线性代数

张神星 (合肥工业大学)

翡翠科教楼 *B1810* 东

zhangshenxing@hfut.edu.cn

https://faculty.hfut.edu.cn/zhangshenxing

第四章　相似和合同

1. 方阵的相似
2. 实对称阵的正交合同
3. 实对称阵的合同

第一节　方阵的相似

- 特征值与特征向量
- 特征值和特征向量的性质
- 方阵的相似
- 相似矩阵的性质
- 相似对角化

特征值与特征向量的定义

设 $f:\mathbb{C}^n\to\mathbb{C}^n$ 是一个线性映射. 由于 $\mathbb{C}^n$ 中的向量由一组基

$$\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$$

唯一线性表示, 因此它完全由它在该组基下的像决定. 若 f 将每个 $\boldsymbol{\alpha}_i$ 映射为它的倍数, 则 f 将会变得很容易研究.

定义. **若常数 λ 和非零向量 $\boldsymbol{x}$ 满足 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$, 称 λ 为 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 为 $\boldsymbol{A}$ 关于 λ 的特征向量.**

设 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, 则

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

该方程有非零解当且仅当

$$|\boldsymbol{A} - \lambda\boldsymbol{E}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

注意到该行列式是 λ 的 n 次多项式, 最高项为 $(-1)^n\lambda^n$. 称之为 $\boldsymbol{A}$ 的**特征多项式**.

特征值和特征向量的性质

在复数域中 n 次多项式总有 n 个根 (计算重数), 也就是说 $\boldsymbol{A}$ 的特征多项式可以写成

$$f(\lambda)=(\lambda_1-\lambda)\cdots(\lambda_n-\lambda).$$

所以 $\boldsymbol{A}$ 的**特征值有 n 个 (计算重数)**.

特征值和特征向量只对方阵存在, 且有如下性质:

(1) 零向量不是特征向量;

(2) 若 $\boldsymbol{x}$ 是对应 λ 的特征向量, 则它的非零倍也是;

(3) 若 $\boldsymbol{x}_1,\boldsymbol{x}_2\neq-\boldsymbol{x}_1$ 是对应 λ 的特征向量, 则 $\boldsymbol{x}_1+\boldsymbol{x}_2$ 也是;

(4) $|\boldsymbol{A}|=0 \iff 0$ 是特征值; $|\boldsymbol{A}|\neq 0 \iff 0$ 不是特征值;

(5) 若 n 阶方阵 $\boldsymbol{A}$ 的各行元素之和为 k, 则 k 是 $\boldsymbol{A}$ 的一个特征值, 且特征向量为 $(1,\ldots,1)^{\mathrm{T}}$.

特征值和特征向量的计算步骤如下:

(1) 求 $\boldsymbol{A}$ 的特征多项式 $f(\lambda) = |\boldsymbol{A} - \lambda\boldsymbol{E}|$;

(2) 解 $f(\lambda) = |\boldsymbol{A} - \lambda\boldsymbol{E}| = 0$ 得到特征值;

(3) 对于每一个特征值 λ_i, 解 $(\boldsymbol{A} - \lambda_i\boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}$, 其非零解就是对应特征向量.

典型例题：求特征值和特征向量

例. 求 $\boldsymbol{A}=\begin{pmatrix}1 & 3\\ 4 & 2\end{pmatrix}$ 的特征值和特征向量.

解.

(1) 特征多项式 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}1-\lambda & 3\\ 4 & 2-\lambda\end{vmatrix}=\lambda^2-3\lambda-10.$

(2) 由 $\lambda^2-3\lambda-10=0$ 解得特征值 $\lambda_1=5,\lambda_2=-2.$

(3) 对于 $\lambda_1=5$, $\boldsymbol{A}-5\boldsymbol{E}=\begin{pmatrix}-4 & 3\\ 4 & -3\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & -3/4\\ 0 & 0\end{pmatrix}$, 得到基础解系 $\begin{pmatrix}3/4\\ 1\end{pmatrix}$. 故对应的所有特征向量为 $k(3,4)^{\mathrm{T}},k\neq 0.$

(4) 对于 $\lambda_2=-2$, $\boldsymbol{A}+2\boldsymbol{E}=\begin{pmatrix}3 & 3\\ 4 & 4\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$, 得到基础解系 $\begin{pmatrix}1\\ -1\end{pmatrix}$. 故对应的所有特征向量为 $k(1,-1)^{\mathrm{T}},k\neq 0.$

例. 求上三角阵 $\boldsymbol{A} = \begin{pmatrix} a_1 & b & c \\ 0 & a_2 & d \\ 0 & 0 & a_3 \end{pmatrix}$ 的特征值.

解. 特征多项式

$$|\boldsymbol{A} - \lambda \boldsymbol{E}| = \begin{vmatrix} a_1 - \lambda & b & c \\ 0 & a_2 - \lambda & d \\ 0 & 0 & a_3 - \lambda \end{vmatrix} = (a_1 - \lambda)(a_2 - \lambda)(a_3 - \lambda).$$

因此特征值为 a_1, a_2, a_3.

上三角阵、下三角阵、对角阵的特征值就是对角元.

典型例题：求特征值和特征向量

例. 求 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$ 的特征值和特征向量.

解. 由特征多项式

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

得到特征值 $\lambda=1,1,2$.

续解. 对于 $\lambda_1 = 1$,

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

得到基础解系 $(-1,-2,1)^{\mathrm{T}}$. 故对应的所有特征向量为 $k(-1,-2,1)^{\mathrm{T}}, k \neq 0$.
对于 $\lambda_2 = 2$,

$$A - 2E = \begin{pmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得到基础解系 $(0,0,1)^{\mathrm{T}}$. 故对应的所有特征向量为 $k(0,0,1)^{\mathrm{T}}, k \neq 0$.

典型例题：求特征值和特征向量

练习. 求 $\boldsymbol{A}=\begin{pmatrix}-2 & 1 & 1\\ 0 & 2 & 0\\ -4 & 1 & 3\end{pmatrix}$ 的特征值和特征向量.

答案.

(1) 特征值 $\lambda=-1,2,2$.

(2) -1 对应的所有特征向量为 $k(1,0,1)^{\mathrm{T}}, k\neq 0$.

(3) 2 对应的所有特征向量为 $k_1(0,1,-1)^{\mathrm{T}}+k_2(1,0,4)^{\mathrm{T}}, k_1,k_2$ 不全为零.

若 λ 是 k 重特征值，则它对应的线性无关的特征向量最多 k 个.

通过

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}$$

的展开可以看出, 若 $i \neq j$, 则 i 行 j 列的代数余子式中最多只会出现 λ^{n-2} 项. 所以

$$\begin{aligned} |\boldsymbol{A}-\lambda\boldsymbol{E}| &= (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) + \text{至多 } n-2 \text{ 次项} \\ &= (-1)^n\left(\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}\right) + \text{至多 } n-2 \text{ 次项}. \end{aligned}$$

根据韦达定理, 特征值之和为 $a_{11}+a_{22}+\cdots+a_{nn}$.

$\lambda=0$ 时, 特征多项式

$$f(\lambda)=(\lambda_1-\lambda)\cdots(\lambda_n-\lambda)$$

的取值 $|\boldsymbol{A}|$ 就是特征值的乘积.

定理 (特征值和特征向量的性质). 定义 A 的迹为对角元之和:

$$\mathrm{Tr}(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn}.$$

(1) 特征值之和等于迹: $\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{Tr}(\boldsymbol{A})$;

(2) 特征值之积等于行列式: $\lambda_1\lambda_2\cdots\lambda_n = |\boldsymbol{A}|$.

例. $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ 的特征值为(C).

(A) $1, 0, 1$ (B) $1, 1, 2$ (C) $-1, 1, 2$ (D) $-1, 1, 1$

定理. 若 λ 是 $\boldsymbol{A}$ 的特征值, $\boldsymbol{x}$ 是对应特征向量, 则下述矩阵有如下对应的特征值与特征向量:

方阵	$k\boldsymbol{A}$	$\boldsymbol{A}^m$	$\boldsymbol{A}^{-1}$	$\boldsymbol{A}^*$	$g(\boldsymbol{A})h(\boldsymbol{A})^{-1}$	$\boldsymbol{A}^{\mathrm{T}}$	$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$
特征值	$k\lambda$	λ^m	λ^{-1}	$\lvert\boldsymbol{A}\rvert/\lambda$	$g(\lambda)/h(\lambda)$	λ	λ
对应特征向量	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	未必是 $\boldsymbol{x}$	$\boldsymbol{P}^{-1}\boldsymbol{x}$

这里 g, h 是多项式, 且满足 $h(\boldsymbol{A})$ 可逆.

