《线性代数》

第七章 线性映射与线性变换

§7.4 实对称阵的对角化

2019秋

杨晶 主讲

内容提要

- 实对称阵的特征值与特征向量的性质
- 实对称阵正交对角化的结论
- 实对称阵正交对角化的方法
- 从内积到正定矩阵

以下哪些是矩阵A可对角化的充分必要条件?

A有n个互异的特征值.

A有n个线性无关的特征向量.

对A的每个特征值λ_i, 均有其几何重数等于代数重数.

在某组基下, A所对应的线性变换σ把每个基均做伸缩变换.

提交

引言：

矩阵(线性变换)的相似对角化问题, 是本课程中一个非常重要的理论问题, 并且在很多实际问题中有应用. 我们在上一讲中给出了n阶方阵可对角化的充要条件, 即

A可对角化 $\Longleftrightarrow$ A有n个线性无关的特征向量.

$\Longleftrightarrow$ 对A的每个特征值λ_i, 均有其几何重数等于代数重数.

然而, 上述两个条件都不太直观. 我们的问题是:

问题： 是否有一批n阶方阵, 我们可以从其直观形式上就可判断它们可对角化? 而无须经过计算才能给出判断.

回答: 有, 例如: 实对称矩阵.

- 对于一般n阶方阵而言, 是否可对角化是一个不容易判断的问题.
- 一个实矩阵要想在实数范围内对角化, 首先特征值必须都是实数, 其次它要有n个线性无关的特征向量.
- 但是, 并非任何一个实矩阵都有实的特征值, 从而不一定可以 (在实数范围内) 对角化.
- n阶(实)对称矩阵的概念, 我们在第三章中已经介绍过, 即n阶方阵 $A = (a_{ij})_{n\times n}$, s.t. $a_{ij} \in \mathbb{R}$, 且$A^{\mathrm{T}} = A$.
- 本节主要的内容就是要说明:

　　任意实对称阵A在实数范围内可对角化, 不仅如此, 还能找到一个正交阵Q, 使得$Q^{-1}AQ$ 为对角阵, 即实对称阵可正交对角化.

　　这就要求实对称阵的特征值均为实数, 并且特征向量可构成一组标准正交基.

(一) 实对称阵的特征值与特征向量的性质

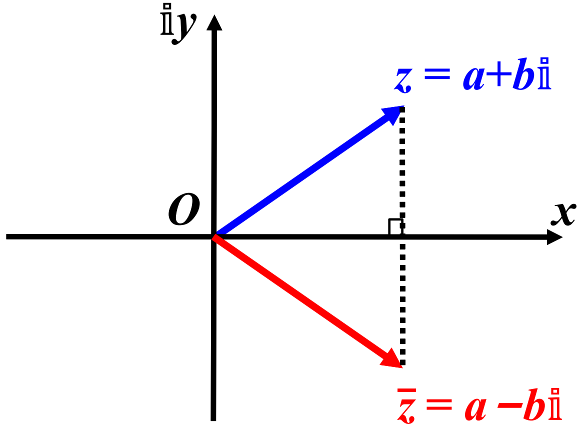
为说明实对称阵的性质, 先简要介绍一些关于复数的基本概念和结论.

定义1 设复数$z = a+b\mathrm{i}$, $(a,b\in\mathbb{R},\ \mathrm{i} = \sqrt{-1}\ )$,

则 z 的复共轭运算记为$\bar{z}$, 定义为:

$$\bar{z} := a - b\mathrm{i}$$

(实部不变, 虚部取反.)



容易验证, 复共轭运算有如下性质:

- 与复数的四则运算(加减乘除) 相容 (可交换运算次序), 即

$$\bar{z} \pm \bar{w} = \overline{z \pm w},\ \ \bar{z} \cdot \bar{w} = \overline{z \cdot w},\ {}^{\bar{z}}/_{\bar{w}} = \overline{({}^{z}/_{w})}\ ;$$

- 正定性: $|z|^2 := z \cdot \bar{z} = a^2 + b^2 \geq 0$, 且等号成立当且仅当 $z = 0$.

进一步, 我们可复数的共轭运算拓展为复矩阵的共轭运算.

定义2 若$A=(a_{ij})_{m\times n}$, 其中$a_{ij}\in\mathbb{C}$, 则称A为$m\times n$型的复矩阵.

复矩阵的**共轭(conjugate)**定义为: $\overline{A}:=\left(\overline{a_{ij}}\right)_{m\times n}$.

根据定义及共轭复数的**运算性质**，容易证明共轭矩阵有以下性质

(1) $\overline{\overline{A}}=A, \overline{A}^{\mathrm{T}}=\overline{A^{\mathrm{T}}}$;

(2) $\overline{kA}=\overline{k}\,\overline{A}\quad(k\in\mathbb{C})$;

(3) $\overline{A}\pm\overline{B}=\overline{A\pm B}$;

(4) $\overline{A}\,\overline{B}=\overline{AB}$;

(5) 若A可逆, 则$\overline{A^{-1}}=\left(\overline{A}\right)^{-1}$;

(6) $\det\overline{A}=\overline{\det A}$;

(7) $A\in M_n(\mathbb{C})$为实对称阵 $\Leftrightarrow$ $A^{\mathrm{T}}=A=\overline{A}$.

有了以上理论准备, 我们来看实对称阵的特征值与特征向量的性质.

定理1 设 A 是 n 阶实对称阵, 则 A 的特征值都是实数.

