

n 元二次函数的最值问题

孙华娟¹, 贺爱娟¹, 陈永义²

(1. 烟台大学 文经学院, 山东 烟台 264005; 2. 烟台大学 数学与信息科学学院, 山东 烟台 264005)

摘要:提出了 n 元二次函数无约束条件最值问题的一般理论和方法, 得到了最大(小)值存在及最值点存在且唯一的简明的充分必要条件; 有 m ($m < n$) 个线性等式约束的问题可化为无约束的低维二次函数的相应问题; 给出了求最大(小)值(点)的具体方法. 对有非线性约束条件的问题, 指出了二次型标准化方法的重要性, 并举了实例.

关键词:二次函数; 二次型的标准型; 最大值; 最小值

中图分类号: O151.2 **文献标志码:** A **文章编号:** 1007-2683(2012)01-0116-04

Maximum and Minimum Problem of Quadratic Function of n Variables

SUN Hua-juan¹, HE Ai-juan¹, CHEN Yong-yi²

(1. Wenjing College, Yantai University, Yantai 264005, China; 2. College of Mathematics and Information, Yantai University, Yantai 264005, China)

Abstract: For maximum and minimum problem of a quadratic function of n variables without restricted conditions, we obtained both concise sufficient and necessary conditions for existence of maximum (or minimum) and existence and uniqueness of maximum (or minimum) point. The problem with m ($m < n$) linear restricted conditions is transformed into a related problem of a quadratic function of less variables without restricted conditions. This paper gives concrete solving method for maximum and minimum. For the problem with nonlinear restricted conditions, we point the importance of standardize method of quadratic form and give examples.

Key words: quadratic function; standard form of quadratic form; maximum; minimum

0 引言

求 n 元二次函数的最值(点)的问题是数学及应用数学中常遇到的问题, 对一些具体的二次函数, 人们已有一些结果^[1-7], 但迄今尚未见到一般性的论述. 本文提出了 n 元二次函数无约束条件最值问题的一般理论, 得到了最值存在和最值点存在且唯一的简明的充分必要条件, 并给出了具体的解法; 有 m ($1 \leq m < n$) 个线性等式约束的问题可化为无约束条件的低维二次函数的相应问题处理; 有非线性约

束的问题需具体问题具体处理, 本文仅讨论了个别特例. 此外, 本文中所讨论函数的定义域是 R^n .

1 无约束条件的最值

略去常数项, 一般的 n 元二次函数可记为

$$f(x_1, x_2, \dots, x_n) = \sum_{i,j=1}^n a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i = X^T A X + b^T X \quad (1)$$

其中: $A = (a_{ij})_{1 \leq i,j \leq n}$ 是 $n \times n$ 实对称矩阵, 即 $A^T = A$, $X = (x_1, \dots, x_n)^T$, $b = (b_1, \dots, b_n)^T$.

收稿日期: 2010-07-02

基金项目: “十一五”国家课题我国高校应用型人才培养模式研究项目(FIB070335-B2-13)

作者简介: 孙华娟(1974—), 女, 硕士, 讲师, E-mail: yingxia1834@sina.com;

陈永义(1931—), 男, 教授.

先将二次型 $X^T A X$ 标准化. 由线性代数理论知, 存在 $n \times n$ 可逆矩阵, 特别地, 存在正交矩阵 P , 令 $X = PY$, 其中 $Y = (y_1, \dots, y_n)^T$, 可使

$$X^T A X = (PY)^T A (PY) = Y^T (P^T A P) Y = Y^T \Lambda Y = \sum_{i=1}^r \lambda_i y_i^2 \quad (2)$$

$$A = P^T \Lambda P = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_r & \\ 0 & & & \ddots \\ & & & & 0 \end{pmatrix},$$

$$\lambda_i \neq 0, 1 \leq i \leq r, 1 \leq r \leq n \quad (3)$$

不妨设 $\lambda_1, \dots, \lambda_p > 0, \lambda_{p+1}, \dots, \lambda_r < 0$. p 称为二次型(或二次型矩阵 A) 的正惯性系数, $r - p$ 称为负惯性系数, r 为二次型(或 A) 的秩.

