线性代数 (Linear Algebra)

第七章 Symmetric Matrices and Quadratic Forms

§7.1 Diagonalization of Symmetric Matrices

对称矩阵的对角化

衡 益

2021 年 12 月 28 日，中山大学南校区

引入

7.1 对称矩阵的对角化

例1：

判断下列矩阵是否为对称矩阵

对称阵：$\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$，$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & -7 \end{bmatrix}$，$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}$

非对称阵：$\begin{bmatrix} 1 & -3 \\ 3 & 0 \end{bmatrix}$，$\begin{bmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{bmatrix}$，$\begin{bmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$

3

7.1 对称矩阵的对角化

定理 对称阵的特征值为实数。

定理 设A为n 阶对称矩阵，则必有正交阵P，使得

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

其中Λ是以A 的n个特征值为对角元素的对角阵。

推论

设A为n 阶对称阵，λ 是A 的特征方程的k 重根，则矩阵A $-$ λI 的秩

$$R(A - \lambda I) = n - k$$

从而对应特征的值λ 恰有k 个线性无关的特征向量。

7.1 对称矩阵的对角化

回顾：第五章将对称阵A对角化的步骤

Step1

求出A 的全部互不相等的特征值$\lambda_1,\cdots,\lambda_s$,它们的重数依次为$k_1,\cdots,k_s$,

$$k_1+\cdots+k_s=n.$$

Next

7.1 对称矩阵的对角化

将对称阵A对角化的步骤

Step2

对每个k_i重特征值λ_i,求方程$(\mathbf{A}-\lambda_i\mathbf{I})x=0$ 的基础解系，得k_i个线性无关得特征向量。再把他们正交化、单位化，得k_i个两两正交的单位特征向量。因为 $k_1+\cdots+k_s=n$，故总共可得n个两两正交的单位特征向量。

Next

7.1 对称矩阵的对角化

将对称阵A对角化的步骤

Step3

把这n个两两正交的单位特征向量构成一个正交阵$\mathbf{P}$，便有

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

注意$\mathbf{\Lambda}$对角元素的排列依次序应与P 中列向量的排列依次序相对应。

End

7

7.1 对称矩阵的对角化

例2： 将矩阵$\mathbf{A}=\begin{bmatrix}6 & -2 & -1\\ -2 & 6 & -1\\ -1 & -1 & 5\end{bmatrix}$对角化。

解 $\mathbf{A}$的特征方程为

$$0=-\lambda^3+17\lambda^2-90\lambda+144=-(\lambda-8)(\lambda-6)(\lambda-3)$$

计算出每个特征子空间的基

$$\lambda=8，\ v_1=\begin{bmatrix}-1\\ 1\\ 0\end{bmatrix};\ \lambda=6，\ v_2=\begin{bmatrix}-1\\ -1\\ 2\end{bmatrix};\ \lambda=8，\ v_1=\begin{bmatrix}-1\\ 1\\ 0\end{bmatrix}$$

这三个向量是R^3空间的一组基，且是正交基。将其单位化：

8

7.1 对称矩阵的对角化

例2： 将矩阵$\mathbf{A}=\begin{bmatrix}6 & -2 & -1\\ -2 & 6 & -1\\ -1 & -1 & 5\end{bmatrix}$对角化。

解

$$u_1=\begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\\ 0\end{bmatrix}, u_2=\begin{bmatrix}-1/\sqrt{6}\\ -1/\sqrt{6}\\ 2/\sqrt{6}\end{bmatrix}, u_3=\begin{bmatrix}1/\sqrt{3}\\ 1/\sqrt{3}\\ 1/\sqrt{3}\end{bmatrix}$$

令$\mathbf{P}=\begin{bmatrix}-1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 0 & 2/\sqrt{6} & 1/\sqrt{3}\end{bmatrix}$, $\mathbf{D}=\begin{bmatrix}8 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & 3\end{bmatrix}$

那么$\mathbf{A}=\mathbf{PDP}^{-1}$.此时，$\mathbf{P}$是正交阵，即$\mathbf{P}^{-1}=\mathbf{P}^{T}$.

9

7.1 对称矩阵的对角化

Question ：为什么例2中的特征向量是正交的?

10

7.1 对称矩阵的对角化

定理1 若**A**是对称矩阵，那么不同特征子空间的任意两个特征向量都是正交的。

证明：

设v_1和v_2是λ_1和λ_2所对应的特征向量。为证$v_1 \cdot v_2 = 0$，计算

$\lambda_1 v_1 \cdot v_2 = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2$　　因为v_1是特征向量

$= (v_1^T A^T) v_2 = v_1^T (Av_2)$　　因为$A^T = A$

$= v_1^T (\lambda_2 v_2)$　　因为v_2是特征向量

$= \lambda_2 v_1^T v_2 = \lambda_2 v_1 \cdot v_2$

因此$(\lambda_1 - \lambda_2) v_1 \cdot v_2 = 0$，但是$\lambda_1 - \lambda_2 \neq 0$，所以$v_1 \cdot v_2 = 0$.

