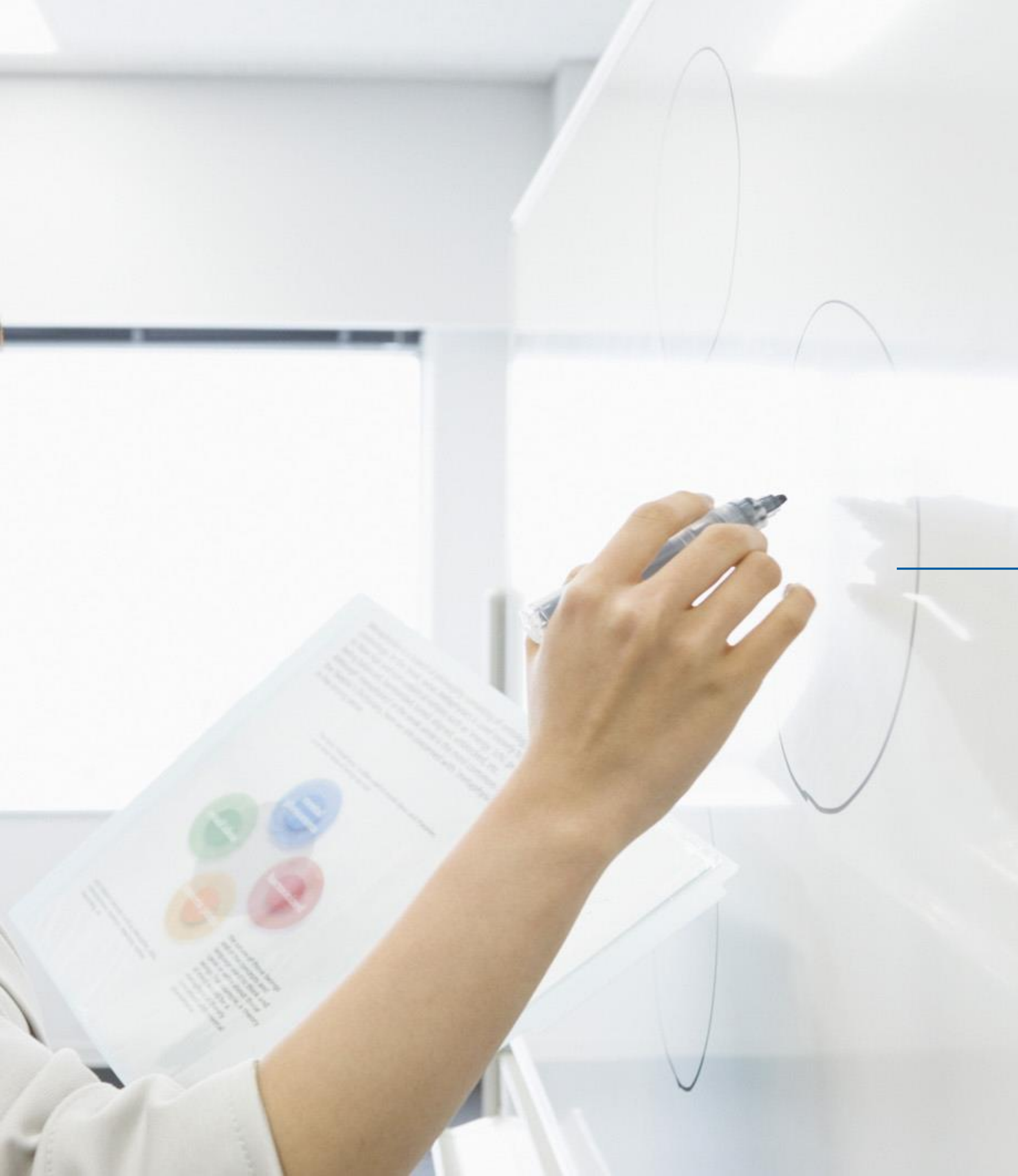

04

相似矩阵及二次型

《线性代数》

目录/Contents

目录/Contents

4.1 向量的内积、长度及正交性

一、向量的内积、长度

定义 1 设有 n 维向量 $\boldsymbol{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}, \boldsymbol{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$， 令

$[\mathrm{x},\mathrm{y}]=\mathrm{x}^{\mathrm{T}}\mathrm{y}=x_1y_1+x_2y_2+\cdots+x_ny_n$， 称 $[\boldsymbol{x},\boldsymbol{y}]$ 为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积.

内积的性质（其中 $\boldsymbol{x},\boldsymbol{y}$ 与 $\boldsymbol{z}$ 都是 n 维列向量， λ 为实数)：

(i) $[\boldsymbol{x},\boldsymbol{y}]=[\boldsymbol{y},\boldsymbol{x}]$；

(ii) $[\lambda\boldsymbol{x},\boldsymbol{y}]=\lambda[\boldsymbol{x},\boldsymbol{y}]=[\boldsymbol{x},\lambda\boldsymbol{y}]$；

(iii) $[\boldsymbol{x}+\boldsymbol{y},\boldsymbol{z}]=[\boldsymbol{x},\boldsymbol{z}]+[\boldsymbol{y},\boldsymbol{z}]$；

(iv) $[\boldsymbol{x},\boldsymbol{x}]\geq 0$， 当且仅当 $\boldsymbol{x}=\boldsymbol{0}$ 时， $[\boldsymbol{x},\boldsymbol{x}]=0$.

利用这些性质，还可以证明著名的柯西-施瓦茨（-Schwarz）不等式

$$[\mathrm{x},\mathrm{y}]^2\leq[\mathrm{x},\mathrm{x}][\mathrm{y},\mathrm{y}].$$

一、向量的内积、长度

定义 2 设有 n 维向量 $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$，令

$$\|x\| = \sqrt{[x,x]} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2},$$

称$\|\boldsymbol{x}\|$为向量 $\boldsymbol{x}$ 的长度（或范数）．

向量的长度具有下述性质：

(i) 非负性 当 $\boldsymbol{x} \neq \boldsymbol{0}$时，$\|\boldsymbol{x}\| > 0$；当 $\boldsymbol{x} = \boldsymbol{0}$时，$\|\boldsymbol{x}\| = 0$；

(ii) 齐次性 $\|\lambda \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$；

(iii) 三角不等式 $\|\boldsymbol{x}+\boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

证明 (i)与(ii)是显然的，下面证明(iii)．因为 $\|x+y\|^2 = [x+y, x+y] = [x,x] + 2[x,y] + [y,y]$,

由施瓦茨不等式,有 $[x,y] \leq \sqrt{[x,x][y,y]}$,

一、向量的内积、长度

从而 $$\|x+y\|^2 \le [x,x] + 2\sqrt{[x,x][y,y]} + [y,y] = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

即 $$\|x+y\| \le \|x\| + \|y\|.$$

当$\|\boldsymbol{x}\|=1$时，称 $\boldsymbol{x}$ 为单位向量. 如果$\boldsymbol{\alpha}\neq\boldsymbol{0}$，取$\boldsymbol{\beta}=\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$，则$\boldsymbol{\beta}$是一个单位向量.

由向量$\boldsymbol{\alpha}$得到单位向量$\boldsymbol{\beta}$的过程称为把向量$\boldsymbol{\alpha}$单位化.

定义 3 当$\boldsymbol{x}\neq\boldsymbol{0}, \boldsymbol{y}\neq\boldsymbol{0}$时,

$$\theta = \arccos\frac{[x,y]}{\|x\|\|y\|}$$ 称为 n 维向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的夹角.

当$[\boldsymbol{x},\boldsymbol{y}]=0$时，称向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 正交.

显然，若 $\boldsymbol{x}=\boldsymbol{0}$，则 $\boldsymbol{x}$ 与任何向量都正交.

二、正交向量组

定义 4 由一组两两正交的非零向量组成的向量组，称为正交向量组.

例如， 向量组

$$\begin{pmatrix}1\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\2\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\3\end{pmatrix}$$

与向量组

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \begin{pmatrix}0\\0\\\frac{\sqrt{2}}{2}\\-\frac{\sqrt{2}}{2}\end{pmatrix},\ \begin{pmatrix}0\\0\\\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{pmatrix}$$

都是正交向量组.