特别地, 当 P 为正交矩阵时, $\lambda_1, \dots, \lambda_r, 0, \dots, 0$ 是 A 的特征值的集合, 又称 A 的谱.

在上述变换下, 二次函数化为

$$f = Y^T \Lambda Y + (b^T P) Y = Y^T \Lambda Y + B^T Y = \sum_{i=1}^r \lambda_i y_i^2 + \sum_{i=1}^n b'_i y_i \quad (4)$$

其中 $B^T = b^T P = (b'_1, \dots, b'_n)$.

作平移变换 $Y = Z + C$, 其中,

$$Z = (z_1, \dots, z_n)^T, C = (c_1, \dots, c_n)^T \quad (5)$$

注意到 $Z^T \Lambda C$ 是 1×1 矩阵, $A^T = A$, 可知 $Z^T \Lambda C = (Z^T \Lambda C)^T = C^T \Lambda Z$, 故

$$f = (Z + C)^T \Lambda (Z + C) + B^T (Z + C) = Z^T \Lambda Z + (Z^T \Lambda C + C^T \Lambda Z + B^T Z) + C^T \Lambda C + B^T C = Z^T \Lambda Z + (2C^T \Lambda + B^T) Z + (C^T \Lambda + B^T) C \quad (6)$$

由此, 立知有如下结论:

命题1 二次函数 f 中一次项可消去的充分必要条件是:

1) 存在常向量 C , 使 $2C^T \Lambda + B^T = 0$. 与 1) 等价的条件是

2) 方程组 $\Lambda C = -\frac{1}{2}B$ (B 为常向量, C 看作未知向量) 有解.

由线性代数的知识知, 条件 2) 等价于如下诸条件中的任何一个:

3) $-\frac{1}{2}B$ 可由 Λ 的列向量组线性表示.

4) $R(\Lambda) = R(\Lambda, -\frac{1}{2}B) = R(\Lambda, B)$.

5) B 属于 Λ 的列向量组张成的子空间.

6) $B = (b'_1, \dots, b'_r, 0, \dots, 0)^T$, 即

$$b'_{r+1} = \dots = b'_n = 0 \quad (7)$$

式(7) 是一个非常简明的判别条件. 由式(7) 判定二次函数是否可消去一次项, 需将二次型及二次函数标准化, 其实不标准化, 直接从式(1) 中的矩阵 A 和向量 b 亦可判定该二次函数是否可消去一次项, 这就是如下的命题 2.

命题2 B 可由 Λ 的列向量组线性表示的充要条件是 b 可由 Λ 的列向量组线性表示.

证明: B 可由 Λ 的列向量组线性表示 $\Leftrightarrow$ 存在 $b^* = (b_1^*, \dots, b_n^*)^T$, 使

$$B = \sum_{i=1}^n b_i^* \Lambda_i = (\Lambda_1, \dots, \Lambda_n) (b_1^*, \dots, b_n^*)^T = \Lambda b^* \quad (8)$$

其中, Λ 的第 i 个列向量 $\Lambda_i = (0, \dots, 0, \lambda_i, 0, \dots, 0)^T$.

注意到 $B = \Lambda b^* = P^T \Lambda P b^*$, $B^T = b^T P$, 立知式(8) 等价于

$$\begin{aligned} P^T b &= B = \Lambda b^* = P^T \Lambda P b^* \\ \Leftrightarrow P P^T b &= P P^T \Lambda P b^* \quad (P P^T = E) \\ \Leftrightarrow b &= \Lambda P b^* = \Lambda b^{**} \quad (b^{**} = P b^*) \end{aligned} \quad (9)$$

式(9) 即表示 b 可由 Λ 的列向量组线性表示, 命题 2 得证.