证明 由A必有复特征值, 设复数 λ 是 A 的特征值, 复向量 $\vec{X} \neq \vec{0}$ 是对应的特征向量, 即$A\vec{X} = \lambda\vec{X}$, 则两边取共轭得 $A\overline{\vec{X}} = \overline{A}\,\overline{\vec{X}} = \overline{A\vec{X}} = \overline{\lambda\vec{X}} = \overline{\lambda}\,\overline{\vec{X}}$. (1)

而对定义式两边左乘$\overline{\vec{X}}^{\mathrm{T}}$, 得 $\overline{\vec{X}}^{\mathrm{T}} A\vec{X} = \overline{\vec{X}}^{\mathrm{T}}\left(A\vec{X}\right) = \overline{\vec{X}}^{\mathrm{T}}\lambda\vec{X} = \lambda\overline{\vec{X}}^{\mathrm{T}}\vec{X}$, (2)

另一方面, 上式中由结合律又得

$$\overline{\vec{X}}^{T} A\vec{X} = \left(\overline{\vec{X}}^{T} A\right)\vec{X} = \left(\overline{\vec{X}}^{T} A^{T}\right)\vec{X} = \left(\overline{A}\,\overline{\vec{X}}\right)^{T}\vec{X} = \left(\overline{\lambda}\,\overline{\vec{X}}\right)^{T}\vec{X} = \overline{\lambda}\,\overline{\vec{X}}^{T}\vec{X}. \qquad (3)$$

(2), (3) 两式相减, 得 $0 = \left(\lambda - \overline{\lambda}\right)\overline{\vec{X}}^{T}\vec{X}$.

$$\because\ \vec{X} \neq \vec{0},\quad \overline{\vec{X}}^{T}\vec{X} = \sum_{i=1}^{n}\overline{x}_i x_i = \sum_{i=1}^{n}\left|x_i\right|^2 \neq 0,\quad \Rightarrow \lambda - \overline{\lambda} = 0.$$

所以 λ 为实数. ■

注: (1) 未必所有的实矩阵对应的特征值都是实数, 例如:

$$A=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}, \quad \Rightarrow \quad |A-\lambda I|=\begin{vmatrix}0-\lambda & -1\\ 1 & 0-\lambda\end{vmatrix}=\lambda^2+1, \text{ 从而}, \lambda=\pm\mathbb{i}.$$

定理1说明, 在特征值方面实对称阵有更好的性质.

注: (2) 因实对称矩阵A的特征值λ_i为实数, 故齐次线性方程组

$$(\lambda_i I-A)\vec{X}=\vec{0}$$

为实系数方程组, 又由$|\lambda_i I-A|=0$, 该齐次线性方程组一定有实的基础解系, 从而属于λ_i的全体特征向量均为实向量, 即 $V_{\lambda_i}\subset\mathbb{R}^n$.

定理2 实对称矩阵A的属于不同特征值的特征向量是正交的.

证明 设λ_1, λ_2是实对称矩阵A的两个不同的特征值, 而$\vec{X}_1, \vec{X}_2$分别为与之对应的特征向量, 则 $A\vec{X}_1 = \lambda_1 \vec{X}_1,\ A\vec{X}_2 = \lambda_2 \vec{X}_2,\ (\lambda_1 \neq \lambda_2).$

λ_2定义式两边左乘$\vec{X}_1^{\mathrm{T}}$, 得:

$$\vec{X}_1^T A\vec{X}_2 = \lambda_2 \vec{X}_1^T \vec{X}_2. \tag{4}$$

λ_1定义式两边转置, 再右乘$\vec{X}_2$, 注意到$A^{\mathrm{T}} = A$, 得:

$$(A\vec{X}_1)^{\mathrm{T}} \vec{X}_2 = \vec{X}_1^{\mathrm{T}} A\vec{X}_2 = \lambda_1 \vec{X}_1^T \vec{X}_2. \tag{5}$$

(4), (5)两式相减, 得 $\left(\lambda_1 - \lambda_2\right) \cdot \vec{X}_1^T \vec{X}_2 = 0.$

由 $\lambda_1 \neq \lambda_2$, 得 $(\vec{X}_1, \vec{X}_2) = \vec{X}_1^T \vec{X}_2 = 0.$

即$\vec{X}_1$与$\vec{X}_2$正交. ■

例如: 在 § 6.4节中的例3(1)中, $A=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}$ 为实对称阵.

经过计算, 我们得到A的特征值为: $\lambda_1=1$ (重数为2); $\lambda_2=4$.

而属于特征值$\lambda_1=1$的两个线性无关的特征向量为:

$$\vec{X}_{11}=(-1,1,0)^{\mathrm{T}};\ \vec{X}_{12}=(-1,0,1)^{\mathrm{T}}.$$

属于特征值$\lambda_2=4$的一个线性无关的特征向量为: $\vec{X}_{21}=(1,1,1)^{\mathrm{T}}$.

容易验证: $\left(\vec{X}_{11},\vec{X}_{21}\right)=-1+1+0=0,\left(\vec{X}_{12},\vec{X}_{21}\right)=-1+0+1=0.$

即: $\vec{X}_{11}\perp\vec{X}_{21},\vec{X}_{12}\perp\vec{X}_{21},\quad\Rightarrow V_{\lambda_1}\perp V_{\lambda_2}.$

推论1 设λ_i,λ_j为实对称阵A的两个互异的特征值, 则 $V_{\lambda_i}\perp V_{\lambda_j}$.

(二) 实对称阵正交对角化的结论

相较于一般矩阵, 实对称阵在特征值与特征向量方面有较好的性质. 于是, 根据特征值、特征向量与矩阵的对角化的密切关系, 我们可以得到本节最主要的结论, 如下:

定理3 对n阶实对称阵A, 总存在正交阵Q, 使得$Q^{-1}AQ=Q^{T}AQ$为对角阵.

证明 对A的阶数n作归纳法.

当n=1时, 对$A=(a_{11})$, $Q = I_1= (1)$, 结论明显成立.

假定结论对 $n-1$成立, 下证 n 的情形.