二、正交向量组

定理 1

若 n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是一个正交向量组，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

设有常数 $\lambda_1,\lambda_2,\cdots,\lambda_m$，使

$$\lambda_1\alpha_1+\lambda_2\alpha_2+\cdots+\lambda_m\alpha_m=0,$$

以 $\boldsymbol{\alpha}_i^{\mathrm{T}}\ (i=1,2,\cdots,m)$ 左乘上式两端，当 $j\neq i$ 时，$\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_j=0$，从而有 $\lambda_i\alpha_i^{\mathrm{T}}\alpha_i=0\ (i=1,2,\cdots,m)$，

因 $\boldsymbol{\alpha}_i\neq\mathbf{0}(i=1,2,\cdots,m)$， 故 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{\alpha}_i\neq 0$，

于是必有 $\lambda_i=0\ (i=1,2,\cdots,m)$，所以向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

例 1 已知 3 维空间 R^3 中的两个向量 $\alpha_1=\begin{pmatrix}1\\-1\\-1\end{pmatrix}$，$\alpha_2=\begin{pmatrix}1\\2\\-1\end{pmatrix}$ 正交，试求一个非零向量 α_3 使 $\alpha_1,\alpha_2,\alpha_3$ 两两正交.

解 记 $A=\begin{pmatrix}\alpha_1^T\\\alpha_2^T\end{pmatrix}=\begin{pmatrix}1&-1&-1\\1&2&-1\end{pmatrix}$，

α_3 应满足齐次线性方程组 $Ax=0$， 即 $\begin{pmatrix}1&-1&-1\\1&2&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$，

对系数矩阵 A 实施初等行变换，有 $A=\begin{pmatrix}1&-1&-1\\1&2&-1\end{pmatrix}\to\begin{pmatrix}1&-1&-1\\0&3&0\end{pmatrix}\to\begin{pmatrix}1&0&-1\\0&1&0\end{pmatrix}$，

得 $\begin{cases}x_1=x_3\\x_2=0\end{cases}$， 从而有基础解系 $\begin{pmatrix}1\\0\\1\end{pmatrix}$.

取 $\alpha_3=\begin{pmatrix}1\\0\\1\end{pmatrix}$， 则 α_3 为所求.

定义 5 设 n 维向量组 $\xi_1,\xi_2,\cdots,\xi_r$ 是向量空间 $V\left(V\subseteq\mathbf{R}^n\right)$ 的一个基，如果 $\xi_1,\xi_2,\cdots,\xi_r$ 两两正交，且都是单位向量， 则称 $\xi_1,\xi_2,\cdots,\xi_r$ 是 V 的一个**规范正交基**.

例如， n **维单位**坐标向量 $e_1,e_2,\cdots,e_n$ 就是 $\mathbf{R}^n$ 的一个规范正交基.

向量组 $\xi_1=\begin{pmatrix}\frac{2}{3}\\ \frac{1}{3}\\ \frac{2}{3}\end{pmatrix}$， $\xi_2=\begin{pmatrix}-\frac{2}{3}\\ \frac{2}{3}\\ \frac{1}{3}\end{pmatrix}$， $\xi_3=\begin{pmatrix}\frac{1}{3}\\ \frac{2}{3}\\ -\frac{2}{3}\end{pmatrix}$ 就是 $\mathbf{R}^3$ 的一个规范正交基.

若 $\xi_1,\xi_2,\cdots,\xi_r$ 是 V 的一个规范正交基， 那么 V 中任一向量 β 都能由 $\xi_1,\xi_2,\cdots,\xi_r$ 线性表示，

设表示式为 $\beta=\lambda\xi_1+\lambda_2\xi_2+\cdots+\lambda_r\xi_r$，

用 $\xi_i^{\mathrm{T}}\left(i=1,\cdots,r\right)$ 左乘上式，有 $\xi_i^{\mathrm{T}}\beta=\lambda_i\xi_i^{\mathrm{T}}\xi_i=\lambda_i\ (i=1,\cdots,r)$， 即 $\lambda_i=\xi_i^{\mathrm{T}}\beta=[\xi_i,\beta](i=1,\cdots,r)$.

三、施密特正交化过程

设$\alpha_1,\alpha_2,\cdots,\alpha_r$是向量空间$V$的一个基，从基$\alpha_1,\alpha_2,\cdots,\alpha_r$出发，找一组两两正交的单位向量，$\xi_1,\xi_2,\cdots,\xi_r$，使$\xi_1,\xi_2,\cdots,\xi_r$与$\alpha_1,\alpha_2,\cdots,\alpha_r$等价，这个过程称为把基$\alpha_1,\alpha_2,\cdots,\alpha_r$规范正交化.

具体步骤如下：

第一步，将基$\alpha_1,\alpha_2,\cdots,\alpha_r$正交化（施密特（Schmidt）正交化过程）.

即取

$$
\begin{aligned}
\beta_1 &= \alpha_1, \\
\beta_2 &= \alpha_2 - \frac{[\beta_1,\alpha_2]}{[\beta_1,\beta_1]}\beta_1, \\
\beta_3 &= \alpha_3 - \frac{[\beta_1,\alpha_3]}{[\beta_1,\beta_1]}\beta_1 - \frac{[\beta_2,\alpha_3]}{[\beta_2,\beta_2]}\beta_2, \\
&\cdots \quad \cdots \quad \cdots \quad \cdots \\
\beta_r &= \alpha_r - \frac{[\beta_1,\alpha_r]}{[\beta_1,\beta_1]}\beta_1 - \frac{[\beta_2,\alpha_r]}{[\beta_2,\beta_2]}\beta_2 - \cdots - \frac{[\beta_{r-1},\alpha_r]}{[\beta_{r-1},\beta_{r-1}]}\beta_{r-1}.
\end{aligned}
$$

第二步，将$\beta_1,\beta_2,\cdots,\beta_r$单位化，得到 $\xi_1=\frac{1}{\|\beta_1\|}\beta_1,\quad \xi_2=\frac{1}{\|\beta_2\|}\beta_2,\quad \cdots,\quad \xi_r=\frac{1}{\|\beta_r\|}\beta_r$

于是，$\xi_1,\xi_2,\cdots,\xi_r$就是V的一个规范正交基.

例 2 设 $\alpha_1=\begin{pmatrix}1\\1\\-1\end{pmatrix},\alpha_2=\begin{pmatrix}0\\4\\1\end{pmatrix},\alpha_3=\begin{pmatrix}-2\\1\\1\end{pmatrix}$ 是 $\mathbf{R}^3$ 的一个基，求一个与 $\alpha_1,\alpha_2,\alpha_3$ 等价的规范正交基.

解 取

$$\beta_1=\alpha_1=\begin{pmatrix}1\\1\\-1\end{pmatrix},$$

$$\beta_2=\alpha_2-\frac{[\beta_1,\alpha_2]}{[\beta_1,\beta_1]}\beta_1=\begin{pmatrix}0\\4\\1\end{pmatrix}-\frac{3}{3}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}-1\\3\\2\end{pmatrix},$$

$$\beta_3=\alpha_3-\frac{[\beta_1,\alpha_3]}{[\beta_1,\beta_1]}\beta_1-\frac{[\beta_2,\alpha_3]}{[\beta_2,\beta_2]}\beta_2$$

$$=\begin{pmatrix}-2\\1\\1\end{pmatrix}-\frac{-2}{3}\begin{pmatrix}1\\1\\-1\end{pmatrix}-\frac{7}{14}\begin{pmatrix}-1\\3\\2\end{pmatrix}=\begin{pmatrix}-\frac{5}{6}\\\frac{1}{6}\\-\frac{2}{3}\end{pmatrix},$$

再将 β_1,β_2,β_3 单位化，得到 $\xi_1=\frac{1}{\|\beta_1\|}\beta_1=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix},\xi_2=\frac{1}{\|\beta_2\|}\beta_2=\frac{1}{\sqrt{14}}\begin{pmatrix}-1\\3\\2\end{pmatrix},\xi_3=\frac{1}{\|\beta_3\|}\beta_3=\frac{1}{\sqrt{42}}\begin{pmatrix}-5\\1\\-4\end{pmatrix},$

ξ_1,ξ_2,ξ_3 即为所求.

例3 已知$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\-1\\1\end{pmatrix}$，求一组非零向量$\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$，使$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$两两正交.

