

二次型极值及其推广的几个结论

董祖引, 周继东

(河海大学数学物理系, 江苏 南京 210098)

摘要: 讨论了二次型在单位球面及其闭子集上的极值问题, 并引入广义二次型的概念, 利用正交变换从两个方面推广了二次型的极值与其对应的特征值之间的关系.

关键词: 二次型; 广义二次型; 极值; 正交变换

中图分类号: O151.21

文献标识码: A

文章编号: 1000-198X(2000)04-0116-03

1 引言

在二次型理论中我们有如下结论^[1]:

命题 设 $f(X) = X^T A X$ 是一个实二次型, 其中 $X^T = (x_1, x_2, \dots, x_n)$, A 是一个实对称阵, 设 A 的最大、最小特征值分别为 λ_n, λ_1 , 则 $f(X)$ 在单位球面 $S: \|X\| = \sqrt{\sum_{i=1}^n x_i^2} = 1$ 上的最大值为 λ_n , 最小值为 λ_1 , 即

$$\max_{X \in S} f(X) = \lambda_n \quad \min_{X \in S} f(X) = \lambda_1$$

本文对此结果进行了进一步的探讨, 并得到更一般的结论. 为此, 先引入两个基本概念:

定义 1 称 $f(X_1, X_2, \dots, X_k) = (X_1, X_2, \dots, X_k)^T A (X_1, X_2, \dots, X_k)$ 为实广义二次型, 其中 $X_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T$ 为 n 维实向量, A 为 $n \times n$ 实对称阵, $k = 1, 2, \dots, m$.

定义 2 设 Q_k 是一个实 $n \times k$ ($k \leq n$) 阵, 将 Q_k 按列分块 $Q_k = (u_1, u_2, \dots, u_k)$, 若满足内积 $(u_i, u_j) = \delta_{ij}$ ($i, j = 1, 2, \dots, k$), 则称 Q_k 为 $n \times k$ 列正交阵.

引理 设 $\alpha_i = (a_{1i}, a_{2i}, \dots, a_{ni})^T$ 是实列向量, 且 $(\alpha_i, \alpha_j) = \delta_{ij}$, $i, j = 1, 2, \dots, r$, 则对于任给 k ($1 \leq k \leq n$), 任给 r ($1 \leq r \leq n$), 有 $a_{1k}^2 + a_{2k}^2 + \dots + a_{rk}^2 \leq 1$.

证明 将 $\alpha_1, \alpha_2, \dots, \alpha_r$ 看作 n 维欧氏空间 R^n 的标准正交向量组, 则存在 R^n 中向量 $\alpha_{r+1}, \dots, \alpha_n$ 使 $\alpha_1, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n$ 是 R^n 的标准正交基, 即对列正交阵 $(\alpha_1, \alpha_2, \dots, \alpha_r)$ 必存在列正交阵 $(\alpha_{r+1}, \dots, \alpha_n)$ 使 $P = (\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}, \dots, \alpha_n)$ 是正交阵, 由正交阵的定义知, 对任给 l , $1 \leq l \leq n$, 有

$$a_{1l}^2 + \dots + a_{rl}^2 + a_{l,r+1}^2 + \dots + a_{ln}^2 = 1$$

所以 $a_{1l}^2 + a_{2l}^2 + \dots + a_{rl}^2 \leq 1$. 证毕.

由引理显然可得, 任一 $n \times k$ 列正交阵其每行元素的平方和 (即行向量的长度平方) 小于等于 1.

2 主要结论

定理 1 设 $f(X) = X^T A X$ 是一个实二次型, A 的特征值分别为 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, 则存在 A 的一组相互正交的单位特征向量 $X_1, X_2, \dots, X_n$ (分别对应于特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$), 使得 $f(X_i) = \lambda_i$ ($1 \leq i \leq n$), 且 $f(X_n)$ 为 $f(X)$ 在单位球面 $S: \|X\| = 1$ 上最大值, $f(X_1)$ 为 $f(X)$ 在 S 上的最小值. 更进一步, 对任一自然数 k ($1 < k < n$), $f(X)$ 在有界闭集 $S_k = \text{span}\{X_1, X_2, \dots, X_k\} \cap S = \{X | (X, X_i) = 0, X \in S, k+1 \leq i \leq n\}$ 上取得最大值 $f(X_k) = \lambda_k$; $f(X)$ 在有界闭集 $S_k^* = \text{span}\{X_k, X_{k+1}, \dots, X_n\} \cap S = \{X | (X, X_i) = 0, X \in S, 1 \leq i \leq k-1\}$ 上取得最小值 $f(X_k) = \lambda_k$.