由此可知, 若 $g(\boldsymbol{A}) = \boldsymbol{O}$, 则 $\boldsymbol{A}$ 的所有特征值 λ 均满足 $g(\lambda) = 0$.

例：特征值和特征向量的性质

例. 设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 $2,2,3$, 则

(1) $\boldsymbol{A}^2$ 的特征值为 $\underline{\quad 4,4,9 \quad}$;

(2) $\boldsymbol{A}^2-2\boldsymbol{A}+\boldsymbol{E}$ 的特征值为 $\underline{\quad 1,1,4 \quad}$;

(3) $|\boldsymbol{A}^2-2\boldsymbol{A}+\boldsymbol{E}|=\underline{\quad 4 \quad}$;

(4) $\boldsymbol{A}^{-1}$ 的特征值为 $\underline{\quad 1/2,1/2,1/3 \quad}$;

(5) $\boldsymbol{A}^*$ 的特征值为 $\underline{\quad 4,6,6 \quad}$;

(6) $A_{11}+A_{22}+A_{33}=\underline{\quad 16 \quad}$, 其中 A_{ij} 表示 $\boldsymbol{A}$ 的代数余子式.

例：特征值和特征向量的性质

例. 设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 $1,-1,2$. 求 $|\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}|$.

解. 由于 $|\boldsymbol{A}|=-2$, $\boldsymbol{A}^*=-2\boldsymbol{A}^{-1}$. 于是

$$\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}=-2\boldsymbol{A}^{-1}+3\boldsymbol{A}-2\boldsymbol{E}$$

的特征值为 $-1,-3,3$, $|\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}|=(-1)\times(-3)\times 3=9$.

练习. 若 4 阶方阵 $\boldsymbol{A}$ 的特征值为 $1,2,-2,0$, 则下列矩阵可逆的是(C).

(A) $\boldsymbol{A}$ (B) $\boldsymbol{A}-2\boldsymbol{E}$ (C) $\boldsymbol{A}+\boldsymbol{E}$ (D) $\boldsymbol{A}-\boldsymbol{E}$

例：特征值和特征向量的性质

例. 若 $\boldsymbol{\alpha}=(1,k,1)^{\mathrm{T}}$ 为 $\boldsymbol{A}=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$ 的逆 $\boldsymbol{A}^{-1}$ 的特征向量，求 k.

解. $\boldsymbol{\alpha}$ 也是 $\boldsymbol{A}$ 的特征向量.

$$\boldsymbol{A}\boldsymbol{\alpha}=\begin{pmatrix}k+3\\2k+2\\k+3\end{pmatrix}=\lambda\begin{pmatrix}1\\k\\1\end{pmatrix},$$

因此 $\lambda=k+3, 2k+2=(k+3)k, k^2+k-2=0, k=-2$ 或 1.

例. 计算 $\boldsymbol{A} = \begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{pmatrix}$ 的行列式.

解. 设 $\boldsymbol{B} = \boldsymbol{A} + (b-a)\boldsymbol{E}$, 则 $\boldsymbol{B}$ 所有元素为 b, 秩为 1, $\boldsymbol{B}\boldsymbol{x} = \mathbf{0}$ 基础解系有 $n-1$ 个向量. 从而 0 是 $\boldsymbol{B}$ 的至少 $n-1$ 重特征值. 由于 $\mathrm{Tr}(\boldsymbol{B}) = nb$, 因此 $\boldsymbol{B}$ 的所有特征值为 $nb, 0, \ldots, 0$, $\boldsymbol{A}$ 的所有特征值为 $nb+a-b, a-b, \ldots, a-b$,

$$|\boldsymbol{A}| = (a-b)^{n-1}(nb-b+a).$$

定理. 若 $\lambda_1, \dots, \lambda_m$ 是 $\boldsymbol{A}$ 的 m 个两两不同的特征值，则其对应的特征向量 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性无关.

即对应于不同特征值的特征向量线性无关.

证明. 设 $k_1\boldsymbol{\alpha}_1 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$. 左乘 $\boldsymbol{A}^k$ 得到 $k_1\lambda_1^k\boldsymbol{\alpha}_1 + \cdots + k_m\lambda_m^k\boldsymbol{\alpha}_m = \mathbf{0}$. 令 $k = 1, 2, \dots, m-1$, 我们得到

$$(k_1\boldsymbol{\alpha}_1, \cdots, k_m\boldsymbol{\alpha}_m)\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{m-1} \end{pmatrix} = \boldsymbol{O}.$$

注意到第二个方阵的行列式是范德蒙行列式，当 $\lambda_1, \dots, \lambda_m$ 两两不同时它非零，从而

$$(k_1\boldsymbol{\alpha}_1, \cdots, k_m\boldsymbol{\alpha}_m) = \boldsymbol{O}, \qquad k_1 = \cdots = k_m = 0.$$ □

设 λ_1 对应线性无关的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$, λ_2 对应线性无关的特征向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$. 若 $\lambda_1 \neq \lambda_2$, 则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 也是线性无关的. 这是因为

$$\boldsymbol{A}(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) = \lambda_1(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2),$$

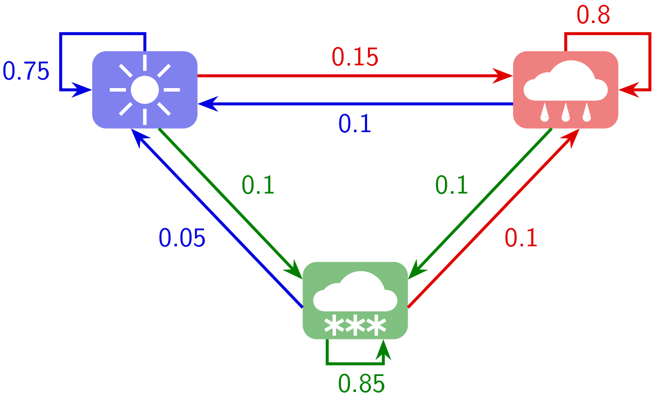
$$\boldsymbol{A}(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2) = \lambda_2(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2).$$

若 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2 = \mathbf{0}$, 同理可证明这些向量都是零向量.

由此也可以知道, 不同特征值的特征向量的线性组合不可能还是特征向量.

应用：天气预报

将天气简化为晴雨雪三种，其它天气由它们组合得到. 根据该地历史的天气信息，得到当天与第二天天气的关系:



应用：天气预报

例. 某地某季节天气仅受前一天天气状态影响：设 k 天后天气为晴天、雨天、雪天概率分别为 a_k, b_k, c_k，则

$$\begin{pmatrix} a_k \\ b_k \\ c_k \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} a_{k-1} \\ b_{k-1} \\ c_{k-1} \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 0.75 & 0.1 & 0.05 \\ 0.15 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.85 \end{pmatrix}.$$

如果今天天气为晴天，请问未来七天的各种天气概率分别是多少？

解. 设 $\boldsymbol{x}_k=(a_k,b_k,c_k)^{\mathrm{T}}$，则 $\boldsymbol{x}_k=\boldsymbol{A}\boldsymbol{x}_{k-1}$. 因此 $x_k=\boldsymbol{A}x_{k-1}=\boldsymbol{A}^2x_{k-2}=\cdots=\boldsymbol{A}^k\boldsymbol{x}_0$. 解特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}0.75-\lambda & 0.1 & 0.05\\ 0.15 & 0.8-\lambda & 0.1\\ 0.1 & 0.1 & 0.85-\lambda\end{vmatrix}=0$$

得到特征值 $\lambda_1=1,\lambda_2=0.75,\lambda_3=0.65$. 解方程 $(\boldsymbol{A}-\lambda_i\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 得到特征向量

$$\boldsymbol{v}_1=(8,13,14)^{\mathrm{T}},\quad \boldsymbol{v}_2=(1,1,-2)^{\mathrm{T}},\quad \boldsymbol{v}_3=(1,-1,0)^{\mathrm{T}}.$$

由于今天天气是晴天，所以 $\boldsymbol{x}_0=(1,0,0)^{\mathrm{T}}$. 注意到 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 线性无关，因此 $\boldsymbol{x}_0$ 可以由它们线性表示. 通过计算得到 $\boldsymbol{x}_0=\dfrac{1}{35}\boldsymbol{v}_1+\dfrac{1}{5}\boldsymbol{v}_2+\dfrac{4}{7}\boldsymbol{v}_3$，因此

$$\boldsymbol{x}_k=\boldsymbol{A}^k\boldsymbol{x}_0=\frac{1}{35}\lambda_1^k\boldsymbol{v}_1+\frac{1}{5}\lambda_2^k\boldsymbol{v}_2+\frac{4}{7}\lambda_3^k\boldsymbol{v}_3.$$

应用：天气预报

续解. 代入计算可得 $\boldsymbol{x}_k$, 从而得到未来七天晴雨雪的概率:

	明天	后天	3 天后	4 天后	5 天后	6 天后	7 天后
晴	0.75	0.58	0.47	0.39	0.34	0.31	0.28
雨	0.15	0.24	0.30	0.33	0.35	0.36	0.37
雪	0.10	0.18	0.23	0.27	0.31	0.33	0.35

基的变换

设 $f:\mathbb{C}^n\to\mathbb{C}^n$ 是一个线性映射. 对于任意 $\boldsymbol{x}\in\mathbb{C}^n$, 它可以唯一表达为

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n.$$

通过将 $f(\boldsymbol{e}_1),\ldots,f(\boldsymbol{e}_n)$ 表达为这组标准基的线性组合, 我们建立了线性映射 f 与一个 n 阶方阵 $\boldsymbol{A}$ 的联系.

