AMS

矩阵特征值及二次型

本章目录

- 向量的内积及正交性
- 方阵的特征值与特征向量
- 相似矩阵
- 实对称矩阵的相似对角化
- 二次型及其标准形
- 正定二次型与正定矩阵

向量的内积及正交性

课前导读

二维和三维空间中的向量有长度，夹角等概念，但是空间$\mathbb{R}^n$中的向量却没有这些概念. 本节我们先将二维和三维向量的数量积(又称内积)推广到$\mathbb{R}^n$中来，再借助于$\mathbb{R}^n$中向量的内积来定义$\mathbb{R}^n$中向量的长度，夹角等概念.

定义 (向量的内积)

设有n维向量

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

定义二元函数$\langle \cdot\,, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^T \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

我们称$\langle \boldsymbol{x}, \boldsymbol{y} \rangle$为$n$**维向量**$x$**与**$y$**的内积**.

定义 (向量的内积)

设有n维向量

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},$$

定义二元函数$\langle \cdot\,, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$:

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^T \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n,$$

我们称$\langle \boldsymbol{x}, \boldsymbol{y} \rangle$为$n$**维向量$\boldsymbol{x}$与$\boldsymbol{y}$的内积**.

注

n维向量的内积是二维,三维向量的数量积的一种推广,其结果是一个实数.

向量的内积

n维向量内积的性质

对任意n维列向量$\boldsymbol{x}$, $\boldsymbol{y}$与$\boldsymbol{z}$,任意$\lambda \in \mathbb{R}$:

(1) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$;

(2) $\langle \lambda\boldsymbol{x}, \boldsymbol{y} \rangle = \lambda\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \lambda\boldsymbol{y} \rangle$;

(3) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$;

(4) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqslant 0$, 当且仅当$\boldsymbol{x} = \boldsymbol{0}$时, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$.

n维向量内积的性质

对任意n维列向量$\boldsymbol{x}$, $\boldsymbol{y}$与$\boldsymbol{z}$,任意$\lambda \in \mathbb{R}$:

(1) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$;

(2) $\langle \lambda \boldsymbol{x}, \boldsymbol{y} \rangle = \lambda \langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \lambda \boldsymbol{y} \rangle$;

(3) $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$;

(4) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqslant 0$, 当且仅当$\boldsymbol{x} = \boldsymbol{0}$时, $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$.

Cauchy-Schwarz不等式

对任意n维列向量$\boldsymbol{x}$, $\boldsymbol{y}$,

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle^2 \leqslant \langle \boldsymbol{x}, \boldsymbol{x} \rangle \langle \boldsymbol{y}, \boldsymbol{y} \rangle.$$

定义 (向量的长度)

设有n维向量

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n,$$

定义函数$\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$:

$$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = x_1^2 + x_2^2 + \cdots + x_n^2,$$

我们称$\|\boldsymbol{x}\|$为n**维向量**$\boldsymbol{x}$**的长度(或范数)**.

向量的长度

n维向量长度的性质

(1) 正定性: 当$\boldsymbol{x} \neq \mathbf{0}$时, $\|\boldsymbol{x}\| > 0$; 当$\boldsymbol{x} = \mathbf{0}$时, $\|\boldsymbol{x}\| = 0$;

(2) 绝对齐次性: $\|\lambda \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|, \forall\, \lambda \in \mathbb{R}$;

(3) 三角不等式: $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

n维向量长度的性质

(1) 正定性: 当$\boldsymbol{x} \neq \mathbf{0}$时, $\|\boldsymbol{x}\| > 0$; 当$\boldsymbol{x} = \mathbf{0}$时, $\|\boldsymbol{x}\| = 0$;

(2) 绝对齐次性: $\|\lambda \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$, $\forall\, \lambda \in \mathbb{R}$;

(3) 三角不等式: $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

注

- 当$\|\boldsymbol{x}\| = 1$时,我们称$\boldsymbol{x}$为单位向量.
- 如果$\boldsymbol{\alpha} \neq \mathbf{0}$, 那么$\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$是一个单位向量.
- 我们称由向量$\boldsymbol{\alpha}$得到单位向量$\dfrac{\boldsymbol{\alpha}}{\|\boldsymbol{\alpha}\|}$的过程称为**向量$\boldsymbol{\alpha}$的单位化**.

向量的夹角

定义 (n维向量之间的夹角)

当n维向量$\boldsymbol{x} \neq \mathbf{0}, \boldsymbol{y} \neq \mathbf{0}$时，我们称

$$\theta := \arccos \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}$$

为n维向量$\boldsymbol{x}$与$\boldsymbol{y}$的夹角.

向量的夹角

定义 (n维向量之间的夹角)

当n维向量$\boldsymbol{x} \neq \mathbf{0}, \boldsymbol{y} \neq \mathbf{0}$时，我们称

$$\theta := \arccos \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\|}$$

为n**维向量**$\boldsymbol{x}$**与**$\boldsymbol{y}$**的夹角**.

注

- 由Cauchy-Schwarz不等式可知向量夹角定义合理.
- 当$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$时，我们称$n$**维向量**$\boldsymbol{x}$**与**$\boldsymbol{y}$**正交**.
- $\boldsymbol{x} = \mathbf{0}$时，我们规定零向量$\boldsymbol{x}$与任何$n$维向量都正交.

定义 (正交向量组)

我们称由一组两两正交的非零n维向量组成的向量组为一个**正交向量组**.

定义 (正交向量组)

我们称由一组两两正交的非零n维向量组成的向量组为一个**正交向量组**.

例

- $\mathbb{R}^3$中一组正交向量组 $\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\2\\0\end{bmatrix},\begin{bmatrix}0\\0\\3\end{bmatrix}.$

定义 (正交向量组)

我们称由一组两两正交的非零n维向量组成的向量组为一个**正交向量组**.

例

- $\mathbb{R}^3$中一组正交向量组$\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\2\\0\end{bmatrix},\begin{bmatrix}0\\0\\3\end{bmatrix}$.
- $\mathbb{R}^4$中一组正交向量组$\begin{bmatrix}1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\\frac{\sqrt{2}}{2}\\-\frac{\sqrt{2}}{2}\end{bmatrix},\begin{bmatrix}0\\0\\\frac{\sqrt{2}}{2}\\\frac{\sqrt{2}}{2}\end{bmatrix}$.

正交向量组

定理

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.

定理

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.

例 1

已知3维空间$\mathbb{R}^3$中的两个向量

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \boldsymbol{\alpha}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix},$$

正交, 试求一个非零向量$\boldsymbol{\alpha}_3$, 使得$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$两两正交.

定义 (向量空间中的标准正交基)

设n维向量组$\xi_1, \xi_2, \cdots, \xi_r$是向量空间$V(V \subseteq \mathbb{R}^n)$的一个基，如果$\xi_1, \xi_2, \cdots, \xi_r$两两正交且都是单位向量，那么我们称$\xi_1, \xi_2, \cdots, \xi_r$是$V$的一个**标准正交基**.

定义 (向量空间中的标准正交基)

设n维向量组$\xi_1,\xi_2,\cdots,\xi_r$是向量空间$V(V\subseteq\mathbb{R}^n)$的一个基，如果$\xi_1,\xi_2,\cdots,\xi_r$两两正交且都是单位向量，那么我们称$\xi_1,\xi_2,\cdots,\xi_r$是$V$的一个**标准正交基**.

例

- n维单位向量$e_1,e_2,\cdots,e^n$是$\mathbb{R}^n$的一个标准正交基.
- 向量组

$$\xi_1=\begin{bmatrix}\frac{2}{3}\\ \frac{1}{3}\\ \frac{2}{3}\end{bmatrix},\xi_2=\begin{bmatrix}-\frac{2}{3}\\ \frac{2}{3}\\ \frac{1}{3}\end{bmatrix},\xi_3=\begin{bmatrix}\frac{1}{3}\\ \frac{2}{3}\\ -\frac{2}{3}\end{bmatrix}$$

是$\mathbb{R}^3$的一个标准正交基.