11

7.1 对称矩阵的对角化

定义 若存在一个正交矩阵 **P** 和对角阵 **D**，使得

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

那么对称矩阵**A**是可正交对角化的。

注

若**A**是正交可对角化矩阵，那么

$$\mathbf{A}^T = (\mathbf{P}\mathbf{D}\mathbf{P}^T)^T = (\mathbf{D}\mathbf{P}^T)^T\mathbf{P}^T$$

$$= (\mathbf{P}^T)^T\mathbf{D}^T\mathbf{P}^T = \mathbf{P}\mathbf{D}\mathbf{P}^T = \mathbf{A}$$

A是对称阵！！

12

7.1 对称矩阵的对角化

定理2 若n 阶矩阵$\mathbf{A}$ 是正交可对角化的，当且仅当 $\mathbf{A}$ 是对称矩阵。

每一个对称矩阵都是可正交对角化的！

13

7.1 对称矩阵的对角化

例3 $\mathbf{A}=\begin{bmatrix}3 & -2 & 4\\ -2 & 6 & 2\\ 4 & 2 & 3\end{bmatrix}$是正交可对角化矩阵，其特征方程为

$$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$$

解 计算出特征子空间的基：

$$\lambda=7，\ v_1=\begin{bmatrix}1\\0\\1\end{bmatrix}, v_2=\begin{bmatrix}-1/2\\1\\0\end{bmatrix};\quad \lambda=-2，\ v_3=\begin{bmatrix}-1\\-1/2\\1\end{bmatrix}$$

尽管v_1和v_2是线性独立的，但不是正交的。v_2 在v_1 上的投影为$\dfrac{v_1\cdot v_2}{v_1\cdot v_1}v_1$，与$v_1$正交的$v_2$的分量是

14

7.1 对称矩阵的对角化

例3 $\mathbf{A}=\begin{bmatrix}3 & -2 & 4\\ -2 & 6 & 2\\ 4 & 2 & 3\end{bmatrix}$是正交可对角化矩阵，其特征方程为

$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$

解

$$z_2=v_2-\frac{v_2\cdot v_1}{v_1\cdot v_1}v_1=\begin{bmatrix}-1/2\\ 1\\ 0\end{bmatrix}-\frac{-1/2}{2}\begin{bmatrix}1\\ 0\\ 1\end{bmatrix}=\begin{bmatrix}-1/4\\ 1\\ 1/4\end{bmatrix}$$

那么$\{v_1,v_2\}$是$\lambda=7$特征子空间的一组正交集合。

将v_1和v_2规范化，可得关于$\lambda=7$特征子空间单位正交基：

$$u_1=\begin{bmatrix}1/\sqrt{2}\\ 0\\ 1/\sqrt{2}\end{bmatrix},\quad u_2=\begin{bmatrix}-1/\sqrt{18}\\ 4/\sqrt{18}\\ 1/\sqrt{18}\end{bmatrix}$$

15

7.1 对称矩阵的对角化

例3 $\mathbf{A}=\begin{bmatrix}3 & -2 & 4\\ -2 & 6 & 2\\ 4 & 2 & 3\end{bmatrix}$是正交可对角化矩阵，其特征方程为

$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$

解 同样的，解出$\lambda=-2$时的特征子空间的一组单位正交基

$$u_3=\frac{1}{\|2v_3\|}2v_3=\frac{1}{3}\begin{bmatrix}-2\\ -1\\ 2\end{bmatrix}=\begin{bmatrix}-2/3\\ -1/3\\ 2/3\end{bmatrix}$$

由定理1可知，u_3与u_1和u_2是正交的，因此$\{u_1,u_2,u_3\}$是一个单位正交基。

16

7.1 对称矩阵的对角化

例3 $\mathbf{A}=\begin{bmatrix}3 & -2 & 4\\ -2 & 6 & 2\\ 4 & 2 & 3\end{bmatrix}$是正交可对角化矩阵，其特征方程为

$$0=-\lambda^3+12\lambda^2-21\lambda-98=-(\lambda-7)^2(\lambda+2)$$

解

令

$$\mathbf{P}=[u_1\quad u_2\quad u_3]=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{18} & -2/3\\ 0 & 4/\sqrt{18} & -1/3\\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3\end{bmatrix},\mathbf{D}=\begin{bmatrix}7 & 0 & 0\\ 0 & 7 & 0\\ 0 & 0 & -2\end{bmatrix}$$

那么$\mathbf{P}$可以将$\mathbf{A}$正交对角化，$\mathbf{A}=\mathbf{PDP}^{-1}$

17

7.1 对称矩阵的对角化

谱定理 (Spectral Theorem)

