第五章　二次型

王征宇

南京大学数学系

二次型——二次齐次多项式

定义 二次型是某个数域（理解为 R、C）上的二次齐次多项式

$$\begin{aligned}
f(x) &= b_{11}x_1x_1 + b_{12}x_1x_2 + \cdots + b_{1n}x_1x_n = x_1(b_{11}x_1 + \cdots + b_{1n}x_n) \\
&+ b_{21}x_2x_1 + b_{22}x_2x_2 + \cdots + b_{2n}x_2x_n = x_2(b_{21}x_1 + \cdots + b_{2n}x_n) \\
&\vdots \qquad \cdots + \cdots + \cdots \qquad\qquad \cdots + \cdots + \cdots \\
&+ b_{n1}x_nx_1 + b_{n2}x_nx_2 + \cdots + b_{nn}x_nx_n = x_n(b_{n1}x_1 + \cdots + b_{nn}x_n)
\end{aligned}$$

$$x = (x_1, x_2, \cdots, x_n)^T, \quad B = (b_{ij})_{n\times n}, \quad f(x) = x^TBx$$

$$\left.\begin{aligned}
f(x) = \; & a_{11}x_1^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\
& + a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\
& + a_{n-1,n-1}x_{n-1}^2 + 2a_{n-1,n}x_{n-1}x_n \\
& + a_{nn}x_n^2
\end{aligned}\right\} = x^TAx$$

$$\underbrace{\frac{1}{2}(b_{ij} + b_{ji})}_{=a_{ij}=a_{ji}}$$

$A = (a_{ij})_{n\times n}$ *symmetric*, $a_{ij} = a_{ji}$

二次型的例子（二次曲面）

二次曲面方程：

$$
\begin{aligned}
0 &= a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3 \\
&\quad -2c_1x_1 - 2c_2x_2 - 2c_3x_3 + d \\
&= \underbrace{\begin{aligned}&a_{11}(x_1-a_1)^2 + a_{22}(x_2-a_2)^2 + a_{33}(x_3-a_3)^2 \\ &+2a_{12}(x_1-a_1)(x_2-a_2) + 2a_{13}(x_1-a_1)(x_3-a_3) \\ &+2a_{23}(x_2-a_2)(x_3-a_3) + d' = y^TAy + d'\end{aligned}}_{translate\ \ y=x-a,\ \ a=(a_1,a_2,a_3)^T,\ \ A=(a_{ij})_{3\times 3}\ \ symmetric}
\end{aligned}
$$

以上曲面什么形状？此处假定以下方程组有解

$$
\overbrace{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}}^{=:c} = \begin{pmatrix} a_{11}a_1 + a_{12}a_2 + a_{13}a_3 \\ a_{22}a_2 + a_{12}a_1 + a_{23}a_3 \\ a_{33}a_3 + a_{13}a_1 + a_{23}a_2 \end{pmatrix} = \overbrace{\begin{pmatrix} a_{11}a_1 + a_{12}a_2 + a_{13}a_3 \\ a_{21}a_1 + a_{22}a_2 + a_{23}a_3 \\ a_{31}a_1 + a_{32}a_2 + a_{33}a_3 \end{pmatrix}}^{=:Aa,\ \ a=(a_1,a_2,a_3)^T}
$$

二次型的例子（Taylor 展式）

$y = g(x) = g(x_1, x_2, x_3)$ 在 $(0,0,0)$ 处的 Taylor 展开式：

$$g(y) = g(x) + \overbrace{\underbrace{g'(x)(y-x)}_{(y-x,\nabla g(x))}}^{linear\ \ form} + \frac{1}{2}\overbrace{\underbrace{(y-x)^T g''(x)(y-x)}_{(y-x,g''(x)(y-x))}}^{quadratic\ \ form} + r$$

$$g'(x) = \underbrace{\left(\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}, \frac{\partial g}{\partial x_3}\right)}_{g'(x)^T = \nabla g(x)\ \ gradient}, \quad g''(x) = \underbrace{\begin{pmatrix} \frac{\partial^2 g}{\partial x_1 \partial x_1} & \frac{\partial^2 g}{\partial x_1 \partial x_2} & \frac{\partial^2 g}{\partial x_1 \partial x_3} \\ \frac{\partial^2 g}{\partial x^2 \partial x_1} & \frac{\partial^2 g}{\partial x^2 \partial x_2} & \frac{\partial^2 g}{\partial x_2 \partial x_3} \\ \frac{\partial^2 g}{\partial x_3 \partial x_1} & \frac{\partial^2 g}{\partial x_3 \partial x_2} & \frac{\partial^2 g}{\partial x_3 \partial x_3} \end{pmatrix}}_{symmetric}$$

- 二次型的性质与函数的局部性质，$\cdots$

线性变换（替换）

定义 给定 $C=(c_{ij})_{n\times n}$，称如下的变换是从变量 $x_1,x_2,\cdots,x_n$ 到 $y_1,y_2,\cdots,y_n$ 的线性变换（替换）：

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

如果 C 可逆，则称此线性变换是非退化的；如果 C 正交，则称此线性变换是正交的。

用非退化线性替换，化简二次型为标准型（只有平方项）：

$$x^TAx=(Cy)^TA(Cy)=y^T\underbrace{\overbrace{(C^TAC)}^{=:D}}_{symmetric}y=\overbrace{d_1y_1^2+d_2y_2^2+\cdots+d_ny_n^2}^{D=diag(d_1,d_2,\cdots,d_n)\ \ diagonal}$$