解 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$应满足方程$\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{x}=0$，即 $x_1-x_2+x_3=0$，

它的基础解系为 $\xi_1=\begin{pmatrix}1\\1\\0\end{pmatrix}$，$\xi_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}$，

令 $\alpha_2=\xi_1=\begin{pmatrix}1\\1\\0\end{pmatrix}$，$\alpha_3=\xi_2-\dfrac{[\alpha_2,\xi_2]}{[\alpha_2,\alpha_2]}\alpha_2=\begin{pmatrix}1\\0\\-1\end{pmatrix}-\dfrac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix}=\dfrac{1}{2}\begin{pmatrix}1\\-1\\-2\end{pmatrix}$，

则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$两两正交.

定义 6 如果 n 阶矩阵 A 满足 $A^TA = E$ (即$A^{-1} = A^T$)， 那么称 A 为正交矩阵，简称正交阵.

定理 2 设矩阵 A 是 n 阶方阵，则下列结论等价：

1 A 是 n 阶正交阵； 2 A 的列向量组是 $\mathbf{R}^n$ 的一个规范正交基；

3 A 的行向量组是 $\mathbf{R}^n$ 的一个规范正交基.

证明 (1) $\Leftrightarrow$ (2)： 将矩阵 A 按列分块 $A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 如果 A 是 n 阶正交阵，

则公式 $A^TA=E$ 可表示为 $$\begin{pmatrix}\alpha_1^T\\ \alpha_2^T\\ \vdots\\ \alpha_n^T\end{pmatrix}(\alpha_1,\alpha_2,\cdots,\alpha_n)=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{pmatrix},$$

亦即 $$\alpha_i^T\alpha_j=\delta_{ij}=\begin{cases}1, & 当 i=j,\\ 0, & 当 i\neq j\end{cases}\quad (i,j=1,2,\cdots,n),$$

这说明 A 的列向量都是 n 维单位向量，且两两正交，从而是 $\mathbf{R}^n$ 的一个规范正交基.

四、正交矩阵

(1) $\Leftrightarrow$ (3)： 因为 $A^TA=E$ 与 $AA^T=E$ 等价，所以将矩阵 A 按行分块 $A=\begin{pmatrix}\beta_1^T\\ \beta_2^T\\ \vdots\\ \beta_n^T\end{pmatrix}$，

于是公式 $AA^T=E$ 可表示为

$$AA^T=\begin{pmatrix}\beta_1^T\\ \beta_2^T\\ \vdots\\ \beta_n^T\end{pmatrix}(\beta_1,\beta_2,\cdots,\beta_n)=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{pmatrix},$$

所以 $$\beta_i^T\beta_j=\delta_{ij}=\begin{cases}1, & 当 i=j,\\ 0, & 当 i\neq j\end{cases}\quad (i,j=1,2,\cdots,n),$$

即：A 的行向量也都是 n 维单位向量，且两两正交，从而是 $\mathbf{R}^n$ 的一个规范正交基.

四、正交矩阵

例4 验证矩阵 $P=\begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ 是正交阵.

证明 容易验证 $\boldsymbol{P}$ 的每个列向量都是单位向量，且两两正交， 所以 $\boldsymbol{P}$ 是正交阵.

正交矩阵具有如下性质：

(i) 若 $\boldsymbol{A}$ 为正交阵,则 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$ 也是正交阵，且 $|\boldsymbol{A}|=1$ 或 -1；

(ii) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都是正交阵，则 $\boldsymbol{AB}$ 也是正交阵.

定义 7 若 $\boldsymbol{P}$ 为正交矩阵，则线性变换 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 称为正交变换. 设 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 为正交变换，

则有 $\|\boldsymbol{y}\|=\sqrt{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x}}=\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}=\|\boldsymbol{x}\|$. 因此正交变换保持向量的长度不变.

目录/Contents

目录/Contents

4.2 方阵的特征值与特征向量

定 义

设 A 是 n 阶矩阵，如果数 λ 和 n 维非零列向量 使关系式

$$A\alpha=\lambda\alpha$$

成立，那么数 λ 称为矩阵 A 的特征值，非零向量 α 称为 A 的对应于特征值 λ 的特征向量.

例如，矩阵 $A=\begin{pmatrix}-1&2&0\\0&3&0\\2&1&-1\end{pmatrix}$，$\alpha=\begin{pmatrix}1\\2\\1\end{pmatrix}$，则有 $A\alpha=\begin{pmatrix}-1&2&0\\0&3&0\\2&1&-1\end{pmatrix}\begin{pmatrix}1\\2\\1\end{pmatrix}=\begin{pmatrix}3\\6\\3\end{pmatrix}=3\begin{pmatrix}1\\2\\1\end{pmatrix}$

所以数3是矩阵 A 的特征值，α 是 A 的对应于特征值3的特征向量.

一、方阵的特征值与特征向量的概念及其求法

一个任意给定的 n 阶矩阵 A 会有多少个特征值？ 对应的特征向量又该如何求呢？假设矩阵 A 有特征值 λ，对应于特征值 λ的特征向量为 α，则有 $A\alpha=\lambda\alpha$.

将 $A\alpha=\lambda\alpha$ 改写成

$$(A-\lambda E)\alpha=0,$$

可见，α 是 n 个未知数 n 个方程的齐次线性方程组 $(A-\lambda E)x=0$ 的非零解.

而方程组有非零解的充分必要条件是系数行列式等于零，即

$$|A-\lambda E|=0$$

一、方阵的特征值与特征向量的概念及其求法

记

$$f(\lambda)=|\mathrm{A}-\lambda\mathrm{E}|=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} & a_{12} & \cdots & a_{nn}-\lambda \end{vmatrix},$$

则 $f(\lambda)$ 是 λ 的 n 次多项式，称为矩阵 A 的特征多项式. 从而公式 $|\mathrm{A}-\lambda\mathrm{E}|=0$ 可以写成 $f(\lambda)=0$，这是以 λ 为未知数的一元 n 次方程，称为 A 的特征方程，而 A 的特征值就是特征方程的根. 我们知道，一元 n 次方程在复数范围内恒有 n 个根 (重根按重数计算). 因此，n 阶矩阵 A 在复数范围内有 n 个特征值，通过解矩阵 A 的特征方程就可以得到这 n 个特征值.

设 $\lambda=\lambda_i$ 为矩阵 A 的一个特征值，则由方程 $(A-\lambda_i E)x=0$ 可求得非零解 $x=\alpha_i$ ，那么 α_i 便是 A 的对应于特征值 λ_i 的特征向量. （若 λ_i 为实数，则 α_i 可取实向量；若 λ_i 为复数，则 α_i 可取复向量.）

例 1

求矩阵 $A=\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}$ 的特征值和特征向量.

解

矩阵 A 的特征多项式为

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

所以 A 的全部特征值为 $\lambda_1=1$ ，$\lambda_2=2$ ， $\lambda_3=3.$

由此例可知，对角矩阵的全部特征值就是它的对角线上的元素.

一、方阵的特征值与特征向量的概念及其求法

当 $\lambda_1 = 1$ 时，解方程 $(A - E)x = 0$ ，

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

得基础解系 $\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

当 $\lambda_2 = 2$ 时，解方程 $(A - 2E)x = 0$ ，

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

得基础解系 $\alpha_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

于是 $k\alpha_1 (k \neq 0)$ 是对应于特征值 $\lambda_1 = 1$ 的全部特征向量.

于是 $k\alpha_2 (k \neq 0)$ 是对应于特征值 $\lambda_2 = 2$ 的全部特征向量.

当 $\lambda_3 = 3$ 时，解方程 $(A - 3E)x = 0$ ，

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

得基础解系 $$\alpha_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

于是 $k\alpha_3 (k \neq 0)$ 是对应于特征值 $\lambda_3 = 3$ 的全部特征向量.

例 2

求矩阵 $B=\begin{pmatrix}-1 & 2 & 0\\ 0 & 3 & 0\\ 2 & 1 & -1\end{pmatrix}$ 的特征值和特征向量

解

B 的特征多项式为 $|B-\lambda E|=\begin{vmatrix}-1-\lambda & 2 & 0\\ 0 & 3-\lambda & 0\\ 2 & 1 & -1-\lambda\end{vmatrix}=(1+\lambda)^2(3-\lambda)$,

所以 B 的全部特征值为 $\boxed{\lambda_1=\lambda_2=-1,\lambda_3=3.}$

一、方阵的特征值与特征向量的概念及其求法

当 $\lambda_1=\lambda_2=-1$ 时，解方程 $(B+E)x=0$ ，

由 $$B+E=\begin{pmatrix}0&2&0\\0&4&0\\2&1&0\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$$

得基础解系 $$\alpha_1=\begin{pmatrix}0\\0\\1\end{pmatrix}$$

当 $\lambda_3=3$ 时，解方程 $(B-3E)x=0$ ，

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

得基础解系 $$\alpha_2=\begin{pmatrix}1\\2\\1\end{pmatrix}$$

对应于 $\lambda_1=\lambda_2=-1$ 的全部特征向量为 $k\alpha_1$（常数 $k\neq0$ ）.