矩阵A的特征值的集合可以称为A的谱

定理4 **对称矩阵的谱定理**

若n 阶数对称矩阵$\mathbf{A}$ 有如下性质：

(1) $\mathbf{A}$有n个实特征值；

(2) λ特征子空间的维数等于特征根λ 的重数；

(3) 特征子空间是相互正交的；

(4) $\mathbf{A}$是可正交对角化矩阵。

18

7.1 对称矩阵的对角化

谱分解

设$\mathbf{A}=\mathbf{PDP}^{-1}$,其中$\mathbf{P}$ 的列向量$u_1,\cdots,u_n$是标准正交的特征向量，其所对应的特征值为$\lambda_1\cdots\lambda_n$，构成了对角阵$\mathbf{D}$. 那么由于$\mathbf{P}^{-1}=\mathbf{P}^T$,

$$\mathbf{A}=\mathbf{PDP}^T=\begin{bmatrix}u_1 & \cdots & u_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{bmatrix}\begin{bmatrix}u_1^T\\ \vdots\\ u_n^T\end{bmatrix}$$

$$=\begin{bmatrix}\lambda_1 u_1 & \cdots & \lambda_n u_n\end{bmatrix}\begin{bmatrix}u_1^T\\ \vdots\\ u_n^T\end{bmatrix}$$

矩阵A的谱分解

可以将A表示为:

$$\mathbf{A}=\lambda_1 u_1 u_1^T+\lambda_2 u_2 u_2^T+\cdots+\lambda_n u_n u_n^T$$

19

7.1 对称矩阵的对角化

例4 构造矩阵A 的谱分解

$$\mathbf{A}=\begin{bmatrix}7 & 2\\ 2 & 4\end{bmatrix}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}8 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

解 设$\mathbf{P}=\begin{bmatrix}u_1 & u_2\end{bmatrix}$,那么

$$\mathbf{A}=8u_1u_1^T+3u_2u_2^T$$

验证A的分解，计算

$$u_1u_1^T=\begin{bmatrix}2/\sqrt{5}\\ 1/\sqrt{5}\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\end{bmatrix}=\begin{bmatrix}4/5 & 2/5\\ 2/5 & 1/5\end{bmatrix}$$

$$u_2u_2^T=\begin{bmatrix}-1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}-1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}=\begin{bmatrix}1/5 & -2/5\\ -2/5 & 4/5\end{bmatrix}$$

20

7.1 对称矩阵的对角化

例4 构造矩阵$\mathbf{A}$ 的谱分解

$$\mathbf{A}=\begin{bmatrix}7 & 2\\ 2 & 4\end{bmatrix}=\begin{bmatrix}2/\sqrt{5} & -1/\sqrt{5}\\ 1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}\begin{bmatrix}8 & 0\\ 0 & 3\end{bmatrix}\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix}$$

解

并且

$$8u_1u_1^T+3u_2u_2^T=\begin{bmatrix}32/5 & 16/5\\ 16/5 & 8/5\end{bmatrix}+\begin{bmatrix}3/5 & -6/5\\ -6/5 & 12/5\end{bmatrix}$$

$$=\begin{bmatrix}7 & 2\\ 2 & 4\end{bmatrix}=\mathbf{A}$$

21

7.1 对称矩阵的对角化

注：数值计算

当A是对称且不太大的矩阵时，现代高性能的计算机程序可以非常精确地计算特征值和特征向量，它们对A使用一系列包含正交矩阵的相似变换，变换后矩阵的对角元素很快收敛于A的特征值。利用正交矩阵常常可避免计算过程的误差积累，当A对称时，正交矩阵序列可形成列向量是A的特征向量的正交矩阵。

一个非对称矩阵没有完全的正交特征向量集，但通过算法仍得到相当精确的特征值，之后，就需要用非正交化方法计算特征向量。

22

Q & A

线性代数 (Linear Algebra)

第七章 Symmetric Matrices and Quadratic Forms

§ 7.2 Quadratic Forms

二次型

衡 益

2021 年 12 月 28 日，中山大学南校区

7.2 二次型

工程和信号处理（标准涉及和优化、输出噪声功率）

物理学（势能、动能）

二次型

微分几何（曲面的法曲率）

经济、统计学（效用函数、置信椭圆体）

25

7.2 二次型

定义在R^n上的函数

在向量x处可由表达式 $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ 计算

R^n上的二次型

$Q(\mathbf{x}) = \mathbf{x}^T I\mathbf{x} = \|\mathbf{x}\|^2$

矩阵A称为关于二次型的矩阵

26

7.2 二次型

例1 设$\mathbf{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}$,计算$\mathbf{x}^T\mathbf{A}\mathbf{x}$.