由定理1, A的所有特征值均为实数, 所有特征向量均为实向量, 所以, 可在欧氏空间$\mathbb{R}^n$中讨论问题. 可设$\lambda_1\in\mathbb{R}$为A的一个实特征值, 而$\vec{X}_1$是属于λ_1的一个单位特征向量. 可以把$\vec{X}_1$扩充成为$\mathbb{R}^n$的一组标准正交基:

$$\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n .$$

令$Q_1=(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n)$, 则$Q_1$为正交矩阵, 作运算

$$AQ_1 = (A\vec{X}_1, A\vec{X}_2, \cdots, A\vec{X}_n) = (\lambda_1\vec{X}_1, A\vec{X}_2, \cdots, A\vec{X}_n)$$

$$= (\vec{X}_1, \vec{X}_2, \cdots, \vec{X}_n)\begin{pmatrix} \lambda_1 & \boxed{\begin{matrix} * \\ * \end{matrix}} \\ \vec{0} & \end{pmatrix} = Q_1\begin{pmatrix} \lambda_1 & \boxed{\begin{matrix} \vec{\alpha}^{\mathrm{T}} \\ A_1 \end{matrix}} \\ \vec{0} & \end{pmatrix}.$$

其中, $\vec{\alpha}^{\mathrm{T}}$为某个$n-1$维行向量, A_1为某个$n-1$阶方阵, 因而得到

$$Q_1^{-1}AQ_1 = Q_1^{\mathrm{T}}AQ_1 = \begin{pmatrix} \lambda_1 & \vec{\alpha}^T \\ \vec{0} & A_1 \end{pmatrix}$$

注意到Q_1为正交阵, $Q_1^{-1} = Q_1^{\mathrm{T}}$, 故上式左边为对称阵, 从而右边也是, 故得$\vec{\alpha}^{\mathrm{T}}$为零向量, A_1为$n-1$阶对称阵. 于是, A_1符合归纳假设, 即存在$n-1$阶正交阵Q_2, 使得

$$Q_2^{\mathrm{T}}A_1Q_2 = \Lambda_1 = \mathrm{diag}(\lambda_2, \ldots, \lambda_n).$$

令 $Q=Q_1\begin{pmatrix}1 & O\\ O & Q_2\end{pmatrix}$, 则由$Q_1$, Q_2均为正交阵知, Q也为正交阵. 因此得

$$Q^{\mathrm{T}}AQ=\begin{pmatrix}1 & O\\ O & Q_2^{\mathrm{T}}\end{pmatrix}Q_1^{\mathrm{T}}AQ_1\begin{pmatrix}1 & O\\ O & Q_2\end{pmatrix}$$

$$=\begin{pmatrix}1 & O\\ O & Q_2^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\lambda_1 & O\\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O\\ O & Q_2\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & O\\ O & Q_2^{\mathrm{T}}A_1Q_2\end{pmatrix}=\begin{pmatrix}\lambda_1 & O\\ O & \Lambda_1\end{pmatrix}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

从而, 结论对n的情形也成立. 我们就完成了定理3的证明. ■

注: 我们对定理3的证明是存在性的, 而非构造性的, 即并没有直接给出对角阵Λ与正交阵Q.

对角阵Λ可由全体特征值决定, 那么, 正交阵Q应该如何确定呢?

当然, 需要从特征向量入手. 由矩阵可对角化的充要条件可知.

推论2 n阶实对称阵必有n个线性无关的特征向量, 且对A的全体互异的特征值$\lambda_1, \lambda_2, \dots, \lambda_s$有, $\sum_{i=1}^{s} m_i = \sum_{i=1}^{s} \dim V_{\lambda_i} = n$.

进一步, 可得实对称阵在特征向量方面的特性.

推论3 n阶实对称阵有n个相互正交的单位特征向量.

证明 由推论2, 取定n个线性无关的特征向量. 在每个特征子空间V_{λ_i}内, 对已有的m_i个线性无关的特征向量做Schimdt正交化, 得V_{λ_i}的单位正交基; 再将上述全体V_{λ_i}的单位正交基并起来, 由推论1 ($V_{\lambda_i} \perp V_{\lambda_j}$)知, 上述所得的即为实对称阵的$n$个相互正交的单位特征向量. ■

(三) 实对称阵正交对角化的方法

Step 1. 分解特征多项式, 求解A的特征值.

设$f_A(\lambda)=(\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s}$, 则$\lambda_1,\lambda_2,\cdots,\lambda_s$为$A$的互异的实特征值, 其代数重数分别为$n_1,n_2,\cdots,n_s$.

Step 2. 对每个特征值λ_i $(1\le i\le s)$计算特征向量. 即求解齐次线性方程组$(\lambda_i I-A)\vec{X}=\vec{0}$, 得到一组基础解系$\vec{X}_{i1},\vec{X}_{i2},\ldots,\vec{X}_{im_i}$, 其中$m_i$为$\lambda_i$的几何重数, 且必有$m_i=n_i$.

Step 3. 求变换矩阵. 若没有“正交”的要求, 则令.

$$P=(\vec{X}_{11},\ldots,\vec{X}_{1n_1},\cdots,\vec{X}_{s1},\ldots,\vec{X}_{sn_s}),$$
$$\Lambda=P^{-1}AP=\mathrm{diag}(\underbrace{\lambda_1,\cdots,\lambda_1}_{n_1\text{个}},\cdots,\underbrace{\lambda_s,\cdots,\lambda_s}_{n_s\text{个}}).$$

若有“正交”的要求, 则分别对每一组特征向量$\vec{X}_{i1},\ldots,\vec{X}_{in_i}$, 用施密特正交化方法进行正交单位化, 得到$\vec{\varepsilon}_{i1},\ldots,\vec{\varepsilon}_{in_i}$, 进而得

正交矩阵 $Q=(\vec{\varepsilon}_{11},\ldots,\vec{\varepsilon}_{1n_1},\cdots,\vec{\varepsilon}_{s1},\ldots,\vec{\varepsilon}_{sn_s})$; $\Lambda=Q^{\mathrm{T}}AQ$ 同前.