定理 2 设 L_k 是 R^n 的任一 k 维子空间, $M_k = L_k \cap S$, 其中 S 为单位球面 $\|X\| = 1$, 二次型 $f(X) = X^TAX$ 的特征值分别为 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, 则

$$\lambda_k = \min_{M_k} \{ \max_{X \in M_k} f(X) \} = \max_{M_{n-k+1}} \{ \min_{X \in M_{n-k+1}} f(X) \}$$

注 1 若 $k = n$, 此时 $M_k = S$, 则由定理 2 的第 1 个等式, $\lambda_n = \min_{M_n} \{ \max_{X \in M_n} f(X) \} = \max_{X \in S} f(X)$;

若 $k = 1$, 由定理 2 的第二个等式, $\lambda_1 = \max_{M_n} \{ \min_{X \in M_n} f(X) \} = \min_{X \in S} f(X)$, 此即前述命题.

定理 3 设 A 是一 n 阶实对称阵, 其特征值 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ (为表写方便, 此处特征值按由大到小顺序排列), 则

$$\begin{aligned} \max_{Q_k \in \alpha(k)} \text{tr}(Q_k^T A Q_k) &= \sum_{i=1}^k \lambda_i \\ \min_{Q_k \in \alpha(k)} \text{tr}(Q_k^T A Q_k) &= \sum_{i=1}^k \lambda_{n-i+1} \end{aligned}$$

这里 $\alpha(k)$ 是由所有 $n \times k$ 列正交阵构成的集合, $\text{tr}(A)$ 代表 A 的迹.

证明 因为 A 为实对称阵, 所以存在正交阵 P , 使得

$$P^T A P = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} = \Lambda$$

将 P 分块, $P = (P_k, P_{n-k})$, 其中 P_k 为 $n \times k$ 列正交阵, P_{n-k} 为 $n \times (n-k)$ 列正交阵, 此时

$$P_k^T A P_k = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_k\}$$

所以

$$\sum_{i=1}^k \lambda_i = \text{tr}(P_k^T A P_k) \leq \max_{Q_k \in \alpha(k)} \text{tr}(Q_k^T A Q_k) \tag{1}$$

另一方面, 设 $Q_k = (u_1, u_2, \dots, u_k)$, 则

$$\text{tr}(Q_k^T A Q_k) = \sum_{i=1}^k u_i^T A u_i = \sum_{i=1}^k (P^T u_i)^T P^T A P (P^T u_i)$$

设 $P^T u_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T$, 则

$$(P^T u_i, P^T u_j) = (P^T u_i)^T P^T u_j = u_i^T P P^T u_j = u_i^T u_j = \delta_{ij}$$

即 $(P^T u_1, P^T u_2, \dots, P^T u_k)$ 构成一个 $n \times k$ 列正交阵. 此时

$$\begin{aligned} \text{tr}(Q_k^T A Q_k) &= \sum_{i=1}^k (\lambda_1 x_{1i}^2 + \lambda_2 x_{2i}^2 + \dots + \lambda_n x_{ni}^2) \leq \\ &\lambda_1 \sum_{i=1}^k x_{1i}^2 + \lambda_2 \sum_{i=1}^k x_{2i}^2 + \dots + \lambda_{k-1} \sum_{i=1}^k x_{k-1,i}^2 + \lambda_k \left(\sum_{i=1}^k x_{ki}^2 + \sum_{i=1}^k x_{k+1,i}^2 + \dots + \sum_{i=1}^k x_{ni}^2 \right) \end{aligned} \tag{2}$$

由于

$$\begin{aligned} &\sum_{i=1}^k x_{ki}^2 + \sum_{i=1}^k x_{k+1,i}^2 + \dots + \sum_{i=1}^k x_{ni}^2 = \sum_{j=k}^n \sum_{i=1}^k x_{ji}^2 = \sum_{i=1}^k \sum_{j=k}^n x_{ji}^2 \\ (2) \text{式} &= \lambda_1 \sum_{i=1}^k x_{1i}^2 + \lambda_2 \sum_{i=1}^k x_{2i}^2 + \dots + \lambda_{k-1} \sum_{i=1}^k x_{k-1,i}^2 + \\ &\lambda_k \left[\left(1 - \sum_{j=1}^{k-1} x_{j1}^2 \right) + \left(1 - \sum_{j=1}^{k-1} x_{j2}^2 \right) + \dots + \left(1 - \sum_{j=1}^{k-1} x_{jk}^2 \right) \right] = \\ &\lambda_1 \sum_{i=1}^k x_{1i}^2 + \lambda_2 \sum_{i=1}^k x_{2i}^2 + \dots + \lambda_{k-1} \sum_{i=1}^k x_{k-1,i}^2 + \lambda_k \left[k - \sum_{i=1}^k \sum_{j=1}^{k-1} x_{ji}^2 \right] = \\ &\lambda_1 \sum_{i=1}^k x_{1i}^2 + \lambda_2 \sum_{i=1}^k x_{2i}^2 + \dots + \lambda_{k-1} \sum_{i=1}^k x_{k-1,i}^2 + \lambda_k \left[k - \sum_{j=1}^{k-1} \sum_{i=1}^k x_{ji}^2 \right] = \\ &(\lambda_1 - \lambda_k) \sum_{i=1}^k x_{1i}^2 + (\lambda_2 - \lambda_k) \sum_{i=1}^k x_{2i}^2 + \dots + (\lambda_{k-1} - \lambda_k) \sum_{i=1}^k x_{k-1,i}^2 + k \lambda_k \end{aligned}$$

