第六章 实对称矩阵

主要内容

6.1 实对称矩阵的谱分解

6.2 正定矩阵

实对称矩阵

6.3 奇异值分解

一般实矩阵

实对称矩阵、奇异值分解等内容在实际生产生活中有重要应用

6.1 实对称矩阵的谱分解

主要内容：

(a) 实对称矩阵的谱分解定理

(b) 实对称矩阵正交对角化的方法

(a) 实对称矩阵的谱定理:

1. 实对称矩阵的特征值都是实数

2. 实对称矩阵可以由**正交矩阵**对角化, 即有正交矩阵Q, 使得 $Q^{-1}SQ = \Lambda$, Λ是实对角矩阵.

   由于正交矩阵Q满足$Q^{-1} = Q^T$, 我们可写为$Q^TSQ = \Lambda$

因此, 实对称矩阵的特征值是实数，特征向量由$\mathbb{R}^n$的一组标准正交基构成.

1. 实对称矩阵的特征值是实数:

设$\lambda \in \mathbb{C}$，$\boldsymbol{x} \neq \boldsymbol{0} \in \mathbb{C}^n$满足$S\boldsymbol{x} = \lambda\boldsymbol{x}$ $\Rightarrow$ $\overline{\boldsymbol{x}}^T S\boldsymbol{x} = \lambda\overline{\boldsymbol{x}}^T \boldsymbol{x}$

对$\overline{\boldsymbol{x}}^T S\boldsymbol{x} = \lambda\overline{\boldsymbol{x}}^T \boldsymbol{x}$ 取共轭转置，由于$\bar{S} = S$,得到$\overline{\boldsymbol{x}}^T S\boldsymbol{x} = \bar{\lambda}\overline{\boldsymbol{x}}^T \boldsymbol{x}$

于是，$\lambda\overline{\boldsymbol{x}}^T \boldsymbol{x} = \bar{\lambda}\overline{\boldsymbol{x}}^T \boldsymbol{x}$.

由于$\overline{\boldsymbol{x}}^T \boldsymbol{x} = |x_1|^2 + \cdots + |x_n|^2 > 0$,，我们有$\bar{\lambda} = \lambda$，即$\lambda \in \mathbb{R}$.

解实系数方程$(\lambda I - S)\boldsymbol{x} = \boldsymbol{0}$可知$S$一定有实特征向量

2. 实对称矩阵可由正交矩阵对角化:

用归纳法。令S为n阶实对称，$(\lambda_1, \boldsymbol{q}_1)$为$S$的一个实特征对，且$\|\boldsymbol{q}_1\| = 1$.

令 $\boldsymbol{q}_2, \dots, \boldsymbol{q}_n$为$\mathcal{N}(\boldsymbol{q}_1^T)$的一组标准正交基。于是$\boldsymbol{q}_1, \dots, \boldsymbol{q}_n$构成$\mathbb{R}^n$的一组标准正交基。

令$Q = [\boldsymbol{q}_1, \dots, \boldsymbol{q}_n]$,

$$SQ = S[\boldsymbol{q}_1, \dots, \boldsymbol{q}_n] = [\boldsymbol{q}_1, \dots, \boldsymbol{q}_n]\begin{bmatrix} \lambda_1 & \boldsymbol{b}^T \\ \boldsymbol{0} & S_1 \end{bmatrix}$$

因此, $Q^T SQ = \begin{bmatrix} \lambda_1 & \boldsymbol{b}^T \\ \boldsymbol{0} & S_1 \end{bmatrix}$.

由于$Q^T SQ$为对称矩阵，得到$\boldsymbol{b} = \boldsymbol{0}$且$S_1$为对称矩阵。

于是， $Q^T SQ = \begin{bmatrix} \lambda_1 & \boldsymbol{0}^T \\ \boldsymbol{0} & S_1 \end{bmatrix}$.

根据归纳假设，有$n-1$阶正交矩阵Q_1使得$Q_1^T S_1 Q_1 = \begin{bmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

令$Q^{new} = Q\begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix}$。

Q^{new}为正交矩阵：

$$(Q^{new})^T Q^{new} = \begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix}^T Q^T Q \begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix} = \begin{bmatrix} 1 & \\ & Q_1^T Q_1 \end{bmatrix} = I$$

且$(Q^{new})^T S Q^{new} = \begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix}^T Q^T S Q \begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix} =$
$\begin{bmatrix} 1 & \\ & Q_1^T \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & S_1 \end{bmatrix} \begin{bmatrix} 1 & \\ & Q_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & Q_1^T S_1 Q_1 \end{bmatrix} = \Lambda.$

