二次型的标准形

定义 1

假定二次型 $Q(x_1,\ldots,x_n)=\boldsymbol{X}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{X}$ 通过某个坐标的可逆线性变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$ 化为不含混合项 (交叉项) 的形式:

$$Q(x_1,\ldots,x_n)|_{\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}}:=\boldsymbol{Y}^{\mathsf{T}}(\boldsymbol{P}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{P})\boldsymbol{Y}=\mu_1y_1^2+\mu_2y_2^2+\cdots+\mu_ny_n^2,$$

即

$$\boldsymbol{P}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\mu_1,\mu_2,\ldots,\mu_n).$$

我们称该形式为二次型 Q 的标准形 (*canonical form*).

注 2

标准形显然不唯一. 例如

$$Q(x_1,x_2)=2x_1^2+3x_2^2=y_1^2+y_2^2,$$

其中 $y_1=\sqrt{2}x_1$, $y_2=\sqrt{3}x_2$.

主轴化方法

很显然, 将二次型化简成标准形的过程, 就是寻找可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角阵的过程. 由于 $\boldsymbol{A}$ 为实对称阵, 我们可以找到正交矩阵 $\boldsymbol{P}$ 将其相似对角化. 由于 $\boldsymbol{P}$ 为正交阵, 对角阵 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{P}$, 且其主对角线上的元素为 $\boldsymbol{A}$ 的所有特征值. 借此, 我们证明了

定理 3 (主轴定理)

任何实二次型都可以通过坐标的正交变换化为标准形

$$Q(x_1, \ldots, x_n)|_{\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2,$$

这儿的 $\lambda_1, \ldots, \lambda_n$ 为 $\boldsymbol{A}$ 的所有特征值.

例 4

考虑二次型 $Q(\gamma)=x_1^2+x_2^2+x_3^2+4x_1x_2+4x_1x_3+4x_2x_3$.

例 5

设二次型 $Q(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+2\alpha x_1x_2+2\beta x_2x_3+2x_1x_3$ 经过正交变换 $\boldsymbol{X}=\boldsymbol{P}\boldsymbol{Y}$ 化为 $Q=y_2^2+2y_3^2$, 其中 $\boldsymbol{X}=(x_1,x_2,x_3)^\mathsf{T}$, $\boldsymbol{Y}=(y_1,y_2,y_3)^\mathsf{T}$, $\boldsymbol{P}$ 是正交阵. 试求 α,β 的值.

例 6

设 $\boldsymbol{A}$ 为 3 阶实对称阵, 特征值为 $1,1,2$, 且有特征向量 $\boldsymbol{\alpha}_1=(1,0,1)^\mathsf{T}$, $\boldsymbol{\alpha}_2=(0,1,1)^\mathsf{T}$. 求正交变换将此实对称阵对应的二次型 $Q=\boldsymbol{X}^\mathsf{T}\boldsymbol{A}\boldsymbol{X}$ 化为标准形.

例 7 (Rayleigh 原理)

设 n 元二次型 $Q(x_1,\ldots,x_n)=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X}$, 其中 $\boldsymbol{X}=(x_1,\ldots,x_n)^{\mathrm{T}}$, 而实对称阵 $\boldsymbol{A}$ 的特征值满足 $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$. 证明: 在条件 $x_1^2+x_2^2+\cdots+x_n^2=1$ 的条件下, $Q(x_1,\ldots,x_n)$ 的最小值为 λ_1, 最大值为 λ_n.

例 8

设 $\boldsymbol{A}=(a_{ij})$ 是 n 阶实对称阵, 它的 n 个特征值为 $\lambda_1\le\lambda_2\le\cdots\le\lambda_n$. 证明: $\lambda_1\le a_{ii}\le\lambda_n$, 其中 $i=1,2,\ldots,n$.

配方法

由于正交阵 $\boldsymbol{P}$ 的计算比较复杂, 在将二次型化为标准形时, 若没有特别要求, 我们有时仅仅通过配方的方法, 消除掉二次型中的交叉项. 这样的好处是计算简便, 不利之处在于由于做了不同方向的不同比例的拉伸, 画出的 (可能是高维的) 几何图形的形状会和原来的形状不一致.