矩阵的合同、二次型的化简

定义 称 $A=(a_{ij})_{n\times n}$合同于$B=(b_{ij})_{n\times n}$，如果存在可逆的 $C=(c_{ij})_{n\times n}$ 使得 $B=C^TAC$，称 C 是 A 到 B 的合同变换矩阵。

- 自反性：A 与 A 合同；
- 对称性：如果 A 与 B 合同，则 B 与 A 合同；
- 传递性：如果 A 与 B 合同，B 与 C 合同，则 A 与 B 合同

$$\begin{array}{rcl} \text{quadratic form } x^TAx & \sim & A \text{ symmetric} \\ X^TAx = y^TBy \ (x = Cy) & \sim & B = C^TAC \\ y^TDy = d_1y_1^2 + \cdots + d_ny_n^2 & \sim & D \text{ diagonal} \end{array}$$

$$\Downarrow$$

$$\underbrace{B = C^TAC \text{ diagonal}}_{\text{find } C \text{ invertible}} \quad v.s. \quad \underbrace{x^TAx = y^TBy \text{ normal}}_{\text{find } C \text{ invertible}, \ x=Cy}$$

实对称矩阵的正交合同对角化（怎么做？）

$$\overbrace{diag(\underbrace{\lambda_1,\cdots,\lambda_1}_{m_1},\ \underbrace{\lambda_2,\cdots,\lambda_2}_{m_2},\ \cdots,\ \underbrace{\lambda_k,\cdots,\lambda_k}_{m_k})}^{=P^TAP=P^{-1}AP=D,\ A\ symmetric,\ P\ ortho-normal}$$

- 计算对称 $A=(a_{ij})_{n\times n}$ 的互异（实）特征值：$\lambda_1,\cdots,\lambda_k$；
- 对每一个 λ_j，计算 $(\lambda_j E-A)x=0$ 基本解组

$$\underbrace{\xi_{j,1},\ \cdots,\ \xi_{j,m_j}}_{m_j=\ algebraic\ index}\ \ \overset{orthonormalization}{\to\cdots\to\cdots\to}\ \ \underbrace{\eta_{j,1},\ \cdots,\ \eta_{j,m_j}}_{orthonormal}$$

$$\underbrace{(\overbrace{\eta_{1,1},\ \cdots,\ \eta_{1,m_1}}^{w.r.t.\ \lambda_1},\ \overbrace{\eta_{2,1},\ \cdots,\ \eta_{2,m_2}}^{w.r.t.\ \lambda_2},\ \cdots\cdots,\ \overbrace{\eta_{k,1},\ \cdots,\ \eta_{j,m_k}}^{w.r.t.\ \lambda_k})}_{=:P\ square\ matrix\ \ (m_1+\cdots+m_k=n),\ \ P^TP=E}$$

实对称矩阵的正交合同对角化（续，概要）

计算正交矩阵 P（既是相似变换矩阵，也是合同变换矩阵），使得

$$\overbrace{diag(\underbrace{\lambda_1,\cdots,\lambda_1}_{m_1},\ \underbrace{\lambda_2,\cdots,\lambda_2}_{m_2},\ \cdots,\ \underbrace{\lambda_k,\cdots,\lambda_k}_{m_k})}^{=P^TAP=P^{-1}AP=D}$$

- A 有互异实特征值 $\lambda_1,\cdots,\lambda_k$；
- 每一个 λ_j 有规范正交的基本解组 $\eta_{j,1},\cdots,\eta_{j,m_j}$，$(m_j=n_j)$；
- $\lambda_i\neq\lambda_j$，$\{\eta_{i,1},\cdots,\eta_{i,m_i}\}\perp\{\eta_{j,1},\cdots,\eta_{j,m_j}\}$；
- k 个规范正交的基本解组拼合为一个方阵，即为正交矩阵 P；
- 正交线性变换 $x=Py$，将二次型化简为标准型：

$$x^TAx=y^TDy=\mu_1y_1^2+\mu_2y_2^2+\cdots+\lambda_ny_n^2$$

实对称矩阵的一般合同对角化（二次型的配平方）

有平方项：不妨设 $a_{11}\neq 0$，关于 x_1 "凑平方"

$$=\sum_{i=1}^{n} a_{ii}x_i^2+\sum_{i<j}2a_{ij}x_ix_j=\sum_{i,j=1}^{n} a_{ij}x_ix_j=x^TAx=f(x)$$

$$\overbrace{a_{11}\left[x_1^2+2\left(\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n\right)x_1+\underbrace{\left(\frac{a_{12}}{a_{11}}x_2+\cdots+\frac{a_{1n}}{a_{11}}x_n\right)^2}_{}\right]}$$

$$+\overbrace{\underbrace{\sum_{i,j=2}^{n}\tilde{a}_{ij}x_ix_j}_{remaining\ \ terms}}^{doesn't\ \ contain\ \ x_1}\quad\leftarrow\ quadratic\ form\ g(x_2,\cdots,x_n)\ in\ x_2,\cdots,x_n$$