实对称矩阵谱分解:

由$Q^T S Q = \Lambda$得到$SQ = Q\Lambda$,

即S的特征向量可由$\mathbb{R}^n$的一组标准正交基构成。

进一步，$S = Q\Lambda Q^{-1} = Q\Lambda Q^T$,即$Q = [\boldsymbol{q}_1, \dots, \boldsymbol{q}_n]$, 那么

$$S = [\boldsymbol{q}_1, \dots, \boldsymbol{q}_n]\Lambda[\boldsymbol{q}_1, \dots, \boldsymbol{q}_n]^T = \lambda_1 \boldsymbol{q}_1 \boldsymbol{q}_1^T + \cdots + \lambda_n \boldsymbol{q}_n \boldsymbol{q}_n^T$$

这也称为S的**谱分解**。

正交相似:

对实方阵A, B，如果存在正交矩阵Q使得$Q^T AQ = B$，则称A, B**正交相似**

例如：实对称矩阵正交相似于对角阵。

命题：实方阵的正交相似关系是等价关系。

正交相似关系是一种特殊的相似关系。因此，正交相似关系下的不变量有特征值，代数重数，几何重数

另外，对称性也是正交相似下的不变量。

实际上，两个实对称矩阵正交相似当且仅当特征多项式相同（习题）

(b) 实对称矩阵正交对角化的方法

命题：实对称矩阵属于不同特征值的特征向量互相正交。

即如果λ_1, λ_2为对称矩阵S两个不同的特征值，$\boldsymbol{x}_1, \boldsymbol{x}_2$分别为属于特征值$\lambda_1, \lambda_2$的特征向量，则$\boldsymbol{x}_1 \perp \boldsymbol{x}_2$.

证明:

由$S\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1$，我们有$\boldsymbol{x}_2^T S\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_2^T\boldsymbol{x}_1$;

由$S\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_2$，我们有$\boldsymbol{x}_1^T S\boldsymbol{x}_2 = \lambda_2\boldsymbol{x}_1^T\boldsymbol{x}_2$.

由于S是对称矩阵且$\boldsymbol{x}_1^T S\boldsymbol{x}_2$为一个数，我们有

$\boldsymbol{x}_1^T S\boldsymbol{x}_2 = \left(\boldsymbol{x}_1^T S\boldsymbol{x}_2\right)^T = \boldsymbol{x}_2^T S^T\boldsymbol{x}_1 = \boldsymbol{x}_2^T S\boldsymbol{x}_1$.

因此，$\lambda_2\boldsymbol{x}_1^T\boldsymbol{x}_2 = \lambda_1\boldsymbol{x}_2^T\boldsymbol{x}_1 = \lambda_1\boldsymbol{x}_1^T\boldsymbol{x}_2$.

由于$\lambda_1 \neq \lambda_2$,必有$\boldsymbol{x}_1^T\boldsymbol{x}_2 = 0$，即$\boldsymbol{x}_1 \perp \boldsymbol{x}_2$

根据上述命题，对称矩阵S的对角化可按如下步骤进行:

(1) 求S的特征多项式，特征值。假设互不相同的特征值为$\lambda_1, \dots, \lambda_k$，代数重数分别为$n_1, \dots, n_k$.

由于S可对角化，代数重数等于几何重数。

(2) 对每个λ_i，求特征子空间$\mathcal{N}(\lambda_i I - S)$的一组基，记为$\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{in_i}$

(3) 利用Gram-Schmidt正交化将$\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{in_i}$化为正交单位向量组$\boldsymbol{q}_{i1}, \dots, \boldsymbol{q}_{in_i}$

(4) 记$Q = \left[\boldsymbol{q}_{11}, \dots, \boldsymbol{q}_{1n_1}, \dots, \boldsymbol{q}_{k1}, \dots, \boldsymbol{q}_{kn_k}\right]$,

那么$SQ = Q\Lambda$, 其中$\Lambda = Diag(\lambda_1, \dots, \lambda_1, \dots, \lambda_k, \dots, \lambda_k)$.