例1 设 $A=\begin{bmatrix}0&1&1&-1\\1&0&-1&1\\1&-1&0&1\\-1&1&1&0\end{bmatrix}$, 用两种方法将其对角化.

解 首先求特征值与代数重数 $|\lambda I-A|=\begin{vmatrix}\lambda&-1&-1&1\\-1&\lambda&1&-1\\-1&1&\lambda&-1\\1&-1&-1&\lambda\end{vmatrix}=(\lambda-1)^3(\lambda+3).$

得到 $\lambda_1=1,n_1=3;\lambda_2=-3,n_2=1.$

其次, 对特征值 $\lambda_1=1$, 解齐次线性方程组 $(\lambda_1 I-A)\vec{X}=\vec{0}$, 求特征向量.

$\begin{pmatrix}1&-1&-1&1\\-1&1&1&-1\\-1&1&1&-1\\1&-1&-1&1\end{pmatrix}\to\begin{pmatrix}1&-1&-1&1\\0&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}$, 得基础解系为: $\vec{X}_{11}=(1,1,0,0)^T$; $\vec{X}_{12}=(1,0,1,0)^T$; $\vec{X}_{13}=(-1,0,0,1)^T$.

再次, 对特征值 $\lambda_2=-3$, 解齐次线性方程组 $(\lambda_2 I-A)\vec{X}=\vec{0}$, 求特征向量.

$$\begin{pmatrix} -3 & -1 & -1 & 1 \\ -1 & -3 & 1 & -1 \\ -1 & 1 & -3 & -1 \\ 1 & -1 & -1 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

得基础解系为: $\vec{X}_{21}=(1,-1,-1,1)^T$.

于是, 给出第一种将A对角化的方法为:

令 $P=(\vec{X}_{11},\vec{X}_{12},\vec{X}_{13};\vec{X}_{21})$

$$=\begin{pmatrix} 1 & 1 & -1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

从而 $\Lambda=P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_1,\lambda_1;\lambda_2)$

$$=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

第二种方法为正交对角化的方法, 分别对两组基础解系做Schimdt正交化方法:

$$\vec{\beta}_{11}=\vec{X}_{11}=(1,1,0,0)^T,\qquad \vec{\varepsilon}_{11}=\frac{\vec{\beta}_{11}}{|\vec{\beta}_{11}|}=\frac{1}{\sqrt{2}}(1,1,0,0)^T,$$

$$\vec{\beta}_{12}=\vec{X}_{12}-\frac{(\vec{X}_{12},\vec{\beta}_{11})}{(\vec{\beta}_{11},\vec{\beta}_{11})}\vec{\beta}_{11}=(\frac{1}{2},-\frac{1}{2},1,0)^T,\qquad \vec{\varepsilon}_{12}=\frac{\vec{\beta}_{12}}{|\vec{\beta}_{12}|}=\frac{1}{\sqrt{6}}(1,-1,2,0)^T,$$

$$\vec{\beta}_{13}=\vec{X}_{13}-\frac{(\vec{X}_{13},\vec{\beta}_{11})}{(\vec{\beta}_{11},\vec{\beta}_{11})}\vec{\beta}_{11}-\frac{(\vec{X}_{13},\vec{\beta}_{12})}{(\vec{\beta}_{12},\vec{\beta}_{12})}\vec{\beta}_{12}=(-\frac{1}{3},\frac{1}{3},\frac{1}{3},1)^T;\qquad \vec{\varepsilon}_{13}=\frac{\vec{\beta}_{13}}{|\vec{\beta}_{13}|}=\frac{1}{2\sqrt{3}}(-1,1,1,3)^T.$$

$$\vec{\beta}_{21}=\vec{X}_{21}=(1,-1,-1,1)^T.\qquad \vec{\varepsilon}_{21}=\frac{\vec{\beta}_{21}}{|\vec{\beta}_{21}|}=\frac{1}{2}(1,-1,-1,1)^T.$$

分别做局部正交化
计算量优于
做整体正交化

将刚刚分别局部正交化所得的单位正交向量组 $\vec{\varepsilon}_{11}, \vec{\varepsilon}_{12}, \vec{\varepsilon}_{13}$ 与 $\vec{\varepsilon}_{21}$ 按列并起来就得到了正交矩阵

$$Q=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{-1}{2\sqrt{3}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2} \\ 0 & 0 & \frac{3}{2\sqrt{3}} & \frac{1}{2} \end{bmatrix},$$

且 $\Lambda = Q^{-1}AQ = Q^{T}AQ$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

正交阵Q是唯一的吗?

例2 已知A为3阶实对称阵,其特征值为1, 1, 2, 且属于2 的特征向量是 $(1, 0, 1)^T$, 求A = ?

分析: 已知特征值与部分特征向量, 求实对称阵A.

由A可正交对角化, 即$Q^TAQ = \Lambda$, $\Rightarrow A=Q\Lambda Q^T$, 其中$\Lambda$与$Q$分别可由$A$的特征值与特征向量来确定, 而剩余的特征向量可由正交性确定.

解 A是3阶实对称阵, 正交相似于对角阵Λ= diag(1, 1, 2). 属于特征值1的特征向量与属于特征值2 的特征向量 $(1, 0, 1)^T$正交, 这等价于求解齐次线性方程组: $(1, 0, 1)(x_1, x_2, x_3)^T= 0$. 求解得到属于1的特征向量为: $(0,1,0)^T$, $(1,0,-1)^T$, 且彼此正交. 故得到相应的正交矩阵为

$$Q = \begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 0 & 0 \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, \quad \Rightarrow A = Q\Lambda Q^T =$$

课后练习

1. 采用正交阵Q可避免矩阵求逆;
2. 只有两个特征值时, 特征子空间V_{λ_1}与V_{λ_2}互为正交补.