(1)$\mathbf{A}=\begin{bmatrix}4&0\\0&3\end{bmatrix}$ (2) $\mathbf{A}=\begin{bmatrix}3&-2\\-2&7\end{bmatrix}$

无交叉项

解

(1) $\mathbf{x}^T\mathbf{A}\mathbf{x}=\begin{bmatrix}x_1&x_2\end{bmatrix}\begin{bmatrix}4&0\\0&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}x_1&x_2\end{bmatrix}\begin{bmatrix}4x_1\\3x_2\end{bmatrix}=\boxed{4x_1^2+3x_2^2}$

(2) A中的元素有两个负的元素，观察这(1, 2)元素如何参与运算

$$\mathbf{x}^T\mathbf{A}\mathbf{x}=\begin{bmatrix}x_1&x_2\end{bmatrix}\begin{bmatrix}3&-2\\-2&7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}x_1&x_2\end{bmatrix}\begin{bmatrix}3x_1-2x_2\\-2x_1+7x_2\end{bmatrix}$$

$$=x_1(3x_1-2x_2)+x_2(-2x_1+7x_2)$$

$$=3x_1^2-2x_1x_2-2x_2x_1+7x_2^2$$

$$=3x_1^2-\boxed{4x_1x_2}+7x_2^2$$

有交叉项

27

7.2 二次型

例2 $\mathbf{x}\in R^3$, 设

$$Q(\mathbf{x})=5x_1^2+3x_2^2+2x_3^2-x_1x_2+8x_2x_3$$

写出其二次型。

解

x_1^2,x_2^2,x_3^2的系数仍在对角线上，为使得**A**对称，当$i\neq j$ 时，x_ix_j的系数必须平均分配给矩阵**A** 中的(i,j)元素和(j,i)元素，x_1x_3的系数为0，容易验证：

$$Q(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{x}=\begin{bmatrix}x_1&x_2&x_3\end{bmatrix}\begin{bmatrix}5&-1/2&0\\-1/2&3&4\\0&4&2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

28

7.2 二次型

例3 令

$$Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$$

计算 $Q(\mathbf{x})$ 在 $\mathbf{x} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ 处的值。

解

$$Q(-3,1) = (-3)^2 - 8(-3)(1) - 5(1)^2 = 28$$
$$Q(2,-2) = (2)^2 - 8(2)(-2) - 5(-2)^2 = 16$$
$$Q(1,-3) = (1)^2 - 8(1)(-3) - 5(-3)^2 = -20$$

注：交叉项可以通过用适当的变量代换来消去

29

7.2 二次型

定义

含有 n 个变量 $x_1, x_2, \dots, x_n$ 的二次齐次函数

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

当 $j > i$ 时，取 $a_{ji} = a_{ij}$，则 $2a_{ij}x_ix_j = a_{ij}x_ix_j + a_{ji}x_jx_i$

$$\Rightarrow f(x_1, x_2, \dots, x_n) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

30

7.2 二次型

$$f(x_1, x_2, \ldots, x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

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

7.2 二次型

用矩阵符号表示二次型 $f(x, y, z) = x^2 - 3z^2 - 4xy + yz$

一一对应

$$f = (x, y, z) \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

对称矩阵

32

7.2 二次型

二次型的变量代换

若$\mathbf{x} \in \mathbb{R}^n$,那么变量代换为下面的等式形式:

$$\mathbf{x} = \mathbf{P}\mathbf{y} \quad 或 \quad \mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$$

此处$\mathbf{P}$是可逆矩阵，$\mathbf{y}$是$\mathbb{R}^n$ 中的另一个变量，此处$\mathbf{P}$ 的列向量可以确定$\mathbf{R}^n$的一个基。y是相对于该基向量x 的坐标向量。

若用变量代换处理二次型，那么

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = (\mathbf{P}\mathbf{y})^T\mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{y}^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{y} = \mathbf{y}^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{y}$$

新的二次型矩阵是

$$\mathbf{P}^T\mathbf{A}\mathbf{P}$$

若$\mathbf{P}$可将$\mathbf{A}$正交对角化，那么$\mathbf{P}^T = \mathbf{P}^{-1}$,且$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D}$,新的二次型矩阵是对角矩阵。

7.2 二次型

$$\mathbf{x}^T\boxed{\mathbf{A}}\mathbf{x} = (\mathbf{P}\mathbf{y})^T\mathbf{A}(\mathbf{P}\mathbf{y}) = \mathbf{y}^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{y} = \mathbf{y}^T\boxed{(\mathbf{P}^T\mathbf{A}\mathbf{P})}\mathbf{y}$$

定义

设$\mathbf{A}$和$\mathbf{B}$是n 阶矩阵，若由可逆矩阵$\mathbf{P}$，使$\mathbf{B} = \mathbf{P}^T\mathbf{A}\mathbf{P}$，则称矩阵$\mathbf{A}$与$\mathbf{B}$合同。