由命题 1 和 2 立知有如下命题 3:

命题3 当且仅当二次函数 $f = X^T A X + b^T X$ 中 A, b 满足条件

$$R(\Lambda) = R(\Lambda, b) \quad (10)$$

f 才能经可逆变换化为无一次项的二次函数 $f =$

$$\sum_{i=1}^r \lambda_i z_i^2 + d. \quad (d \text{ 为常数})$$

命题4 若 $R(\Lambda) < R(\Lambda, b)$, 等价地, $R(\Lambda) < R(\Lambda, B)$, 则二次函数 f 无最值.

证明: 由命题 1, 2, 3 知 $f = \sum_{i=1}^r \lambda_i z_i^2 + \sum_{i=1}^n b'_i z_i$, 其中 $b'_{r+1}, \dots, b'_n$ 不全为零.

不妨设 $b'_{r+1} > 0$ (其他情形同理可证). 令 $z_i = 0, \forall i \neq r+1$,

当 $z_{r+1} \rightarrow +\infty$ 时, $f \rightarrow +\infty$; 当 $z_{r+1} \rightarrow -\infty$ 时, $f \rightarrow -\infty$. 由是可知 f 无最值.

命题5 若 $R(\Lambda) = R(\Lambda, b)$, 且 f 标准化为 $f =$

$$\sum_{i=1}^r \lambda_i z_i^2 + d \quad (d \text{ 为常数}), \text{ 则}$$

1) 若 λ_i 全大于零, $i = 1, \dots, r$ (二次型的正惯性系数 $p = r$), 则 f 有最小值(点), 无最大值(点). 当

且仅当 $z_1 = \cdots = z_r = 0$ 时, $f = d$ 最小.

2) 若 λ_i 全小于零, $i = 1, \cdots, r$ (二次型的负惯性系数为 r), 则 f 有最大值(点), 无最小值(点). 当且仅当 $z_1 = \cdots = z_r = 0$ 时, $f = d$ 最大.

3) 若 $\{\lambda_i\}_{1 \leq i \leq r}$ 中有正有负, 二次型为不定型, 则 f 无最值(点).

证明: (1, 2) 显然是正确的. 为证 3) 成立, 不妨设 $\lambda_1, \cdots, \lambda_p > 0, \lambda_{p+1}, \cdots, \lambda_r < 0$,

令 $z_1 = \cdots = z_p = 0, z_{p+1} = \cdots = z_r \rightarrow +\infty$, 则 $f \rightarrow -\infty$,

令 $z_1 = \cdots = z_p = -\rightarrow +\infty, z_{p+1} = \cdots = z_r = 0$, 则 $f \rightarrow +\infty$. 故知 f 无最值.

命题 6 若 $R(A) = R(A, b)$, 且经可逆变换得

$$f = \sum_{i=1}^r \lambda_i z_i^2 + d, \text{ 则}$$

1) 当 λ_i 全大于零时, $i = 1, \cdots, r$; 若 $r = n$, 则 f 有唯一最小值点; 若 $r < n$, 则 f 有无穷个最小值点.

2) 当 λ_i 全小于零时, $i = 1, \cdots, r$; 若 $r = n$, 则 f 有唯一最大值点; 若 $r < n$, 则 f 有无穷个最大值点.

证明: 由于所作变换均可逆, 因而点 $(x_1, \cdots, x_n)$ 与其对应点 $(z_1, \cdots, z_n)$ 是一一对应的, 由命题 5 立知结论成立.

2 有线性等式约束条件的最值

设 $m (1 \leq m < n)$ 个线性约束方程为

$$\left. \begin{aligned} C_{11}x_1 + \cdots + C_{1n}x_n &= d_1 \\ &\vdots \\ C_{m1}x_1 + \cdots + C_{mn}x_n &= d_m \end{aligned} \right\} \quad (11)$$

其矩阵形式为 $CX = D$. 系数矩阵 C 是 $m \times n$ 实数矩阵.