对应于 $\lambda_3=3$ 的全部特征向量为 $k\alpha_2$（常数 $k\neq0$ ）.

例3

求矩阵$C=\begin{pmatrix}1&0&2\\0&3&0\\2&0&1\end{pmatrix}$的特征值和特征向量.

解

矩阵C的特征多项式为

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

所以C的全部特征值为$\lambda_1=\lambda_2=3$，$\lambda_3=-1$.

一、方阵的特征值与特征向量的概念及其求法

当 $\lambda_1=\lambda_2=3$ 时，解方程 $(\boldsymbol{C}-3\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，

由 $C-3E=\begin{pmatrix}-2 & 0 & 2\\ 0 & 0 & 0\\ 2 & 0 & -2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$

得基础解系 $\alpha_1=\begin{pmatrix}0\\1\\0\end{pmatrix}$，$\alpha_2=\begin{pmatrix}1\\0\\1\end{pmatrix}$，

当 $\lambda_3=-1$ 时，解方程 $(\boldsymbol{C}+\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，

由 $C+E=\begin{pmatrix}2 & 0 & 2\\ 0 & 4 & 0\\ 2 & 0 & 2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}$

得基础解系 $\alpha_3=\begin{pmatrix}1\\0\\-1\end{pmatrix}$，

$\boldsymbol{\alpha}_1$、$\boldsymbol{\alpha}_2$ 就是对应于 $\lambda_1=\lambda_2=3$ 的两个线性无关的特征向量，对应于 $\lambda_1=\lambda_2=3$ 的全部特征向量为 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2$（$k_1$、$k_2$ 不同时为零). 对应于 $\lambda_3=-1$ 的全部特征向量为 $k\boldsymbol{\alpha}_3\ (k\neq 0)$.

性质1

设 n 阶矩阵 $A=(a_{ij})$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ ，则

(i) $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$,

(ii) $\lambda_1\lambda_2\cdots\lambda_n=|A|$.

由此可见， n 阶方阵 A 可逆的充分必要条件是 A 的特征值全不为零.

二、方阵的特征值与特征向量的性质

性质2

若 λ 是方阵 A 的特征值， α 为对应于特征值 λ 的特征向量，则

01 OPTION

λ^k 是方阵 A^k 的特征值（ k 为非负整数），对应于特征值 λ^k 的特征向量是 α；

02 OPTION

$k\lambda$ 是方阵 kA 的特征值（ k 为任意常数），对应于特征值 $k\lambda$ 的特征向量是 α；

03 OPTION

当 A 可逆时，λ^{-1} 是方阵 A^{-1} 的特征值，对应于特征值 λ^{-1} 的特征向量是 α；

04 OPTION

若矩阵 A 的多项式是 $\varphi(A)=a_mA^m+\cdots+a_1A+a_0E$， 则方阵 $\varphi(A)$ 的特征值是 $\varphi(\lambda)$（其中 $\varphi(\lambda)=a_m\lambda^m+\cdots+a_1\lambda+a_0$ 是关于 λ 的多项式），对应于特征值 $\varphi(\lambda)$ 的特征向量是 α .

证

因 λ 是方阵 A 的特征值，α 为对应于特征值 λ 的特征向量，故有 $A\alpha=\lambda\alpha$. 于是

(i) $A^k\alpha=A^{k-1}(A\alpha)=A^{k-1}(\lambda\alpha)=\lambda\left(A^{k-1}\alpha\right)=\lambda A^{k-2}(A\alpha)=\lambda^2A^{k-2}\alpha=\cdots=\lambda^k\alpha,$

所以 λ^k 是方阵 A^k 的特征值，对应于特征值 λ^k 的特征向量是 α .

(ii) $(kA)\alpha=k(A\alpha)=k(\lambda\alpha)=(k\lambda)\alpha,$

所以 $k\lambda$ 是方阵 kA 的特征值，对应于特征值 $k\lambda$ 的特征向量是 α .

当 A 可逆时，特征值均不为零，于是

(iii) $A^{-1}A = E \Rightarrow A^{-1}(A\alpha) = E\alpha \Rightarrow \lambda A^{-1}\alpha = \alpha \Rightarrow A^{-1}\alpha = \lambda^{-1}\alpha,$

所以 λ^{-1} 是方阵A^{-1} 的特征值，对应于特征值 λ^{-1}的特征向量是 α .

由(i)可知，

(iv)

$$\varphi(A)\alpha = (a_m A^m + \cdots + a_1 A + a_0 E)\alpha = a_m A^m \alpha + \cdots + a_1 A\alpha + a_0 E\alpha$$

$$= a_m \lambda^m \alpha + \cdots + a_1 \lambda \alpha + a_0 \alpha = (a_m \lambda^m + \cdots + a_1 \lambda + a_0)\alpha = \varphi(\lambda)\alpha,$$

所以方阵 $\varphi(A)$ 的特征值是 $\varphi(\lambda)$ ，对应于特征值 $\varphi(\lambda)$ 的特征向量是 α .

例 4

设3阶矩阵的特征值为 $1,2,3$ ，求 $2A^* - 3A + 2E$ 的特征值.

解

因 A 的特征值全不为0，知 A 可逆，故 $\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{A}^{-1}$. 而 $|A| = \lambda_1\lambda_2\lambda_3 = 6$ ，记

$$\varphi(A) = 2A^* - 3A + 2E = 12A^{-1} - 3A + 2E.$$

这里，$\varphi(A)$虽不是矩阵多项式，但也具有矩阵多项式的特性，从而可利用性质2(iv)来计算 $\varphi(A)$ 的特征值. 由 $\varphi(\lambda) = 12\lambda^{-1} - 3\lambda + 2$ $\varphi(A)$的特征值为

$$\varphi(1) = 6 - 3 + 2 = 5,\ \varphi(2) = \frac{6}{2} - 3 \times 2 + 2 = -1,\ \varphi(3) = \frac{6}{3} - 3 \times 3 + 2 = -5,$$

性质3

如果 α_1与 α_2 是方阵 A 的同一特征值 λ 所对应的特征向量，则 $k_1\alpha_1+k_2\alpha_2$ (k_1、k_2 不同时为零)也是特征值 λ 所对应的特征向量.

证明 由 $A\alpha_1=\lambda\alpha_1$ ， $A\alpha_2=\lambda\alpha_2$ 得

$$A(k_1\alpha_1+k_2\alpha_2)=A(k_1\alpha_1)+A(k_2\alpha_2)=k_1(A\alpha_1)+k_2(A\alpha_2)=k_1\lambda\alpha_1+k_2\lambda\alpha_2=\lambda(k_1\alpha_1+k_2\alpha_2),$$

所以 $k_1\alpha_1+k_2\alpha_2$ (k_1、k_2 不同时为零)也是特征值 λ 所对应的特征向量.

性质4

设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是方阵 A 的 m 个互不相同的特征值，$\alpha_1,\alpha_2,\cdots,\alpha_m$ 是依次与之对应的特征向量，则

$\alpha_1,\alpha_2,\cdots,\alpha_m$ 线性无关.

性质5

设 λ_1 和 λ_2 是矩阵 A 的两个不同的特征值，$\alpha_1,\alpha_2,\cdots,\alpha_s$ 和 $\beta_1,\beta_2,\cdots,\beta_t$ 是分别对应于 λ_1 和 λ_2 的线性无关的特征向量，则

$\alpha_1,\alpha_2,\cdots,\alpha_s,\beta_1,\beta_2,\cdots,\beta_t$ 线性无关.