向量空间中的标准正交基

设$\xi_1, \xi_2, \cdots, \xi_r$是$V$的一个标准正交基. 对$V$中任一向量$\beta$, 我们有线性表示式

$$\boldsymbol{\beta} = \lambda_1 \boldsymbol{\xi}_1 + \lambda_2 \boldsymbol{\xi}_2 + \cdots + \lambda_r \boldsymbol{\xi}_r,$$

用$\boldsymbol{\xi}_i^T (i = 1, 2, \cdots, r)$左乘上式,我们有

$$\boldsymbol{\xi}_i^T \boldsymbol{\beta} = \lambda_i \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i = \lambda_i (i = 1, 2, \cdots, r),$$

即

$$\lambda_i = \boldsymbol{\xi}_i^T \boldsymbol{\beta} = \langle \boldsymbol{\xi}_i, \boldsymbol{\beta} \rangle \ (i = 1, 2, \cdots, r),$$

向量空间中的标准正交基

设$\xi_1, \xi_2, \cdots, \xi_r$是$V$的一个标准正交基. 对$V$中任一向量$\beta$, 我们有线性表示式

$$\boldsymbol{\beta} = \lambda_1 \boldsymbol{\xi}_1 + \lambda_2 \boldsymbol{\xi}_2 + \cdots + \lambda_r \boldsymbol{\xi}_r,$$

用$\boldsymbol{\xi}_i^T (i = 1, 2, \cdots, r)$左乘上式,我们有

$$\boldsymbol{\xi}_i^T \boldsymbol{\beta} = \lambda_i \boldsymbol{\xi}_i^T \boldsymbol{\xi}_i = \lambda_i (i = 1, 2, \cdots, r),$$

即

$$\lambda_i = \boldsymbol{\xi}_i^T \boldsymbol{\beta} = \langle \boldsymbol{\xi}_i, \boldsymbol{\beta} \rangle \, (i = 1, 2, \cdots, r),$$

向量β在标准正交基$\xi_1, \xi_2, \cdots, \xi_r$中的坐标$\lambda_i$为内积$\langle \boldsymbol{\xi}_i, \boldsymbol{\beta} \rangle (i = 1, 2, \cdots, r)$, 因此我们从给定的向量空间中取一组基时常常取标准正交基.

向量空间中的标准正交基

设$\xi_1,\xi_2,\cdots,\xi_r$是V的一个标准正交基. 对V中任一向量β, 我们有线性表示式

$$\boldsymbol{\beta}=\lambda_1\boldsymbol{\xi}_1+\lambda_2\boldsymbol{\xi}_2+\cdots+\lambda_r\boldsymbol{\xi}_r,$$

用$\boldsymbol{\xi}_i^T(i=1,2,\cdots,r)$左乘上式,我们有

$$\boldsymbol{\xi}_i^T\boldsymbol{\beta}=\lambda_i\boldsymbol{\xi}_i^T\boldsymbol{\xi}_i=\lambda_i(i=1,2,\cdots,r),$$

即

$$\lambda_i=\boldsymbol{\xi}_i^T\boldsymbol{\beta}=\langle\boldsymbol{\xi}_i,\boldsymbol{\beta}\rangle\;(i=1,2,\cdots,r),$$

向量β在标准正交基$\xi_1,\xi_2,\cdots,\xi_r$中的坐标λ_i为内积$\langle\boldsymbol{\xi}_i,\boldsymbol{\beta}\rangle(i=1,2,\cdots,r)$, 因此我们从给定的向量空间中取一组基时常常取标准正交基.

问题

从向量空间V的一个基出发,我们如何找到V的一组标准正交基?

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, 我们需要找到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.

施密特正交化过程

设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$是向量空间$V$的一个基，从基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$出发，我们需要找到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$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$等价.

(1) 将基$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$正交化. 即取

$$
\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_1, \\
\boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_2 - \frac{\langle \boldsymbol{\beta}_1, \boldsymbol{\alpha}_2 \rangle}{\langle \boldsymbol{\beta}_1, \boldsymbol{\beta}_1 \rangle}\boldsymbol{\beta}_1, \\
\boldsymbol{\beta}_3 &= \boldsymbol{\alpha}_3 - \frac{\langle \boldsymbol{\beta}_1, \boldsymbol{\alpha}_3 \rangle}{\langle \boldsymbol{\beta}_1, \boldsymbol{\beta}_1 \rangle}\boldsymbol{\beta}_1 - \frac{\langle \boldsymbol{\beta}_2, \boldsymbol{\alpha}_3 \rangle}{\langle \boldsymbol{\beta}_2, \boldsymbol{\beta}_2 \rangle}\boldsymbol{\beta}_2, \\
&\cdots\cdots\cdots\cdots \\
\boldsymbol{\beta}_r &= \boldsymbol{\alpha}_r - \frac{\langle \boldsymbol{\beta}_1, \boldsymbol{\alpha}_r \rangle}{\langle \boldsymbol{\beta}_1, \boldsymbol{\beta}_1 \rangle}\boldsymbol{\beta}_1 - \frac{\langle \boldsymbol{\beta}_2, \boldsymbol{\alpha}_r \rangle}{\langle \boldsymbol{\beta}_2, \boldsymbol{\beta}_2 \rangle}\boldsymbol{\beta}_2 - \cdots - \frac{\langle \boldsymbol{\beta}_{r-1}, \boldsymbol{\alpha}_r \rangle}{\langle \boldsymbol{\beta}_{r-1}, \boldsymbol{\beta}_{r-1} \rangle}\boldsymbol{\beta}_{r-1}.
\end{aligned}
$$

我们可以验证$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$两两正交，且与$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$等价.

施密特正交化过程

定义 (施密特正交化过程)

我们称上述从线性无关向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$导出正交向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$的过程为**施密特(Schmidt)**正交化过程.

定义 (施密特正交化过程)

我们称上述从线性无关向量组$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_r$导出正交向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$的过程为**施密特(Schmidt)正交化过程**.

(2) 我们将$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_r$单位化, 得到

$$\boldsymbol{\xi}_1 = \frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}, \boldsymbol{\xi}_2 = \frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}, \cdots, \boldsymbol{\xi}_r = \frac{\boldsymbol{\beta}_r}{\|\boldsymbol{\beta}_r\|}.$$

因此向量组$\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_r$就是$V$的一个标准正交基.

例 2

设

$$\boldsymbol{\alpha}_1=\begin{bmatrix}1\\1\\-1\end{bmatrix},\boldsymbol{\alpha}_2=\begin{bmatrix}0\\4\\1\end{bmatrix},\boldsymbol{\alpha}_3=\begin{bmatrix}-2\\1\\1\end{bmatrix}$$

是$\mathbb{R}^3$的一个基，求一个与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$等价的标准正交基.

例 2

设

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{\alpha}_2 = \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}, \boldsymbol{\alpha}_3 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

是$\mathbb{R}^3$的一个基，求一个与$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$等价的标准正交基.

例 3

已知 $\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$，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.

定义 (正交矩阵)

如果n阶矩阵$\mathbf{A}$满足

$$\mathbf{A}^T\mathbf{A} = \mathbf{E}_n(\text{即}\mathbf{A}^{-1} = \mathbf{A}^T),$$

那么我们称$\mathbf{A}$为**正交矩阵**, 简称**正交阵**.

定义 (正交矩阵)

如果n阶矩阵$\mathbf{A}$满足

$$\mathbf{A}^T\mathbf{A} = \mathbf{E}_n (\text{即}\mathbf{A}^{-1} = \mathbf{A}^T),$$

那么我们称$\mathbf{A}$为**正交矩阵**, 简称**正交阵**.

定理 (判定方阵为正交矩阵)

设矩阵$\mathbf{A}$是n阶方阵, 则下列结论等价:

(1) $\mathbf{A}$是n阶正交矩阵;

(2) $\mathbf{A}$的列向量组是$\mathbb{R}^n$的一个标准正交基;

(3) $\mathbf{A}$的行向量组是$\mathbb{R}^n$的一个标准正交基.

例 4

验证矩阵

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

是正交阵.

正交矩阵的性质

(1) 若$\mathbf{A}$为正交矩阵，则$\mathbf{A}^{-1}=\mathbf{A}^{T}$也是正交矩阵，且$|\mathbf{A}|=1$或$-1$.