例: 正交对角化$S=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & -1\\ 1 & -1 & 0\end{bmatrix}$

特征多项式$p_S(\lambda)=(\lambda-1)^2(\lambda+2)$。

特征值为1（重数2），-2（重数1）。

解方程$(-2I-S)\boldsymbol{x}=\boldsymbol{0}$，得到$\boldsymbol{x}_1=[1,-1,-1]^T$; $\boldsymbol{q}_1=\left[1/\sqrt{3},-1/\sqrt{3}-1/\sqrt{3}\right]^T$;

解方程$(I-S)\boldsymbol{x}=\boldsymbol{0}$，得到$\boldsymbol{x}_2=[1,1,0]^T$, $\boldsymbol{x}_3=[1,0,1]^T$;

对$\boldsymbol{x}_2,\boldsymbol{x}_3$进行Gram-Schmidt正交化得到

$$\boldsymbol{q}_2=\left[1/\sqrt{2},1/\sqrt{2},0\right]^T.$$

$$\boldsymbol{x}_3'=\boldsymbol{x}_3-\boldsymbol{q}_2\boldsymbol{q}_2^T\boldsymbol{x}_3=[1/2,-1/2,1]^T$$

$$\boldsymbol{q}_3=\frac{\boldsymbol{x}_3'}{\|\boldsymbol{x}_3'\|}=[1/\sqrt{6},-1/\sqrt{6},\sqrt{\frac{2}{3}}].$$

令$Q=[\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3]=\begin{bmatrix}1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6}\\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\\ -1/\sqrt{3} & 0 & \sqrt{\frac{2}{3}}\end{bmatrix}$

$$S=Q\begin{bmatrix}-2 & & \\ & 1 & \\ & & 1\end{bmatrix}Q^T.$$

本节小结:

1. 实对称矩阵的谱定理:

   特征值都是实数

   可由正交矩阵对角化, 即有正交矩阵Q, 使得

   $Q^{-1}SQ = Q^TSQ = \Lambda$, Λ是实对角矩阵.

2. 对角化实对称矩阵的方法:

   属于不同特征值的特征向量互相正交

   对同一特征值对应的特征子空间进行Gram-Schmidt正交化

6.2 正定矩阵

主要内容：

(a) 正定矩阵的定义

(b) 正定实对称矩阵判定

(c) Rayleigh商与特征值

(d) 合同标准形与惯性定理

(a) 正定矩阵的定义:

实矩阵A称为**正定**，如果$\boldsymbol{x}^T A\boldsymbol{x} > 0$对所有非零向量$\boldsymbol{x} \in \mathbb{R}^n$.

A称为**半正定**，如果$\boldsymbol{x}^T A\boldsymbol{x} \geq 0$对所有$\boldsymbol{x} \in \mathbb{R}^n$.

$\boldsymbol{x}^T A\boldsymbol{x}$称为关于$\boldsymbol{x}$的能量函数。

类似定义，实矩阵A称为**负定**，如果$\boldsymbol{x}^T A\boldsymbol{x} < 0$对所有非零向量$\boldsymbol{x} \in \mathbb{R}^n$.

A称为**半负定**，如果$\boldsymbol{x}^T A\boldsymbol{x} \leq 0$对所有$x \in \mathbb{R}^n$.

A称为**不定**，如果A即不正定、半正定也不负定、半负定。

正定、半正定等条件常用在实对称矩阵上。

例:

如果$D = Diag(d_1, \dots, d_n)$为对角阵。$D$何时正定?

$$\boldsymbol{x}^T D \boldsymbol{x} = d_1 \boldsymbol{x}_1^2 + \cdots + d_n \boldsymbol{x}_n^2.$$

D为正定当且仅当$d_1, d_2, \dots, d_n > 0$.

例:

设$A \in M_{m\times n}(\mathbb{R})$。从$A$出发可以得到对称矩阵$S = A^TA$（与$AA^T$）。

$S = A^TA$是半正定的实对称矩阵。

$$\boldsymbol{x}^T S\boldsymbol{x} = \boldsymbol{x}^T A^T A\boldsymbol{x} = \|A\boldsymbol{x}\|^2 \geq 0,$$

因此，$S = A^TA$是半正定的实对称矩阵。

S正定 $\Leftrightarrow$ 对任意非零向量$\boldsymbol{x}$, $A\boldsymbol{x} \neq \boldsymbol{0}$

$\Leftrightarrow$ A列满秩

在6.3节，半正定对称矩阵A^TA在对于A的研究中发挥重要作用

问题：为什么考虑正定性，以及半正定、负定、半负定?

回顾一元函数极值问题： 如何判断$f(x)$在x_0取局部极小值?

不妨设$f(x_0) = 0$

$f(x)$在驻点x_0取极小值如果$f''(x_0) > 0$.

