二次型及其意义

Eason

目录

1 定义

任意一个方阵A的二次型$\mathbf{x}^T A\mathbf{x}$ 是一个标量。举个例子，假设

$$A = \begin{bmatrix} 1 & 4 & 2 \\ -1 & 7 & 5 \\ -1 & 6 & 3 \end{bmatrix} \tag{1}$$

为例。

其二次型为：

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= [x_1, x_2, x_3] \begin{bmatrix} 1 & 4 & 2 \\ -1 & 7 & 5 \\ -1 & 6 & 3 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \qquad (2) \\ &= x_1^2 + 7x_2^2 + 3x_3^2 + 3x_1x_2 + x_1x_3 + 11x_2x_3 \qquad (3) \end{aligned}$$

上式的计算可以简单的做出来，可以观察到，对于A上对角线的元素对应的分别是x_1^2, x_2^2, x_3^2的系数，$A_{ij} + A_{ji}$对应的是$x_i x_j$的系数。另外可以看到上式的计算结果中最高的幂次是 2，称这是变元x的二次型函数，称$\mathbf{x}^T A\mathbf{x}$为矩阵A的二次型。

推而广之，若$\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$，且 $n\times n$矩阵A的元素a_{ij}，则二次型：

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij} \tag{4}$$

进一步计算：

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j a_{ij} &= \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i=1,i\neq j}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j & (5)\\ &= \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}(a_{ij}+a_{ji})x_i x_j & (6)\end{aligned}$$

根据式 (5) 我们发现，同一个二次型函数可以对应不同的矩阵，比如：

$$B = \begin{bmatrix} 1 & 2 & 1\\ 1 & 7 & 3\\ 1 & 8 & 3 \end{bmatrix} \tag{7}$$

和A具有相同的二次型函数。即，对于一个二次型函数：

$$f(x_1, x_2, \ldots, x_n) = \sum_{i}^{n} \alpha_{ii}x_i^2 + \sum_{i=1,i\neq j}^{n}\sum_{j=1}^{n} \alpha_{ij}x_i x_j \tag{8}$$

存在许多矩阵A，它们的二次型$\mathbf{x}^T A\mathbf{x}$相同。但是只有一个对称矩阵A满足$\mathbf{x}^T A\mathbf{x} = f(x_1, \ldots, x_n)$，其元素$a_{ii} = \alpha_{ii}, a_{ij} = a_{ji} = \frac{1}{2}(\alpha_{ij} + \alpha_{ji})$ 其中$i = 1, \ldots, n; j = 1, \ldots, n, i \neq j$，因此讨论二次型的时候通常假定矩阵$A$为对称矩阵或者Hermitian矩阵。Hermitian矩阵是复对称矩阵，所以实对称矩阵是 Hermitian矩阵的特殊形式。

2 正定矩阵

如果将大于零的二次型$\mathbf{x}^T A\mathbf{x}$称为正定的二次型，则与之对应的Hermitian矩阵称为正定矩阵。类似的还可以定义Hermitian矩阵的半正定性，负定性和半负定性。

若$\mathbf{x}^T A\mathbf{x}$ 的取值即可能为正，也可能为负，则对应的 Hermitian矩阵为不定矩阵。

表 1：矩阵定性定义

条件	定义
$\mathbf{x}^T A\mathbf{x} > 0, \forall \mathbf{x} \neq 0$	正定矩阵
$\mathbf{x}^T A\mathbf{x} \geq 0, \forall \mathbf{x} \neq 0$	半正定矩阵
$\mathbf{x}^T A\mathbf{x} < 0, \forall \mathbf{x} \neq 0$	负定矩阵
$\mathbf{x}^T A\mathbf{x} \leq 0, \forall \mathbf{x} \neq 0$	半负定矩阵

3 Hermitian矩阵的正定性

我们之前提到过，对于一个二次型，有多个矩阵A与之对应。但是，只有一个矩阵是对称的。我们这么做不仅仅是因为这个Hermitian矩阵的唯一性，还因为这个矩阵具有非常多的特性。另外，我们在实际的工程应用中也有很多矩阵是 Hermitian矩阵，所以深入的研究Hermitian矩阵很有必要。

定义Hermitian二次型为：

$$H(\mathbf{x},\mathbf{x}) = \langle \mathbf{x}, A\mathbf{x} \rangle = \mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} r_{ij} x_i^T x_j \tag{9}$$

称A为二次型的核或者矩阵。核为单位阵时，二次型退化为向量 $\mathbf{x}$和它自己的内积$\mathbf{x}^T\mathbf{x}$。因此，二次型可以视为内积的推广。

考虑酉变换之后的二次型。定义线性变换$\mathbf{x} = R\mathbf{y}$，其中R是酉矩阵。此时，可以将原来的二次型改写为：

$$H(\mathbf{x},\mathbf{x}) = \mathbf{y}^T R^T A R\mathbf{y} = \mathbf{y}^T P\mathbf{y} \tag{10}$$

式中，$P = R^T AR = R^{-1}AR$ 。特别地，若选择酉矩阵R使得 P为对角矩阵，则称：

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T (R^T AR)\mathbf{y} = y^T \mathrm{diag}(\lambda_1, \ldots, \lambda_n)\mathbf{y} \tag{11}$$

为标准二次型。因为这样的二次型，只有核对角线上的元素非零。因为R是酉矩阵，则有$R^{-1} = R^T$ 。令$A = U\Sigma U^T$是A的特征值分解，并将它带入二次型$H(x,x)$，得到：

$$H(x,x) = \mathbf{x}^T A\mathbf{x} = \mathbf{x}^T U\Sigma U^T\mathbf{x} \tag{12}$$

令$R = U$, 则$\mathbf{x} = R\mathbf{y} = U\mathbf{y}$, $\mathbf{y} = U^T x$, 所以式 (12) 变为：

$$H(x,x) = \mathbf{y}^T \Sigma\mathbf{y} = \langle \mathbf{x}, A x \rangle = \sum_{i=1}^{n} \lambda_i |y_i|^2 \tag{13}$$

经过酉变换$\mathbf{x} = R\mathbf{y}$之后，二次型 $\mathbf{x}^T A\mathbf{x}$的值等于二次型核矩阵A的特征值与 y_i的模值平方的乘积之和。

通过观察式 (13)，我们可以很容易的判定Hermitian 的正定性。一个$n\times n$的Hermitian矩阵是正定的，当且仅当其满足如下条件：

1. 二次型$\mathbf{x}^T A\mathbf{x} > 0, \forall \mathbf{x} \neq 0$
2. 矩阵A的所有特征值都大于零。
3. 所有主子矩阵A_k都具有正的行列式，$A_k = A(1:k, 1:k)$由矩阵 A的第$1\sim k$行和第$1\sim k$列组成。
4. 存在一个非奇异的$n\times n$矩阵R使得 $A = R^T R$
5. 存在一个非奇异的$n\times n$矩阵P使得$P^T AP$是正定的。

4 正定矩阵的意义