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

证明 按题设，有$\boldsymbol{A\alpha}_1=\lambda_1\boldsymbol{\alpha}_1$，$\boldsymbol{A\alpha}_2=\lambda_2\boldsymbol{\alpha}_2$. 假设$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$是$\boldsymbol{A}$的特征向量，则应该存在数$\lambda$，

使

$$A(\alpha_1+\alpha_2)=\lambda(\alpha_1+\alpha_2).$$

另一方面，

$$A(\alpha_1+\alpha_2)=\lambda_1\alpha_1+\lambda_2\alpha_2,$$

于是 $\lambda(\alpha_1+\alpha_2)=\lambda_1\alpha_1+\lambda_2\alpha_2,$ 即 $(\lambda_1-\lambda)\alpha_1+(\lambda_2-\lambda)\alpha_2=0.$

因$\lambda_1\neq\lambda_2$，所以$\boldsymbol{\alpha}_1$和$\boldsymbol{\alpha}_2$线性无关，从而由上式得$\lambda_1-\lambda=\lambda_2-\lambda=0$，

即$\lambda_1=\lambda_2$，与题设矛盾. 因此$\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2$不是$\boldsymbol{A}$的特征向量.

目录/Contents

目录/Contents

4.3 相似矩阵

一、方阵相似的定义与性质

定义 1　设 A,B 都是 n 阶矩阵,若有可逆矩阵 $\boldsymbol{P}$,使

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

则称 $\boldsymbol{B}$ 是 $\boldsymbol{A}$ 的相似矩阵,或者说矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似.对 $\boldsymbol{A}$ 进行运算 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 称为对 $\boldsymbol{A}$ 进行相似变换,可逆矩阵 $\boldsymbol{P}$ 称为把 $\boldsymbol{A}$ 变成 $\boldsymbol{B}$ 的相似变换矩阵.

定理 1

若 n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似,则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式,从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值.

证明　因 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，即有可逆矩阵 P，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$,故

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

推论　若 n 阶矩阵 $\boldsymbol{A}$ 与对角阵

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

相似,则 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 即是 $\boldsymbol{A}$ 的 n 个特征值.

一、方阵相似的定义与性质

若 n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，即 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$，则 $\boldsymbol{A}^k=\left(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\right)^k=\boldsymbol{P}\boldsymbol{B}^k\boldsymbol{P}^{-1}$，并且 $\boldsymbol{A}$ 的多项式

$$\varphi(\boldsymbol{A})=a_m\boldsymbol{A}^m+\cdots+a_1\boldsymbol{A}+a_0\boldsymbol{E}=a_m\left(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\right)^m+\cdots+a_1\left(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\right)+a_0E=a_m\left(\boldsymbol{P}\boldsymbol{B}^m\boldsymbol{P}^{-1}\right)+\cdots+a_1\left(\boldsymbol{P}\boldsymbol{B}\boldsymbol{P}^{-1}\right)+a_0\boldsymbol{E}$$

$$=\boldsymbol{P}\left(a_m\boldsymbol{B}^m\right)\boldsymbol{P}^{-1}+\cdots+\boldsymbol{P}\left(a_1\boldsymbol{B}\right)\boldsymbol{P}^{-1}+\boldsymbol{P}\left(a_0\boldsymbol{E}\right)\boldsymbol{P}^{-1}=\boldsymbol{P}\left(a_m\boldsymbol{B}^m+\cdots+a_1\boldsymbol{B}+a_0\boldsymbol{E}\right)$$

特别地，若有可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 为对角阵，则

$$\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1},\quad \varphi(\boldsymbol{A})=\boldsymbol{P}\varphi(\boldsymbol{\Lambda})\boldsymbol{P}^{-1}.$$

而对于对角阵 $\boldsymbol{\Lambda}=diag\left(\lambda_1,\lambda_2,\cdots,\lambda_n\right)$，有

$$\boldsymbol{\Lambda}^k=\begin{pmatrix}\lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k\end{pmatrix},\quad \varphi(\boldsymbol{\Lambda})=\begin{pmatrix}\varphi(\lambda_1) & & & \\ & \varphi(\lambda_2) & & \\ & & \ddots & \\ & & & \varphi(\lambda_n)\end{pmatrix},$$

由此可方便地计算 $\boldsymbol{A}$ 的高次幂 $\boldsymbol{A}^k$ 及 $\boldsymbol{A}$ 的多项式 $\varphi(\boldsymbol{A})$.

设 $f(\lambda)$是矩阵 $\boldsymbol{A}$ 的特征多项式,则

$$f(\boldsymbol{A})=\boldsymbol{O}.$$

这个结论的证明比较困难，但若 $\boldsymbol{A}$ 与对角阵相似，则容易证明此结论.

这是因为：若 $\boldsymbol{A}$ 与对角阵相似，即有可逆矩阵 $\boldsymbol{P}$ ，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$，其中

λ_i 为 $\boldsymbol{A}$ 的特征值，有 $f(\lambda_i)=0$.

于是由上面的讨论可得：

$$f(\boldsymbol{A})=\boldsymbol{P}f(\Lambda)\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{pmatrix} f(\lambda_1) & & & \\ & f(\lambda_2) & & \\ & & \ddots & \\ & & & f(\lambda_n) \end{pmatrix}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{O}\boldsymbol{P}^{-1}=\boldsymbol{O}.$$

二、方阵的相似对角化

把矩阵 P 列分块为 $P=(p_1,p_2,\cdots,p_n)$,

由 $P^{-1}AP=\Lambda$ ，得 $AP=P\Lambda$ ，即

$$A(p_1,p_2,\cdots,p_n)=(p_1,p_2,\cdots,p_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=(\lambda_1 p_1,\lambda_2 p_2,\cdots,\lambda_n p_n).$$

于是有 $Ap_i=\lambda_i p_i \quad (i=1,2,\cdots,n).$

可见 λ_i 为 A 的特征值，而 P 的列向量 p_i 就是 A 对应于特征值 λ_i 的特征向量.

反之，如果 n 阶矩阵 A 恰好有 n 个特征向量，则这 n 个特征向量即可构成矩阵 P，使得 $AP = P\Lambda$. 并且这 n 个特征向量必定是线性无关的，从而 P 可逆，因此有 $P^{-1}AP = \Lambda$.

由上面的讨论即有：

定理2

n 阶矩阵 A 与对角阵相似(即 A 能对角化)的充分必要条件是 A 有 n 个线性无关的特征向量.

推论

如果 n 阶矩阵 A 的 n 个特征值互不相等，则 A 与对角阵相似.

例 1

设 $A=\begin{pmatrix}0&0&1\\x&1&y\\1&0&0\end{pmatrix}$ 有三个线性无关的特征向量，求 x 与 y 应满足的条件.

解

因为矩阵 A 是3阶矩阵，又有三个线性无关的特征向量， 所以A 可以相似对角化. 由

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

得到 A 的特征值为 $\lambda_1 = \lambda_2 = 1,\ \lambda_3 = -1.$

对应单根 $\lambda_3 = -1$，可求得线性无关的特征向量恰好有1 个， 故对应重根 $\lambda_1 = \lambda_2 = 1$ 应有2个线性无关的特征向量， 即方程 $(A - E)x = 0$ 有 2个线性无关的解， 亦即系数矩阵 $A - E$ 的秩 $R(A - E) = 1$.

由 $$A - E = \begin{pmatrix} -1 & 0 & 1 \\ x & 0 & y \\ 1 & 0 & -1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & x+y \\ 0 & 0 & 0 \end{pmatrix}$$

可知，要使系数矩阵 $A - E$ 的秩 $R(A - E) = 1$ ，必须 $x + y = 0$.

目录/Contents

目录/Contents

4.4 实对称矩阵的相似对角化

一、实对称矩阵的特征值和特征向量的性质

二、实对称矩阵的相似对角化

性质1 实对称矩阵的特征值为实数.

证明 先介绍一个记号. 设复数矩阵 $X=(x_{ij})$，复数 x_{ij} 的共轭复数为 $\overline{x}_{ij}$，记 $\overline{X}=(\overline{x}_{ij})$，则矩阵 $\overline{X}$ 称为矩阵 X 的共轭矩阵.

设复数 λ 为对称阵 $\boldsymbol{A}$ 的特征值，复向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 为对应的特征向量，即 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$.用 $\overline{\lambda}$ 表示 λ 的共轭复数，$\overline{\boldsymbol{x}}=(\overline{x}_1,\overline{x}_2,\cdots,\overline{x}_n)^{\mathrm{T}}$ 表示 $\boldsymbol{x}$ 的共轭复向量，而 $\boldsymbol{A}$ 为实对称矩阵，有 $\overline{\boldsymbol{A}}=\boldsymbol{A}$ 及 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$， 于是

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

且

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

两式相减，得

$$\left(\bar{\lambda}-\lambda\right)\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=0,$$

由 $\boldsymbol{x}\neq\boldsymbol{0}$ 可知

$$\bar{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=\sum_{i=1}^{n}\bar{x}_i x_i=\sum_{i=1}^{n}\left|\boldsymbol{x}\right|^2\neq 0,$$

故 $\bar{\lambda}-\lambda=0$，即 $\lambda=\bar{\lambda}$，这就说明 λ 为实数.