对 $g(x_2,\cdots,x_n)$（关于 $n-1$ 个变量 $x_2,\cdots,x_n$）"凑平方"，$\cdots$。

实对称矩阵的一般合同对角化（二次型的配平方）

例 5.1.5

$$\begin{aligned}
f(x) &= x_1^2 - 2x_1x_2 + 8x_1x_3 + x_2^2 + 8x_2x_3 - 4x_3^2 \\
&= x_1^2 - 2x_1(x_2 - 4x_3) + (x_2 - 4x_3)^2 \\
&\quad -(x_2 - 4x_3)^2 + x_2^2 + 8x_2x_3 - 4x_3^2 \\
&= (x_1 - x_2 + 4x_3)^2 + {\color{red}16x_2x_3 - 20x_3^2 \quad (x_1 \ \ vanishing)} \\
&= (\underbrace{x_1 - x_2 + 4x_3}_{y_1})^2 - 20(\underbrace{x_3 - \frac{2}{5}x_2}_{y_2})^2 + \frac{16}{5}\underbrace{x_2}^{2}_{y_3}
\end{aligned}$$

$$\underbrace{\left\{\begin{array}{lcl} y_1 &=& x_1 - x_2 + 4x_3 \\ y_2 &=& -\frac{2}{5}x_2 + x_3 \\ y_3 &=& x_2 \end{array}\right.}_{\Rightarrow \quad x = C^{-1}y}, \quad C = \begin{pmatrix} 1 & -1 & 4 \\ 0 & -\frac{2}{5} & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

实对称矩阵的一般合同对角化（二次型的配平方，续）

例 5.1.6 没有平方项，产生平方项：

$$\begin{aligned} f(x) &= 2x_1x_2+2x_1x_3-2x_2x_3 & x_1 &= y_1+y_2 \\ &= 2(y_1+y_2)(y_1-y_2)+4y_2y_3 & x_2 &= y_1-y_2,\ x_3=y_3 \\ &= 2y_1^2-2y_2^2+4y_2y_3 & z_1 &= y_1,\ z_2=y_2-y_3 \\ &= 2y_1^2-2(y_2-y_3)^2+2y_3^2 & z_3 &= y_3 \end{aligned}$$

$$\overbrace{\begin{aligned} x_1 &= y_1+y_2 \\ x_2 &= y_1-y_2 \\ x_3 &= y_3 \end{aligned}}^{\text{first transform } x=C_1y} \qquad \overbrace{\begin{aligned} z_1 &= y_1 \\ z_2 &= y_2-y_3 \\ z_3 &= y_3 \end{aligned}}^{\text{second transform } z=C_2y} \Rightarrow \overbrace{x=C_1y}^{y=C_2^{-1}z} = \overbrace{C_1C_2^{-1}}^{\text{invertible}} z$$

$$C_1=\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ invertible}, \quad C_2=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \text{ invertible}$$

初等变换与合同对角化

- $P_1^T A P_1$：交换 A 的第 i, j 列且交换 A 的第 i, j 行；

$$P_1 = \begin{array}{cc} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow \quad i-th\ row \\ \\ \leftarrow \quad j-th\ row \\ \\ \\ \end{array} \end{array}$$

$P_1 = P_1^T$ symmetric

- P_1 为用数 s 乘以 E 的第 i 列（行）所得到的初等矩阵，
$P_1^T A P_1$：用数 s 乘以 A 的第 i 列，用数 s 乘以 A 的第 i 行；

初等变换与合同对角化（续）

- 将第 j 列的 s 倍加到第 i 列，得到初等矩阵

$$P_1=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&0&&\\&&&\ddots&&&\\&&s&&1&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}\begin{array}{l}\\\\\leftarrow\quad i-th\ row\\\\\leftarrow\quad j-th\ row\\\\\\\end{array}$$

P_1^T：将 E 的第 j 行的 s 倍加到第 i 行
$P_1^TAP_1$：将 A 的 j 列的 s 倍加到 i 列，j 行的 s 倍加到 i 行。

初等变换与合同对角化（续）

计算可逆 $C=P_1P_2\cdots P_k$，使得 C^TAC 对角：

$$\underbrace{\overbrace{(P_1P_2\cdots P_k)^T}^{=C^T}A\overbrace{(P_1P_2\cdots P_k)}^{=C}}_{=C^TAC\ diagonal}=P_k^T\cdots\overbrace{P_2^T\underbrace{(P_1^TAP_1)}_{=:B_1}P_2}^{=:B_2}\cdots P_k$$

- 对 A 实施一次初等列变换 AP_1，再实施一次“**相同**”的初等行变换，得到 $P_1^TAP_1=B_1$；
- 对 B_1 实施一次初等列变换 B_1P_2，再实施一次“**相同**”的初等行变换，得到 $P_2^TB_1P_2=B_2$$=P_2^TP_1^TAP_1P_2=B_2$，$\cdots$；
- 对 B_{k-1} 实施一次初等列变换 $B_{k-1}P_k$，再实施一次“**相同**”的初等行变换，得到

$$P_k^TB_{k-1}P_k=B_k=P_k^T\cdots P_2^TP_1^TAP_1P_2\cdots P_k=C^TAC$$

初等变换与合同对角化（续）

$$\underbrace{\overbrace{(P_1P_2\cdots P_k)^T}^{=C^T} A \overbrace{(P_1P_2\cdots P_k)}^{=C}}_{=C^TAC\ diagonal} = \underbrace{P_k^T\cdots P_2^TP_1^T\ \overbrace{(AP_1P_2\cdots P_k)}^{=:B\ row\ transforms}}_{same\ column\ transforms}$$