例:

$\mathbb{R}^3$中二次曲面$F(x,y) = ax^2 + 2bxy + cy^2$。考查$F(x,y)$在(0,0)点的极值问题。

$\frac{\partial F}{\partial x} = 2ax + 2by, \frac{\partial F}{\partial y} = 2bx + 2cy$

在(0,0)点，$F(0,0) = 0$, 且$\left.\frac{\partial F}{\partial x}\right|_{(0,0)} = \left.\frac{\partial F}{\partial y}\right|_{(0,0)} = 0$, (0,0)点称为$F(x,y)$的驻点。

问题：什么条件保证$F(x,y)$在(0,0)处取得最小值?

令$S = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$为对称矩阵，将$F(x,y)$写成矩阵形式$F(x,y) = [x,y]\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{v}^T S \boldsymbol{v}$

$F(x,y)$在(0,0)处取最小即是要求$\boldsymbol{v}^T S \boldsymbol{v} \geq 0$，$\forall \boldsymbol{v} \in \mathbb{R}^2$.满足条件的$S$称为半正定矩阵

(0,0)是唯一的最小值点当且仅当$\boldsymbol{v}^T S \boldsymbol{v} > 0, \forall \boldsymbol{v} \neq \boldsymbol{0} \in \mathbb{R}^2$。满足条件的$S$称为正定。

$F(x,y)$在$(0,0)$点处的二阶导数给出的矩阵（Hesse矩阵）为

$$\begin{bmatrix} \frac{\partial^2 F}{\partial x \partial x} & \frac{\partial^2 F}{\partial x \partial y} \\ \frac{\partial^2 F}{\partial y \partial x} & \frac{\partial^2 F}{\partial y \partial y} \end{bmatrix} = \begin{bmatrix} 2a & 2b \\ 2b & 2c \end{bmatrix} = 2S$$

因此，S的正定性推广了一元函数时的极值条件$f''(x_0) > 0$.

合同:

对方阵A, B，如果存在可逆矩阵X使得$X^T AX = B$，则称A, B**合同。**

方阵的合同关系是等价关系。正交相似是一种特殊的合同。

合同变换下的不变量：实矩阵的正定，半正定，负定，半负定，不定

命题:

如果A为正定矩阵，X可逆，那么$X^T AX$为正定矩阵。

证明：设$\boldsymbol{x} \in \mathbb{R}^n$为非零向量。$\boldsymbol{x}^T \left(X^T AX\right) \boldsymbol{x} = (X\boldsymbol{x})^T A (X\boldsymbol{x})$.

由于X可逆，$X\boldsymbol{x} \neq \boldsymbol{0}$. 由于$A$正定，$(X\boldsymbol{x})^T A (X\boldsymbol{x}) > 0$.

于是, $\boldsymbol{x}^T \left(X^T AX\right) \boldsymbol{x} = (X\boldsymbol{x})^T A (X\boldsymbol{x}) > 0$. 得到$X^T AX$的正定性。

(b) 正定实对称矩阵判定

回顾顺序主子阵:

方阵A左上角$k \times k$子块A_k称为A的第k个顺序主子阵。

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix} \quad A_1 = [1] \quad A_2 = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \quad A_3 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 9 & 7 \end{bmatrix}$$

回顾：A有LU分解当且仅当$A_1, A_2 \dots, A_n$可逆

A_k的行列式$\det A_k$称为A的第k个**顺序主子式。**

正定实对称矩阵的判定定理:

对于实对称矩阵S，以下陈述等价：

(1) S正定，即对所有非零向量$\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}^T S \boldsymbol{x} > 0$（能量判定）

(2) S的所有特征值> 0.（特征值判定）

(3) 存在可逆矩阵A使得$S = A^T A$.

(4) S有LDL^T分解，且D的对角元均为正数。（主元判定）

回顾：L为单位下三角矩阵，D为对角阵，对角线元素为S的主元。

(5) S的n个顺序主子式均> 0.（顺序主子式判定）

注意: (3)中写法不唯一

证明采用: $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (1)$

(1)对所有非零向量$\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}^T S \boldsymbol{x} > 0 \Rightarrow$ (2) S的所有特征值> 0。

设$\lambda \in \mathbb{R}$是特征值，$\boldsymbol{x} \in \mathbb{R}^n$为属于$\lambda$的特征向量，由$S\boldsymbol{x} = \lambda\boldsymbol{x}$得到

$$\boldsymbol{x}^T S \boldsymbol{x} = \lambda \boldsymbol{x}^T \boldsymbol{x}$$

因此 $\lambda = \frac{\boldsymbol{x}^T S \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}} > 0$ ⟸ $\frac{\boldsymbol{x}^T S \boldsymbol{x}}{\boldsymbol{x}^T \boldsymbol{x}}$称为$\boldsymbol{x}$对$S$的Rayleigh商