当特征值 λ_i 为实数时，齐次线性方程组

$$\left(\boldsymbol{A}-\lambda_i\boldsymbol{E}\right)\boldsymbol{x}=\boldsymbol{0}$$

是实系数方程组，由 $\left|\boldsymbol{A}-\lambda_i\boldsymbol{E}\right|=0$ 知必有实的基础解系，所以对应的特征向量可以取实向量.

性质 2 设 λ_1,λ_2 是对称阵 $\boldsymbol{A}$ 的两个特征值，$\boldsymbol{p}_1,\boldsymbol{p}_2$ 是对应的两个特征向量. 若 $\lambda_1\neq\lambda_2$，则 $\boldsymbol{p}_1$ 与 $\boldsymbol{p}_2$ 正交.

证明 已知 $\boldsymbol{A}\boldsymbol{p}_1=\lambda_1\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2=\lambda_2\boldsymbol{p}_2,\lambda_1\neq\lambda_2$，且 $\boldsymbol{A}$ 对称，于是

$$\lambda_1\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=\left(\lambda_1\boldsymbol{p}_1^{\mathrm{T}}\right)\boldsymbol{p}_2=\left(\lambda_1\boldsymbol{p}_1\right)^{\mathrm{T}}\boldsymbol{p}_2=\left(\boldsymbol{A}\boldsymbol{p}_1\right)^{\mathrm{T}}\boldsymbol{p}_2=\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{p}_2=\boldsymbol{p}_1^{\mathrm{T}}\left(\boldsymbol{A}\boldsymbol{p}_2\right)=\boldsymbol{p}_1^{\mathrm{T}}\left(\lambda_2\boldsymbol{p}_2\right)=\lambda_2\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2\,,$$

即

$$\left(\lambda_1-\lambda_2\right)\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=0\,.$$

但 $\lambda_1\neq\lambda_2$，故 $\boldsymbol{p}_1^{\mathrm{T}}\boldsymbol{p}_2=0$，即 $\boldsymbol{p}_1$ 与 $\boldsymbol{p}_2$ 正交.

二、实对称矩阵的相似对角化

定理 n阶实对称阵A必定正交相似于实对角阵Λ，即存在正交阵$\boldsymbol{P}$，使

$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\Lambda$，其中$\Lambda$的对角线上的元素是$\boldsymbol{A}$的$n$个特征值.

推论 设$\boldsymbol{A}$为n阶实对称阵，λ是A的特征方程的k重根，则矩阵$\boldsymbol{A}-\lambda\boldsymbol{E}$的秩

$R(\boldsymbol{A}-\lambda\boldsymbol{E})=n-k$，从而对应特征值$\lambda$有$k$个线性无关的特征向量.

证明 按定理知对称阵$\boldsymbol{A}$与对角阵$\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$相似，从而$\boldsymbol{A}-\lambda\boldsymbol{E}$与$\Lambda-\lambda\boldsymbol{E}=\mathrm{diag}(\lambda_1-\lambda,\lambda_2-\lambda,\cdots,\lambda_n-\lambda)$相似. 当$\lambda$是$\boldsymbol{A}$的$k$重特征根时，$\lambda_1,\lambda_2,\cdots,\lambda_n$这$n$个特征值中有$k$个等于$\lambda$，有$n-k$个不等于$\lambda$，从而对角阵$\Lambda-\lambda\boldsymbol{E}$的对角元恰有$k$个等0，有$n-k$个不等于0，因此$R(\Lambda-\lambda\boldsymbol{E})=n-k$.于是有$R(\boldsymbol{A}-\lambda\boldsymbol{E})=R(\Lambda-\lambda\boldsymbol{E})=n-k$.

将对称阵 $\boldsymbol{A}$ 对角化的步骤如下：

(i) 求出 $\boldsymbol{A}$ 的全部互不相等的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_s$，它们的重数依次为 $k_1,k_2,\cdots,k_s\ (k_1+k_2+\cdots+k_s=n)$；

(ii) 对于每个 k_i 重特征值 λ_i，求方程 $(\boldsymbol{A}-\lambda_i\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 的基础解系，得 k_i 个线性无关的特征向量，再把它们正交化、单位化，得 k_i 个两两正交的单位特征向量．因 $k_1+k_2+\cdots+k_s=n$，故总共可得 n 个两两正交的单位特征向量．

(iii) 把这 n 个两两正交的单位特征向量构成正交阵 $\boldsymbol{P}$，便有 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{AP}=\boldsymbol{\Lambda}$．注意 $\boldsymbol{\Lambda}$ 中对角元的排列次序应与 $\boldsymbol{P}$ 中列向量的排列次序相对应．

例1 设矩阵 $A=\begin{pmatrix}1&0&2\\0&-1&0\\3&0&2\end{pmatrix}$，求正交阵 P ，使得 $P^{-1}AP=P^{\mathrm{T}}AP$ 为对角阵.

解 由

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

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

对特征值 $\lambda_1=4$，解齐次线性方程组 $(A-4E)x=\mathbf{0}$，由

$$A-4E=\begin{pmatrix}-3&0&2\\0&-5&0\\3&0&-2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&0&-\frac{2}{3}\\0&1&0\\0&0&0\end{pmatrix},$$

取特征向量为 $\alpha_1=\begin{pmatrix}2\\0\\3\end{pmatrix}$，单位化，得 $\eta_1=\frac{1}{\|\alpha_1\|}\alpha_1=\begin{pmatrix}\frac{2}{\sqrt{13}}\\0\\\frac{3}{\sqrt{13}}\end{pmatrix}$.

对特征值$\lambda_2=\lambda_3=-1$，解齐次线性方程组$(\boldsymbol{A}+\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，由

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

取特征向量为$\boldsymbol{\alpha}_2=\begin{pmatrix}0\\1\\0\end{pmatrix}$，$\boldsymbol{\alpha}_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}$. 由于$\boldsymbol{\alpha}_2$与$\boldsymbol{\alpha}_3$已经正交，所以只需将这两个向量单位化，得

$$\boldsymbol{\eta}_2=\boldsymbol{\alpha}_2=\begin{pmatrix}0\\1\\0\end{pmatrix},\quad \boldsymbol{\eta}_3=\frac{1}{\|\boldsymbol{\alpha}_3\|}\boldsymbol{\alpha}_3=\begin{pmatrix}-\frac{\sqrt{2}}{2}\\0\\\frac{\sqrt{2}}{2}\end{pmatrix}.$$

令矩阵$\boldsymbol{P}=(\eta_1,\eta_2,\eta_3)=\begin{pmatrix}\frac{2}{\sqrt{13}}&0&-\frac{\sqrt{2}}{2}\\0&1&0\\\frac{3}{\sqrt{13}}&0&\frac{\sqrt{2}}{2}\end{pmatrix}$，则$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}4&&\\&-1&\\&&-1\end{pmatrix}$.

例2 设 $A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}$，求 A^{10}.

解 因为 A 是实对称阵，从而可求一个正交阵 P，使得 $P^{-1}AP=\Lambda=\begin{pmatrix}\lambda_1&&\\&\lambda_2&\\&&\lambda_3\end{pmatrix}$，其中 $\lambda_1,\lambda_2,\lambda_3$

是 A 的全部特征值.于是

$$A^{10}=\left(P\Lambda P^{-1}\right)^{10}=P\Lambda^{10}P^{-1}=P\Lambda^{10}P^{\mathrm{T}}.$$

由

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

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

对特征值 $\lambda_1 = 4$，解齐次线性方程组 $(\boldsymbol{A}-4\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，由

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

取特征向量为 $\boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$，单位化，得 $\boldsymbol{p}_1 = \dfrac{1}{\|\boldsymbol{\alpha}_1\|}\boldsymbol{\alpha}_1 = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix}$.

求特征向量可视作求基础解系

对特征值 $\lambda_2 = \lambda_3 = 1$，解齐次线性方程组 $(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$，由

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

取特征向量为 $\boldsymbol{\alpha}_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$，$\boldsymbol{\alpha}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$.