不失一般性, 不妨设 $R(C) = m$, 从而 $R(C, D) = m$, 由此可得同解方程

$$\left. \begin{aligned} x_1 &= e_{1,m+1}x_{m+1} + \cdots + e_{1n}x_n + r_1 \\ &\vdots \\ x_m &= e_{m,m+1}x_{m+1} + \cdots + e_{mn}x_n + r_m \end{aligned} \right\} \quad (12)$$

将式(12)代入 f 后成为 $x_{m+1}, \cdots, x_n$ 的 $n-m$ 元的二次函数, 于是求 $f(x_1, \cdots, x_n)$ 在约束式(11)成立条件下的最大(小)值问题, 转化为 $n-m$ 元的二次函数的无约束条件的最值问题.

下述例子中二次型的标准化用的是比较实用的配方法, 而不是理论证明中常用的正交变换法.

例 1 求 $f_n(a_1, \cdots, a_n) = \sum_{i=1}^n i(n-i+1)a_i^2 +$

$2 \sum_{1 \leq i < j \leq n} i(n-j+1)a_i a_j$ 在 $\sum_{i=1}^n i a_i = n+1$ 条件下的最小值(点).

将 $\sum_{i=1}^n i a_i = n+1$ 变形为

$$a_n = \frac{1}{n} \left(n+1 - \sum_{i=1}^{n-1} i a_i \right) \quad (13)$$

把式(13)代入 $f_n(a_1, \cdots, a_n)$ 中, 得

$$f_n \left(a_1, \cdots, a_{n-1}, \frac{1}{n} \left(n+1 - \sum_{i=1}^{n-1} i a_i \right) \right) = \frac{n+1}{n} f_{n-1}(a_1, \cdots, a_{n-1}) + \frac{(n+1)^2}{n}.$$

用配方法可证 $f_{n-1}(a_1, \cdots, a_{n-1}) = \sum_{r=1}^{n-1} A_r$, 其中 $A_r = \frac{n(n-r)}{n-r+1} \left(a_r + \frac{1}{n-r} \sum_{j=r+1}^{n-1} (n-j)a_j \right)^2$.

由是即知, 当且仅当 $a_1 = \cdots = a_{n-1} = 0$ 时,

$f_{n-1}(a_1, \cdots, a_{n-1})$ 最小, 从而 $f_n(a_1, \cdots, a_n)$ 在 $\sum_{i=1}^n i a_i = n+1$ 条件下的最小值点为 $a_1 = \cdots = a_{n-1} = 0, a_n = \frac{n+1}{n}$.

3 非线性约束下 n 元二次函数的最值问题的实例

此类问题比较复杂, 尚无统一的理论, 需具体问题具体处理, 但二次型的标准化方法应是其关键技术. 下面是一个经典的例子.

例 2 若 $\sum_{i=1}^n x_i^2 = 1$, 求二次型 $u = \sum_{i,j=1}^n a_{ij} x_i x_j$ ($a_{ij} = a_{ji}$) 的最值.

此例是众所周知的文 [8] 中的第 3671 题, 被许多教材或参考书采用, 如文 [9] 和文 [10].

该问题的矩阵形式是: 求二次型 $f(X) =$

$X^T A X$ ($A^T = A$) 在 $\|X\| = \sqrt{\sum_{i=1}^n x_i^2} = 1$ 时的最值.