(四) 从内积到正定矩阵

回顾: 在欧氏空间$V(\mathbb{R})$中, 设有内积$(\vec{\alpha},\vec{\beta})$. 设 $\vec{\alpha}_1, \vec{\alpha}_2,\ldots, \vec{\alpha}_n$ 是 V 的一组基. 则对

$\vec{\alpha},\vec{\beta}\in V$, 可设

$$\vec{\alpha} = x_1\vec{\alpha}_1+\ldots+x_n\vec{\alpha}_n= (\vec{\alpha}_1, \vec{\alpha}_2,\ldots, \vec{\alpha}_n)\vec{x},\ \ \vec{\beta} = y_1\vec{\alpha}_1+\ldots+y_n\vec{\alpha}_n= (\vec{\alpha}_1, \vec{\alpha}_2,\ldots, \vec{\alpha}_n)\vec{y}$$

并计算

$$(\vec{\alpha},\vec{\beta}) = \left(\sum_{i=1}^{n} x_i\vec{\alpha}_i, \sum_{j=1}^{n} y_j\vec{\alpha}_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}(\vec{\alpha}_i,\vec{\alpha}_j)x_i y_j$$

$$=(x_1,x_2,\cdots,x_n)\begin{bmatrix} (\vec{\alpha}_1,\vec{\alpha}_1) & (\vec{\alpha}_1,\vec{\alpha}_2) & \cdots & (\vec{\alpha}_1,\vec{\alpha}_n) \\ (\vec{\alpha}_2,\vec{\alpha}_1) & (\vec{\alpha}_2,\vec{\alpha}_2) & \cdots & (\vec{\alpha}_2,\vec{\alpha}_n) \\ \cdots & \cdots & \cdots & \cdots \\ (\vec{\alpha}_n,\vec{\alpha}_1) & (\vec{\alpha}_n,\vec{\alpha}_2) & \cdots & (\vec{\alpha}_n,\vec{\alpha}_n) \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \triangleq \vec{x}^{\mathrm{T}} A\vec{y}$$

抽象内积的具体坐标表示.

其中, n阶实方阵A称为该内积的度量矩阵. 由内积的对称性知, A为实对称阵.
另外, 容易验证: $\vec{x}^{\mathrm{T}}A\vec{y}$对于列向量$\vec{x},\vec{y}\in\mathbb{R}^n$满足双线性.

问题1： 内积定义中的第四条，即**正定性**，如何用坐标与度量矩阵表示？

回顾: 内积的正定性—— $(\vec{\alpha},\vec{\alpha})\geq 0$，且等号成立当且仅当 $\vec{\alpha}=\vec{0}$.

定义3. 设A为实对称矩阵.

(1) 若 $\vec{x}^{\mathrm{T}}A\vec{x}>0,\forall\vec{x}\neq\vec{0}$，则称$A$为正定矩阵(**positive definite matrix**);

(2) 若 $\vec{x}^{\mathrm{T}}A\vec{x}<0,\forall\vec{x}\neq\vec{0}$，则称$A$为负定矩阵(**negative definite matrix**);

(3) 若 $\vec{x}^{\mathrm{T}}A\vec{x}\geq 0,\forall\vec{x}\neq\vec{0}$，则称$A$为半正定矩阵(**positive semidefinite matrix**);

(4) 若 $\vec{x}^{\mathrm{T}}A\vec{x}\leq 0,\forall\vec{x}\neq\vec{0}$，则称$A$为半负定矩阵(**negative semidefinite matrix**);

(5) 若存在 $\vec{\alpha}\neq\vec{0},\vec{\beta}\neq\vec{0}$，使得 $\vec{\alpha}^{\mathrm{T}}A\vec{\alpha}>0,\vec{\beta}^{\mathrm{T}}A\vec{\beta}<0$，则称$A$为不定矩阵.

问题2：如何选择空间V中合适的一组基, 使得在该基下内积的坐标表示较为简单?

设$\vec{\beta}_1,\dots,\vec{\beta}_n$是$V(\mathbb{R})$的另一组基，并令$\vec{\alpha}_1,\dots,\vec{\alpha}_n$到$\vec{\beta}_1,\dots,\vec{\beta}_n$的过渡矩阵为$P$, 即 $(\vec{\beta}_1,\dots,\vec{\beta}_n)=(\vec{\alpha}_1,\dots,\vec{\alpha}_n)P$. 于是, 在新基下的度量矩阵为:

$$B=\left[\left(\vec{\beta}_i,\vec{\beta}_j\right)\right]_{1\le i,j\le n}$$

代入得,

$$\left(\vec{\beta}_i,\vec{\beta}_j\right)=\left(\sum_{s=1}^{n}p_{si}\vec{\alpha}_s,\sum_{t=1}^{n}p_{tj}\vec{\alpha}_t\right)=\sum_{s=1}^{n}\sum_{t=1}^{n}p_{si}(\vec{\alpha}_s,\vec{\alpha}_t)p_{tj}=\sum_{s=1}^{n}\sum_{t=1}^{n}p_{si}a_{st}p_{tj}\quad\Rightarrow\quad B=P^{\mathrm{T}}AP$$

于是在新基下, 内积$(\vec{\alpha},\vec{\beta})$的坐标表示为 $\vec{x}'^{T}\left(P^{T}AP\right)\vec{y}'=(P\vec{x}')^{\mathrm{T}}A\,(P\vec{y}')=\vec{x}^{\mathrm{T}}A\vec{y}$.

定义4 给定两个 n 阶对称阵A和B, 如果存在可逆方阵 P, 使得 $B=P^{\mathrm{T}}AP$, 则称 B 与 A **合同(相合, congruent).**

- **课后验证:** 相合是n阶对称矩阵的一种等价关系.

问题2：如何选择空间V中合适的一组基, 使得在该基下内积的坐标表示较为简单?

