{n-1}$ 的特征值均大于零, 且有 $|\mathbf{A}_{n-1}|>0$. 令

$$\mathbf{P}=\begin{pmatrix} \mathbf{E}_{n-1} & -\mathbf{A}_{n-1}^{-1}\alpha \\ \mathbf{0} & 1 \end{pmatrix},$$

则

$$\mathbf{P}^{\mathrm{T}}\mathbf{A}\mathbf{P} = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{0} \\ \mathbf{0} & -\alpha^{\mathrm{T}}\mathbf{A}_{n-1}^{-1}\alpha + a_{nn} \end{pmatrix} \xlongequal{\text{def}} \mathbf{B}.$$

记 $-\alpha^{\mathrm{T}}\mathbf{A}_{n-1}^{-1}\alpha + a_{nn} = b$, 则由于 $|\mathbf{A}_{n-1}| > 0$, $|\mathbf{A}| > 0$, 有 $b > 0$. 因为矩阵 $\mathbf{B}$ 的特征值由 $\mathbf{A}_{n-1}$ 的特征值及正数 b 组成, 所以 $\mathbf{B}$ 的特征值均大于零. 从而 $\mathbf{A}$ 合同于一个正定矩阵, 因此 $\mathbf{A}$ 为正定矩阵.

实对称矩阵 $\mathbf{A}$ 为半正定矩阵的充要条件是 $\mathbf{A}$ 的各阶主子式非负, 但至少有一个主子式等于零.

推论 实对称阵 $\mathbf{A} = (a_{ij})_{n\times n}$ 负定的充要条件是它的顺序主子式满足：

$$A_1 = a_{11} < 0, A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0, \cdots,$$

$$(-1)^k A_k > 0, \cdots, A_n = (-1)^n |\mathbf{A}| > 0.$$

例 4　判别二次型 $f = -5x^2 - 6y^2 - 4z^2 + 4xy + 4xz$ 的正定性.

例 4　判别二次型 $f=-5x^2-6y^2-4z^2+4xy+4xz$ 的正定性.

解　f 的矩阵为 $\mathbf{A}=\begin{pmatrix}-5 & 2 & 2\\ 2 & -6 & 0\\ 2 & 0 & -4\end{pmatrix}$. 因为 A 的各阶顺序主子式的符号为

$$A_1=-5<0,\ A_2=\begin{vmatrix}5 & -2\\ -2 & 6\end{vmatrix}=26>0,$$

$$A_3=|A|=\begin{vmatrix}-5 & 2 & 2\\ 2 & -6 & 0\\ 2 & 0 & -4\end{vmatrix}=-80<0.$$

所以原二次型是负定二次型.

例 5 判别 $f=2x_1^2+3x_2^2+4x_3^2-12x_1x_2+8x_1x_3+2x_2x_3$ 的正定性.

例 5 判别 $f=2x_1^2+3x_2^2+4x_3^2-12x_1x_2+8x_1x_3+2x_2x_3$ 的正定性.

解 f 的矩阵为 $\mathbf{A}=\begin{pmatrix} 2 & -6 & 4 \\ -6 & 3 & 1 \\ 4 & 1 & 4 \end{pmatrix}$, 因为 $\mathbf{A}$ 的各阶顺序主子式的符号为

$$A_1=2>0,\quad A_2=\begin{vmatrix} 2 & -6 \\ -6 & 3 \end{vmatrix}=-30<0,$$

A_3 无需计算, 即可判定原二次型是不定二次型.

例 6　当 λ 取何值时，$f=\lambda x_1^2+\lambda x_2^2+2x_3^2+2x_1x_2+2x_1x_3-2x_2x_3$ 为正定二次型？

例 6 当 λ 取何值时, $f=\lambda x_1^2+\lambda x_2^2+2x_3^2+2x_1x_2+2x_1x_3-2x_2x_3$ 为正定二次型?

解 $A=\begin{pmatrix}\lambda & 1 & 1\\ 1 & \lambda & -1\\ 1 & -1 & 2\end{pmatrix}$, 因为 $\mathbf{A}$ 是正定矩阵, 所以 $\mathbf{A}$ 的各阶顺序主子式 A_k 皆大于 0.

$$A_1=\lambda>0,\ |A_2|=\begin{vmatrix}\lambda & 1\\ 1 & \lambda\end{vmatrix}=\lambda^2-1>0 \quad \Rightarrow \lambda>1;$$

$$|A_3| = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & -1 \\ 1 & -1 & 2 \end{vmatrix} = 2(\lambda-2)(\lambda+1) > 0 \quad \Rightarrow$$

$\lambda > 2$,

综上可得：当 $\lambda > 2$ 时 f 是正定二次型.