(2) 若$\mathbf{A}$和$\mathbf{B}$都是正交矩阵，则$\mathbf{AB}$也是正交矩阵.

正交矩阵

正交矩阵的性质

(1) 若$\mathbf{A}$为正交矩阵, 则$\mathbf{A}^{-1} = \mathbf{A}^T$也是正交矩阵, 且$|\mathbf{A}| = 1$或$-1$.

(2) 若$\mathbf{A}$和$\mathbf{B}$都是正交矩阵, 则$\mathbf{AB}$也是正交矩阵.

定义 (正交变换)

如果$\mathbf{P}$为正交矩阵, 那么我们称变换$\boldsymbol{y} = \mathbf{P}\boldsymbol{x}$为**正交变换**.

正交矩阵的性质

(1) 若$\mathbf{A}$为正交矩阵，则$\mathbf{A}^{-1}=\mathbf{A}^T$也是正交矩阵，且$|\mathbf{A}|=1$或$-1$.

(2) 若$\mathbf{A}$和$\mathbf{B}$都是正交矩阵，则$\mathbf{AB}$也是正交矩阵.

定义 (正交变换)

如果$\mathbf{P}$为正交矩阵，那么我们称变换$\boldsymbol{y}=\mathbf{P}\boldsymbol{x}$为**正交变换**.

设$\boldsymbol{y}=\mathbf{P}\boldsymbol{x}$为正交变换，则

$$\|\boldsymbol{y}\|=\sqrt{\boldsymbol{y}^T\boldsymbol{y}}=\sqrt{\boldsymbol{x}^T\mathbf{P}^T\mathbf{P}\boldsymbol{x}}=\sqrt{\boldsymbol{x}^T\boldsymbol{x}}=\|\boldsymbol{x}\|.$$

即正交变换保持向量的长度不变.

方阵的特征值与特征向量

课前导读

工程技术中的一些问题，常可归结为求一个方阵的特征值和特征向量:

- 振动问题和稳定性问题.
- 方阵的对角化及解微分方程组等问题.

工程技术中的一些问题，常可归结为求一个方阵的特征值和特征向量:

- 振动问题和稳定性问题.
- 方阵的对角化及解微分方程组等问题.

本节我们就来介绍方阵的特征值理论.

定义 (特征值与特征向量)

设$\mathbf{A}$是n阶矩阵, 如果存在数λ和n维非零向量$\boldsymbol{\alpha}$,使得

$$\mathbf{A}\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha},$$

那么我们称

- λ为矩阵$\mathbf{A}$的**特征值**,
- 非零向量α为$\mathbf{A}$对应于特征值λ的**特征向量**.

例

设矩阵$\mathbf{A}=\begin{bmatrix}-1 & 2 & 0\\ 0 & 3 & 0\\ 2 & 1 & -1\end{bmatrix}$, 向量$\boldsymbol{\alpha}=\begin{bmatrix}1\\2\\1\end{bmatrix}$,

例

设矩阵$\mathbf{A}=\begin{bmatrix}-1 & 2 & 0\\ 0 & 3 & 0\\ 2 & 1 & -1\end{bmatrix}$, 向量$\boldsymbol{\alpha}=\begin{bmatrix}1\\2\\1\end{bmatrix}$,由于

$$\mathbf{A}\boldsymbol{\alpha}=\begin{bmatrix}-1 & 2 & 0\\ 0 & 3 & 0\\ 2 & 1 & -1\end{bmatrix}\begin{bmatrix}1\\2\\1\end{bmatrix}=\begin{bmatrix}3\\6\\3\end{bmatrix}=3\begin{bmatrix}1\\2\\1\end{bmatrix},$$

所以数3是矩阵$\mathbf{A}$的一个特征值, α是$\mathbf{A}$的对应于特征值3的特征向量.

问题

一个任意给定的n阶矩阵$\mathbf{A}$有多少个特征值？我们如何计算矩阵$\mathbf{A}$的特征向量？

方阵的特征多项式与特征方程

问题

一个任意给定的n阶矩阵$\mathbf{A}$有多少个特征值？我们如何计算矩阵$\mathbf{A}$的特征向量？

我们先假设矩阵$\mathbf{A}$有特征值λ, 对应于特征值λ的特征向量$\boldsymbol{\alpha}$, 那么λ与$\boldsymbol{\alpha}$满足$\mathbf{A}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha}$, 将上式改写成

$$(\mathbf{A}-\lambda\mathbf{E})\boldsymbol{\alpha}=0,$$

所以$\boldsymbol{\alpha}$是n个未知数n个方程的齐次方程组$(\mathbf{A}\boldsymbol{\alpha}-\lambda\mathbf{E})\boldsymbol{x}=0$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$|\ \mathbf{A}-\lambda\mathbf{E}\ |=0$. 我们记

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

定义 (方阵的特征多项式与特征方程)

- 我们称λ的n次多项式

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

  为矩阵$\mathbf{A}$的**特征多项式**.

- $|\mathbf{A}-\lambda\mathbf{E}|=0$是以$\lambda$为未知数的一元$n$次方程, 我们称其为矩阵$\mathbf{A}$的**特征方程**,而$\mathbf{A}$的特征值就是特征方程$f(\lambda)=0$的根.

定义 (方阵的特征多项式与特征方程)

- 我们称λ的n次多项式

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

为矩阵$\mathbf{A}$的**特征多项式**.

- $|\mathbf{A}-\lambda\mathbf{E}|=0$是以$\lambda$为未知数的一元$n$次方程, 我们称其为矩阵$\mathbf{A}$的**特征方程**,而$\mathbf{A}$的特征值就是特征方程$f(\lambda)=0$的根.

定理 (代数基本定理(2.0))

一元n次方程在复数域$\mathbb{C}$范围内恒有n个根*(重根按重数计算)*.

注

- n阶矩阵$\mathbf{A}$在复数范围内有n个特征值, 通过求解矩阵$\mathbf{A}$的特征方程我们就可以得到这n个特征值.
- 设$\lambda=\lambda_i$为矩阵$\mathbf{A}$的一个特征值,由方程$(\mathbf{A}-\lambda_i\mathbf{E})\boldsymbol{x}=0$可求得非零解$\boldsymbol{x}=\boldsymbol{\alpha}_i$,因此$\boldsymbol{\alpha}_i$是矩阵$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.

注

- n阶矩阵$\mathbf{A}$在复数范围内有n个特征值，通过求解矩阵$\mathbf{A}$的特征方程我们就可以得到这n个特征值.
- 设$\lambda = \lambda_i$为矩阵$\mathbf{A}$的一个特征值,由方程$(\mathbf{A} - \lambda_i \mathbf{E})\boldsymbol{x} = 0$可求得非零解$\boldsymbol{x} = \boldsymbol{\alpha}_i$,因此$\boldsymbol{\alpha}_i$是矩阵$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.

例 1

求矩阵

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

的特征值和特征向量.

例 2

求矩阵

$$\mathbf{B} = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 3 & 0 \\ 2 & 1 & -1 \end{bmatrix}$$

的特征值和特征向量.

例 2

求矩阵

$$\mathbf{B}=\begin{bmatrix}-1 & 2 & 0\\ 0 & 3 & 0\\ 2 & 1 & -1\end{bmatrix}$$

的特征值和特征向量.

例 3

求矩阵

$$\mathbf{C}=\begin{bmatrix}1 & 0 & 2\\ 0 & 3 & 0\\ 2 & 0 & 1\end{bmatrix}$$

的特征值和特征向量.

性质1

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

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

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

方阵的特征值与特征向量的性质

性质1

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

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

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

性质2

如果λ是方阵$\mathbf{A}$的特征值且$\boldsymbol{\alpha}$为对应于特征值λ的特征向量,那么

(1) λ^k是方阵$\mathbf{A}^k$的特征值(k为非负整数), 对应于特征值λ^k的特征向量是$\boldsymbol{\alpha}$;

(2) $k\lambda$是方阵$k\mathbf{A}$的特征值(k为任意常数), 对应于特征值$k\lambda$的特征向量是$\boldsymbol{\alpha}$;

(3) 当$\mathbf{A}$可逆时, λ^{-1}是方阵$\mathbf{A}^{-1}$的特征值, 对应于特征值λ^{-1}的特征向量是$\boldsymbol{\alpha}$.