等价的问题：如何选取恰当的可逆阵P, 使得P^TAP形式较为简单?

回答: 对角阵, 即相合对角化问题.

实际上, 利用正交矩阵$Q^{-1}=Q^T$, 我们已经有结论:

定理3' n维欧氏空间$V(\mathbb{R})$中, 设在某组基下内积对应的度量矩阵为A, 则存在V的一组标准正交基, 使得在该基下, 内积运算$(\vec{\alpha},\vec{\beta})$的坐标表示为:

$$\lambda_1 x_1' y_1' + \cdots + \lambda_n x_n' y_n'$$

其中, $\lambda_1, \cdots, \lambda_n$为$A$的特征值, $(x_1', \cdots x_n')^T$, $(y_1', \cdots y_n')^T$,分别为$\vec{\alpha}$与$\vec{\beta}$的新坐标.

注：相合对角化问题还有其他计算方法, 以及相合意义下的不变量与标准形的相关结论, 参见中文教材第七章. (下学期详细分解)

单选题 1分

设置

用来表示内积的实对称阵需要是正定的, 然而并不是所有实对称阵都是正定的, 例如:

例3. 设n阶矩阵$A=\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, $r<n$, 则A的正定性为:

正定

半正定

负定

半负定

不定

提交

性质1. 相合变换, 不改变实对称阵的正定性.

证明. 令$\vec{x}=P\vec{y}$, P可逆, 则

$$\vec{x}\neq\vec{0}\Leftrightarrow\vec{y}\neq\vec{0}$$

$$\vec{x}^{\mathrm{T}}A\vec{x}=\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}.$$

故

$\vec{x}^{\mathrm{T}}A\vec{x}$正定 $\Leftrightarrow\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}$正定;

$\vec{x}^{\mathrm{T}}A\vec{x}$半正定 $\Leftrightarrow\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}$半正定;

$\vec{x}^{\mathrm{T}}A\vec{x}$负定 $\Leftrightarrow\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}$负定;

$\vec{x}^{\mathrm{T}}A\vec{x}$半负定 $\Leftrightarrow\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}$半负定;

$\vec{x}^{\mathrm{T}}A\vec{x}$不定 $\Leftrightarrow\vec{y}^{\mathrm{T}}(P^{\mathrm{T}}AP)\vec{y}$不定.

即

A正定 $\Leftrightarrow P^{\mathrm{T}}AP$正定;

A半正定 $\Leftrightarrow P^{\mathrm{T}}AP$半正定;

A负定 $\Leftrightarrow P^{\mathrm{T}}AP$负定;

A半负定 $\Leftrightarrow P^{\mathrm{T}}AP$半负定;

A不定 $\Leftrightarrow P^{\mathrm{T}}AP$不定. □

对实对称阵$\boldsymbol{A}$,存在正交矩阵$\boldsymbol{Q}$, $s.t.$

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{Q}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n).$$

令$\vec{\boldsymbol{x}}=\boldsymbol{Q}\vec{\boldsymbol{y}}$,则

$$\vec{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}\vec{\boldsymbol{x}}=\vec{\boldsymbol{y}}^{\mathrm{T}}\boldsymbol{\Lambda}\vec{\boldsymbol{y}}=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2.$$

性质2. 设A为实对称阵,则

A正定 $\Leftrightarrow$ A的特征值全>0;

A半正定 $\Leftrightarrow$ A的特征值全$\geq \mathbf{0}$;

A负定 $\Leftrightarrow$ A的特征值全$< \mathbf{0}$;

A半负定 $\Leftrightarrow$ A的特征值全$\leq \mathbf{0}$;

A不定 $\Leftrightarrow$ A既有正特征值又有负特征值. □

单选题 1分

设置

例4. 判断实对称阵$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -3\\ 0 & -3 & 1\end{pmatrix}$的正定性.

正定

半正定

负定

半负定

不定

提交

例5. 已知$\boldsymbol{A}$为实对称矩阵，$\lambda_1,\lambda_2,\cdots,\lambda_n$是它的特征值，问$t$为何值时，$\boldsymbol{A}+t\boldsymbol{I}$为负定矩阵.

解：$\boldsymbol{A}$为实对称矩阵，则$\boldsymbol{A}+t\boldsymbol{I}$也为实对称矩阵，于是

$$\boldsymbol{A}+t\boldsymbol{I}\text{为负定矩阵} \Leftrightarrow \boldsymbol{A}+t\boldsymbol{I}\text{的特征值都}<0.$$

已知$\boldsymbol{A}$的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，则$\boldsymbol{A}+t\boldsymbol{I}$的特征值为

$$t+\lambda_1, t+\lambda_2, \cdots, t+\lambda_n.$$

故$t+\lambda_i<0, i=1,2,\cdots,n$，即

$$t<\min\{-\lambda_1,-\lambda_2,\cdots,-\lambda_n\}$$

时$\boldsymbol{A}+t\boldsymbol{I}$为负定矩阵.□

性质3. 实对称阵A正定 $\Leftrightarrow$ 存在可逆矩阵P, $s.t. A = P^{\mathrm{T}}P$.

性质4. 实对称阵A正定 $\Leftrightarrow$ A与I相合.

性质5. 实对称阵A正定 $\Rightarrow$ $\det A > 0$.

半正定，（半）负定如何？

证明：只证(3)：对实对称阵A, 存在正交矩阵Q, $s.t.$

$$A = Q\Lambda Q^{\mathrm{T}} = Q\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}^{\mathrm{T}}\underbrace{\begin{bmatrix}\sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n}\end{bmatrix}Q^{\mathrm{T}}}_{P} = P^{T}P$$

有定二次型的判断

特征值的正负是判断有定二次型的一个重要准则, 但求全体特征值涉及到n阶含参行列式的计算与多项式求根(因式分解), 有时比较复杂. 是否有其它判断有定二次型的判则?