一般的向量空间并没有这样的标准正交基, 或者 $\mathbb{C}^n$ 本身我们也可以选择其它基. 则这种对应会有什么变化呢? 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 是 $\mathbb{C}^n$ 中一组线性无关的向量,

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n).$$

则

$$f(\boldsymbol{\alpha}_i)=\boldsymbol{A}\boldsymbol{\alpha}_i=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_i,$$

所以 $f(\boldsymbol{\alpha}_i)$ 表达为 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 线性组合的系数形成的向量是 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_i$. 它们构成矩阵

$$(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_1,\ldots,\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_n)=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}.$$

方阵的相似

也就是说，若我们将向量空间 $\mathbb{C}^n$ 换一组基表达，线性映射对应的矩阵就会变成 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 我们称 f 在不同基下矩阵为相似的.

定义. 若存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 则称方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似**.

注意，相似只对方阵有定义.

命题. 方阵的相似满足

(1) 自反性: $\boldsymbol{A}$ 与自身相似;

(2) 对称性: 若 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$, 则 $\boldsymbol{B}$ 相似于 $\boldsymbol{A}$;

(3) 传递性: 若 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$, $\boldsymbol{B}$ 相似于 $\boldsymbol{C}$, 则 $\boldsymbol{A}$ 相似于 $\boldsymbol{C}$.

若 A, B 相似 $(B = P^{-1}AP)$, 则 A, B 等价 $(B = PAQ)$. 反之未必成立, 例如 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. 这是因为和 $A = E$ 相似的只有它自己.

例. 若 3 阶方阵 $A \overset{r_1 \leftrightarrow r_3}{\sim} B \overset{c_1 \leftrightarrow c_3}{\sim} C$, 则 A 与 C (C).

(A) 等价但不相似　　(B) 相似但不等价

(C) 等价而且相似　　(D) 既不等价也不相似

定理 (相似矩阵的性质). 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则二者的特征多项式相同, 从而特征值也相同.

这是因为若 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则

$$\boldsymbol{P}^{-1}(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{P}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}-\lambda\boldsymbol{P}^{-1}\boldsymbol{E}\boldsymbol{P}=\boldsymbol{B}-\lambda\boldsymbol{E}.$$

两边取行列式并利用行列式的可乘性得到 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=|\boldsymbol{B}-\lambda\boldsymbol{E}|$.

注意反过来未必成立, 例如 $\boldsymbol{A}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$. 的特征多项式相同, 但它们不相似.

相似矩阵具有相同的特征值, 但对应的特征向量未必相同.

定理 (相似矩阵的性质). 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则

(1) $\boldsymbol{A}, \boldsymbol{B}$ 特征值相同 (包括重数);

(2) $|\boldsymbol{A}| = |\boldsymbol{B}|$, $\mathrm{Tr}(\boldsymbol{A}) = \mathrm{Tr}(\boldsymbol{B})$;

(3) $\boldsymbol{A} \sim \boldsymbol{B}$, 即 $\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{B})$;

(4) 对于多项式 g, h, 若 $h(\boldsymbol{A})$ 可逆, 则 $h(\boldsymbol{B})$ 也可逆, 且 $g(\boldsymbol{A})/h(\boldsymbol{A})$ 与 $g(\boldsymbol{B})/h(\boldsymbol{B})$ 相似. 特别地, $\boldsymbol{A} - \lambda\boldsymbol{E}$ 与 $\boldsymbol{B} - \lambda\boldsymbol{E}$ 相似.

推论. 若 $\boldsymbol{A}$ 与对角阵 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ 相似, 则 $\lambda_1, \ldots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值.

例：相似矩阵的性质

例. 若 $\boldsymbol{A}=\begin{pmatrix}2&-1&4\\0&a&7\\0&0&3\end{pmatrix}$ 与 $\boldsymbol{\Lambda}=\begin{pmatrix}1&0&0\\0&2&0\\0&0&b\end{pmatrix}$ 相似，则 $a=$ __1__，$b=$ __3__.

例. 若 3 阶可逆阵 $\boldsymbol{A},\boldsymbol{B}$ 相似，$\boldsymbol{A}^{-1}$ 的特征值为 $1/2,1/3,1/4$，则 $|\boldsymbol{E}-\boldsymbol{B}|=$ __-6__.

练习. 若 3 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，且存在非零矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{AC}=2\boldsymbol{C}$, $|\boldsymbol{A}+2\boldsymbol{E}|=0$, $|2\boldsymbol{A}-\boldsymbol{E}|=0$，则 $|\boldsymbol{B}^{-1}-\boldsymbol{E}|=$ __$3/4$__.

若 $\boldsymbol{AB}=k\boldsymbol{B}$，则 $\boldsymbol{B}$ 的每个非零列向量均为 $\boldsymbol{A}$ 的属于特征值 k 的特征向量.

定义. 若 n 阶方阵 $\boldsymbol{A}$ 相似于某个对角阵

$$\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

则称 $\boldsymbol{A}$ 可**(相似) 对角化**.

设 $\boldsymbol{P} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$, 则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} \iff \boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{\Lambda} \iff (\boldsymbol{A}\boldsymbol{p}_1, \ldots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1\boldsymbol{p}_1, \ldots, \lambda_n\boldsymbol{p}_n),$$

即 $\boldsymbol{A}\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i$. 由于 $\boldsymbol{P}$ 可逆, $\boldsymbol{A}$ 拥有 n 个线性无关的特征向量 $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$. 反之, 若 $\boldsymbol{A}$ 拥有 n 个线性无关的特征向量, 则选择 $\boldsymbol{P}$ 以它们为列向量即可使 $\boldsymbol{A}$ 对角化.

定理 (相似对角化的等价刻画). n 阶矩阵 $\boldsymbol{A}$ 可对角化 $\iff$ $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

推论. 若 $\boldsymbol{A}$ 的特征值两两不同, 则 $\boldsymbol{A}$ 可对角化.

反之未必成立.

例. 设 3 阶方阵 $\boldsymbol{B}$ 的特征值为 $1, 2, -2$, $\boldsymbol{A} = \boldsymbol{B}^3 - 4\boldsymbol{B} + \boldsymbol{E}$, 求 $\boldsymbol{A}$ 的相似对角阵.

解. 由于 $\boldsymbol{B}$ 特征值两两不同, 因此存在 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{\Lambda} = \operatorname{diag}(1, 2, -2)$. 于是

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}^3 - 4\boldsymbol{\Lambda} + \boldsymbol{E} = \operatorname{diag}(-2, 1, 1).$$

回忆 k 重特征值对应的线性无关的特征向量最多 k 个.

定理 (相似对角化的等价刻画). 若 λ 是 $\boldsymbol{A}$ 的 k 重特征值, 则 $\boldsymbol{A}$ 可对角化 $\iff$ $\forall\lambda, \mathrm{R}(\boldsymbol{A}-\lambda\boldsymbol{E})=n-k$, 即 λ 对应的线性无关的特征向量恰有 k 个.

相似对角化的步骤如下:

(1) 求出 $\boldsymbol{A}$ 的所有特征值 λ_i 和特征向量 $\boldsymbol{p}_i$;

(2) 根据上述判定方法判断 $\boldsymbol{A}$ 是否可以相似对角化;

(3) 若能, 将 n 个对应的线性无关的特征向量 $\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n$ 组成方阵 $\boldsymbol{P}=(\boldsymbol{p}_1,\ldots,\boldsymbol{p}_n)$,

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\ldots,\lambda_n).$$

例: 对角化的计算

例. $\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}$ 能否对角化? 若能, 求 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角阵.

解.