性质3

如果矩阵$\mathbf{A}$的多项式是

$$\varphi(\mathbf{A}) = a_m\mathbf{A}^m + \cdots + a_1\mathbf{A} + a_0\mathbf{E},$$

那么方阵$\varphi(\mathbf{A})$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$\boldsymbol{\alpha}$.

性质3

如果矩阵$\mathbf{A}$的多项式是

$$\varphi(\mathbf{A}) = a_m\mathbf{A}^m + \cdots + a_1\mathbf{A} + a_0\mathbf{E},$$

那么方阵$\varphi(\mathbf{A})$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$\boldsymbol{\alpha}$.

性质4

如果$\boldsymbol{\alpha}_1$与$\boldsymbol{\alpha}_2$是方阵$\mathbf{A}$的同一特征值λ所对应的特征向量，那么$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2$($k_1, k_2$不同时为零)也是特征值$\lambda$所对应的特征向量.

性质5

设$\lambda_1, \lambda_2, \cdots, \lambda_m$是方阵$\mathbf{A}$的$m$个互不相同的特征值,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是依次与之对应的特征向量,那么$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$线性无关.

性质5

设$\lambda_1, \lambda_2, \cdots, \lambda_m$是方阵$\mathbf{A}$的$m$个互不相同的特征值,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是依次与之对应的特征向量,那么$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$线性无关.

性质6

设λ_1, λ_2是矩阵$\mathbf{A}$的两个不同的特征值,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$和$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$是分别对应于$\lambda_1$和$\lambda_2$的线性无关的特征向量, 则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$线性无关.

性质5

设$\lambda_1, \lambda_2, \cdots, \lambda_m$是方阵A的$m$个互不相同的特征值,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是依次与之对应的特征向量,那么$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$线性无关.

性质6

设λ_1, λ_2是矩阵A的两个不同的特征值,$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s$和$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$是分别对应于$\lambda_1$和$\lambda_2$的线性无关的特征向量,则$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_s, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_t$线性无关.

例 4

设3阶矩阵的特征值为$1, 2, 3$, 求$2\boldsymbol{A}^* - 3\boldsymbol{A} + 2\boldsymbol{E}_3$的特征值.

性质5

设$\lambda_1,\lambda_2,\cdots,\lambda_m$是方阵A的$m$个互不相同的特征值,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是依次与之对应的特征向量,那么$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$线性无关.

性质6

设λ_1,λ_2是矩阵A的两个不同的特征值,$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$和$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$是分别对应于$\lambda_1$和$\lambda_2$的线性无关的特征向量, 则$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_t$线性无关.

例 4

设3阶矩阵的特征值为$1,2,3$, 求$2\boldsymbol{A}^*-3\boldsymbol{A}+2\boldsymbol{E}_3$的特征值.

例 5

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

相似矩阵

课前导读

本节我们介绍矩阵相似的概念和性质，并给出矩阵与对角矩阵相似的充要条件.

定义 (相似矩阵)

设n阶方阵$\mathbf{A}, \mathbf{B}$.

- 如果存在可逆矩阵$\mathbf{P}$,使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 那么我们称$\mathbf{B}$是$\mathbf{A}$的**相似矩阵**,或者说**矩阵$\mathbf{A}$与$\mathbf{B}$相似**.
- 我们称对矩阵$\mathbf{A}$的运算$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$为对$\mathbf{A}$进行**相似变换**.
- 我们称可逆矩阵$\mathbf{P}$为把$\mathbf{A}$变成$\mathbf{B}$的**相似变换矩阵**.

方阵相似的定义

定义 (相似矩阵)

设n阶方阵$\mathbf{A}, \mathbf{B}$.

- 如果存在可逆矩阵$\mathbf{P}$,使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{B}$, 那么我们称$\mathbf{B}$是$\mathbf{A}$的**相似矩阵**,或者说**矩阵$\mathbf{A}$与$\mathbf{B}$相似**.
- 我们称对矩阵$\mathbf{A}$的运算$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$为对$\mathbf{A}$进行**相似变换**.
- 我们称可逆矩阵$\mathbf{P}$为把$\mathbf{A}$变成$\mathbf{B}$的**相似变换矩阵**.

定理

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

推论

若n阶矩阵$\mathbf{A}$与对角阵

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

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

推论

若n阶矩阵$\mathbf{A}$与对角阵

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

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

定理

设$f(\lambda)$是矩阵$\mathbf{A}$的特征多项式, 则$f(\mathbf{A}) = \mathbf{0}$.

定义 (方阵的对角化)

我们称对n阶方阵$\mathbf{A}$去寻求相似变换矩阵$\mathbf{P}$,使得$\Lambda = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$为对角阵的过程为矩阵$\mathbf{A}$的**相似对角化**.

方阵的相似对角化

定义 (方阵的对角化)

我们称对n阶方阵$\mathbf{A}$去寻求相似变换矩阵$\mathbf{P}$,使得$\Lambda = \mathbf{P}^{-1}\mathbf{AP}$为对角阵的过程为矩阵$\mathbf{A}$的**相似对角化**.

问题

如果n阶矩阵$\mathbf{A}$可相似对角化, 我们如何寻找相似变换矩阵$\mathbf{P}$?

- 假设n阶矩阵$\mathbf{A}$可以相似对角化, 即存在可逆矩阵$\mathbf{P}$, 使得$\mathbf{P}^{-1}\mathbf{AP} = \Lambda$为对角阵.

定义 (方阵的对角化)

我们称对n阶方阵$\mathbf{A}$去寻求相似变换矩阵$\mathbf{P}$,使得$\Lambda = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$为对角阵的过程为矩阵$\mathbf{A}$的**相似对角化**.

问题

如果n阶矩阵$\mathbf{A}$可相似对角化, 我们如何寻找相似变换矩阵$\mathbf{P}$?

- 假设n阶矩阵$\mathbf{A}$可以相似对角化, 即存在可逆矩阵$\mathbf{P}$, 使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda$为对角阵.我们把矩阵$\mathbf{P}$列分块为$\mathbf{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$, 由$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \Lambda$, 得$\mathbf{A}\mathbf{P} = \mathbf{P}\Lambda$, 即

$$\mathbf{A}(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n) = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = (\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \cdots, \lambda_n\boldsymbol{p}_n).$$

方阵的相似对角化

- 于是$\mathbf{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$, 即$\lambda_i$为$\mathbf{A}$的特征值, 而$\mathbf{P}$的列向量$\boldsymbol{p}_i$就是$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.

- 于是$\mathbf{A}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i(i=1,2,\cdots,n)$，即$\lambda_i$为$\mathbf{A}$的特征值，而$\mathbf{P}$的列向量$\boldsymbol{p}_i$就是$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.
- 反之，如果n阶矩阵$\mathbf{A}$恰好有n个特征向量，那么这n个特征向量即可构成矩阵$\mathbf{P}$，使得

$$\mathbf{AP}=\mathbf{P}\Lambda.$$

这n个特征向量必定线性无关,从而$\mathbf{P}$可逆,因此有$\mathbf{P}^{-1}\mathbf{AP}=\Lambda$.

- 于是$\mathbf{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$，即$\lambda_i$为$\mathbf{A}$的特征值，而$\mathbf{P}$的列向量$\boldsymbol{p}_i$就是$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.
- 反之，如果n阶矩阵$\mathbf{A}$恰好有n个特征向量，那么这n个特征向量即可构成矩阵$\mathbf{P}$，使得

$$\mathbf{AP} = \mathbf{P}\Lambda.$$

这n个特征向量必定线性无关,从而$\mathbf{P}$可逆,因此有$\mathbf{P}^{-1}\mathbf{AP} = \Lambda$.

定理 (方阵对角化的充要条件)

n阶矩阵$\mathbf{A}$与对角阵相似($\mathbf{A}$能对角化) $\Leftrightarrow$ $\mathbf{A}$有n个线性无关的特征向量.