首先, 需给出如下(顺序)主子式的概念.

定义5. 设$A=(a_{ij})\in M_n$,称子式

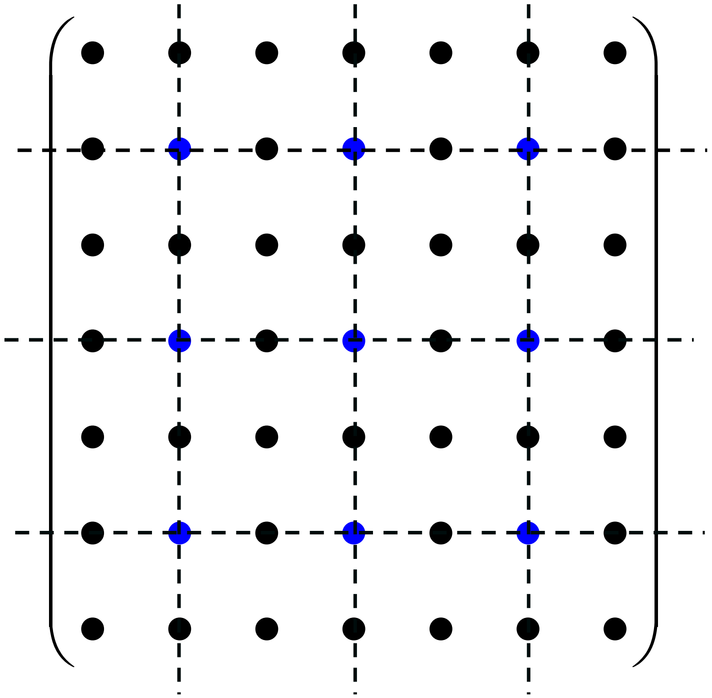
$$P_i=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{21} & a_{22} & \cdots & a_{2i} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{vmatrix} \quad (i=1,2,\cdots,n)$$

为矩阵A的i 阶顺序主子式(Sequential Principal Minor).

回忆：设$A=(a_{ij})\in M_n$,若A的i阶子式的行指标与列指标相同,则称该子式为A的一个i阶主子式.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

顺序主子式



主子式

定理4 (1) 实二次型 $\vec{x}^{\mathrm{T}}A\vec{x}$ 正定 $\Longleftrightarrow$ (2) A的各阶顺序主子式 $P_i > 0 \Longleftrightarrow$ (3) A的各阶主子式 > 0.

证明思路：(2)是(3)的子情形, 则必有(3)⇒(2),下只证(1) $\Longleftrightarrow$ (2).
(1) $\Longleftrightarrow$ (3)的证明留作习题(见教材习题7.28).

必要性 设 n 元实二次型正定, 如果令$\vec{X}_i = (x_1,\ldots, x_i, 0,\ldots, 0)^{\mathrm{T}}$, 则 $\vec{X}_i^{\mathrm{T}}A\vec{X}_i$ 相当于一个 i 元正定二次型, 其矩阵的行列式正好是 P_i, 由正定二次型性质6知道$P_i > 0$, $i = 1, 2,\ldots, n$.

充分性 若对所有i, 有$P_i > 0$, 对A的阶数用归纳法.
当n=1时，$A=(a_{11})$， $a_{11}>0$, 则$x_1a_{11}x_1=a_{11}x_1^2>0$.
假设$n-1$时命题成立, 下考虑n个变元的二次型, 令:

$$A = \begin{pmatrix} A_{n-1} & \vec{\beta} \\ \vec{\beta}^{\mathrm{T}} & a_{nn} \end{pmatrix}$$

于是A_{n-1}的各阶主子式均大于0, 由归纳假设A_{n-1}正定.

由性质4, 存在$n-1$阶可逆阵C_{n-1}, 使得

$$C_{n-1}^{\mathrm{T}}A_{n-1}C_{n-1}=I_{n-1}$$

令 $C=\begin{pmatrix} C_{n-1} & \triangle \\ \vec{0}^{\mathrm{T}} & 1 \end{pmatrix}$, 于是有

$$C^{\mathrm{T}}AC=\begin{pmatrix} C_{n-1}^{\mathrm{T}} & \vec{0} \\ \triangle^{\mathrm{T}} & 1 \end{pmatrix}\begin{pmatrix} A_{n-1} & \vec{\beta} \\ \vec{\beta}^{\mathrm{T}} & a_{nn} \end{pmatrix}\begin{pmatrix} C_{n-1} & \triangle \\ \vec{0}^{\mathrm{T}} & 1 \end{pmatrix}$$

$$=\begin{pmatrix} C_{n-1}^{\mathrm{T}}A_{n-1}C_{n-1} & \vec{0} \\ \vec{0}^{T} & **** \end{pmatrix}=\begin{pmatrix} I_{n-1} & \vec{0} \\ \vec{0}^{\mathrm{T}} & **** \end{pmatrix}$$

$$\because (\det C)^2\det A>0 \Rightarrow d=a_{nn}-\vec{\beta}^{\mathrm{T}}A_{n-1}^{-1}\vec{\beta}>0$$

令$D=\begin{pmatrix} I_{n-1} & \vec{0} \\ \vec{0}^{\mathrm{T}} & 1/\sqrt{d} \end{pmatrix}C$, 则有$D^{\mathrm{T}}AD=I_n$, 由性质4知, A及其对应的二次型是正定的.

例6.判断下述矩阵

$$\vec{x}^{\mathrm{T}} A\vec{x} = 2x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + 2x_2x_3$$

的正定性.