(2)S的所有特征值$> 0 \Rightarrow$ (3)存在可逆矩阵A使得$S = A^T A$

$$S = Q \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} Q^T, \lambda_i > 0. \quad \text{令} A = \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix} Q^T$$

$$S = A^T A$$

(3)存在可逆矩阵A使得$S = A^TA \Rightarrow$ (4)S有LDL^T分解且D对角元素> 0

令$S = A^TA$，A可逆。令$A = QR$为A的QR分解。

$S = R^TQ^TQR = R^TR.$ Cholesky分解

R^T为下三角矩阵，令 $R^T = LD_1, D_1$为可逆对角矩阵。

则$A = LD_1D_1^TL^T$. 令$D = D_1D_1^T = D_1^2, D$的对角元$> 0.$

(4)S有LDL^T分解, D的对角元$> 0. \Rightarrow$ (5)S的n个顺序主子式均> 0

假设$S = LDL^T$，D的对角元> 0。记$L = \begin{bmatrix} L_k & \\ C & L' \end{bmatrix}, D = \begin{bmatrix} D_k & \\ & D' \end{bmatrix}$.

$$S = \begin{bmatrix} L_k & \\ C & L' \end{bmatrix} \begin{bmatrix} D_k & \\ & D' \end{bmatrix} \begin{bmatrix} L_k^T & C^T \\ & L'^T \end{bmatrix} = \begin{bmatrix} L_kD_kL_k^T & L_kD_kC^T \\ CD_kL_k^T & CD_kC^T + L'D'L'^T \end{bmatrix}$$

S的第k个顺序主子阵$S_k = L_kD_kL_k^T$.

因此，$|S_k| = |D_k| > 0.$

(5)S的n个顺序主子式均$> 0 \Rightarrow$ (1)$\boldsymbol{x}^T S\boldsymbol{x} > 0, \forall \boldsymbol{x} \neq \boldsymbol{0}$

由S的n个顺序主子式均> 0知，S的n个顺序主子阵可逆，

因此S有LU分解，由于S对称，$S = LDL^T$

$$S = \begin{bmatrix} L_k & \\ C & L' \end{bmatrix} \begin{bmatrix} D_k & \\ & D' \end{bmatrix} \begin{bmatrix} L_k^T & C^T \\ & L'^T \end{bmatrix} = \begin{bmatrix} L_k D_k L_k^T & L_k D_k C^T \\ C D_k L_k^T & C D_k C^T + L' D' L'^T \end{bmatrix}$$

由于S的n个顺序主子式> 0，D的对角元均> 0，

于是D是正定矩阵。$S = LDL^T$也是正定矩阵（正定性在合同下不变）。

注记:

主元　　　特征值

如果是S正定实对称矩阵，那么S有两个分解

$S = LDL^T = Q\Lambda Q^T$.

D, Λ均为对角矩阵且对角线元素均> 0.

这个等式连接了这门课程两个重要概念：主元和特征值！

例: $S=\begin{bmatrix}1 & a & b\\ a & 1 & c\\ b & c & 1\end{bmatrix}$何时正定?

利用主子式判别: $\begin{vmatrix}1 & a\\ a & 1\end{vmatrix}>0$，得到$1-a^2>0$；

$|S|>0$,得到$1+2abc>a^2+b^2+c^2$。

因此，两个条件$1-a^2>0$，$1+2abc>a^2+b^2+c^2$，保证S正定。

例: 如果S,T为正定对称矩阵，那么$S+T$也为正定对称矩阵。

证明：对非零向量$\boldsymbol{x}$, $\boldsymbol{x}^TS\boldsymbol{x}>0, \boldsymbol{x}^TT\boldsymbol{x}>0$，得到$\boldsymbol{x}^T(S+T)\boldsymbol{x}>0.$

定义：A一个k阶主子式是指去掉$i_1, \ldots, i_{n-k}$行与$i_1, \ldots, i_{n-k}$列剩下方阵的行列式。

半正定对称矩阵的判定定理:

对于对称实矩阵，下列叙述等价:

(1)$\forall \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x}^T S \boldsymbol{x} \geq 0$;

(2)S的特征值均≥ 0;

(3)$S = A^T A$;

(4)S所有主元非负;

(5)S的所有主子式均≥ 0;

注意：在半正定下需要考虑所有主子式，而不仅仅是顺序主子式。

例如：$\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$半负定，顺序主子式$\geq 0$

小结:

1. 正定实对称矩阵的判定

   五个等价判定性质

2. 对称性、正定性在合同变换下保持不变