设 n 元函数 $f(x_1, x_2, \cdots, x_n)$ 在点 $P_0(x_1^0, x_2^0, \cdots, x_n^0)$ 的某个邻域内有二阶连续偏导数, 由多元函数的泰勒 (Taylor) 公式得

$$f(x_1^0+\Delta x_1, x_2^0+\Delta x_2, \cdots, x_n^0+\Delta x_n) - f(x_1^0, x_2^0, \cdots, x_n^0)$$

$$= \left(\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \cdots + \Delta x_n \frac{\partial}{\partial x_n}\right) f(x_1^0, x_2^0, \cdots, x_n^0)$$

$$+ \frac{1}{2!}\left(\Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \cdots + \Delta x_n \frac{\partial}{\partial x_n}\right)^2 f(x_1^0, x_2^0, \cdots, x_n^0) + R,$$

简写为矩阵表达式

$$f\left((x_0+\Delta x)^{\mathrm{T}}\right) - f\left((x_0)^{\mathrm{T}}\right) = \xi^{\mathrm{T}} \Delta x + \frac{1}{2}(\Delta x)^{\mathrm{T}} A (\Delta x) + R,\ R = o\left(\|\Delta x\|^2\right),$$

其中

$$x_0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \\ x_n^0 \end{pmatrix}, \Delta x = \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{pmatrix}, \xi = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}_{P_0},$$

$$A = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}_{P_0},$$

显然 A 是实对称矩阵, 由 “微积分” 课程知, 函数 $f(x_1,x_2,\cdots,x_n)$ 在 $P_0\left(x_1^0,x_2^0,\cdots,x_n^0\right)$ 处有极值的必要条件是 ξ 为零向量, 即

$$\xi=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\cdots,\frac{\partial f}{\partial x_n}\right)^{\mathrm{T}}\Big|_{P_0}=(0,0,\cdots,0)^{\mathrm{T}}.$$

在此条件下，点 P_0 为函数 $f(x_1, x_2, \cdots, x_n)$ 的驻点，这时有

$$f\left((x_0+\Delta x)^{\mathrm{T}}\right)-f\left((x_0)^{\mathrm{T}}\right)=\frac{1}{2}(\Delta x)^{\mathrm{T}}A(\Delta x)+R\ ,$$

$$R=o\left(\|\Delta x\|^2\right),$$

当 $\|\Delta x\|$ 足够小时，上式右端的符号完全由二次型 $(\Delta x)^{\mathrm{T}}A(\Delta x)$ 来决定，若 $R(A)=n$，则

(1) 当 $(\Delta x)^{\mathrm{T}} A (\Delta x)$ 为正定 (即 A 为正定) 时, P_0 为 f 的一个极小值点.

(2) 当 $(\Delta x)^{\mathrm{T}} A (\Delta x)$ 为负定 (即 A 为负定) 时, P_0 为 f 的一个极大值点.

(3) 当 $(\Delta x)^{\mathrm{T}} A (\Delta x)$ 为不定 (即 A 为不定) 时, P_0 不是 f 的极值点.

当 $R(A) < n$ 时, 要决定 f 在点 P_0 的性态, 还需研究余项 R, 这里不讨论.

当 $n2$ 时, 就得到熟知的二元函数在 P0 有极值的充分条件, 即若

$$\left.\frac{\partial f}{\partial x_1}\right|_{P_0} = \left.\frac{\partial f}{\partial x_2}\right|_{P_0} = 0,$$

记 $a = \left.\frac{\partial^2 f}{\partial x_1^2}\right|_{P_0}, b = \left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{P_0}, c = \left.\frac{\partial^2 f}{\partial x_2^2}\right|_{P_0}$, 得 $\mathbf{A} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$.

于是当 $R(\mathbf{A})=2$ 时, 有

(1) A 为正定矩阵时, 即 $A_1=a>0, A_2=|A|=ac-b^2>0$ 时, $f(P_0)$ 为极小值;

(2) A 为负定矩阵时, 即 $A_1=a<0, A_2=|A|=ac-b^2>0$ 时, $f(P_0)$ 为极大值;

(3) A 为不定矩阵时, 即 $A_2=|A|=ac-b^2<0$ 时, $f(P_0)$ 不是极值.

例 7 求函数 $f(x,y)=3xy-x^3-y^3$ 的极值.

例 7 求函数 $f(x,y)=3xy-x^3-y^3$ 的极值.

解 解方程组 $\begin{cases} \dfrac{\partial f}{\partial x}=3y-3x^2=0, \\ \dfrac{\partial f}{\partial y}=3x-3y^2=0 \end{cases}$

得驻点 $P_1(0,0)$, $P_2(1,1)$. 因 $\dfrac{\partial^2 f}{\partial x^2}=-6x$, $\dfrac{\partial^2 f}{\partial x\partial y}=3$, $\dfrac{\partial^2 f}{\partial y^2}=-6y$, 所以在 P_1 处有 $\mathbf{A}=\begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}$, A 为不定矩阵, 于是 $f(P_1)$ 不是极值;

在 P_2 处, 有 $A=\begin{pmatrix} -6 & 3 \\ 3 & -6 \end{pmatrix}$, A 为负定矩阵, 于是 $f(P_2)=1$ 为极大值.