- 对 A 实施一系列（s 次）初等列变换 $AP_1P_2\cdots P_s = A_1$，再实施一系列（s 次）“相同”的初等行变换，得到 $P_s^T\cdots P_2^TP_1^TA_1 = B_1$；
- 对 B_1 实施一系列（m 次）初等列变换 $AQ_1Q_2\cdots Q_m = A_2$，再实施一系列（m 次）“相同”的初等行变换，得到 $Q_m^T\cdots Q_2^TQ_1^TA_2 = B_2$；
- ……，直至将 A 对化为对角矩阵。

实对称矩阵的一般合同对角化（初等变换，续）

例 5.1.5 第一步：以 a_{11} 为基准，用列变换将第一行其余元素消成 0：c_2+c_1 对应初等矩阵 P_1，c_3-4c_1 对应初等矩阵 P_2

$$\begin{array}{c}\overbrace{\begin{pmatrix} 1 & -1 & 4 \\ -1 & 1 & 4 \\ 4 & 4 & -4 \\ \cdots & \cdots & \cdots \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{A}\\ \underbrace{\phantom{\begin{pmatrix} 1 & -1 & 4 \end{pmatrix}}}_{=E}\end{array} \xrightarrow[c_3-4c_1]{c_2+c_1} \begin{array}{c}\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 8 \\ 4 & 8 & -20 \\ \cdots & \cdots & \cdots \\ 1 & 1 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{=AP_1P_2}\\ \underbrace{\phantom{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix}}}_{=P_1P_2}\end{array}$$

实对称矩阵的一般合同对角化（初等变换，续）

例 5.1.5 第一步：已经用列变换将第一行其余（除对角元以外的）元素消成 0，再实施两次“相同”的初等行变换

$$\underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 8 \\ 4 & 8 & -20 \\ \cdots & \cdots & \cdots \\ 1 & 1 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{=AP_1P_2}}_{=P_1P_2} \xrightarrow[r_3-4r_1]{r_2+r_1} \underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 8 \\ 0 & 8 & -20 \\ \cdots & \cdots & \cdots \\ 1 & 1 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{P_2^TP_1^TAP_1P_2=B_1 \ \text{symmetric}}}_{=P_1P_2}$$

实对称矩阵的一般合同对角化（初等变换，续）

例 5.1.5 第二步：以 $(B_1)_{22}$ 为基准，用列变换将第二行除对角元以外的元素消成 0；产生非 0 对角元（$c_2 + c_3$ 对应初等矩阵 P_3）

$$\underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 8 \\ 0 & 8 & -20 \\ \cdots & \cdots & \cdots \\ 1 & 1 & -4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{P_2^T P_1^T A P_1 P_2 = B_1}}_{=P_1 P_2} \xrightarrow[c_2+c_3]{r_2+r_3} \underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -12 \\ 0 & -12 & -20 \\ \cdots & \cdots & \cdots \\ 1 & -3 & -4 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}^{=P_3^T P_2^T P_1^T A P_1 P_2 P_3 = B_2}}_{=P_1 P_2 P_3}$$

实对称矩阵的一般合同对角化（初等变换，续）

例 5.1.5 第二步：以 $(B_1)_{22}$ 为基准，用列变换将第二行除对角元以外的元素消成 0（c_3-3c_2 对应初等矩阵 P_4）

$$\underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & -12 \\ 0 & -12 & -20 \\ \cdots & \cdots & \cdots \\ 1 & -3 & -4 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}}^{P_3^TP_2^TP_1^TAP_1P_2P_3=B_2}}_{=P_1P_2P_3} \xrightarrow[c_3-3c_2]{r_3-3r_2} \underbrace{\overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 16 \\ \cdots & \cdots & \cdots \\ 1 & -3 & 5 \\ 0 & 1 & -3 \\ 0 & 1 & -2 \end{pmatrix}}^{P_4^TP_3^TP_2^TP_1^TAP_1P_2P_3P_4=B_3}}_{=P_1P_2P_3P_4=C}$$

二次型的规范型（从标准型讲起。。。）

- 标准型不唯一：$C^TAC = D = diag(d_1, \cdots, d_n)$ 对角，C 可逆，$P = diag(c_1, \cdots, c_n)$ 为任意可逆对角矩阵，

$$\overbrace{(CP)^TA(CP) = P^TDP = \tilde{D} = diag(\tilde{d}_1, \cdots, \tilde{d}_n)}^{diagonal}$$

$$f(x) = x^TAx \xrightarrow[x=Cy]{x=(CP)z} \underbrace{\overbrace{x^TAx = y^TDy = z^T\tilde{D}z}^{d_1y_1^2+\cdots+d_ny_n^2}}_{\tilde{d}_1y_1^2+\cdots+\tilde{d}_ny_n^2}$$

- 不同标准型有一些共性，例如 $r(A) = r$，

$$\underbrace{C \ invertible, \ D = C^TAC \quad \Rightarrow \quad r = r(A) = r(D)}_{d_1y_1^2+\cdots+d_ry_r^2 \ \ only \ \ r \ \ square \ \ terms, \ \ d_{r+1}=\cdots=d_n=0}$$