(1) 上三角阵 $\boldsymbol{A}$ 特征值为 $1,0,0$.

(2) 对于 $\lambda_1=1$, $\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}0&1&1\\0&-1&0\\0&0&-1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$, 取特征向量 $\boldsymbol{p}_1=(1,0,0)^{\mathrm{T}}$.

(3) 对于 $\lambda_2=\lambda_3=0$, $\boldsymbol{A}$ 对应的基础解系可以取 $\boldsymbol{p}_2=(-1,1,0)^{\mathrm{T}}, \boldsymbol{p}_3=(-1,0,1)^{\mathrm{T}}$.

(4) 因此 $\boldsymbol{A}$ 可对角化, 取 $\boldsymbol{P}=\begin{pmatrix}1&-1&-1\\0&1&0\\0&0&1\end{pmatrix}$, 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(1,0,0)$.

例. 判断下列矩阵 $\boldsymbol{A}$ 能否相似对角化： (1) $\boldsymbol{A}$ 是二阶实矩阵且 $|\boldsymbol{A}|<0$;

(2) $\boldsymbol{A}=\begin{pmatrix}1&2&1\\0&3&0\\0&0&0\end{pmatrix}$; (3) $\boldsymbol{A}=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}$; (4) $\boldsymbol{A}=\begin{pmatrix}-1&1&0\\-4&3&0\\1&0&2\end{pmatrix}$.

解. (1) 特征值一正一负，能对角化; (2) 特征值为 $1,3,0$, 能对角化;

(3) $(2-\lambda)(\lambda^2-\lambda-2)=0 \implies \lambda=-1,2,2.$ $\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}-4&1&1\\0&0&0\\-4&1&1\end{pmatrix}$ 秩 1, 能对角化;

(4) $(2-\lambda)(\lambda^2-2\lambda+1)=0 \implies \lambda=2,1,1.$ $\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}-2&1&0\\-4&2&0\\1&0&1\end{pmatrix}$ 秩 2, 不能对角化.

练习.

(1) 若 $\boldsymbol{A}=\begin{pmatrix}0&0&1\\a&1&1\\1&0&0\end{pmatrix}$ 可对角化，则 $a=$ $\underline{\quad -1 \quad}$.

(2) 设 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\boldsymbol{Q}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2)$. 若 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(1,2,3)$，则 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\underline{\mathrm{diag}(1,3,2)}$.

(3) 若 3 阶方阵 $\boldsymbol{A}$ 的特征值互不相同且 $|\boldsymbol{A}|=0$，则 $\mathrm{R}(\boldsymbol{A})=$ $\underline{\quad 2 \quad}$.

(4) 若 $\boldsymbol{A}=\begin{pmatrix}2&0&0\\0&2&1\\0&0&1\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix},\boldsymbol{C}=\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}$，则(C).

(A) $\boldsymbol{A},\boldsymbol{C}$ 相似，$\boldsymbol{B},\boldsymbol{C}$ 相似

(B) $\boldsymbol{A},\boldsymbol{C}$ 不相似，$\boldsymbol{B},\boldsymbol{C}$ 相似

(C) $\boldsymbol{A},\boldsymbol{C}$ 相似，$\boldsymbol{B},\boldsymbol{C}$ 不相似

(D) $\boldsymbol{A},\boldsymbol{C}$ 不相似，$\boldsymbol{B},\boldsymbol{C}$ 不相似

例: 对角化的计算

例. 设 $\boldsymbol{A}=\begin{pmatrix}1 & -1 & 1\\ x & 4 & y\\ -3 & -3 & 5\end{pmatrix}$ 有 3 个线性无关特征向量, $\lambda=2$ 是 $\boldsymbol{A}$ 的二重特征值. 求可逆阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角阵.

解. 由 $\mathrm{Tr}(\boldsymbol{A})=10$ 可知特征值为 $2,2,6$. 由 $\mathrm{R}(\boldsymbol{A}-2\boldsymbol{E})=1$ 可知 $x=2, y=-2$.

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}-1 & -1 & 1\\ x & 2 & y\\ -3 & -3 & 3\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 1 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow \boldsymbol{p}_1=\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}, \boldsymbol{p}_2=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}.$$

$$\boldsymbol{A}-6\boldsymbol{E}=\begin{pmatrix}-5 & -1 & 1\\ 2 & -2 & -2\\ -3 & -3 & -1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & -1/3\\ 0 & 1 & 2/3\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow \boldsymbol{p}_3=\begin{pmatrix}1\\ -2\\ 3\end{pmatrix}. \ \boldsymbol{P}=\begin{pmatrix}-1 & 1 & 1\\ 1 & 0 & -2\\ 0 & 1 & 3\end{pmatrix}.$$

例：对角化的计算

例. 若 $\boldsymbol{A}$ 的各行元素之和为 2，$\boldsymbol{A}\begin{pmatrix}1 & 2\\ 0 & 1\\ -1 & 1\end{pmatrix}=\begin{pmatrix}-1 & 2\\ 0 & 1\\ 1 & 1\end{pmatrix}$，则 $\boldsymbol{A}$ 相似于(C).

(A) $\mathrm{diag}(1,1,2)$　(B) $\mathrm{diag}(2,1,1)$　(C) $\mathrm{diag}(2,1,-1)$　(D) $\mathrm{diag}(2,-1,-1)$

练习. 设 $\boldsymbol{A}=\begin{pmatrix}2 & a & 2\\ 5 & b & 3\\ -1 & 1 & -1\end{pmatrix}$ 有特征值 ±1，问 $\boldsymbol{A}$ 能否相似对角化？

答案. $|\boldsymbol{A}\pm\boldsymbol{E}|=0 \implies a=-1, b=-3.$ $\mathrm{Tr}(\boldsymbol{A})=-2 \implies \lambda_3=-2$，可对角化.

例：对角化的计算

例. 设 $\boldsymbol{A}$ 为三阶方阵, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 是线性无关的三维列向量且

$$\boldsymbol{A}\boldsymbol{\alpha}_1 = 2\boldsymbol{\alpha}_1, \boldsymbol{A}\boldsymbol{\alpha}_2 = 3\boldsymbol{\alpha}_2 + 2\boldsymbol{\alpha}_3, \boldsymbol{A}\boldsymbol{\alpha}_3 = 2\boldsymbol{\alpha}_2 + 3\boldsymbol{\alpha}_3.$$

证明 $\boldsymbol{A}$ 可对角化.

解. 设 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$, 则

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)\boldsymbol{B} = \boldsymbol{P}\boldsymbol{B}, \quad \text{其中 } \boldsymbol{B} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 2 \\ 0 & 2 & 3 \end{pmatrix}.$$

由于 $\boldsymbol{B}$ 的特征值是 $2, 1, 5$, 因此 $\boldsymbol{B}$ 能对角化, 从而 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}$ 也可以.

一般地, 实对称矩阵一定能对角化. 我们将在下一节中解释这为何成立.

任何方阵都可以相似于**约当标准形**

$$\mathrm{diag}(\boldsymbol{J}_{k_1}(\lambda_1), \ldots, \boldsymbol{J}_{k_t}(\lambda_t)),$$

其中 $\boldsymbol{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$ 是 k 阶方阵. 当 $k_1 = \cdots = k_t = 1$ 时这就是对角阵.

特别地, 任何方阵都可以相似于一个上三角阵.

若 f 没有重根, 且 $f(\boldsymbol{A}) = \boldsymbol{O}$, 则可以看出每个约当块必须是一阶的, 从而 $\boldsymbol{A}$ 可对角化.

伴随矩阵的特征值

容易知道上三角阵的伴随矩阵的对角元. 因此 $\boldsymbol{A}^*$ 的所有特征值的就是 $\boldsymbol{A}$ 的 n 个特征值中 $n-1$ 个相乘得到的.

(1) 当 $|\boldsymbol{A}| \neq 0$ 时, $\boldsymbol{A}^*$ 的所有特征值为 $|\boldsymbol{A}|/\lambda$.

(2) 当 $|\boldsymbol{A}| = 0$ 且 $\lambda_1 = 0$ 是一重特征值, 则 $\boldsymbol{A}^*$ 唯一的非零特征值为 $\boldsymbol{A}$ 非零特征值之乘积.

(3) 当 $|\boldsymbol{A}| = 0$ 且 $\lambda_1 = 0$ 是 $\geqslant 2$ 重特征值, 则 $\boldsymbol{A}^* = \boldsymbol{O}$.

练习. 若 4 阶实矩阵 $\boldsymbol{A}^*$ 的特征值为 $-1, 1, 2, 4$, 则下列矩阵可逆的是(D).