- 于是$\mathbf{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$，即$\lambda_i$为$\mathbf{A}$的特征值，而$\mathbf{P}$的列向量$\boldsymbol{p}_i$就是$\mathbf{A}$对应于特征值$\lambda_i$的特征向量.
- 反之，如果n阶矩阵$\mathbf{A}$恰好有n个特征向量，那么这n个特征向量即可构成矩阵$\mathbf{P}$，使得

$$\mathbf{AP} = \mathbf{P}\Lambda.$$

这n个特征向量必定线性无关,从而$\mathbf{P}$可逆,因此有$\mathbf{P}^{-1}\mathbf{AP} = \Lambda$.

定理 (方阵对角化的充要条件)

n阶矩阵$\mathbf{A}$与对角阵相似($\mathbf{A}$能对角化) $\Leftrightarrow$ $\mathbf{A}$有n个线性无关的特征向量.

推论

如果n阶矩阵$\mathbf{A}$的n个特征值互不相等，则$\mathbf{A}$与对角阵相似.

注

矩阵的特征方程有重根时不一定有n个线性无关的特征向量,从而不一定能对角化.

注

矩阵的特征方程有重根时不一定有n个线性无关的特征向量,从而不一定能对角化.

例 1

设

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

有3个线性无关的特征向量, 求x与y应满足的条件.

实对称矩阵的相似对角化

课前导读

根据方阵可对角化的理论内容可知, 要判断一个n阶矩阵$\mathbf{A}$是否可对角化, 关键在于判断这个矩阵是否有n个线性无关的特征向量.

课前导读

根据方阵可对角化的理论内容可知，要判断一个n阶矩阵$\mathbf{A}$是否可对角化，关键在于判断这个矩阵是否有n个线性无关的特征向量.

但这不是一件容易的事情(我们对此不进行一般性的讨论)，以下我们仅讨论当方阵$\mathbf{A}$是实对称矩阵的情形(关于实对称矩阵的对角化问题有确定的结果)：实对称矩阵总是可对角化.

根据方阵可对角化的理论内容可知，要判断一个n阶矩阵$\mathbf{A}$是否可对角化，关键在于判断这个矩阵是否有n个线性无关的特征向量.

但这不是一件容易的事情(我们对此不进行一般性的讨论)，以下我们仅讨论当方阵$\mathbf{A}$是实对称矩阵的情形(关于实对称矩阵的对角化问题有确定的结果): 实对称矩阵总是可对角化.

本节我们来具体讨论实对称矩阵的对角化.

实对称矩阵的定义

定义 (实对称矩阵)

如果一个n阶实矩阵$\mathbf{A}$满足$\mathbf{A} = \mathbf{A}^T$,那么我们称$\mathbf{A}$为一个**实对称矩阵**.

定义 (哈密顿矩阵)

如果一个n阶复矩阵$\mathbf{A}$满足$\mathbf{A} = \mathbf{A}^T$,那么我们称$\mathbf{A}$为一个**哈密顿矩阵**.

定义 (复矩阵及其共轭矩阵)

设复矩阵$\mathbf{X} = (x_{ij}), x_{ij} \in \mathbb{C}$. 我们称$\overline{\mathbf{X}} = (\overline{x}_{ij})$,$\overline{x}_{ij} \in \mathbb{C}$为矩阵$\mathbf{X} = (x_{ij})$的**共轭矩阵**.

实对称矩阵的特征值和特征向量的性质

性质1

实对称矩阵的特征值为实数.

- 设对称矩阵$\mathbf{A}$的特征值为$\lambda \in \mathbb{C}$, 对应λ的特征向量为$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$, 即$\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$.

实对称矩阵的特征值和特征向量的性质

性质1

实对称矩阵的特征值为实数.

- 设对称矩阵$\mathbf{A}$的特征值为$\lambda \in \mathbb{C}$, 对应λ的特征向量为$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$, 即$\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$.
- $\mathbf{A}$为实对称矩阵,有$\mathbf{A} = \overline{\mathbf{A}}$及$\mathbf{A}^T = \mathbf{A}$.

性质1

实对称矩阵的特征值为实数.

- 设对称矩阵$\mathbf{A}$的特征值为$\lambda \in \mathbb{C}$, 对应λ的特征向量为$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$, 即$\mathbf{A}\boldsymbol{x} = \lambda\boldsymbol{x}$.
- $\mathbf{A}$为实对称矩阵,有$\mathbf{A} = \overline{\mathbf{A}}$及$\mathbf{A}^T = \mathbf{A}$.
- 注意到

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

且

$$\overline{\boldsymbol{x}}^T\mathbf{A}\boldsymbol{x} = (\overline{\boldsymbol{x}}^T\mathbf{A})\boldsymbol{x} = (\mathbf{A}\overline{\boldsymbol{x}})^T\boldsymbol{x} = (\overline{\mathbf{A}}\overline{\boldsymbol{x}})^T\boldsymbol{x} = (\overline{\mathbf{A}\boldsymbol{x}})^T\boldsymbol{x} = (\overline{\lambda\boldsymbol{x}})^T\boldsymbol{x} = \overline{\lambda}\overline{\boldsymbol{x}}^T\boldsymbol{x}.$$

两式相减,我们得到$(\lambda - \overline{\lambda})\overline{\boldsymbol{x}}^T\boldsymbol{x} = 0$,$\boldsymbol{x} \neq \mathbf{0}$.

实对称矩阵的特征值和特征向量的性质

- $\boldsymbol{x} \neq \mathbf{0} \Rightarrow \overline{\boldsymbol{x}}^T \boldsymbol{x} = \sum_{i=1}^{n} \overline{\boldsymbol{x}}_i \boldsymbol{x}_i = \sum_{i=1}^{n} |x_i|^2 \neq 0$,所以$\lambda - \overline{\lambda} = 0$,即$\lambda = \overline{\lambda}$,这说明$\lambda \in \mathbb{R}$.
- 特征值$\lambda_i \in \mathbb{R} \Rightarrow$实系数齐次线性方程组$(\mathbf{A} - \lambda_i \mathbf{E}\boldsymbol{x} = \mathbf{0})$必有实的基础解系,所以对应特征值$\lambda_i$的特征向量可以取实向量.

性质2

设对称矩阵$\mathbf{A}$的特征值为λ_1, λ_2,其对应特征向量$\boldsymbol{p}_1, \boldsymbol{p}_2$.若$\lambda_1 \neq \lambda_2$,则$\boldsymbol{p}_1$与$\boldsymbol{p}_2$正交.

实对称矩阵的特征值和特征向量的性质

- $\boldsymbol{x} \neq \mathbf{0} \Rightarrow \overline{\boldsymbol{x}}^T \boldsymbol{x} = \sum_{i=1}^{n} \overline{\boldsymbol{x}}_i \boldsymbol{x}_i = \sum_{i=1}^{n} |x_i|^2 \neq 0$,所以$\lambda - \overline{\lambda} = 0$,即$\lambda = \overline{\lambda}$,这说明$\lambda \in \mathbb{R}$.
- 特征值$\lambda_i \in \mathbb{R} \Rightarrow$实系数齐次线性方程组$(\mathbf{A} - \lambda_i \mathbf{E}\boldsymbol{x} = \mathbf{0})$必有实的基础解系,所以对应特征值$\lambda_i$的特征向量可以取实向量.

性质2

设对称矩阵$\mathbf{A}$的特征值为λ_1, λ_2,其对应特征向量$\boldsymbol{p}_1, \boldsymbol{p}_2$.若$\lambda_1 \neq \lambda_2$,则$\boldsymbol{p}_1$与$\boldsymbol{p}_2$正交.