先将$\boldsymbol{\alpha}_2$与$\boldsymbol{\alpha}_3$正交化：令

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad \boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2=\begin{pmatrix}-1\\0\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}-1\\1\\0\end{pmatrix}=\frac{1}{2}\begin{pmatrix}-1\\-1\\2\end{pmatrix},$$

再将$\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$单位化，得

$$\boldsymbol{p}_2=\frac{1}{\|\boldsymbol{\beta}_2\|}\boldsymbol{\beta}_2=\begin{pmatrix}-\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\\0\end{pmatrix},\quad \boldsymbol{p}_3=\frac{1}{\|\boldsymbol{\beta}_3\|}\boldsymbol{\beta}_3=\begin{pmatrix}-\frac{\sqrt{6}}{6}\\-\frac{\sqrt{6}}{6}\\\frac{\sqrt{6}}{3}\end{pmatrix}。$$

令矩阵$\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3)=\begin{pmatrix}\frac{\sqrt{3}}{3}&-\frac{\sqrt{2}}{2}&-\frac{\sqrt{6}}{6}\\\frac{\sqrt{3}}{3}&\frac{\sqrt{2}}{2}&-\frac{\sqrt{6}}{6}\\\frac{\sqrt{3}}{3}&0&\frac{\sqrt{6}}{3}\end{pmatrix}$，则$\boldsymbol{P}$为所求正交阵，且$\boldsymbol{\Lambda}=\begin{pmatrix}4&&\\&1&\\&&1\end{pmatrix}$，

从而

$$A^{10}=P\begin{pmatrix}4&&\\&1&\\&&1\end{pmatrix}^{10}P^{\mathrm{T}}=P\begin{pmatrix}4^{10}&&\\&1&\\&&1\end{pmatrix}P^{\mathrm{T}}=\frac{1}{3}\begin{pmatrix}4^{10}+2&4^{10}-1&4^{10}-1\\4^{10}-1&4^{10}+2&4^{10}-1\\4^{10}-1&4^{10}-1&4^{10}+2\end{pmatrix}.$$

目录/Contents

目录/Contents

4.5 二次型及其标准形

一、二次型及其标准形的定义

定义 1 含有 n 个变量 $x_1,x_2,\cdots,x_n$ 的二次齐次多项式

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+2a_{12}x_1x_2+2a_{13}x_1x_3+\cdots+2a_{1,n}x_1x_n$$
$$+a_{22}x_2^2+2a_{23}x_2x_3+\cdots+2a_{2,n}x_2x_n+\cdots+a_{n-1,n-1}x_{n-1}^2+2a_{n-1,n}x_{n-1}x_n+a_{nn}x_n^2$$

称为二次型. 如果所有系数 $a_{ij}\ (1\le i,j\le n)$ 均为实数，则称二次型为实二次型. 特别地，

如果 n 元二次型 $f(x_1,x_2,\cdots,x_n)$ 只含有平方项，即

$$f(x_1,x_2,\cdots,x_n)=k_1x_1^2+k_2x_2^2+\cdots+k_nx_n^2,$$

称这样的二次型为二次型的标准形. 如果标准形的系数 $k_1,k_2,\cdots,k_n$ 只在 $1,-1,0$ 三个数中取值，也就是

$$f(x_1,x_2,\cdots,x_n)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2,$$

就称其为二次型的规范形.

一、二次型及其标准形的定义

对 $j>i$ 取 $a_{ji}=a_{ij}$，则 $2a_{ij}x_ix_j=a_{ij}x_ix_j+a_{ji}x_jx_i$，于是二次型可写成

$$f(x_1,x_2,\cdots,x_n)=a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1,n}x_1x_n$$
$$+a_{21}x_2x_1+a_{22}x_2^2+a_{23}x_2x_3+\cdots+a_{2,n}x_2x_n+\cdots$$
$$+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2$$
$$=\sum_{i,j=1}^{n}a_{ij}x_ix_j.$$

利用矩阵，二次型可以表示为

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

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

一、二次型及其标准形的定义

记

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

则二次型可记作

$$f(x)=x^{\mathrm{T}}Ax,$$

其中 A 为对称阵.

例如，二次型 $f=x_1^2-3x_3^2-4x_1x_2+x_2x_3$ 用矩阵记号写出来，就是

$$f(x_1,x_2,x_3)=(x_1,x_2,x_3)\begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

一、二次型及其标准形的定义

任给一个二次型，就唯一地确定一个对称阵；反之，任给一个对称阵，也可唯一的确定一个二次型. 这样，二次型与对称阵之间存在一一对应的关系.因此，我们把对称阵 $\boldsymbol{A}$ 叫做二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的矩阵，也把 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 叫做对称阵 $\boldsymbol{A}$ 的二次型. 对称阵 $\boldsymbol{A}$ 的秩就叫做二次型 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的秩.

标准形的矩阵是对角阵.

定义 2 设 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是 n 阶矩阵，若有可逆矩阵 $\boldsymbol{C}$，使 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$，则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同.

矩阵间的合同关系是一个等价关系，满足

（1）反身性：每一个方阵都与它自身合同. 这是因为 $\boldsymbol{A}=\boldsymbol{E}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{E}$.

（2）对称性：如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 也合同. 这是因为由 $\boldsymbol{B}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{C}$ 及矩阵 $\boldsymbol{C}$ 可逆可得 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{P}$，其中 $\boldsymbol{P}=\boldsymbol{C}^{-1}$.

（3）传递性：如果 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同，$\boldsymbol{B}$ 与 $\boldsymbol{C}$ 合同，则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 也合同. 这是因为由 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 及 $\boldsymbol{C}=\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{Q}$ 可得 $\boldsymbol{C}=(\boldsymbol{P}\boldsymbol{Q})^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Q})$.

二、用正交变换化二次型为标准形

定理 任给二次型 $f=\sum_{i,j=1}^{n}a_{ij}x_ix_j\ (a_{ij}=a_{ji})$，总有正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，使 f 化为标准形

$$f=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2,$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 f 的矩阵 $\boldsymbol{A}=(a_{ij})$ 的特征值.

推论 任给 n 元二次型 $f(x)=x^{\mathrm{T}}Ax\ (A^{\mathrm{T}}=A)$，总有可逆变换 $x=Cz$，使 $f(Cz)$ 为规范形.

证明 按定理，有

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

设二次型 f 的秩为 r，则特征值中恰有 r 个不为0，不妨设

$\lambda_1\neq 0,\lambda_2\neq 0,\cdots,\lambda_r\neq 0,\lambda_{r+1}=\cdots=\lambda_n=0$，令

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

则 K 可逆，变换 $y=Kz$ 把 $f(Py)$ 化为

$$f(PKz)=(Kz)^{\mathrm{T}}\Lambda(Kz)=z^{\mathrm{T}}(K^{\mathrm{T}}\Lambda K)z,$$

而

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

记 $\boldsymbol{C}=\boldsymbol{PK}$，即知可逆变换 $\boldsymbol{x}=\boldsymbol{Cz}$ 把 f 化成规范形

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

例1 求一个正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$，把二次型

$$f=2x_1^2+2x_2^2+2x_3^2+2x_1x_2+2x_1x_3+2x_2x_3$$ 化为标准形.