注：如果 A 是对称矩阵，$B^T = (C^TAC)^T = C^TA^TC = C^TAC = B$，$R(B) = R(A)$

7.2 二次型

二次型的变量代换

对于二次型，我们讨论的主要问题是：寻求可逆的线性变换

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

使二次型只含平方项，即

$$f = k_1y_1^2 + k_2y_2^2 + \cdots + k_ny_n^2$$

这种只含平方项的二次型，称为标准形。若能使标准形的系数 $k_1, k_2 \cdots, k_n$ 只在1,-1,0三个数中取值，也就是使

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

则称上式为二次型的规范形。

35

7.2 二次型

定理1

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 $x = Py$，使

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

是对应矩阵 (a_{ij}) 的特征值。

推论

任意给出 n 元二次型 $f(\mathrm{x}) = \mathrm{x}^T A \mathrm{x} \ (A^T = A)$，总有可逆变换

$$\mathrm{x} = C\mathrm{z}$$

使 $f(\mathrm{x})$ 为规范形。

36

7.2 二次型

例4 求一个变量代换将例3中的二次型变成一个没有交叉项的二次型。

解 例3中二次型对应的矩阵是

$$\mathbf{A}=\begin{bmatrix}1 & -4\\ -4 & -5\end{bmatrix}$$

首先将矩阵A正交对角化，$\mathbf{A}$ 的特征值是λ=3 和λ=-7，相应的单位特征向量是：

$$\lambda=3\ 时：\begin{bmatrix}2/\sqrt{5}\\ -1/\sqrt{5}\end{bmatrix},\quad \lambda=-7时：\begin{bmatrix}1/\sqrt{5}\\ 2/\sqrt{5}\end{bmatrix}$$

37

7.2 二次型

例4 求一个变量代换将例3中的二次型变成一个没有交叉项的二次型。

解 这些特征向量自动正交（因为它们属于不同的特征值）且构成R^2的一个单位正交基。取

$$\mathbf{P}=\begin{bmatrix}2/\sqrt{5} & 1/\sqrt{5}\\ -1/\sqrt{5} & 2/\sqrt{5}\end{bmatrix},\mathbf{D}=\begin{bmatrix}3 & 0\\ 0 & -7\end{bmatrix}$$

那么$\mathbf{A}=\mathbf{PDP}^{-1}$,且$\mathbf{D}=\mathbf{P}^{-1}\mathbf{AP}=\mathbf{P}^{T}\mathbf{AP}$，像前面指出的那样，一个适当的变换是：

$$\mathbf{x}=\mathbf{Py},\ 此处\ \mathbf{x}=\begin{bmatrix}x_1\\ x_2\end{bmatrix},\ \mathbf{y}=\begin{bmatrix}y_1\\ y_2\end{bmatrix}$$

38

7.2 二次型

例4 求一个变量代换将例3中的二次型变成一个没有交叉项的二次型。

解

那么

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^T\mathbf{A}\mathbf{x} \\ &= (\mathbf{P}\mathbf{y})^T\mathbf{A}(\mathbf{P}\mathbf{y}) \\ &= \mathbf{y}^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{y} = \mathbf{y}^T\mathbf{D}\mathbf{y} \\ &= 3y_1^2 - 7y_2^2 \end{aligned}$$

39

7.2 二次型

注:

为了说明例4中的二次型是相等的，我们可以利用新的二次型计算$Q(\mathbf{x})$在$\mathbf{x} = (2,\text{-}2)$处的值，首先，由于$\mathbf{x} = \mathbf{P}\mathbf{y}$，我们得到

$$\mathbf{y} = \mathbf{P}^{-1}\mathbf{x} = \mathbf{P}^T\mathbf{x}$$

则有

$$\mathbf{y} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}\begin{bmatrix} 2 \\ -2 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

因此

$$\begin{aligned} 3y_1^2 - 7y_2^2 &= 3(6/\sqrt{5})^2 - 7(-2/\sqrt{5})^2 = 3(36/5) - 7(4/5) \\ &= 80/5 = 16 \end{aligned}$$

40

7.2 二次型

这就是例3中$Q(\mathbf{x})$在$x=(2,-2)$处的值，如所示：

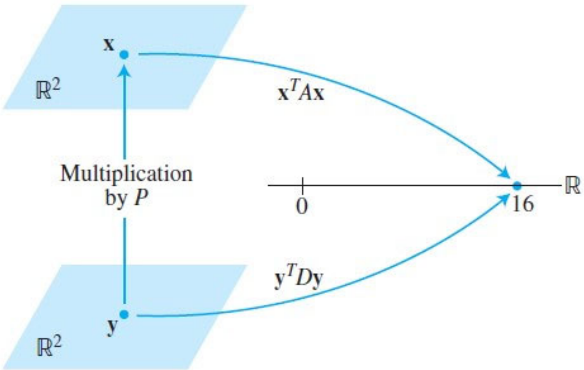


FIGURE 1 Change of variable in $\mathbf{x}^T A\mathbf{x}$.