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

$$
\begin{aligned}
\lambda_1 \boldsymbol{p}_1^T \boldsymbol{p}_2 &= (\lambda_1 \boldsymbol{p}_1^T) \boldsymbol{p}_2 = (\lambda_1 \boldsymbol{p}_1)^T \boldsymbol{p}_2 = (\mathbf{A}\boldsymbol{p}_1)^T \boldsymbol{p}_2 \\
&= \boldsymbol{p}_1^T \mathbf{A}^T \boldsymbol{p}_2 = \boldsymbol{p}_1^T (\mathbf{A}\boldsymbol{p}_2) = \boldsymbol{p}_1^T (\lambda_2 \boldsymbol{p}_2) = \lambda_2 \boldsymbol{p}_1^T \boldsymbol{p}_2.
\end{aligned}
$$

- $\boldsymbol{x} \neq \mathbf{0} \Rightarrow \overline{\boldsymbol{x}}^T\boldsymbol{x} = \sum_{i=1}^{n} \overline{\boldsymbol{x}}_i \boldsymbol{x}_i = \sum_{i=1}^{n} |x_i|^2 \neq 0$,所以$\lambda - \overline{\lambda} = 0$,即$\lambda = \overline{\lambda}$,这说明$\lambda \in \mathbb{R}$.
- 特征值$\lambda_i \in \mathbb{R} \Rightarrow$实系数齐次线性方程组$(\mathbf{A} - \lambda_i \mathbf{E}\boldsymbol{x} = \mathbf{0})$必有实的基础解系,所以对应特征值$\lambda_i$的特征向量可以取实向量.

性质2

设对称矩阵$\mathbf{A}$的特征值为λ_1, λ_2,其对应特征向量$\boldsymbol{p}_1, \boldsymbol{p}_2$.若$\lambda_1 \neq \lambda_2$,则$\boldsymbol{p}_1$与$\boldsymbol{p}_2$正交.

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

$$
\begin{aligned}
\lambda_1 \boldsymbol{p}_1^T \boldsymbol{p}_2 =& (\lambda_1 \boldsymbol{p}_1^T)\boldsymbol{p}_2 = (\lambda_1 \boldsymbol{p}_1)^T \boldsymbol{p}_2 = (\mathbf{A}\boldsymbol{p}_1)^T \boldsymbol{p}_2 \\
=& \boldsymbol{p}_1^T \mathbf{A}^T \boldsymbol{p}_2 = \boldsymbol{p}_1^T (\mathbf{A}\boldsymbol{p}_2) = \boldsymbol{p}_1^T (\lambda_2 \boldsymbol{p}_2) = \lambda_2 \boldsymbol{p}_1^T \boldsymbol{p}_2.
\end{aligned}
$$

即$(\lambda_1 - \lambda_2)\boldsymbol{p}_1^T \boldsymbol{p}_2 = 0$但$\lambda_1 \neq \lambda_2$, 所以$\langle \boldsymbol{p}_1, \boldsymbol{p}_2 \rangle = \boldsymbol{p}_1^T \boldsymbol{p}_2 = 0$, 即$\boldsymbol{p}_1$与$\boldsymbol{p}_2$正交.

定理

n阶实对称阵$\mathbf{A}$必定正交相似于实对角阵$\mathbf{\Lambda}$:即存在正交矩阵$\mathbf{P}$,使得

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

其中$\mathbf{\Lambda}$的对角线上的元素是$\mathbf{A}$的n个特征值.

定理

n阶实对称阵$\mathbf{A}$必定正交相似于实对角阵$\mathbf{\Lambda}$:即存在正交矩阵$\mathbf{P}$,使得

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

其中$\mathbf{\Lambda}$的对角线上的元素是$\mathbf{A}$的n个特征值.

推论

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

$$R(\mathbf{A}-\lambda\mathbf{E}) = n-k,$$

从而对应特征值λ有k个线性无关的特征向量.

算法

我们有如下将实对称矩阵$\mathbf{A}$对角化的步骤:

(1) 计算$\mathbf{A}$的全部特征值$\lambda_1 \neq \lambda_2 \neq \cdots \neq \lambda_s$, λ_i的重数依次为$k_i, i = 1, 2, \cdots, s$且$k_1 + k_2 + \cdots + k_s = n$;

(2) 对于每个k_i重特征值λ_i, 我们求方程$(\mathbf{A} - \lambda_i \mathbf{E})\boldsymbol{x} = \mathbf{0}$的基础解系,得到$k_i$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. 由于$k_1 + k_2 + \cdots + k_s = n$, 所以我们总共可得$n$个两两正交的单位特征向量;

(3) 将这n个两两正交的单位特征向量构成正交阵$\mathbf{P}$,有$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{\Lambda}$.

(4) 需要注意:对角矩阵$\mathbf{\Lambda}$中对角元的排列次序应与$\mathbf{P}$中列向量的排列次序相对应.

例 1

设矩阵$\mathbf{A}=\begin{bmatrix}1 & 0 & 2\\ 0 & -1 & 0\\ 3 & 0 & 2\end{bmatrix}$, 求正交阵$\mathbf{P}$, 使$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\mathbf{P}^{T}\mathbf{A}\mathbf{P}$为对角阵.

例 1

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

例 2

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

二次型及其标准形

课前导读

已知平面$\mathbb{R}^2$上一条曲线方程为$3x^2+3y^2+4xy=1$，为了求曲线上到原点的距离最长和最短的点，可以先选择适当的坐标旋转变换

$$\begin{cases} x = x'\cos\theta - y'\sin\theta, \\ y = x'\sin\theta + y'\cos\theta, \end{cases} \quad \theta = \frac{\pi}{4},$$

即

$$\begin{cases} x = \dfrac{\sqrt{2}}{2}x' - \dfrac{\sqrt{2}}{2}y', \\ y = \dfrac{\sqrt{2}}{2}x' + \dfrac{\sqrt{2}}{2}y', \end{cases}$$

将曲线方程化为标准方程: $x'^2+5y'^2=1$.

课前导读

已知平面$\mathbb{R}^2$上一条曲线方程为$3x^2+3y^2+4xy=1$，为了求曲线上到原点的距离最长和最短的点，可以先选择适当的坐标旋转变换

$$\begin{cases} x = x'\cos\theta - y'\sin\theta, \\ y = x'\sin\theta + y'\cos\theta, \end{cases} \quad \theta = \frac{\pi}{4},$$

即

$$\begin{cases} x = \dfrac{\sqrt{2}}{2}x' - \dfrac{\sqrt{2}}{2}y', \\ y = \dfrac{\sqrt{2}}{2}x' + \dfrac{\sqrt{2}}{2}y', \end{cases}$$

将曲线方程化为标准方程: $x'^2+5y'^2=1$.

显然，这是一条椭圆曲线，从而曲线上到原点的距离最长和最短的点分别可取$(\pm 1, 0)$和$(0, \pm\frac{1}{\sqrt{5}})$.

课前导读

从代数学的观点来看，上述化曲线的一般方程为标准方程的过程，就是通过变量间非退化的线性替换把一个二次齐次多项式化简为只含有平方项的过程．我们会经常在许多实际问题或理论问题中遇到这样的问题．

从代数学的观点来看, 上述化曲线的一般方程为标准方程的过程, 就是通过变量间非退化的线性替换把一个二次齐次多项式化简为只含有平方项的过程. 我们会经常在许多实际问题或理论问题中遇到这样的问题.

本节我们讨论一般的含n个变量的二次齐次多项式并研究如何利用变量间非退化的线性替换将二次齐次多项式化简为只含有平方项的二次多项式.

定义 (二次型及其标准形)

- 我们称数域$\mathbb{F}$上含有n个变量$x_1, x_2, \cdots, x_n$的二次齐次多项式

$$
\begin{aligned}
f(x_1, x_2, \cdots, x_n) = \\
a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 \quad &+ \cdots + \quad 2a_{1,n-1}x_1x_{n-1} + 2a_{1n}x_1x_n \\
+a_{22}x_2^2 + 2a_{23}x_2x_3 \quad &+ \cdots + \quad 2a_{2,n-1}x_2x_{n-1} + 2a_{2n}x_2x_n \\
&\cdots\cdots \\
+a_{n-1,n-1}x_{n-1}^2 \quad &+ \cdots + \quad 2a_{n-1,n}x_{n-1}x_n \\
&+a_{n,n}x_n^2
\end{aligned}
\tag{5-1}
$$

  为一个二次型.

- 如果所有系数$a_{ij} \in \mathbb{R}(1 \leqslant i, j \leqslant n)$, 那么我们称该二次型为**实二次型**.