二次型的规范型（续）

A 是 n 阶实对称矩阵，特征值分组

$$\underbrace{\begin{array}{cccccccccc} \lambda_1, & \cdots, & \lambda_p > 0, & \lambda_{p+1}, & \cdots, & \lambda_r < 0, & \lambda_{r+1} = & \cdots = & \lambda_n = 0 \\ \eta_1, & \cdots, & \eta_p, & \eta_{p+1}, & \cdots, & \eta_r, & \eta_{r+1}, & \cdots, & \eta_n \end{array}}_{ortho-normal\ \ eigenvectors,\ giving\ square\ P=(\eta_1,\cdots,\eta_n),\quad P^TP=E_n}$$

经过非退化的实线性变换 $x = Py$ 得到如下形式的标准型

$$\begin{aligned} & x^TAx \quad linear\ transform\ y_{r+1} = z_{r+1}, \ \cdots, \ y_n = z_n \\ = \ & \lambda_1 y_1^2 + \cdots + \lambda_p y_p^2 + \lambda_{p+1} y_{p+1}^2 + \cdots + \lambda_r y_r^2 \\ = \ & (\underbrace{\sqrt{\lambda_1} y_1}_{z_1})^2 + \cdots + (\underbrace{\sqrt{\lambda_p} y_p}_{z_p})^2 - (\underbrace{\sqrt{-\lambda_{p+1}} y_{p+1}}_{z_{p+1}})^2 - \cdots - (\underbrace{\sqrt{-\lambda_r} y_r}_{z_r})^2 \end{aligned}$$

实二次型的规范型：$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2$，$r = r(A)$。

二次型的规范型（续，规范型唯一吗？）

$$\underbrace{\overbrace{z_1^2+\cdots+z_p^2-z_{p+1}^2-\cdots-z_r^2}^{=x^TAx,\ x=C_1z,\ C_1\ \ invertible}=\overbrace{y_1^2+\cdots+y_s^2-y_{s+1}^2-\cdots-y_r^2}^{=x^TAx,\ x=C_2y,\ C_2\ \ invertible}}_{z=Cy,\ C\ \ invertible,\ if\ p<s,\ what\ would\ happen?}$$

考察线性方程组

$$\begin{cases} c_{11}y_1+\cdots+c_{1s}y_s+c_{1,s+1}y_{s+1}+\cdots+c_{1n}y_n=z_1\\ \cdots\cdot\\ c_{p1}y_1+\cdots+c_{ps}y_s++c_{p,s+1}y_{s+1}+\cdots+c_{pn}y_n=z_p\\ c_{p+1,1}y_1+\cdots+c_{p+1,s}y_s+c_{p+1,s+1}y_{s+1}+\cdots+c_{p+1,n}y_n=z_{p+1}\\ \cdots\cdot\\ c_{n1}y_1+\cdots+c_{ns}y_s+c_{n,s+1}y_{s+1}+\cdots+c_{nn}y_n=z_n \end{cases}$$

二次型的规范型（续，规范型唯一吗？）

$$\begin{cases} c_{11}y_1+\cdots+c_{1s}y_s+c_{1,s+1}\overbrace{y_{s+1}}^{:=0}+\cdots+c_{1n}\overbrace{y_n}^{:=0}=\overbrace{z_1}^{:=0} \\ c_{21}y_1+\cdots+c_{2s}y_s+c_{2,s+1}\overbrace{y_{s+1}}^{:=0}+\cdots+c_{2n}\overbrace{y_n}^{:=0}=\overbrace{z_2}^{:=0} \\ \cdots\cdots \\ c_{p1}y_1+\cdots+c_{ps}y_s+c_{p,s+1}\underbrace{y_{s+1}}_{:=0}+\cdots+c_{pn}\underbrace{y_n}_{:=0}=\overbrace{z_p}^{:=0} \end{cases}$$

$p<s$，故有不全为零的 $\bar{y}_1,\cdots,\bar{y}_s$ 满足（$\bar{y}_{s+1}=\cdots=\bar{y}_n=0$），

$$\begin{aligned} c_{11}\bar{y}_1+\cdots+c_{1s}\bar{y}_s&=\bar{z}_1=0 \\ \cdots\cdots& \\ c_{p1}\bar{y}_1+\cdots+c_{ps}\bar{y}_s&=\bar{z}_p=0 \end{aligned}, \quad \begin{aligned} \bar{z}_k \quad := \quad & c_{n1}\bar{y}_1+\cdots+c_{ns}\bar{y}_s\ (k>p) \\ & \underbrace{+c_{n,s+1}\bar{y}_{s+1}+\cdots+c_{nn}\bar{y}_n}_{=0} \end{aligned}$$