41

7.2 二次型

例5 **二次型化标准型**

求一个正交变换 $x = Py$，把二次型 $f = -2x_1x_2 + 2x_1x_3 + 2x_2x_3$ 化为标准型

解 二次型对应矩阵为 $A = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

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

正交矩阵

42

7.2 二次型

例5 **二次型化标准型**

求一个正交变换 $x = Py$，把二次型 $f = -2x_1x_2 + 2x_1x_3 + 2x_2x_3$ 化为标准型

解

$$P^TAP = \Lambda = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix},$$

把二次型化成标准型 $\boxed{f = -2y_1^2 + y_2^2 + y_3^2.}$

43

7.2 二次型

例6 化二次型 $f = x_1^2 + 2x_2^2 + 5x_3^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$ 为标准型，并求所用的变换矩阵。

解

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

拉格朗日配方法

令 $\begin{cases} y_1 = x_1 + x_2 + x_3 \\ y_2 = x_2 + 2x_3 \\ y_3 = x_3 \end{cases}$，即 $\begin{cases} x_1 = y_1 - y_2 + y_3 \\ x_2 = y_2 - 2y_3 \\ x_3 = y_3 \end{cases}$，将 f 化成标准型

$f = y_1^2 + y_2^2$，所用变换矩阵为 $C = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$，$|C| = 1$

44

7.2 二次型

定理3 **主轴定理**

设$\mathbf{A}$是一个n阶对称阵，那么存在一个正交变换$\mathbf{x}=\mathbf{P}\mathbf{y}$，它将二次型$\mathbf{x}^T\mathbf{A}\mathbf{x}$变换成不含交叉项的二次型$\mathbf{y}^T\mathbf{D}\mathbf{y}$.

注:

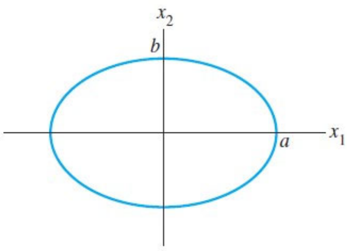
定理中矩阵$\mathbf{P}$的列称为二次型$\mathbf{x}^T\mathbf{A}\mathbf{x}$的主轴。向量$\mathbf{y}$是$\mathbf{x}$在这些主轴构造的$\mathbb{R}^n$空间的单位正交基下的坐标向量。

45

7.2 二次型

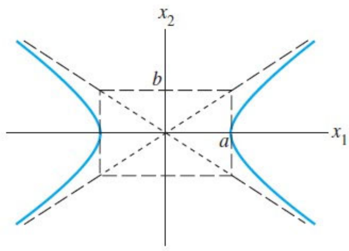
主轴的几何意义

若$\mathbf{A}$是对角阵，如下图所示$\mathbf{x}^T\mathbf{A}\mathbf{x}=\mathbf{c}$的图像是标准位置。



$\frac{x_1^2}{a^2}+\frac{x_2^2}{b^2}=1,\ a>b>0$

ellipse



$\frac{x_1^2}{a^2}-\frac{x_2^2}{b^2}=1,\ a>b>0$

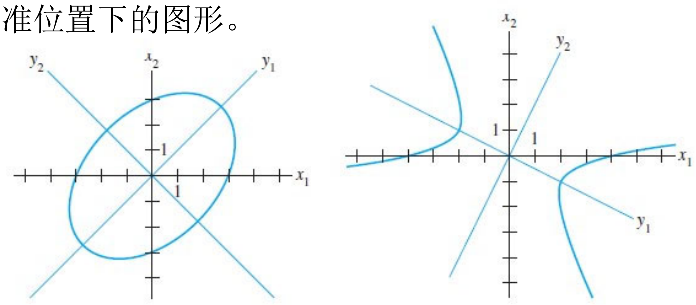
hyperbola

46

7.2 二次型

主轴的几何意义

若$\mathbf{A}$是不是对角阵，如下图所示$\mathbf{x}^T\mathbf{A}\mathbf{x}=\mathbf{c}$的图像是标准位置的旋转。找到主轴（由A的特征向量确定）等同于找到一个新的坐标系统，在该坐标系统下其图形是在标准位置下的图形。



(a) $5x_1^2-4x_1x_2+5x_2^2=48$ (b) $x_1^2-8x_1x_2-5x_2^2=16$

FIGURE 3 An ellipse and a hyperbola *not* in standard position.