(A) $\boldsymbol{A} + 2\boldsymbol{E}$　(B) $\boldsymbol{A} - 2\boldsymbol{E}$　(C) $2\boldsymbol{A} + \boldsymbol{E}$　(D) $\boldsymbol{A} - \boldsymbol{E}$

例：方阵的相似的应用

利用约当标准形还可以计算任意矩阵的幂:

例. 设 $\boldsymbol{A}=\begin{pmatrix}5 & -6\\ 1 & 0\end{pmatrix}$, 计算 $\boldsymbol{A}^k$.

解. $\boldsymbol{A}$ 特征值为 $2,3$, 对应的特征向量可以取 $\boldsymbol{p}_1=\begin{pmatrix}2\\ 1\end{pmatrix}, \boldsymbol{p}_2=\begin{pmatrix}3\\ 1\end{pmatrix}$. 设 $\boldsymbol{P}=\begin{pmatrix}2 & 3\\ 1 & 1\end{pmatrix}$, 则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\begin{pmatrix}2 & \\ & 3\end{pmatrix},$$

因此

$$\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}=\begin{pmatrix}2 & 3\\ 1 & 1\end{pmatrix}\begin{pmatrix}2^k & \\ & 3^k\end{pmatrix}\begin{pmatrix}-1 & 3\\ 1 & -2\end{pmatrix}=\begin{pmatrix}3^{k+1}-2^{k+1} & 6(2^k-3^k)\\ 3^k-2^k & 6(2^{k-1}-3^{k-1})\end{pmatrix}.$$

对于约当标准型也可以计算:

$$\boldsymbol{J}_k(\lambda)^m=(\lambda\boldsymbol{E}+\boldsymbol{N})^m=\sum_{i=0}^{m}\mathrm{C}_m^i\lambda^{m-i}\boldsymbol{N}^i,\quad 其中\ \boldsymbol{N}=\boldsymbol{J}_k(0).$$

利用矩阵的相似可以计算线性递推数列

$$a_{n+k}-c_{k-1}a_{n+k-1}-\cdots-c_1a_{n+1}-c_0a_n=0,$$

的通项. 例如

$$a_{n+2}=5a_{n+1}-6a_n,\quad a_0=a_1=1.$$

设 $\boldsymbol{x}_n=\begin{pmatrix}a_{n+1}\\ a_n\end{pmatrix}$, 则 $\boldsymbol{x}_n=\boldsymbol{A}\boldsymbol{x}_{n-1}\ =\boldsymbol{A}^n\boldsymbol{x}_0=\begin{pmatrix}2^{n+2}-3^{n+1}\\ 2^{n+1}-3^n\end{pmatrix}$, 故 $a_n=2^{n+1}-3^n$.

第二节　实对称阵的正交合同

- 实二次型
- 实对称阵和实二次型的合同

我们知道，

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \quad \frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$

分别表示椭圆和双曲线. 对于二次曲线

$$Ax^2+Bxy+Cy^2=1,$$

它又表示什么图形呢?

实二次型的定义

定义. 若 n 元多项式 $f(x_1,\dots,x_n)$ 满足

$$f(\lambda x_1,\dots,\lambda x_n)=\lambda^2 f(x_1,\dots,x_n),\quad \forall\lambda\in\mathbb{R},$$

则称 f 为二次齐次多项式或实二次型. 它可以写成

$$f(x)=a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2+2a_{12}x_1x_2+\cdots+2a_{n-1,n}x_{n-1}x_n.$$

本课程仅讨论实二次型. 根据定义, f 不能包含一次项和常数项. 若 f 的交叉项 $x_ix_j(i<j)$ 系数均为零, 则称 f 为实二次型的标准形.

设实二次型 f 的 x_i^2 项的系数为 a_{ii}, $x_ix_j(i<j)$ 项的系数为 $2a_{ij}$. 设 $a_{ji}=a_{ij}$, 对称阵 $\boldsymbol{A}=(a_{ij})_n$, 则

$$(x_1,\dots,x_n)\boldsymbol{A}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=(\sum_{i=1}^n a_{i1}x_i,\dots,\sum_{i=1}^n a_{in}x_i)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

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

即 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, 其中 $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{T}}$, $\boldsymbol{A}$ 为实对称阵.

反过来, 任给一个实对称阵 $\boldsymbol{A}$, 多项式 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 显然满足

$$f(\lambda\boldsymbol{x})=(\lambda\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}(\lambda\boldsymbol{x})=\lambda^2f(\boldsymbol{x}),$$

故 f 是实二次型. 因此实二次型 f 与对称阵 $\boldsymbol{A}$ 之间存在一一对应的关系.

例. 写出实二次型 $f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 4x_2^2 + 2x_1x_3 + x_3^2 + 4x_2x_3$ 对应的矩阵.

解. $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$.

若 f 是标准形, 则 f 对应矩阵 $\boldsymbol{A}$ 是对角阵.

定义.

(1) 若存在可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 使得实二次型 f 在变量 $\boldsymbol{x},\boldsymbol{y}$ 下的矩阵分别为 $\boldsymbol{A},\boldsymbol{B}$，则称矩阵 $\boldsymbol{A}$ 合同或相合于 $\boldsymbol{B}$.

(2) 若 $\boldsymbol{P}$ 是正交阵，则称矩阵 $\boldsymbol{A}$ 正交合同或正交相合于 $\boldsymbol{B}$.

若 $\boldsymbol{A}$ 是对称阵, $\boldsymbol{P}$ 可逆, 则 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 也是对称阵. 由

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P})\boldsymbol{y}$$

可知 $\boldsymbol{A}$ (正交) 合同于 $\boldsymbol{B}$ 当且仅当存在可逆 (正交) 方阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$.

命题. 对称方阵的 (正交) 合同满足

(1) 自反性: A 与自身 (正交) 合同;

(2) 对称性: 若 A (正交) 合同于 B, 则 B (正交) 合同于 A;

(3) 传递性: 若 A (正交) 合同于 B, B (正交) 合同于 C, 则 A (正交) 合同于 C.

合同、等价、相似有如下关系:

(1) 若 $\boldsymbol{A}, \boldsymbol{B}$ 合同, 则 $\boldsymbol{A}, \boldsymbol{B}$ 等价, $\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{B})$. 反之未必.

(2) 若 $\boldsymbol{A}, \boldsymbol{B}$ 正交合同, 则 $\boldsymbol{A}, \boldsymbol{B}$ 相似. 反之, **若实对称阵 $\boldsymbol{A}, \boldsymbol{B}$ 相似, 则二者正交合同**.

定理. 对于实对称阵 $\boldsymbol{A}$, 存在正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 是对角阵. 从而 $\boldsymbol{A}$ 对应的实二次型在线性变换 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 下变为标准形.

命题. 实对称阵的特征值一定是实数, 从而其特征向量均可取实向量.

证明. 设 $\boldsymbol{A}$ 是实对称阵, 非零向量 $\boldsymbol{x}$ 满足 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$. 两边取转置和共轭并右乘 $\boldsymbol{x}$ 得到

$$\bar{\lambda}\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=\overline{\boldsymbol{x}}^{\mathrm{T}}\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{x}=\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\lambda\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}.$$

显然 $\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=|x_1|^2+\cdots+|x_n|^2>0$, 因此 λ 是实数. 由于特征向量是方程 $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 的解, 系数矩阵是实方阵, 因此特征向量可取实向量. □

实二次型的对角化的证明

定理的证明. 归纳证明. 假设我们已找到 $0 \leqslant k < n$ 个两两正交的特征向量 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$, 分别对应特征值 $\lambda_1, \dots, \lambda_k$. 设 V 是与这些向量均正交的实向量全体, 即方程 $(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k)^{\mathrm{T}} \boldsymbol{x} = \mathbf{0}$ 的解空间. 由于正交的向量组线性无关, 系数秩为 $k < n$, 解空间 V 非零. 对任意 $v \in V$,

$$[\boldsymbol{\alpha}_i, \boldsymbol{A}\boldsymbol{v}] = \boldsymbol{\alpha}_i^{\mathrm{T}} \boldsymbol{A}\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{\alpha}_i)^{\mathrm{T}} \boldsymbol{v} = \lambda_i \boldsymbol{\alpha}_i^{\mathrm{T}} \boldsymbol{v} = 0 \implies \boldsymbol{A}\boldsymbol{v} \in V.$$

因此非零空间 V 上的线性映射 $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ 有一个特征值和对应的特征向量: $\boldsymbol{A}\boldsymbol{\alpha}_{k+1} = \lambda_{k+1} \boldsymbol{\alpha}_{k+1}$. 最后将它们标准化后构成的正交阵 $\boldsymbol{P} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ 满足题述要求. $\square$

由于特征值 λ 对应的实特征向量就是 $\boldsymbol{P}$ 中 λ 对应的那些列向量的线性组合, 因此:

推论. 实对称阵的不同特征值对应的实特征向量正交.