接下来, 设

$$Q(x_1, x_2, \ldots, x_n) = a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots + 2a_{1n}x_1x_n + (\text{与 } x_1 \text{ 无关的项}).$$

配方中的两种基本情形

1. 假设 $a_{11} \neq 0$. 此时,

$$Q(x_1, x_2, \ldots, x_n) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + (\text{配方后新的项})$$
$$+ (\text{之前与 } x_1 \text{ 无关的项})$$
$$= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + (\text{与 } x_1 \text{ 无关的项}).$$

在这种情况下, 我们会设

$$y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n, \quad y_2 = x_2, \quad \ldots, \quad y_n = x_n,$$

从而

$$x_1 = y_1 - \frac{a_{12}}{a_{11}}y_2 - \cdots - \frac{a_{1n}}{a_{11}}y_n, \quad x_2 = y_2, \quad \ldots, \quad x_n = y_n.$$

相应的坐标变换的矩阵形式为

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & -\frac{a_{12}}{a_{11}} & \cdots & -\frac{a_{1n}}{a_{11}} \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

2. 假设 $a_{12} \neq 0$, 而 $a_{11} = a_{22} = \cdots = a_{nn} = 0$. 我们利用等式

$$x_1 x_2 = \frac{1}{4}\left[(x_1 + x_2)^2 - (x_1 - x_2)^2\right] = \left(\frac{x_1 + x_2}{2}\right)^2 - \left(\frac{x_1 - x_2}{2}\right)^2,$$

令

$$y_1 = \frac{x_1 + x_2}{2}, \quad y_2 = \frac{x_1 - x_2}{2}, \quad y_3 = x_3, \quad \ldots, \quad y_n = x_n,$$

从而

$$x_1 = y_1 + y_2, \quad x_2 = y_1 - y_2, \quad x_3 = y_3, \quad \ldots, \quad x_n = y_n.$$

这种情形下的坐标的基本变换的矩阵形式为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = \left(\begin{array}{cc|ccc} 1 & 1 & & & \\ 1 & -1 & & & \\ \hline & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{array}\right) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}.$$

在此之后，我们可以化归为第一种情形，其中 y_1^2 前的系数为原式

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = (a_{11} + 2a_{12} + a_{22})y_1^2 + \cdots$$

中 y_1^2 的系数，即 $2a_{12}$，不为零. 注意，这里仅仅 $a_{11} = 0$, $a_{12} \neq 0$, 是不够的.

例 9

考察 $Q(x_1, x_2, x_3) = x_1^2 + 2x_3^2 + x_1x_2$.

例 10

考察 $Q(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3$.

矩阵的初等变换法

在主轴化定理 3 中我们已经见到, 存在可逆矩阵 $\boldsymbol{P}$ 将二次型的矩阵 $\boldsymbol{A}$ 相合对角化. 由于 $\boldsymbol{P}$ 可以表示成一系列初等矩阵的乘积, 于是我们有

定理 11

对每个实对称阵 $\boldsymbol{A}$, 存在同阶的初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \ldots, \boldsymbol{P}_s$ 使得

$$\boldsymbol{P}_S^{\mathsf{T}} \cdots \boldsymbol{P}_2^{\mathsf{T}} \boldsymbol{P}_1^{\mathsf{T}} \boldsymbol{A} \boldsymbol{P}_1 \boldsymbol{P}_2 \cdots \boldsymbol{P}_s$$

为对角阵.

由此出发, 为了计算二次型的标准形, 我们将对实矩阵采取如下的操作 (其中矩阵 $\boldsymbol{A}$ 为对称阵)

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \longrightarrow \begin{pmatrix} \boldsymbol{AP} \\ \boldsymbol{BP} \end{pmatrix} \longrightarrow \begin{pmatrix} \boldsymbol{P}^{\mathsf{T}}\boldsymbol{AP} \\ \boldsymbol{BP} \end{pmatrix},$$

即, 我们先对整个矩阵作列变换 (右乘 $\boldsymbol{P}$), 再对分块矩阵中上面的方阵做相应的行变换 (左乘 $\boldsymbol{P}^{\mathsf{T}}$).

注 12

1. 若 $\boldsymbol{P} = \boldsymbol{S}_{ij} : c_i \leftrightarrow c_j$, 则 $\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{S}_{ij} : r_i \leftrightarrow r_j$.
2. 若 $\boldsymbol{P} = \boldsymbol{D}_i(\lambda) : \lambda c_i$, 则 $\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{D}_i(\lambda) : \lambda r_i$.
3. 若 $\boldsymbol{P} = \boldsymbol{T}_{ij}(\lambda) : \lambda c_i \to c_j$, 则 $\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{T}_{ji}(\lambda) : \lambda r_i \to r_j$.

若最开始 $\boldsymbol{B} = \boldsymbol{I}$, 而经过一系列这样的操作后, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 为对角阵, 则分块矩阵中下面的矩阵 $\boldsymbol{B}\boldsymbol{P} = \boldsymbol{P}$ 为所求.

例 13

用初等变换的方法把二次型 $Q(x_1, x_2, x_3) = x_1^2 + 4x_1x_3 - x_2^2 + 2x_2x_3 + x_3^2$ 化为标准形.

例 14

用初等变换的方法把二次型 $Q(x_1, x_2, x_3) = 4x_1x_2 + 2x_1x_3 + x_2^2 + 2x_2x_3 + x_3^2$ 化为标准形.