47

7.2 二次型

例7 将上图椭圆方程$5x_1^2-4x_1x_2+5x_2^2=48$ 中的交叉项消去。

二次型对应的矩阵为$\mathbf{A}=\begin{bmatrix}5 & -2\\ -2 & 5\end{bmatrix}$, $\mathbf{A}$ 的特征值是3 和7，对应的单位特征向量为

$$u_1=\begin{bmatrix}1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix},\quad u_2=\begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$$

令$\mathbf{P}=[u_1 \quad u_2]=\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\ 1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$，那么$\mathbf{P}$可将$\mathbf{A}$正交对角化，所以变量代换$\mathbf{x}=\mathbf{P}\mathbf{y}$得到的二次型为

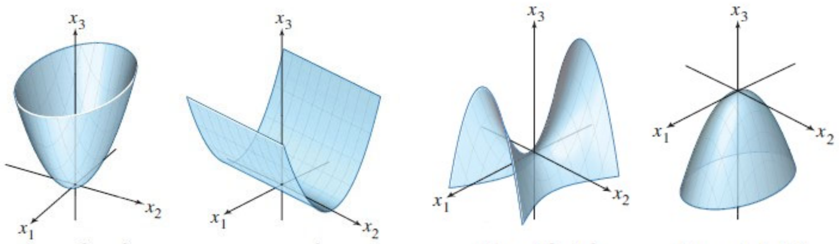
$$\mathbf{y}^T\mathbf{D}\mathbf{y}=3y_1^2+7y_2^2$$

48

7.2 二次型

二次型的分类

当$\mathbf{A}$是一个$n\times n$矩阵时，二次型$Q(\mathbf{x})=\mathbf{x}^T\mathbf{A}\mathbf{x}$是一个定义域为$\mathbb{R}^n$的实值函数，对二次型$Q(\mathbf{x})$定义域中的每一个点对应$x=(x_1,x_2)$可以画出点$(x_1,x_2,z)$，其中$z=Q(\mathbf{x})$.



(a) $z=3x_1^2+7x_2^2$ (b) $z=3x_1^2$ (c) $z=3x_1^2-7x_2^2$ (d) $z=-3x_1^2-7x_2^2$

7.2 二次型

定义

一个二次型Q是：

(1) 正定的，如果对所有的$\mathbf{x}\neq 0$，有$Q(\mathbf{x})>0$.

(2) 负定的，如果对所有的$\mathbf{x}\neq 0$，有$Q(\mathbf{x})<0$.

(3) 不定的，如果$Q(\mathbf{x})$既有正值又有负值。

Q被称为半正定的，如果对所有$\mathbf{x}$，$Q(\mathbf{x})\geq 0$;

Q被称为半负定的，如果对所有$\mathbf{x}$，$Q(\mathbf{x})\leq 0$.

7.2 二次型

定理4

设**A**是n 阶对称矩阵，那么一个二次型是：

(1) 正定，当且仅当**A** 的所有特征值都是正数；

(2) 负定，当且仅当**A** 的所有特征值都是负数；

(3) 不定，当且仅当**A** 既有正特征值，又有负特征值。

由主轴定理，存在一个正交变换$x = Py$, 使得

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

此处$\lambda_1 \cdots \lambda_n$是**A** 的特征值，由于$P$ 是可逆的，非零向量**x** 和非零向量**y**之间存在一个一一映射，这样，$\mathbf{x} \neq 0$ 时，$Q(\mathbf{x})$ 的值与上式右边的表达式的值完全对应。显然，像定理所描述的三类方式一样，它由特征值$\lambda_1 \cdots \lambda_n$的符号所确定。

51

7.2 二次型

例8 $Q(\mathbf{x}) = 3x_1^2 + x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$是正定的吗？

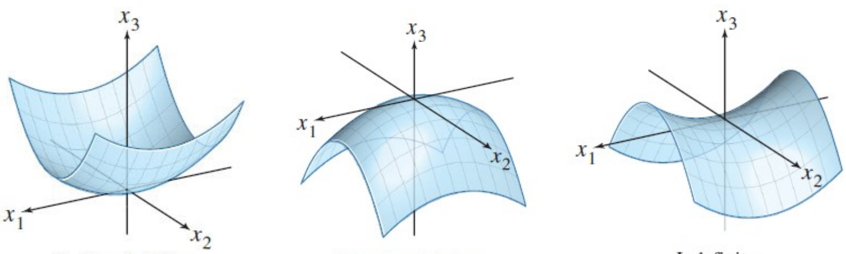
由于所有项的系数是正数，二次型表面上看是正定的。但二次型对应的矩阵为

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

且**A**的特征值是5,2和-1，所以$Q(\mathbf{x})$是不定二次型，而不是正定二次型。

52

7.2 二次型



Positive definite　　Negative definite　　Indefinite

数值计算的注解

确定对称矩阵 **A** 是正定的最快的方式，是尝试将矩阵**A**分解为 $\mathbf{A}=\mathbf{R}^T\mathbf{R}$，此处 **R** 是具有正对角线元素的上三角阵（可以采用一个略微修改的*LU* 分解算法），这样的乔雷斯分解算法可行充分必要条件是为**A** 是正定的。