解: 设有正交矩阵 P , 使

$$P^T A P = \Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & 0 & \\ & & & \ddots \\ 0 & & & & \lambda_n \end{pmatrix}. \text{ 令 } X = PY,$$

$$f(X) = X^T A X = f(PY) = Y^T (P^T A P) Y =$$

$$Y^T \Lambda Y = \sum_{i=1}^n \lambda_i y_i^2.$$

其中, $\{\lambda_i\}_{1 \leq i \leq n}$ 为 A 的特征值集合, 设其最大特征值为 λ_k .

$$\text{注意: } X^T X = \|X\|^2 = \sum_{i=1}^n x_i^2 = Y^T (P^T P) Y =$$

$$Y^T Y = \sum_{i=1}^n y_i^2 = 1, \text{ 可知}$$

$$f(X) = f(PY) = \sum_{i=1}^n \lambda_i y_i^2 \leq \lambda_k \sum_{i=1}^n y_i^2 = \lambda_k.$$

即在 $\|X\| = 1$ 条件下, $f(X)$ 的最大值不超过 λ_k , 令 $Y_0 = e_k = (0, \dots, 0, 1, 0, \dots, 0)^T$ (第 k 个分量为 1, 其余分量为 0 的单位向量), 则知 Y_0 在单位球面 $Y^T Y = 1$ 上, 且 $Y_0^T \Lambda Y_0 = \lambda_k$. 令 $X_0 = P Y_0$, 则知

$$X_0^T X_0 = Y_0^T Y_0 = 1,$$

$$f(X_0) = f(PY_0) = (PY_0)^T A (PY_0) = Y_0^T \Lambda Y_0 = \lambda_k.$$

由此证明 $f(X)$ 在 X_0 ($\|X_0\| = 1$) 达到最大值, 故 λ_k 是 $f(X)$ 在 $\|X\| = 1$ 条件下的最大值, X_0 是一个最大值点.

同理可证, $f(X) = X^T A X$ 在 $\|X\| = 1$ 上的最小值即为 A 的最小特征值.

例 2 中的二次型是一般形式的二次型, 用正交变换将二次型标准化, 结果非常简洁漂亮; 例 1 中的二次型是具体的二次型, 需求出标准型的具体结果, 若用正交变换, 计算量大, 且表达式也很复杂, 因此采用了较实用的配方法.

当二次函数的定义域不是 $\mathbb{R}^n$, 而是其子集时, 上述结论需作相应修改. 事实上, 定义域的边界通常是一些超曲面, 求函数的最大(小)值通常还需计算函数在这些超曲面上的最大(小)值, 这正是本文第 3 节中所论述的问题. 此类问题较复杂, 超出了本文讨论的范围.

参考文献:

- [1] 彭学梅. 关于多元二次函数的最值问题 [J]. 高等数学研究, 2005, 8(2): 44-46.
- [2] 刘锋, 王讲书. 基于矩阵初等变换的 n 元函数极值的快速判别法 [J]. 数学的实践与认识, 2004, 34(5): 170-174.
- [3] 秦桂香, 高纯一. n 元函数极值的探讨 [J]. 数学理论与应用, 2003, 23(4): 105-107.
- [4] 黎敦云. 关于 n 元二次函数的最值性研究 [J]. 柳州师专学报, 2001, 16(3): 100-102.
- [5] 王朝霞, 张庆. 多元二次函数的最值问题 [J]. 数学通报, 2001(2): 40-41.
- [6] 卢小宁, 萧振纲. 多元二次函数的一个性质及其应用 [J]. 数学理论与应用, 2001, 21(4): 101-104.
- [7] 赵国清, 田广悦. 二元函数的外推算法及其应用 [J]. 哈尔滨理工大学学报, 1986, 1(2): 90-95.
- [8] Б. П. 吉米多维奇. 数学分析习题集 [M]. 李荣涑译. 北京: 人民教育出版社, 1958: 355.
- [9] Б. П. 吉米多维奇著. 数学分析习题集题解 [M]. 费定晖, 周学圣编. 山东: 山东科学技术出版社, 1999: 425-526.
- [10] 毛纲源. 线性代数解题方法技巧归纳 [M]. 北京: 华中理工大学出版社, 2000: 409-410.

(编辑: 温泽宇)