二次型的规范型（续，规范型唯一吗？）

$$\begin{array}{l} c_{11}\bar{y}_1+\cdots+c_{1s}\bar{y}_s=\bar{z}_1=0 \\ \cdots\cdots \\ c_{p1}\bar{y}_1+\cdots+c_{ps}\bar{y}_s=\bar{z}_p=0 \end{array}, \quad \bar{z}_k \;:=\; \begin{array}{l} c_{n1}\bar{y}_1+\cdots+c_{ns}\bar{y}_s \ (k>p) \\ +\underbrace{c_{n,s+1}\bar{y}_{s+1}+\cdots+c_{nn}\bar{y}_n}_{=0} \end{array}$$

$$\underbrace{\bar{y}=(\overbrace{\bar{y}_1,\cdots,\bar{y}_s}^{(\bar{y}_1,\cdots,\bar{y}_s)^T\neq 0},\overbrace{0,\cdots,0}^{n-s})^T, \quad \bar{z}=(\overbrace{0,\cdots,0}^{p},\bar{z}_{p+1},\cdots,\bar{z}_n)^T}_{\bar{z}=C\bar{y}}$$

$$\underbrace{\underbrace{\overbrace{\bar{z}_1^2+\cdots+\bar{z}_p^2}^{=0}-\bar{z}_{p+1}^2-\cdots-\bar{z}_r^2}_{\leq 0}=\overbrace{\bar{y}_1^2+\cdots+\bar{y}_s^2}^{>0,\ (\bar{y}_1,\cdots,\bar{y}_s)^T\neq 0}\overbrace{-\bar{y}_{s+1}^2-\cdots-\bar{y}_r^2}^{=0}}_{\Rightarrow\ \textit{constradiction!}\ \Rightarrow\ p\geq s,\ \textit{similarly showing}\ s\geq p\ \Rightarrow\ p=s}$$

二次型的惯性指数

实二次型 $f(x)=x^TAx$ 在实的非退化线性替换下有唯一的规范型

$$x^TAx \xlongequal[|C|\neq 0]{x=Cz} z_1^2+\cdots+z_s^2-z_{s+1}^2-\cdots-z_r^2$$

$$x^TAx \xlongequal[|P|\neq 0]{x=Py} c_1y_1^2+\cdots+c_sy_s^2-c_{s+1}y_{s+1}^2-\cdots-c_ry_r^2$$

二次型的几个指数（常数）

- $r=r(A)$：二次型的秩（标准型的平方项数）；
- s：二次型的正惯性指数（标准型的正平方项数）；
- $q=r-p$：二次型的负惯性指数（标准型的负方项数）。

二次型的惯性指数（续）

实对称矩阵 A，存在可逆矩阵 C，正交矩阵 P，使得

$$C^TAC = diag(\overbrace{1, \cdots, 1}^{p}, \overbrace{-1, \cdots, -1}^{q}, \overbrace{0, \cdots, 0}^{n-r})$$

$$P^TAP = diag(\underbrace{\overbrace{\lambda_1, \cdots, \lambda_p}^{positive}, \overbrace{\lambda_{p+1}, \cdots, \lambda_r}^{negative}, \overbrace{0, \cdots, 0}^{n-r}}_{eigenvalue})$$

二次型的几个指数（常数）

- $r = r(A)$： A 非零特征值的个数（重数计算在内）；
- s： A 正特征值的个数（重数计算在内）；
- $q = r - p$： A 负特征值的个数（重数计算在内）。

二次型的惯性指数（续，例子）

例 5.1.8

$$
\begin{aligned}
f(x_1,x_2,x_3) &= 2x_1^2-4x_1x_2-12x_1x_3+8x_1x_4\\
&\quad +9x_2^2+18x_2x_3-10x_2x_4+23x_3^2-20x_3x_4+4x_4^2
\end{aligned}
$$

$$
A=\begin{pmatrix} 2 & -2 & -6 & 4\\ -2 & 9 & 9 & -5\\ -6 & 9 & 23 & -10\\ 4 & -5 & -10 & 4\end{pmatrix} \xrightarrow[c_2+c_1,\ \cdots]{r_2+r_1,\ \cdots} \begin{pmatrix} 2 & 0 & 0 & 0\\ 0 & 7 & 0 & 0\\ 0 & 0 & \frac{26}{7} & 0\\ 0 & 0 & 0 & -\frac{149}{26}\end{pmatrix}
$$

$$
f(x_1,x_2,x_3,x_4) \overset{x=Cy}{\underset{|C|\neq 0}{=\!=\!=}} 2y_1^2+7y_2^2+\frac{26}{7}y_3^2-\frac{149}{26}y_4^2
$$

$$
\lambda_1\approx -1.5436,\ \lambda_2\approx 1.2088,\ \lambda_3\approx 4.7560,\ \lambda_4\approx 33.5788
$$

正惯性指数 $p=3$，负惯性指数 $q=1$，秩 $r=p+q=4$。

复二次型的规范型

复二次型，经过复非退化线性变换，得到标准型：

$$x^TAx \xrightarrow[|C|\neq 0]{x=Cy} \overbrace{c_1y_1^2+\cdots+c_ry_r^2}^{r=r(A)} \xrightarrow[|C|\neq 0]{\sqrt{c_k}y_k=z_k} z_1^2+\cdots+z_r^2$$

例 5.1.8

$$\begin{aligned} f(x_1,x_2,x_3,x_4) &\overset{x=Cy}{\underset{|C|\neq 0}{=\!=\!=}} 2y_1^2+7y_2^2+\frac{26}{7}y_3^2-\frac{149}{26}y_4^2 \\ &\overset{\sqrt{c_k}y_k=z_k}{\underset{\pm i\sqrt{\frac{149}{26}}y_4=z_4}{=\!=\!=\!=\!=}} z_1^2+z_2^2+z_3^2+z_4^2 \end{aligned}$$