解: 对应的矩阵为$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. 各阶顺序主子式为

$$P_1 = 2 > 0, P_2 = \begin{vmatrix} 2 & -1 \\ -1 & 2 \end{vmatrix} = 3 > 0, P_3 = \begin{vmatrix} 2 & -1 & 0 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{vmatrix} = 1 > 0.$$

因此二次型正定.□

例7. 求参数t的范围,使下列矩阵正定

$$\vec{x}^{\mathrm{T}}\boldsymbol{A}\vec{x}=2x_1^2+2x_2^2+2x_3^2-2tx_1x_2-2tx_1x_3-2tx_2x_3.$$

解: 对应的矩阵为$\boldsymbol{A}=\begin{pmatrix}2 & -t & -t\\ -t & 2 & -t\\ -t & -t & 2\end{pmatrix}.$

二次型正定 $\Leftrightarrow$ 各阶顺序主子式 >0

$$\Leftrightarrow\begin{cases}P_1=2>0,P_2=\begin{vmatrix}2 & -t\\ -t & 2\end{vmatrix}=4-t^2>0\\ P_3=\begin{vmatrix}2 & -t & -t\\ -t & 2 & -t\\ -t & -t & 2\end{vmatrix}=2(1-t)(2+t)^2>0\end{cases}\Leftrightarrow\begin{cases}4-t^2>0,\\ 1-t>0.\end{cases}$$

$$\Leftrightarrow -2<t<1.\square$$

Remark: A的顺序主子式都非负 $\not\Rightarrow$ 二次型$\vec{x}^{\mathrm{T}}A\vec{x}$半正定.

思考：*若用定理3的证明方法来推导，哪一个步骤会有问题？*

例8. $\vec{x}^{\mathrm{T}}A\vec{x}=x_1^2-x_3^2$是不定的,但$A=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}$的顺序主子式都非负：$P_1=1>0, P_2=P_3=0.$

定理5. 二次型$\vec{x}^{\mathrm{T}}A\vec{x}$半正定 $\Leftrightarrow$ A的各阶主子式$\geq 0,\ i=1,2,\cdots n.$

上例中$A=\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}$有一个1阶主子式$-1<0$,因此$A$不是半正定的.

定理4’. 二次型$\vec{x}^{\mathrm{T}}A\vec{x}$负定

$\Leftrightarrow$ 二次型$\vec{x}^{\mathrm{T}}(-A)\vec{x}$正定

$\Leftrightarrow -A$的各阶顺序主子式>0.

$\Leftrightarrow A$的所有奇数阶顺序主子式<0,

且A的所有偶数阶顺序主子式>0.

定理4”. 二次型$\vec{x}^{\mathrm{T}}A\vec{x}$半负定

$\Leftrightarrow$ 二次型$\vec{x}^{\mathrm{T}}(-A)\vec{x}$半正定

$\Leftrightarrow -A$的各阶主子式≥ 0.

$\Leftrightarrow A$的所有奇数阶主子式≤ 0,

且A的所有偶数阶主子式≥ 0.

例9. 判断下述$\vec{x}^T A\vec{x} = -2x_1^2 - 2x_2^2 - x_3^2 + 2x_1x_2 - 2x_2x_3$的正定性.

解: 对应的矩阵为$A = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & -1 \\ 0 & -1 & -1 \end{pmatrix}$. 其顺序主子式

$$P_1 = -2 < 0, P_2 = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = 3 > 0,$$

$$P_3 = \begin{vmatrix} -2 & 1 & 0 \\ 1 & -2 & -1 \\ 0 & -1 & -1 \end{vmatrix} = -1 < 0.$$

因此$\vec{x}^{\mathrm{T}} A\vec{x}$为负定二次型.□

例10. 设$\boldsymbol{A}\in M_{m\times n}(\mathbb{R}), \mathrm{r}(\boldsymbol{A})=n$, 证明$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$为正定阵.

证明: $\boldsymbol{A}\in M_{m\times n}(\mathbb{R}), \mathrm{r}(\boldsymbol{A})=n$, 则$\boldsymbol{A}$的列向量组线性无关. 于是

$\forall \vec{\boldsymbol{x}}\neq\vec{\boldsymbol{0}}(\in\mathbb{R}^n)$, 有$\boldsymbol{A}\vec{\boldsymbol{x}}\neq\vec{\boldsymbol{0}}$, 从而$\vec{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\vec{\boldsymbol{x}}=(\boldsymbol{A}\vec{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{A}\vec{\boldsymbol{x}}>0$, 即二次型$\vec{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\vec{\boldsymbol{x}}$正定, 因此$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$为正定阵. □

别忘了定义法判定正定性
这个结论很重要，下一讲将从这里讲起

本讲小结

- 实对称阵的特征值 ——$\lambda_1, \cdots, \lambda_n$必为实数;
- 实对称阵的特征向量 ——$\lambda_i \neq \lambda_j \Rightarrow V_{\lambda_i} \perp V_{\lambda_j}$;
- 实对称阵正交对角化的结论——$\exists$正交阵Q, $Q^{\mathrm{T}}AQ = \Lambda$;
- 实对称阵正交对角化的方法

——局部做施密特正交化

本讲小结　　实对称阵的正定性

▲ 概念

*实对称阵A是正定(半正定、负定、半负定)的：

$$\vec{x}^{\mathrm{T}} A\vec{x} > (\ge,<,\le)0,\quad \forall \vec{x} \ne \vec{0}$$

注：A为(半)负定矩阵 $\Longleftrightarrow$ $-A$为(半)正定矩阵

▲ 性质

1. 相合变换不改变二次型的正定性.
2. A正定$\Leftrightarrow A$的特征值都大于0.
3. A正定 $\Leftrightarrow$ A与I相合.
4. A正定$\Leftrightarrow$ $A=P^{\mathrm{T}}P$, 其中P可逆.
5. A正定$\Leftrightarrow$各阶顺序主子式>0 $\Leftrightarrow$所有主子式>0.

▲ 判定方法

1. 定义法.
2. 特征值法.
3. 相合标准形.
4. (顺序)主子式法.