解 二次型的矩阵为

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

按照第四节例 2 所给的结果，有正交阵

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

二、用正交变换化二次型为标准形

使

$$P^{\mathrm{T}}AP=\Lambda=\begin{pmatrix}4&&\\&1&\\&&1\end{pmatrix}.$$

于是有正交变换

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}\frac{\sqrt{3}}{3}&-\frac{\sqrt{2}}{2}&-\frac{\sqrt{6}}{6}\\\frac{\sqrt{3}}{3}&\frac{\sqrt{2}}{2}&-\frac{\sqrt{6}}{6}\\\frac{\sqrt{3}}{3}&0&\frac{\sqrt{6}}{3}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}$$

把二次型 f 化成标准形

$$f=4y_1^2+y_2^2+y_3^2.$$

如果要把二次型 f 化成规范形，只需令

$$\begin{cases} y_1 = \dfrac{1}{2} z_1, \\ y_2 = z_2, \\ y_3 = z_3, \end{cases}$$

即得 f 的规范形

$$f = z_1^2 + z_2^2 + z_3^2 .$$

例 2 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 由于 f 中含变量 x_1 的平方项，故把含 x_1 的项归并起来，配方可得

$$\begin{aligned} f&=x_1^2+2x_1x_2+2x_1x_3+2x_2^2+5x_3^2+6x_2x_3 \\ &=\left(x_1+x_2+x_3\right)^2-x_2^2-x_3^2-2x_2x_3+2x_2^2+5x_3^2+6x_2x_3 \\ &=\left(x_1+x_2+x_3\right)^2+x_2^2+4x_3^2+4x_2x_3, \end{aligned}$$

上式右端除第一项外已不再含 x_1.继续配方,可得

$$f=\left(x_1+x_2+x_3\right)^2+\left(x_2+2x_3\right)^2.$$

三、用配方法化二次型为标准形

令 $$\begin{cases} y_1 = x_1 + x_2 + x_3, \\ y_2 = \quad\ \ x_2 + 2x_3, \\ y_3 = \qquad\qquad x_3, \end{cases}$$

即 $$\begin{cases} x_1 = y_1 - y_2 + y_3, \\ x_2 = y_2 - 2y_3, \\ x_3 = y_3, \end{cases}$$

就把 f 化成标准形(规范形) $f = y_1^2 + y_2^2$,所用变换矩阵为

$$\boldsymbol{C} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \left(|\boldsymbol{C}| = 1 \neq 0\right).$$

例 3 化二次型

$$f=2x_1x_2+4x_1x_3-6x_2x_3$$

成规范形,并求所用的变换矩阵.

解 在 f 中不含平方项. 由于含有 x_1,x_2 乘积项，故令

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

代入可得 $f=2y_1^2-2y_2^2-2y_1y_3+10y_2y_3.$

再配方,得

$$f=2\left(y_1-\frac{1}{2}y_3\right)^2-2\left(y_2-\frac{5}{2}y_3\right)^2+12y_3^2.$$

令 $\begin{cases} z_1 = y_1 - \dfrac{1}{2}y_3, \\ z_2 = y_2 - \dfrac{5}{2}y_3, \\ z_3 = y_3, \end{cases}$

于是

$$\begin{cases} y_1 = z_1 + \dfrac{1}{2}z_3, \\ y_2 = z_2 + \dfrac{5}{2}z_3, \\ y_3 = z_3, \end{cases}，\quad 即 \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dfrac{1}{2} \\ 0 & 1 & \dfrac{5}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix},$$

于是，二次型化为标准形 $f = 2z_1^2 - 2z_2^2 + 12z_3^2$.

再另

$$\begin{cases} w_1 = \sqrt{2} z_1, \\ w_2 = \sqrt{2} z_2, \\ w_3 = \sqrt{12} z_3, \end{cases}，\text{即} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \\ 0 & \frac{1}{\sqrt{2}} & \\ 0 & 0 & \frac{1}{\sqrt{12}} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix},$$

就把二次型化为了规范形

$$f = w_1^2 - w_2^2 + w_3^2,$$

所用变换矩阵为

$$\boldsymbol{C} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{5}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \\ 0 & \frac{1}{\sqrt{2}} & \\ 0 & 0 & \frac{1}{\sqrt{12}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{12}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & -\frac{2}{\sqrt{12}} \\ 0 & 0 & \frac{1}{\sqrt{12}} \end{pmatrix} \left(|\boldsymbol{C}| = -\frac{1}{\sqrt{12}} \neq 0 \right).$$

目录/Contents

目录/Contents

一、惯性定理

定理1

设有二次型 $f(x)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，它的秩为 r，有两个可逆变换

$$\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y} \quad 及 \quad \boldsymbol{x}=\boldsymbol{P}\boldsymbol{z}$$

使

$$f=k_1y_1^2+k_2y_2^2+\cdots+k_ry_r^2 \quad (k_i\neq 0),$$

及

$$f=\lambda_1z_1^2+\lambda_2z_2^2+\cdots+\lambda_rz_r^2 \quad (k_i\neq 0),$$

则 $k_1,\cdots,k_r$ 中正数的个数与 $\lambda_1,\cdots,\lambda_r$ 中正数的个数相等.

这个定理称为惯性定理，这里不予证明.

二、正定二次型与正定阵

定义　设有二次型 $f(x)=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$，如果对于任何 $\boldsymbol{x}\neq\mathbf{0}$，都有 $f(\boldsymbol{x})>0$(显然 $f(\mathbf{0})=0$)，则称二次型 f 为正定二次型，并称对称阵 $\boldsymbol{A}$ 是正定的；如果对任何 $\boldsymbol{x}\neq\mathbf{0}$ 都有 $f(\boldsymbol{x})<0$，则称二次型 f 为负定二次型，并称对称阵 $\boldsymbol{A}$ 是负定的.

定理 2 n 元二次型 $f(x)=x^{\mathrm{T}}Ax$ 为正定的充分必要条件是它的正惯性指数等于 n，即它的规范形的 n 个系数全为1.

证明 设可逆变换 $\boldsymbol{x}=\boldsymbol{C}\boldsymbol{y}$ 使

$$f(\boldsymbol{x})=f(\boldsymbol{C}\boldsymbol{y})=\sum_{i=1}^{n}k_i y_i^2 .$$

先证充分性. 设 $k_i>0\,(i=1,2,\cdots,n)$. 任给 $\boldsymbol{x}\neq\boldsymbol{0}$,则 $\boldsymbol{y}=\boldsymbol{C}^{-1}\boldsymbol{x}\neq\boldsymbol{0}$,故

$$f(\boldsymbol{x})=\sum_{i=1}^{n}k_i y_i^2>0.$$

再证必要性. 用反证法. 假设有 $k_i\leq 0$,则当 $\boldsymbol{y}=\boldsymbol{e}_s$ (单位坐标向量)时, $f(\boldsymbol{C}\boldsymbol{e}_s)=k_s\leq 0$.显然 $\boldsymbol{C}\boldsymbol{e}_s\neq\boldsymbol{0}$,

这与 f 为正定相矛盾.

这就证明了 $k_i>0\,(i=1,2,\cdots,n)$.

推论 1 对称阵 A 为正定的充分必要条件是： A 与单位矩阵 E 合同.

推论 2 对称阵 A 为正定的充分必要条件是： A 的特征值全为正.

定理 3 对称阵 A 为正定的充分必要条件是： A 的各阶主子式都为正，即

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

对称阵为负定的充分必要条件是:奇数阶主子式为负，偶数阶主子式为正,即

$$(-1)^r\begin{vmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots \\ a_{r1} & \cdots & a_{rr} \end{vmatrix}>0\ (r=1,2,\cdots,n).$$

定理 3 称为赫尔维茨定理.

例1

例 1 判别二次型 $f(x_1,x_2,x_3)=-2x_1^2-6x_2^2-4x_3^2+2x_1x_2+2x_1x_3$ 的正定性.

解 此二次型的矩阵为

$$\boldsymbol{A}=\begin{pmatrix}-2 & 1 & 1\\ 1 & -6 & 0\\ 1 & 0 & -4\end{pmatrix},$$

它的各阶顺序主子式为：

$$a_{11}=-2<0,\quad \begin{vmatrix}-2 & 1\\ 1 & -6\end{vmatrix}=11>0,\quad \begin{vmatrix}-2 & 1 & 1\\ 1 & -6 & 0\\ 1 & 0 & -4\end{vmatrix}=-38<0,$$

所以，该二次型是负定的.

例2 设 $\boldsymbol{A}$ 为正定矩阵，证明 $\boldsymbol{A}^{-1}$ 也是正定矩阵.

证明 因为 $\boldsymbol{A}$ 正定，所以 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，从而

$$\left(\boldsymbol{A}^{-1}\right)^{\mathrm{T}}=\left(\boldsymbol{A}^{\mathrm{T}}\right)^{-1}=\boldsymbol{A}^{-1},$$

即 $\boldsymbol{A}^{-1}$ 为实对称矩阵.

又由于 $\boldsymbol{A}$ 正定，存在可逆阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{E}$. 等式两端求逆，得到

$$\boldsymbol{P}^{-1}\boldsymbol{A}^{-1}\left(\boldsymbol{P}^{\mathrm{T}}\right)^{-1}=\boldsymbol{E}.$$

令 $\left(\boldsymbol{P}^{\mathrm{T}}\right)^{-1}=\boldsymbol{Q}$，则 $\boldsymbol{Q}$ 为可逆矩阵，且满足

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{Q}=\boldsymbol{E},$$

所以 $\boldsymbol{A}^{-1}$ 也是正定矩阵.