练习.

(1) 设 $\boldsymbol{\alpha}_1=(1,-3,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,a,2)^{\mathrm{T}}$ 是实对称阵 $\boldsymbol{A}$ 对应特征值 $\lambda_1=1$ 和 $\lambda_2=2$ 的特征向量, 则 $a=$ __1__.

(2) 若 3 阶实对称阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{A},\mathrm{R}(\boldsymbol{A})=1$, 则 $\boldsymbol{A}$ 的特征值为 __$0,0,1$__.

例：对阵矩阵的性质

例. 设 3 阶实对称阵 $\boldsymbol{A}$ 的特征值为 $6,3,3$, 与特征值 6 对应的特征向量为 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$, 求 $\boldsymbol{A}$.

解. 由于 $\boldsymbol{A}$ 有两个对应特征值 3 的线性无关特征向量, 因此与 $\boldsymbol{\alpha}_1$ 正交的向量都是与特征值 3 对应的特征向量. 由 $\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{x}=0$ 得 $\boldsymbol{\alpha}_2=(-1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(-1,0,1)^{\mathrm{T}}$. 故

$$\boldsymbol{A}=\boldsymbol{P}\mathrm{diag}(6,3,3)\boldsymbol{P}^{-1}=\begin{pmatrix}1&-1&-1\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}6&&\\&3&\\&&3\end{pmatrix}\begin{pmatrix}1&-1&-1\\1&1&0\\1&0&1\end{pmatrix}^{-1}=\begin{pmatrix}4&1&1\\1&4&1\\1&1&4\end{pmatrix}.$$

另解. 根据 $\boldsymbol{A}$ 的行和、迹和对称性可设 $\boldsymbol{A}=\begin{pmatrix}a&b&6-a-b\\b&9-a-b&a\\6-a-b&a&3+b\end{pmatrix}$. 再由 $\mathrm{R}(\boldsymbol{A}-3\boldsymbol{E})=1$ 可知 $a=4,b=1$.

对称阵正交合同对角化, 或求正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 将实二次型 f 化为标准形的步骤:

(1) 写出 f 对应的对称阵 $\boldsymbol{A}$.
(2) 求出 $\boldsymbol{A}$ 的特征值.
(3) **若特征值是 k 重的, 求出 k 个特征向量后, 用格拉姆-施密特方法将其正交单位化.**
(4) 这些特征向量构成正交阵 $\boldsymbol{P}$, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$.
(5) 写出正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 以及对应的实二次型

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

例. 设 $\boldsymbol{A}=\begin{pmatrix}2&0&0\\0&0&1\\0&1&x\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}2&0&0\\0&y&0\\0&0&-1\end{pmatrix}$ 相似，求 x,y 以及正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

解. 由 $\boldsymbol{A},\boldsymbol{B}$ 相似得 $|\boldsymbol{A}|=-2=|\boldsymbol{B}|=-2y$, $\mathrm{Tr}(\boldsymbol{A})=2+x=\mathrm{Tr}(\boldsymbol{B})=1+y$, 故 $x=0,y=1$.

- 对于 $\lambda_1=2$, $\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}0&0&0\\0&-2&1\\0&1&-2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}\implies\boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$.

续解.

- 对于 $\lambda_2 = 1$, $\boldsymbol{A} - \boldsymbol{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \implies \boldsymbol{\alpha}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.
- 对于 $\lambda_3 = -1$, $\boldsymbol{A} + \boldsymbol{E} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \implies \boldsymbol{\alpha}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.
- 将特征向量单位化得到 $\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$.

典型例题：实二次型的对角化

例. 求正交变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$ 化 $f = 4x_1^2 + 4x_2^2 + 4x_3^2 + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 为标准形.

解.

- f 对应 $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$. 由 $|\boldsymbol{A} - \lambda\boldsymbol{E}| = -(\lambda-2)^2(\lambda-8)$ 得到特征值 $8, 2, 2$.
- 对于 $\lambda_1 = 8$, $\boldsymbol{A} - 8\boldsymbol{E} = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. 将其单位

  化得到 $\boldsymbol{e}_1 = \dfrac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

续解.

- 对于 $\lambda_2, \lambda_3 = 2$, $\boldsymbol{A} - 2\boldsymbol{E} = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \boldsymbol{\alpha}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$

  将其正交单位化得到 $\boldsymbol{\beta}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{e}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$

$$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_3 - \frac{[\boldsymbol{\alpha}_3, \boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2, \boldsymbol{\beta}_2]} \boldsymbol{\beta}_2 = \boldsymbol{\alpha}_3 - \frac{1}{2} \boldsymbol{\beta}_2 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1 \end{pmatrix}, \quad \boldsymbol{e}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

- 因此经过正交变换 $\boldsymbol{x} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \boldsymbol{y}$, f 化为标准形 $f = 8y_1^2 + 2y_2^2 + 2y_3^2$.

例. 设实二次型 $f = ax_1^2 + x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 经过正交变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$ 化为 $f = 5y_1^2 - y_2^2 - y_3^2$. 求常数 a 和正交阵 $\boldsymbol{P}$.

解. f 对应 $\boldsymbol{A} = \begin{pmatrix} a & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, $\mathrm{Tr}(\boldsymbol{A}) = a + 2 = 5 - 1 - 1, a = 1$.

同上例可得 $\boldsymbol{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$.

典型例题：实二次型的对角化

例. 设实二次型 $f = a(x_1^2 + x_2^2 + x_3^2) + 4x_1x_2 + 4x_2x_3 + 4x_3x_1$ 经正交变换化成标准形 $f(y_1, y_2, y_3) = 6y_1^2$，则 $a =$ __2__.

练习. 设实二次型 $f = x_1^2 + 2x_2^2 + ax_3^2 - 4x_1x_2 - 4x_2x_3$ 经过正交变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$ 化为 $f = 2y_1^2 + 5y_2^2 + by_3^2$. 求常数 a, b 和正交阵 $\boldsymbol{P}$.

答案. $a = 3, b = -1, \boldsymbol{P} = \dfrac{1}{3}\begin{pmatrix} -2 & 1 & 2 \\ 1 & -2 & 2 \\ 2 & 2 & 1 \end{pmatrix}$.

例：正定的性质和判定

例. 设二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. 证明：当 $\|x\|=1$ 时，$f(\boldsymbol{x})$ 的最大 (小) 值为实对称阵 $\boldsymbol{A}$ 的最大 (小) 特征值.

证明. 将 $\boldsymbol{A}$ 的特征值排序为 $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. 存在正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 使得

$$\boldsymbol{\Lambda}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}(\lambda_1,\ldots,\lambda_n), \quad f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2.$$

由于正交变换保持长度，因此 $\|\boldsymbol{x}\|=1 \iff \|\boldsymbol{y}\|=1$.

$$f=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2 \leqslant \lambda_1(y_1^2+\cdots+y_n^2)=\lambda_1,$$

且等式可在 $\boldsymbol{y}=(1,0,\ldots,0)$ 处取得.

$$f=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2 \geqslant \lambda_n(y_1^2+\cdots+y_n^2)=\lambda_n,$$

且等式可在 $\boldsymbol{y}=(0,\ldots,0,1)$ 处取得. 故 f 的最大值为 λ_1, 最小值为 λ_n. $\square$

第三节　实对称阵的合同

- 惯性指数
- 正定二次型

引例：二次曲线的分类

设 A, B, C 是不全为零的实数. 二次曲线 $Ax^2 + Bxy + Cy^2 = 1$ 左侧的实二次型对应方阵 $\boldsymbol{A} = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$. 由

$$|\boldsymbol{A} - \lambda \boldsymbol{E}| = \begin{vmatrix} A-\lambda & B/2 \\ B/2 & C-\lambda \end{vmatrix} = \lambda^2 - (A+C)\lambda + (AC - B^2/4)$$

可知,

(1) 当 $B^2 - 4AC > 0$ 时, $\boldsymbol{A}$ 特征值一正一负, 从而通过正交变换 $\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} s \\ t \end{pmatrix}$ 可知该曲线为双曲线.

(2) 同理, $B^2 - 4AC < 0$ 时该曲线为椭圆 (或空集);

(3) $B^2 - 4AC = 0$ 时该曲线为两条直线 (若有一次项则为抛物线).

可以看出我们有时候只关心实二次型对应的矩阵的特征值的符号.