- 复二次型的规范型取决于：秩；
- 实二次型的规范型取决于：秩 + 正惯性指数；

二次型的正定性

A 是 n 阶实矩阵，二次型 $f(x) = x^T Ax$ 称为是

- 正定的，如果对于任意非零实向量 $x \neq 0$ 有 $x^T Ax > 0$；
- 负定的，如果对于任意非零实向量 $x \neq 0$ 有 $x^T Ax < 0$；
- 半正定的，如果对于任意非零实向量 $x \neq 0$ 有 $x^T Ax \geq 0$；
- 半负定的，如果对于任意非零实向量 $x \neq 0$ 有 $x^T Ax \leq 0$；
- 不定的，如果有实向量 x 与 y 使得 $x^T Ax > 0$ 而 $y^T Ay < 0$。

如果 $x^T Ax$ 是正（负）定的、半正（负）定的、不定的，则相应地称 A 是正（负）定的、半正（负）定的、不定的。

(1) 怎样判断二次型 $x^T Ax$/实对称矩阵 A 的正定性？

(2) 怎么理解二次型 $x^T Ax$/实对称矩阵 A 正定性？

二次型的正定性（续，标准型的正定性）

$A = diag(c_1, \cdots, c_n)$，二次型 $f(x) = x^T A x$ 是标准型：

- 正定的，如果 $c_1, \cdots, c_n > 0$：

$$x = (x_1, \cdots, x_n)^T \neq 0 \ \Rightarrow \ c_1 x_1^2 + \cdots + c_n x_n^2 > 0$$

- 负定的，如果 $c_1, \cdots, c_n < 0$；
- 半正定的，如果 $c_1, \cdots, c_n \geq 0$（有某个 $c_k = 0$）；

$$\underbrace{x = (x_1, \cdots, x_n)^T}_{\bar{x}=e_k} \neq 0 \ \Rightarrow \ \underbrace{c_1 x_1^2 + \cdots + c_n x_n^2}_{c_1\bar{x}_1^2+\cdots+c_n\bar{x}_n^2=c_k=0} \geq 0$$

- 半负定的，如果 $c_1, \cdots, c_n \leq 0$（有某个 $c_k = 0$）；
- 不定的，如果有某个 $c_k > 0$ 且有某个 $c_m < 0$。

二次型的正定性（续，正定性与特征值的关系）

$$f(x) = x^T Ax \xlongequal[C^TC=E_n]{x=Cy} \lambda_1 y_1^2 + \cdots \lambda_n y_n^2$$

- 二次型 $f(x) = x^T Ax$ 是正定的，当且仅当其特征值均为正：

$$C = (\eta_1, \cdots, \eta_n), \quad A\eta_j = \lambda_j \eta_j$$

必要性：

$$y = e_j,\ x = Ce_j = \eta_j \neq 0 \ \Rightarrow\ x^T Ax = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = \lambda_j > 0$$

充分性：

$$x \neq 0 \iff y = C^{-1}x \neq 0 \Rightarrow\ x^T Ax = \underbrace{\lambda_1}_{>0} y_1^2 + \cdots + \underbrace{\lambda_n}_{>0} y_n^2 > 0$$

- 二次型 $f(x) = x^T Ax$ 是正定的，当且仅当其正惯性指数是 n。

二次型的正定性（续，正定性与特征值的关系）

$$f(x) = x^T Ax \xlongequal[C^TC=E_n]{x=Cy} \lambda_1 y_1^2 + \cdots \lambda_n y_n^2$$

- 二次型 $f(x) = x^T Ax$ 是半正定的，当且仅当其特征值均非负：

$$C = (\eta_1, \cdots, \eta_n), \quad A\eta_j = \lambda_j \eta_j$$

必要性：

$$y = e_j,\ x = Ce_j = \eta_j \neq 0 \ \Rightarrow\ x^T Ax = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = \lambda_j \geq 0$$

充分性：

$$x \neq 0,\ x^T Ax \xlongequal[C^TC=E_n]{x=Cy} \underbrace{\lambda_1}_{\geq 0} y_1^2 + \cdots + \underbrace{\lambda_n}_{\geq 0} y_n^2 \geq 0$$

- 二次型 $x^T Ax$ 半正定，当且仅当其正惯性指数是 $r(A)$。

二次型的正定性（续，正定性与特征值的关系）

$A = (a_{ij})_{n\times n}$ 实对称：

- 当 A 的特征值均小于零时，二次型 $f(x) = x^T Ax$ 负定；
- 当 A 的特征值均非正时，二次型 $f(x) = x^T Ax$ 半负定；
- 当 A 的特征值有正有负时，二次型 $f(x) = x^T Ax$ 不定；
- A 负定，当且仅当其负惯性指数为 n；
- A 半负定，当且仅当其负惯性指数为 $r(A)$；
- A 不定，当且仅当其正、负惯性指数均非零；
- 当 A 正定时，A^{-1} 也正定；
- 当 A 负定时，A^{-1} 也负定；
- 当 A 不定且可逆时，A^{-1} 也不定。。。