定义 (二次型及其标准形)

- 如果n元实二次型$f(x_1, x_2, \cdots, x_n)$只含有平方项, 即

$$f(x_1, x_2, \cdots, x_n) = k_1 x_1^2 + k_2 x_2^2 + \cdots + k_n x_n^2,$$

  那么我们称这样的二次型为**二次型的标准形**.

- 如果二次型标准形的系数$k_1, k_2, \cdots, k_n$只在$1, -1, 0$三个数中取值, 即

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

  那么我们称这样的二次型为**二次型的规范形**.

注

在式(5-1)中，对$j>i$取$a_{ji}=a_{ij}$，那么

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

于是式(5-1)中二次型有如下对称多项式求和表示:

$$\begin{aligned} f(x_1,x_2,\cdots,x_n) &= a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n \\ &+ a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n \\ &+ \cdots\cdots\cdots\cdots \\ &+ a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2 \\ &= \sum_{i,j=1}^{n} a_{ij}x_ix_j. \end{aligned} \tag{5-2}$$

注

式(5-2)中二次型有如下矩阵乘积表示:

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

二次型及其标准形的定义

注

记

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

那么上述二次型又可记作

$$f(x)=(x_1,x_2,\cdots,x_n)\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x},$$

其中$\mathbf{A}$为一个对称矩阵.

二次型及其标准形的定义

例

二次型$f = x_1^2 - 3x_3^2 - 4x_1x_2 + x_2x_3$有如下矩阵乘积表示:

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

注

上述推理说明二次型与对称矩阵之间存在一一对应的关系. 因此,

- 矩阵$\mathbf{A}$叫做二次型$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$的矩阵.
- $f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$叫做对称阵$\mathbf{A}$的二次型.

注

对于二次型$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,如何寻求可逆的线性变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n, \\ \qquad \cdots\cdots\cdots\cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases} \tag{5-3}$$

使得二次型$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$可以化为标准形，其中$c_{ij}\in\mathbb{R}(1\leqslant i,j\leqslant n)$.

注

对于二次型$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,如何寻求可逆的线性变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n, \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n, \\ \qquad \cdots\cdots\cdots\cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases} \tag{5-3}$$

使得二次型$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$可以化为标准形，其中$c_{ij} \in \mathbb{R}(1 \leqslant i, j \leqslant n)$.

记$\mathbf{C} = (c_{ij})$，$\boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$，我们可以把可逆变换(5-3)记作$\boldsymbol{x} = \mathbf{C}\boldsymbol{y}$.

用正交变换化二次型为标准形

将$\boldsymbol{x}=\mathbf{C}\boldsymbol{y}$代入$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,我们有

$$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}=(\mathbf{C}\boldsymbol{y})^T\mathbf{A}\mathbf{C}\boldsymbol{y}=\boldsymbol{y}^T(\mathbf{C}^T\mathbf{A}\mathbf{C})\boldsymbol{y}=\boldsymbol{y}^T\mathbf{B}\boldsymbol{y}=g(y).$$

将$\boldsymbol{x}=\mathbf{C}\boldsymbol{y}$代入$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,我们有

$$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}=(\mathbf{C}\boldsymbol{y})^T\mathbf{A}\mathbf{C}\boldsymbol{y}=\boldsymbol{y}^T(\mathbf{C}^T\mathbf{A}\mathbf{C})\boldsymbol{y}=\boldsymbol{y}^T\mathbf{B}\boldsymbol{y}=g(y).$$

因此, 如果二次型$g(y)=\boldsymbol{y}^T\mathbf{B}\boldsymbol{y}$是标准形, 那么矩阵$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}$就是对角阵.

用正交变换化二次型为标准形

将$\boldsymbol{x} = \mathbf{C}\boldsymbol{y}$代入$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,我们有

$$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x} = (\mathbf{C}\boldsymbol{y})^T\mathbf{A}\mathbf{C}\boldsymbol{y} = \boldsymbol{y}^T(\mathbf{C}^T\mathbf{A}\mathbf{C})\boldsymbol{y} = \boldsymbol{y}^T\mathbf{B}\boldsymbol{y} = g(y).$$

因此，如果二次型$g(y) = \boldsymbol{y}^T\mathbf{B}\boldsymbol{y}$是标准形，那么矩阵$\mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}$就是对角阵.

定义 (合同矩阵)

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

用正交变换化二次型为标准形

将$\boldsymbol{x} = \mathbf{C}\boldsymbol{y}$代入$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$,我们有

$$f(x) = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x} = (\mathbf{C}\boldsymbol{y})^T\mathbf{A}\mathbf{C}\boldsymbol{y} = \boldsymbol{y}^T(\mathbf{C}^T\mathbf{A}\mathbf{C})\boldsymbol{y} = \boldsymbol{y}^T\mathbf{B}\boldsymbol{y} = g(y).$$

因此, 如果二次型$g(y) = \boldsymbol{y}^T\mathbf{B}\boldsymbol{y}$是标准形, 那么矩阵$\mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}$就是对角阵.

定义 (合同矩阵)

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

合同关系

矩阵间的合同关系是一个等价关系, 其满足

(1) 反身性:每一个方阵都与它自身合同:$\mathbf{A} = \mathbf{E}^T\mathbf{A}\mathbf{E}$

(2) 对称性:如果$\mathbf{A}$与$\mathbf{B}$合同,那么$\mathbf{B}$与$\mathbf{A}$也合同:
$\mathbf{B} = \mathbf{C}^T\mathbf{A}\mathbf{C}$,$\mathbf{C}$可逆$\Rightarrow\ \mathbf{A} = \mathbf{P}^T\mathbf{A}\mathbf{P}$,$\mathbf{P} = \mathbf{C}^{-1}$.

(3) 传递性:如果$\mathbf{A}$与$\mathbf{B}$合同,如果$\mathbf{B}$与$\mathbf{C}$合同,那么$\mathbf{A}$与$\mathbf{C}$也合同:
$\mathbf{B} = \mathbf{P_1}^T\mathbf{A}\mathbf{P_1}, \mathbf{C} = \mathbf{P_2}^T\mathbf{B}\mathbf{P_2} \Rightarrow\ \mathbf{C} = \mathbf{P_2}^T\mathbf{P_1}^T\mathbf{A}\mathbf{P_1}\mathbf{P_2} = (\mathbf{P_1}\mathbf{P_2})^T\mathbf{A}(\mathbf{P_1}\mathbf{P_2})$.

命题

如果$\mathbf{A}$为对称阵，那么$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}$也为对称阵，且$R(\mathbf{B})=R(\mathbf{A})$.

命题

如果$\mathbf{A}$为对称阵，那么$\mathbf{B}=\mathbf{C}^T\mathbf{A}\mathbf{C}$也为对称阵，且$R(\mathbf{B})=R(\mathbf{A})$.

注

- 二次型$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$经可逆变换$\boldsymbol{x}=\mathbf{C}\boldsymbol{y}$变成标准形，就是要使矩阵$\mathbf{C}^T\mathbf{A}\mathbf{C}$是对角阵.
- 我们称通过寻求可逆矩阵$\mathbf{C}$,将对称矩阵$\mathbf{A}$转化为$\mathbf{C}^T\mathbf{A}\mathbf{C}$并使其称为对角阵的过程称为对称矩阵$\mathbf{A}$的**合同对角化**.
- 由实对称矩阵对角化的理论可知:任给对称阵$\mathbf{A}$，总有正交阵$\mathbf{P}$，使得

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

用正交变换化二次型为标准形

定理

任给实二次型$f = \sum_{i,j=1}^{n} a_{ij}x_ix_j (a_{ij} = a_{ji})$, 总有正交变换$\boldsymbol{x} = \mathbf{P}\boldsymbol{y}$, 使得实二次型$f$化为标准形$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$, 其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是$f$的矩阵$\mathbf{A} = (a_{ij})$的特征值.