定理 (惯性定理).

(1) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 为合同的对角阵, 则 $\boldsymbol{A}, \boldsymbol{B}$ 对角元中正数的个数相同.

(2) 若可逆线性变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y} = \boldsymbol{Q}\boldsymbol{z}$ 分别将 f 变为

$$f = k_1 y_1^2 + \cdots + k_r y_r^2, \qquad f = \ell_1 z_1^2 + \cdots + \ell_r z_r^2,$$

其中这些系数非零, 则 $k_1, \ldots, k_r$ 中正的个数和 $\ell_1, \ldots, \ell_r$ 中正的个数相同.

证明. 设

$$\begin{aligned} \boldsymbol{A} &= \mathrm{diag}(k_1, \ldots, k_r, 0, \ldots, 0), \\ \boldsymbol{B} &= \mathrm{diag}(\ell_1, \ldots, \ell_r, 0, \ldots, 0) = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P}, \end{aligned}$$

其中 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)$ 为可逆矩阵. 不难得到 $\boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P}$ 的对角元:

$$\boldsymbol{B} = \boldsymbol{P}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{P} = \mathrm{diag}(k_1 \boldsymbol{\alpha}_1^{\mathrm{T}} \boldsymbol{\alpha}_1, \ldots, k_r \boldsymbol{\alpha}_r^{\mathrm{T}} \boldsymbol{\alpha}_r, 0, \ldots, 0).$$

从而 $\ell_i = k_i \|\boldsymbol{\alpha}_i\|^2$, 二者符号相同. (2) 由(1) 得到. □

注意到实对称阵都可正交合同于对角阵, 且对角元可以排序为

$$\Lambda = \operatorname{diag}\{\lambda_1, \ldots, \lambda_p, \mu_1, \ldots, \mu_q, 0, \ldots, 0\},$$

其中 λ_i 为正特征值, μ_j 为负特征值. 设

$$\boldsymbol{P} = \operatorname{diag}\left\{\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_p}}, \frac{1}{\sqrt{-\mu_1}}, \ldots, \frac{1}{\sqrt{-\mu_q}}, 0, \ldots, 0\right\},$$

则

$$\boldsymbol{P}^{\mathrm{T}} \Lambda \boldsymbol{P} = \operatorname{diag}\{1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0\}.$$

合同规范形

定理. 任意 n 阶实对称矩阵 $\boldsymbol{A}$ 合同于对角矩阵

$$\begin{pmatrix} \boldsymbol{E}_p & & \\ & -\boldsymbol{E}_q & \\ & & \boldsymbol{O}_{n-p-q} \end{pmatrix},$$

其中 p, q 分别为正负特征值个数 (计算重数), $\mathrm{R}(\boldsymbol{A}) = p + q$.

推论. n 阶实对称阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同 $\iff$ $\boldsymbol{A}, \boldsymbol{B}$ 的正负特征值个数均相同.

相应地,

定义. 若实二次型的标准形的系数只在 $-1, 0, 1$ 三个数中取值, 则称此实二次型为**规范形**.

定义. 把实二次型 f 标准形系数中为正数的个数称为 f 的**正惯性指数** p, 为负数的个数称为 f 的**负惯性指数** q, $r = p + q$ 为 f 的**秩**(即对应实对称阵的秩).

定理. 任意一个实二次型总可经过适当的可逆线性变换化为规范形, 且规范形是唯一的 (可任意交换变量顺序):

$$f = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2,$$

其中 p, q 分别为正负惯性指数.

例：惯性指数的应用

例. 设 $\boldsymbol{A}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$，则 $\boldsymbol{A}$ 合同于(D).

(A) $\begin{pmatrix}-2&1\\1&-2\end{pmatrix}$ (B) $\begin{pmatrix}2&-1\\-1&2\end{pmatrix}$ (C) $\begin{pmatrix}2&1\\1&2\end{pmatrix}$ (D) $\begin{pmatrix}1&-2\\-2&1\end{pmatrix}$

练习. 矩阵 $\begin{pmatrix}1&1\\1&1\end{pmatrix}$ 与 $\begin{pmatrix}3&0\\0&0\end{pmatrix}$(B).

(A) 合同且相似 (B) 合同但不相似 (C) 不合同但相似 (D) 既不合同也不相似

例. 若实对称矩阵 $\boldsymbol{A}$ 合同于 $\begin{pmatrix}1&0&0\\0&0&2\\0&2&0\end{pmatrix}$，则通过可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 可将二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 化为规范形 $\underline{\quad y_1^2+y_2^2-y_3^2\quad}$.

例：二次曲面的分类

对于三元实二次型，正负惯性指数确定了二次曲面的类别.

(1) $p=3, q=0$ 为椭球面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1.$

(2) $p=2, q=1$ 为单叶双曲面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1.$

(3) $p=1, q=2$ 为双叶双曲面 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1.$

(4) $p=2, q=0$ 为椭圆柱面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1.$

(5) $p=1, q=1$ 为双曲柱面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1.$

正定和负定

定义. 设 $f = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$ 是二次型.

(1) 若对任意 $\boldsymbol{x} \neq \boldsymbol{0}$, 都有 $f(\boldsymbol{x}) > 0$, 则称 f 为正定二次型, 并称实对称阵 $\boldsymbol{A}$ 为正定矩阵, 也记作 $\boldsymbol{A} > 0$.

(2) 若对任意 $\boldsymbol{x}$, 都有 $f(\boldsymbol{x}) \geqslant 0$, 则称 f 为半正定二次型, 并称实对称阵 $\boldsymbol{A}$ 为半正定矩阵, 也记作 $\boldsymbol{A} \geqslant 0$.

(3) 若 $-f$ (半) 正定, 则称 f 为(半) 负定二次型, 并称实对称阵 $\boldsymbol{A}$ 为(半) 负定矩阵, 也记作 $\boldsymbol{A} < 0$ ($\boldsymbol{A} \leqslant 0$).

(4) 除此之外, 称 f 不定.

例: 正定和负定

例.

(1) $f(x_1,x_2,x_3)=x_1^2+x_2^2$ 半正定.

(2) $f(x_1,x_2,x_3)=x_1^2-2x_2^2+x_3^2$ 不定.

(3) $f(x_1,x_2,x_3)=(x_1+x_2)^2+(x_2+x_3)^2+(x_3+x_1)^2$ 正定.

(4) $f(x_1,x_2,x_3)=(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2$ 半正定.

(5) 椭球面 $f(x,y,z)=1$ 对应的二次型正定.

(6) 单叶/双叶双曲面 $f(x,y,z)=1$ 对应的二次型不定.

定理. 设 $\boldsymbol{A}$ 是 n 阶实对称阵, $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. 如下命题等价:

(1) $\boldsymbol{A}>0$ 正定, 即 f 正定.

(2) f 的正惯性指数为 n, 即 $\boldsymbol{A}$ 特征值全为正.

(3) 存在可逆阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$.

(4) (赫尔维茨定理) $\boldsymbol{A}$ 的各阶顺序主子式都为正, 即

$$a_{11}>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,\quad \ldots,\quad |\boldsymbol{A}|>0.$$

推论. 若 A 正定，则 $|A|>0$ 且对角元全为正.

例. 若 $f(x_1,x_2,x_3)=2x_1^2+x_2^2+x_3^2+2x_1x_2+tx_2x_3$ 正定，求 t 的取值范围.

解. f 对应 $A=\begin{pmatrix}2&1&0\\1&1&t/2\\0&t/2&1\end{pmatrix}$，顺序主子式

$$2>0,\quad \begin{vmatrix}2&1\\1&1\end{vmatrix}=1>0,\quad |A|=1-\frac{t^2}{2}>0$$

得到 $-\sqrt{2}<t<\sqrt{2}$.

例. 求可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 二次型 $f=-5x_1^2+6x_2^2-2x_3^2+6x_1x_2+4x_1x_3$ 化为规范形.

解.

$$
\begin{aligned}
f &= -2(x_3-x_1)^2-3x_1^2+6x_2^2+6x_1x_2 \\
&= -2(x_3-x_1)^2-3(x_1-x_2)^2+9x_2^2
\end{aligned}
$$

因此 $\boldsymbol{x}=\boldsymbol{y}\boldsymbol{P},\boldsymbol{P}=\begin{pmatrix}-\sqrt{2} & \sqrt{3} & 0\\ 0 & -\sqrt{3} & 3\\ \sqrt{2} & 0 & 0\end{pmatrix}^{-1}=\begin{pmatrix}0 & 0 & \dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{3}} & 0 & \dfrac{1}{\sqrt{3}}\\ \dfrac{1}{3} & \dfrac{1}{3} & \dfrac{1}{3}\end{pmatrix}$，将 f 化为规范形 $f=-y_1^2-y_2^2+y_3^2$，

例：正定的性质和判定

例. 若实对称阵 $\boldsymbol{A}$ 正定，证明 $|\boldsymbol{A}+\boldsymbol{E}|>1$.