二次型的正定性（续，正定性与顺序主子式的关系）

A 实对称，其顺序主子矩阵 A_k 的行列式 $|A_k|$ 称为顺序主子式

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$$

命题：A 正定，则其顺序主子式均大于零。

证明：A_k 也正定

$$x = (\overbrace{x_1, \cdots, x_k}^{=:y\neq 0}, 0, \cdots, 0) \neq 0, \; x^T A x = y^T A_k y > 0$$

A_k 正定，其特征值 $\mu_1, \cdots, \mu_k$ 均为正，所以

$$|A_k| = \mu_1 \times \cdots \times \mu_k > 0$$

二次型的正定性（续，正定性与顺序主子式的关系）

命题：如果 A 的顺序主子式均大于零，则 A 正定。（归纳法）

证明：$A=(a_{11})$ 的顺序主子式大于零，$a_{11}>0$，A 正定。

假定命题对于 $n\times n$ 实对称矩阵成立，$B=(b_{ij})_{n\times n}$ 实对称，令

$$A=\begin{pmatrix} B & u^T \\ u & s \end{pmatrix}$$

如果 A 的顺序主子式均大于零，则 A 的顺序主子式也均大于零，由归纳假定可知 A 正定，$P^TBP=D=diag(\lambda_1,\cdots,\lambda_n)$，$P$ 正交

$$C^TAC=diag(D,s-u^TB^{-1}u),\quad C=\begin{pmatrix} P^T & 0 \\ -u^TB^{-1} & 1 \end{pmatrix}\ (invertible)$$

$$|C^TAC|=\underbrace{|C|^2|A|}_{>0}=\underbrace{|\lambda_1\cdots\lambda_n|}_{=|B|>0}\underbrace{(s-u^TB^{-1}u)}_{\Rightarrow\ >0}$$

二次型的正定性（续，充分必要条件）

定理：A 是 n 阶实对称矩阵，则以下命题等价

- A 正定；
- A 的特征值均大于零；
- A 的正惯性指数为 n
- A 的顺序主子式均大于零。

定理：A 是 n 阶实对称矩阵，则以下命题等价

- A 半正定；
- A 的特征值均非负；
- A 的正惯性指数为 $r(A)$
- A 的顺序主子式均非负。

二次型的正定性（续，一阶最优性条件）

$g(y)=g(y_1,\cdots,y_n)$ 在 $x=(x_1,\cdots,x_n)^T$ 处的 Taylor 展开式：

$$g(y)=g(x)+\overbrace{\underbrace{g'(x)(y-x)}_{(y-x,\nabla g(x))}}^{linear\ \ form}+r$$

如果 x 是局部极大（小）值点，那么

$$0=\nabla g(x)=\begin{pmatrix}\frac{\partial g}{\partial x_1}\\ \vdots\\ \frac{\partial g}{\partial x_n}\end{pmatrix}\underbrace{\xrightarrow[y=x+t\nabla g(x),\ t>0]{if\ \ otherwise} g(y)=\underbrace{g(x)+\overbrace{g'(x)(y-x)}^{(t\|\nabla g(x)\|^2>0)}+r}_{>g(x)\ \ t\ \ small\ \ enough}}_{x\ \ not\ \ maximizer}$$

极小值点情形：$y=x+t\nabla g(x)$，$t>0$。

二次型的正定性（续，二阶最优性条件）

$g(y)=g(y_1,\cdots,y_n)$ 在驻点 $x=(x_1,\cdots,x_n)^T$ 处展开：

$$g(y)=g(x)+\underbrace{\overbrace{g'(x)(y-x)}^{=0}}_{(y-x,\nabla g(x))}+\frac{1}{2}\underbrace{\overbrace{(y-x)^Tg''(x)(y-x)}^{quadratic\ \ form}}_{(y-x,g''(x)(y-x))}+r$$

如果 Hessian 矩阵正定（负定）：

$$g''(x)=\begin{pmatrix} \frac{\partial^2 g}{\partial x_1\partial x_1} & \cdots & \frac{\partial^2 g}{\partial x_1\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial^2 g}{\partial x_n\partial x_1} & \cdots & \frac{\partial^2 g}{\partial x_n\partial x_n} \end{pmatrix}$$

则驻点 $x=(x_1,\cdots,x_n)^T$ 是局部极小（大）值点。

面试试题

- 给定线性相关的向量组 $\alpha_1, \cdots, \alpha_n$，$A = (\alpha_1, \cdots, \alpha_n) = (a_{ij})_{m\times n}$，什么是向量组的线性相关性？线性相关性与 $Ax = 0$ 的解的情况有什么联系？
- 给定向量组 $\alpha_1, \cdots, \alpha_n$，什么是向量组的"极大线性无关组"，什么是"秩"？
- 向量 β 可由向量组 $\alpha_1, \cdots, \alpha_n$ 线性表出时什么意思？$Ax = \beta$ 什么时候有解？何时解唯一？

面试试题

- $A=(a_{ij})_{n\times n}$，叙述特征值与特征向量的定义，什么叫做“A 可以相似对角化”，叙述相似对角化的条件。
- $A=(a_{ij})_{n\times n}$ 实对称，什么叫做“合同对角化”，什么叫做“正交对角化”，A 可以合同对角化吗？A 可以相似对角化吗？
- $A=(a_{ij})_{n\times n}$ 实对称，叙述“正定矩阵”的含义，简述 A 正定的充分必要条件。