用正交变换化二次型为标准形

定理

任给实二次型$f = \sum_{i,j=1}^{n} a_{ij}x_i x_j (a_{ij} = a_{ji})$, 总有正交变换$\boldsymbol{x} = \mathbf{P}\boldsymbol{y}$, 使得实二次型$f$化为标准形$f = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$, 其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是$f$的矩阵$\mathbf{A} = (a_{ij})$的特征值.

推论

任给n元实二次型$f(x) = \boldsymbol{x}^T \mathbf{A}\boldsymbol{x} (\mathbf{A}^T = \mathbf{A})$, 总有可逆变换$\boldsymbol{x} = \mathbf{C}\boldsymbol{z}$, 使得实二次型$f(\mathbf{C}\boldsymbol{z})$可化为规范形.

定理

任给实二次型$f=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j(a_{ij}=a_{ji})$, 总有正交变换$\boldsymbol{x}=\mathbf{P}\boldsymbol{y}$, 使得实二次型$f$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$\mathbf{A}=(a_{ij})$的特征值.

推论

任给n元实二次型$f(x)=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}(\mathbf{A}^T=\mathbf{A})$, 总有可逆变换$\boldsymbol{x}=\mathbf{C}\boldsymbol{z}$, 使得实二次型$f(\mathbf{C}\boldsymbol{z})$可化为规范形.

例 1

求一个正交变换$\boldsymbol{x}=\mathbf{P}\boldsymbol{y}$, 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.

用配方法化二次型为标准形

命题

- 用正交变换化二次型成标准形具有保持几何形状不变的优点.
- 对于研究二次型正定性来说, 我们不仅可以用正交变换, 还可以用可逆的线性变换$x = \mathbf{P}y$把二次型化成标准形, 有时我们称这种方法称为**配方法**.

用配方法化二次型为标准形

命题

- 用正交变换化二次型成标准形具有保持几何形状不变的优点.
- 对于研究二次型正定性来说, 我们不仅可以用正交变换, 还可以用可逆的线性变换$\boldsymbol{x} = \mathbf{P}\boldsymbol{y}$把二次型化成标准形, 有时我们称这种方法称为**配方法**.

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.

命题

- 用正交变换化二次型成标准形具有保持几何形状不变的优点.
- 对于研究二次型正定性来说, 我们不仅可以用正交变换, 还可以用可逆的线性变换$\boldsymbol{x} = \mathbf{P}\boldsymbol{y}$把二次型化成标准形, 有时我们称这种方法称为**配方法**.

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.

例 3

化二次型$f = 2x_1x_2 + 4x_1x_3 - 6x_2x_3$成标准形, 并求所用的变换矩阵.

正定二次型与正定矩阵

课前导读

一个实二次型$f = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$总可以经过可逆的变量替换化为标准形, 但是标准形并不是唯一确定的.

一个实二次型$f = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$总可以经过可逆的变量替换化为标准形, 但是标准形并不是唯一确定的. 例如二次型

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

一个实二次型$f = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$总可以经过可逆的变量替换化为标准形，但是标准形并不是唯一确定的. 例如二次型

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

如果我们令

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 1 \\ \frac{1}{3} & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

那么我们有标准形$f = \frac{4}{3}y_1^2 - 6y_2^2 + 6y_3^2$.

一个实二次型$f=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$总可以经过可逆的变量替换化为标准形, 但是标准形并不是唯一确定的. 例如二次型

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

如果我们令

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 1 \\ \frac{1}{3} & 1 & -1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

那么我们有标准形$f=\frac{4}{3}y_1^2-6y_2^2+6y_3^2$.

进一步我们可以发现, 虽然用不同的可逆变量替换, 二次型的标准形不同, 但在不同的标准形中, 正系数的个数相同, 负系数的个数也相同.

一个实二次型$f = \boldsymbol{x}^T \mathbf{A}\boldsymbol{x}$总可以经过可逆的变量替换化为标准形, 但是标准形并不是唯一确定的. 例如二次型

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

如果我们令

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{2}{3} & 1 & 1 \\ \frac{1}{3} & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

那么我们有标准形$f = \frac{4}{3}y_1^2 - 6y_2^2 + 6y_3^2$.

进一步我们可以发现, 虽然用不同的可逆变量替换, 二次型的标准形不同, 但在不同的标准形中, 正系数的个数相同, 负系数的个数也相同.

本节我们来讨论一般二次型标准形的正系数和负系数.

定理 (惯性定理)

设有二次型$f = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$, 它的秩为$r$. 如果我们有两个可逆变换$\boldsymbol{x} = \mathbf{C}\boldsymbol{y}$及$\boldsymbol{x} = \mathbf{P}\boldsymbol{z}$,使得

$$f = k_1 y_1^2 + k_2 y_2^2 + \cdots + k_r y_r^2 \, (k_i \neq 0),$$

及

$$f = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_r z_r^2 \, (\lambda_i \neq 0),$$

那么$k_1, k_2, \cdots, k_r$中正数的个数与$\lambda_1, \lambda_2, \cdots, \lambda_r$中正数的个数相等.

定理 (惯性定理)

设有二次型$f = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$, 它的秩为$r$. 如果我们有两个可逆变换$\boldsymbol{x} = \mathbf{C}\boldsymbol{y}$及$\boldsymbol{x} = \mathbf{P}\boldsymbol{z}$,使得

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

及

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

那么$k_1, k_2, \cdots, k_r$中正数的个数与$\lambda_1, \lambda_2, \cdots, \lambda_r$中正数的个数相等.

定义 (二次型的正惯性指数和负惯性指数)

- 二次型标准形中正系数的个数称为**二次型的正惯性指数**,
- 二次型标准形中负系数的个数称为**二次型的负惯性指数**.

注

若二次型f的正惯性指数为p, 秩为r, 则f的规范形便可确定为

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

注

若二次型f的正惯性指数为p, 秩为r, 则f的规范形便可确定为

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

定义 (正定二次型与负定二次型)

设有二次型$f = \boldsymbol{x}^T \mathbf{A} \boldsymbol{x}$.

- 如果对于任何$\boldsymbol{x} \neq 0$, 都有$f(\boldsymbol{x}) > 0$, 那么我们称二次型f为**正定二次型**, 并称对称矩阵$\mathbf{A}$**是正定的**;
- 如果对于任何$\boldsymbol{x} \neq 0$, 都有$f(\boldsymbol{x}) < 0$, 那么我们称二次型f为**负定二次型**, 并称对称矩阵$\mathbf{A}$**是负定的**.

正定二次型与正定阵

定理

n元二次型$f = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$为正定的$\Leftrightarrow$二次型$f$的正惯性指数等于$n$, 即二次型$f$的规范形的$n$个系数全为$1$.

正定二次型与正定阵

定理

n元二次型$f=\boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$为正定的$\Leftrightarrow$二次型$f$的正惯性指数等于$n$, 即二次型$f$的规范形的$n$个系数全为$1$.

推论

- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow\mathbf{A}$与单位矩阵$\mathbf{E}$合同.
- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow\mathbf{A}$的特征值全为正.

正定二次型与正定阵

定理

n元二次型$f = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$为正定的$\Leftrightarrow$二次型$f$的正惯性指数等于$n$, 即二次型$f$的规范形的$n$个系数全为$1$.

推论

- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow \mathbf{A}$与单位矩阵$\mathbf{E}$合同.
- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow \mathbf{A}$的特征值全为正.

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.

正定二次型与正定阵

定理

n元二次型$f = \boldsymbol{x}^T\mathbf{A}\boldsymbol{x}$为正定的$\Leftrightarrow$二次型$f$的正惯性指数等于$n$, 即二次型$f$的规范形的$n$个系数全为$1$.

推论

- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow$ $\mathbf{A}$与单位矩阵$\mathbf{E}$合同.
- 对称矩阵$\mathbf{A}$为正定的$\Leftrightarrow$ $\mathbf{A}$的特征值全为正.

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.

例 2

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

定理 (Hurwitz)

- 对称阵A为正定的 $\Rightarrow$ A的各阶顺序主子式都为正,即

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

- 对称阵A为负定的 $\Rightarrow$ 奇数阶顺序主子式为负,偶数阶顺序主子式为正,即

$$(-1)^r \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0 (r = 1, 2, \cdots, n).$$

The End