证明. 由 $\boldsymbol{A}$ 正定可知其特征值均为正，从而 $\boldsymbol{A}+\boldsymbol{E}$ 特征值都大于 1. 从而 $|\boldsymbol{A}+\boldsymbol{E}|>1$. □

例. 设 3 阶实对称阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2+2\boldsymbol{A}=\boldsymbol{O}, \mathrm{R}(\boldsymbol{A})=2$. 当 k 为何值时，矩阵 $\boldsymbol{A}+k\boldsymbol{E}$ 为正定矩阵.

解. 由 $\boldsymbol{A}^2+2\boldsymbol{A}=\boldsymbol{O}$ 可知 $\boldsymbol{A}$ 的特征值满足 $\lambda^2+2\lambda=0, \lambda=0,-2$. 由 $\mathrm{R}(\boldsymbol{A})=2$ 可知 $\boldsymbol{A}$ 特征值为 $0,-2,-2$, $\boldsymbol{A}+k\boldsymbol{E}$ 特征值为 $k, k-2, k-2$. 因此 $k>2$.

例. 设 $\boldsymbol{A}$ 是 $m \times n$ 实矩阵且 $\mathrm{R}(\boldsymbol{A}) = n$. 证明 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 正定.

证明. 显然 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是对称的. 注意到

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x} = (\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \|\boldsymbol{A}\boldsymbol{x}\|^2.$$

由于 $\mathrm{R}(\boldsymbol{A}) = n$, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 只有零解. 因此当 $\boldsymbol{x} \neq \boldsymbol{0}$ 时, $\boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{0}$, 从而

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x} = \|\boldsymbol{A}\boldsymbol{x}\|^2 > 0.$$

□

例. 设 $\boldsymbol{A}$ 为 n 阶实对称矩阵. 证明 $\mathrm{R}(\boldsymbol{A})=n \iff$ 存在一个 n 阶实方阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 正定.

证明. 显然 $\boldsymbol{AB}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 是对称的, 且

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{AB}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}=(\boldsymbol{Ax})^{\mathrm{T}}\boldsymbol{Bx}+(\boldsymbol{Bx})^{\mathrm{T}}(\boldsymbol{Ax})=2[\boldsymbol{Ax},\boldsymbol{Bx}].$$

若 $\mathrm{R}(\boldsymbol{A})=n$, 令 $\boldsymbol{B}=\boldsymbol{A}$, 则当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时, $\boldsymbol{Ax}\neq\boldsymbol{0}$, 从而

$$[\boldsymbol{Ax},\boldsymbol{Bx}]=\|\boldsymbol{Ax}\|^2>0.$$

若 $\mathrm{R}(\boldsymbol{A})<n$, 则存在 $\boldsymbol{x}\neq\boldsymbol{0}$ 使得 $\boldsymbol{Ax}=\boldsymbol{0}$, 从而 $[\boldsymbol{Ax},\boldsymbol{Bx}]=0$, $\boldsymbol{AB}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 不正定. □

例. 设 x 是实数，证明 $\begin{vmatrix} 10 & 2 & 3 & 5 \\ 2 & 10 & 1 & -1 \\ 3 & 1 & 10 & x \\ 5 & -1 & x & 10 \end{vmatrix} \leqslant 10000.$

证明. 设 $\boldsymbol{A}$ 为题述方阵. 它的前三个顺序主子式

$$10 > 0, \quad \begin{vmatrix} 10 & 2 \\ 2 & 10 \end{vmatrix} = 96 > 0, \quad \begin{vmatrix} 10 & 2 & 3 \\ 2 & 10 & 1 \\ 3 & 1 & 10 \end{vmatrix} = 872 > 0,$$

若 $|\boldsymbol{A}| \leqslant 0$, 命题显然成立. 若 $|\boldsymbol{A}| > 0$, 则 $\boldsymbol{A}$ 正定, 从而特征值全正. 因此

$$|\boldsymbol{A}| = \lambda_1\lambda_2\lambda_3\lambda_4 \leqslant \left(\frac{\lambda_1+\lambda_2+\lambda_3+\lambda_4}{4}\right)^4 = 10000.$$

□

正定的应用：函数极值

实对称阵可用于判断多元函数的极值. 设 $f(\boldsymbol{x})=f(x_1,\dots,x_n)$ 是一个 n 元实函数, $\boldsymbol{a}$ 是其定义域内一点, 且 f 在 $\boldsymbol{a}$ 附近具有连续的二阶偏导数. 记 $f''_{ij}=\dfrac{\partial^2 f}{\partial x_i\,\partial x_j}$, 则 $f''_{ij}=f''_{ji}$. 于是 $\boldsymbol{A}=(f''_{ij}(\boldsymbol{a}))$ 是 n 阶实对称阵.

定理. 设 f 在 $\boldsymbol{a}$ 处各阶偏导均为零.

(1) 若 $\boldsymbol{A}$ 正定, 则 f 在 $\boldsymbol{a}$ 处取极小值;

(2) 若 $\boldsymbol{A}$ 负定, 则 f 在 $\boldsymbol{a}$ 处取极大值.

若 $\boldsymbol{A}$ 不定, 则无法判断 $\boldsymbol{a}$ 是否是极值点.

对于一般的 $m\times n$ 实矩阵 $\boldsymbol{A}$, 我们有奇异值分解

$$\boldsymbol{A}=\boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\boldsymbol{V}_{n\times n}^{\mathrm{T}},$$

其中 $\boldsymbol{U},\boldsymbol{V}$ 分别是 m,n 阶正交阵, $\boldsymbol{\Sigma}$ 是 $m\times n$ 型对角阵, 对角元非负且按降序排列.

首先对 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 这一半正定对称阵做正交合同对角化

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{V},\quad \boldsymbol{V}=(\boldsymbol{v}_1,\dots,\boldsymbol{v}_n),\quad \boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\dots,\lambda_n),\lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant 0.$$

奇异值是指 $\sigma_i=\sqrt{\lambda_i}$. 令 $\boldsymbol{\Sigma}\in M_{m\times n}$ 为对角阵, 对角元为 $\sigma_1,\dots,\sigma_n$.

对于 $1\leqslant j\leqslant r=\mathrm{R}(\boldsymbol{A})$, 令 $\boldsymbol{u}_j=\dfrac{1}{\sigma_j}\boldsymbol{A}\boldsymbol{v}_j$, 设 $\boldsymbol{u}_{r+1},\dots,\boldsymbol{u}_m$ 是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 的一组标准正交基础解系, 则 $\boldsymbol{U}=(\boldsymbol{u}_1,\dots,\boldsymbol{u}_m)$ 是正交阵, 且 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$.

注意到, 如果 $\boldsymbol{U}'$ 是 $\boldsymbol{U}$ 的前 r 列, $\boldsymbol{V}'$ 是 $\boldsymbol{V}$ 的前 r 行, $\boldsymbol{\Sigma}'$ 是 $\boldsymbol{\Sigma}$ 的前 r 行 r 列, 则

$$\boldsymbol{A} = \boldsymbol{U}'_{m\times r}\boldsymbol{\Sigma}'_{r\times r}\boldsymbol{V}'_{r\times n},$$

其中 $\boldsymbol{\Sigma}'$ 是 $\boldsymbol{A}$ 奇异值降序的对角阵, $\boldsymbol{U}', \boldsymbol{V}'$ 为列/行为标准向量且两两正交的矩阵.

这意味着当 r 相比 m, n 较小时, 只需存储 $(m+n+1)r$ 个元素即可还原 $\boldsymbol{A}$. 这是一种无损压缩算法. 如果我们只截取前 $k < r$ 个奇异值以及对应的 $\boldsymbol{U}, \boldsymbol{V}$ 部分, 则可以对 $\boldsymbol{A}$ 进行有损压缩到 $\boldsymbol{A}'$. 例如 $\boldsymbol{A}$ 表示一张图像的像素信息, 保留它较大的奇异值往往对它的信息影响很小. 有时候, 我们甚至需要主动舍弃较小的奇异值, 只保留较大的奇异值来实现信号降噪.

矩阵还有诸如 LU 分解, QR 分解, 科列斯基分解等. 这些分解往往都在压缩或降噪中发挥着作用.

当 $m = n = 3$ 时, 可以看出线性变换可以分解为旋转、放缩、旋转的复合.