由引理知 $\sum_{i=1}^k x_{ji}^2 \leq 1 (j = 1, 2, \dots, k-1)$, 故上式小于或等于

$$(\lambda_1 - \lambda_k) + (\lambda_2 - \lambda_k) + \dots + (\lambda_{k-1} - \lambda_k) + k \lambda_k = \lambda_1 + \lambda_2 + \dots + \lambda_k$$

即 $\text{tr}(\mathbf{Q}_k^T \mathbf{A} \mathbf{Q}_k) \leq \sum_{i=1}^k \lambda_i$, 由 $\mathbf{Q}_k \in \mathcal{O}(k)$ 的任意性, 所以

$$\max_{\mathbf{Q}_k \in \mathcal{O}(k)} \text{tr}(\mathbf{Q}_k^T \mathbf{A} \mathbf{Q}_k) \leq \sum_{i=1}^k \lambda_i \quad (3)$$

由式 (1) 和 (3) 即证得定理 3 的第一个等式, 对第二个等式可类似证明, 或对 $-\mathbf{A}$ 利用上述证明过程即可.

注 2 当 $k=1$ 时, 即得前述命题.

注 3 当 $k=n$ 时, 对任意正交阵 $\mathbf{Q}_n$, $\text{tr}(\mathbf{Q}_n^T \mathbf{A} \mathbf{Q}_n) = \sum_{i=1}^n \lambda_i$, 因为此时 $\mathbf{Q}_n^T \mathbf{A} \mathbf{Q}_n$ 与 $\mathbf{A}$ 相似, 从而有相同的特征值, 相同的迹^[2].

注 4 一般地, 若记矩阵 $\mathbf{A}$ 的属于特征值 $\lambda_{i_1}, \lambda_{i_2}, \dots, \lambda_{i_k}$ ($1 \leq i_1 < i_2 < \dots < i_k \leq n$) 的所有特征向量所张成的特征子空间为 $S_{i_1, i_2, \dots, i_k}$ (显然它是 R^n 的 k 维子空间), 又记 $\mathcal{O}_{i_1, i_2, \dots, i_k}$ 为所有列向量均属于 $S_{i_1, i_2, \dots, i_k}$ 的 $n \times k$ 列正交阵的全体, 则

$$\text{tr}(\mathbf{Q}_k^T \mathbf{A} \mathbf{Q}_k) = \sum_{j=1}^k \lambda_{i_j} \quad \mathbf{Q}_k \in \mathcal{O}_{i_1, i_2, \dots, i_k}$$

3 举 例

设二次型 $f(x_1, x_2, x_3) = 11x_1^2 + 5x_2^2 + 2x_3^2 + 16x_1x_2 + 4x_1x_3 - 20x_2x_3$, 其所对应的矩阵为

$$\mathbf{A} = \begin{bmatrix} 11 & 8 & 2 \\ 8 & 5 & -10 \\ 2 & -10 & 2 \end{bmatrix}$$

容易算出 $\mathbf{A}$ 的特征值分别为 $\lambda_1 = -9, \lambda_2 = 9, \lambda_3 = 18$, 对应的单位正交特征向量分别为

$$\mathbf{X}_1 = \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^T \quad \mathbf{X}_2 = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)^T \quad \mathbf{X}_3 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right)^T$$

由定理 1, $f(x_1, x_2, x_3)$ 在 $S: x_1^2 + x_2^2 + x_3^2 = 1$ 上的最大值为 18, 最小值为 -9;

$$f(x_1, x_2, x_3) \text{ 在 } S_2: \begin{cases} x_1^2 + x_2^2 + x_3^2 = 1 \\ 2x_1 + 2x_2 - x_3 = 0 \end{cases} \text{ 上的最大值为 9;}$$

$$f(x_1, x_2, x_3) \text{ 在 } S_2^*: \begin{cases} x_1^2 + x_2^2 + x_3^2 = 1 \\ -x_1 + 2x_2 + 2x_3 = 0 \end{cases} \text{ 上的最小值为 9.}$$

参考文献:

- [1] Birkhoff G, Mac Lane S. A survey of modern algebra[M]. New York: Macmillan Publishing Co. Inc, 1977. 292 ~ 295.
[2] 丘维声. 高等代数[M]. 北京: 高等教育出版社, 1996. 258 ~ 268.

Some Propositions on the Extreme Value of Quadratic Form

DONG Zu-yin, ZHOU Ji-dong

(Dept. of Mathematics and Physics, Hohai Univ., Nanjing 210098, China)

Abstract: The extreme value of quadratic form on a unit sphere and its closed sets are discussed in this paper. By introducing the extended quadratic form and by use of orthogonal transform, a more general relation between the extreme value and the eigenvalue of the quadratic form is obtained.

Key words: quadratic form, extended quadratic form, extreme value, orthogonal transform