53

7.2 二次型

推论 对称矩阵 A 为正定的充分必要条件是 A 的特征值全为正。

定理6

对称矩阵 A 为正定的充分必要条件是：A 的各阶主子式都为正，

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

为负定的充分必要条件是

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

赫尔维茨定理

7.2 二次型

例9 判定二次型$f=-5x^2-6y^2-4z^2+4xy+4xz$的正定性。

解 f矩阵为

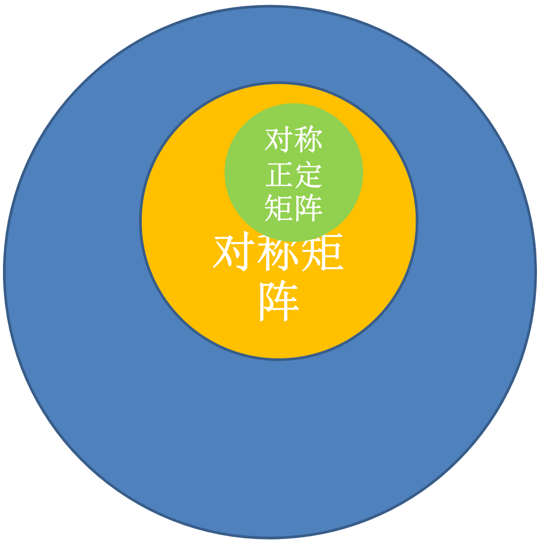
$$A=\begin{bmatrix}-5 & 2 & 2\\ 2 & -6 & 0\\ 2 & 0 & -4\end{bmatrix}$$

$$a_{11}=-5<0,\quad \begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}=\begin{vmatrix}-5 & 2\\ 2 & -6\end{vmatrix}=26>0,\quad |A|=-80<0$$

由定理可知f负定。

55

7.2 二次型



矩阵

⇩

对称矩阵

⇩

对称正定矩阵

56

7.2 二次型

如果 S 和 T 是对称正定矩阵，S + T 也是对称正定矩阵。

测试：对于所有非零向量，$x^T(S+T)x = x^TSx + x^TTx > 0$

如果一个对称矩阵 S 满足以下任一性质，则满足其它全部性质

(1) S 矩阵所有的主元为正数

(2) S 矩阵所有 n 个左上行列式为正数

(3) S 矩阵所有 n 个特征值为正数

(4) 针对所有非零向量，x^TSx 为正数

(5) $S = A^TA$，A 的所有列向量线性无关

判断方法

57

7.2 二次型

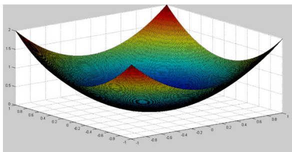
$$A = \begin{pmatrix} 5 & 2 \\ 2 & 3 \end{pmatrix} \longrightarrow \textbf{正定}$$

$$A = \begin{pmatrix} 2 & 6 \\ 6 & ? \end{pmatrix} \xrightarrow{? = 19} \textbf{正定}$$

58

7.2 二次型

函数 $F(x, y)$ 是否有最小值？



一阶偏导数 → $\frac{\partial F}{\partial x} = 0$? $\frac{\partial F}{\partial y} = 0$?

二阶偏导数 → $S = \begin{pmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} \end{pmatrix}$ 正定？

n 个自变量 → nxn 矩阵

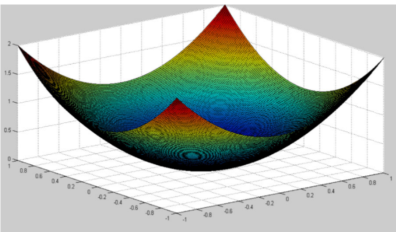
59

7.2 二次型

$F(x, y) = x^2 + y^2$

解：$\frac{\partial F}{\partial x} = 2x = 0, \frac{\partial F}{\partial y} = 2y = 0 \Rightarrow (x, y) = (0, 0)$

$$S = \begin{pmatrix} \frac{\partial^2 F}{\partial x^2} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$



$\Rightarrow$ S 正定，函数有最小值

60

7.2 二次型

正定二次型

二次型的**标准形显然不是唯一的**，只是标准形中所含项数是确定的，不仅如此，在限定变换为实变换时，标准形中正系数的个数是不变的，也就有接下来的定理。

Next

惯性定理

61

7.2 二次型

正定二次型

定理5

设二次型 $f = \mathrm{x}^T A\mathrm{x}$ 的秩是 r，且有两个可逆变换 $\mathrm{x} = C\mathrm{y}$ 及 $\mathrm{x} = P\mathrm{z}$ 使 $f = k_1 y_1^2 + k_2 y_2^2 + \cdots + k_r y_r^2 \ (k_i \neq 0)$ 及 $f = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_r z_r^2 \ (\lambda_i \neq 0)$ 则 k_i 中的正数个数与 λ_i 中的正数个数相同

惯性定理

正系数的个数 → **正惯性指数**

负系数的个数 → **负惯性指数**

62

Q & A