Pure Mathematics 理论数学, 2023, 13(6), 1610-1618
Published Online June 2023 in Hans. https://www.hanspub.org/journal/pm
https://doi.org/10.12677/pm.2023.136164

剖析二次型理论在多元函数极值和条件极值中的应用

贾瑞玲，孙铭娟，韩艺兵

信息工程大学，河南 郑州

收稿日期：2023年5月9日；录用日期：2023年6月12日；发布日期：2023年6月20日

摘 要

二次型理论是高等代数中的一个重要分支，其在多元函数极值和条件极值中具有重要的地位。由于二次型理论涉及到代数学、分析学等方面的知识，因此，对学生而言学习难度较大。实际上，运用矩阵的正定性来判断多元函数的极值和条件极值，是一种极为便捷的方法，但其适用范围存在一定的限制。因为驻点处对应的矩阵的正定、负定仅是该点成为极值点的充分条件。鉴于此，本文首先探讨了多元函数的极值与矩阵特征值之间的相互联系，其次深入研究二次型理论在多元函数极值、最值问题中的应用。同时还讨论了二次型理论在多元函数条件极值中的应用及其存在的不足。当判别法失效时，通常需要在驻点处判断Lagrange函数的二阶微分的符号。通过这几方面的介绍，旨在加深学生对该理论的理解和应用。

关键词

二次型，多元函数，极值，条件极值，特征值

Analysis of the Application of Quadratic Form Theory in Extremum and Conditional Extremum of Multivariate Function

Ruiling Jia, Mingjuan Sun, Yibing Han

Information Engineering University, Zhengzhou Henan

Received: May 9th, 2023; accepted: Jun. 12th, 2023; published: Jun. 20th, 2023

Abstract

Quadratic theory is an important branch of advanced algebra, and it plays an important role in the

文章引用: 贾瑞玲, 孙铭娟, 韩艺兵. 剖析二次型理论在多元函数极值和条件极值中的应用[J]. 理论数学, 2023, 13(6): 1610-1618. DOI: 10.12677/pm.2023.136164

extreme value and conditional value of multivariate function. Because quadratic theory involves algebra and analytic knowledge, it is difficult for students to learn. In fact, it is an extremely convenient method to determine the extremum and conditional extremum of multivariate function by the positive properties of matrix, but its application scope is limited. Because the positive definite and negative definite of the matrix corresponding to the stagnation point are only sufficient conditions for the point to become the extreme point. In view of this, this paper first discusses the correlation between the extreme value of multivariate function and the eigenvalue of matrix, and then studies the application of quadratic theory in the extreme value and optimal value of multivariate function. At the same time, the application of quadratic theory in conditional extreme value of multivariate function and its shortcomings are discussed. When the discriminant fails, it is usually necessary to judge the sign of the second derivative of Lagrange function at the stagnation point. Through the introduction of these aspects, the aim is to deepen students' understanding and application of the theory.

Keywords

Quadratic Form, Multivariate Function, The Extremum, Conditional Extremum, Eigenvalue

1. 引言

二次型作为一种重要的数学工具既能把二次函数、双曲线和椭圆这样的曲线表达成一般线性方程组，又能把多项式方程和复变函数这样的非线性方程组表达成线性方程组的集合，在一定条件下也能把多项式方程或复变函数这些线性方程组化为二次型问题来求解。所以二次型理论在多元函数极值与条件极值问题中有重要地位与广泛应用。

关于这方面的内容，目前已有很多学者进行了相关研究。早在 1997 年，米少田[1]就讨论了二次型理论在判定多元函数极值中的应用以及存在的局限性。随后董丽华、梁伟秋等人[2] [3] [4]也分别研究过类似问题。卢里举[5]从二次型多项式的正定性判别作为切入点，研讨得到极值判定的充分条件，并给出该方法在多元线性回归中计算回归系数的应用。刘爱德[6]等人从矩阵角度出发阐述了多元函数在稳定点取得极值的充分条件，并用代数方法解决了两类多元函数的极值问题。陈荣群和黄勇[7]着重研究用二次型理论求两种特殊类型函数的条件极值，并讨论一般元二次式的极值和判别法。黄朝军[8]探讨了二次型矩阵在多元函数的最值和条件极值中的应用。杨廷鸿等人[9]从 Lagrange 条件极值出发，讨论二项式条件极值的 Lagrange 乘数法求解，并将其迁移到矩阵特征值问题，进一步拓展到 n 维空间。本文在前人研究的基础上，全面阐述二次型理论、矩阵特征值在多元函数极值和条件极值中的应用并解析二次型理论在判定条件极值时存在的不足。当判别法失效时，利用二次型与多元函数间的联系，求解多元函数极值时一般要对驻点邻域内函数二阶微分的符号进行判断。但是限于多个变量等因素，计算函数的二阶微分时会产生很大的计算量。

安排如下：第二部分给出二次型理论、矩阵特征值和极值存在的充分条件，为后面内容提供理论支撑。第三部分是本文的核心，主要介绍多元二次齐次函数的极值与矩阵特征值之间的关系、二次型理论在多元函数极值和最值中的应用、二次型理论在多元函数条件极值中的应用及存在的局限性。第四部分对全文内容进行总结和概括。

2. 准备知识

有关二次型的概念、性质及相关定理以及极值存在的充分条件，详细内容可参看文献[10] [11]。

定义 1 设 A 是 n 阶矩阵，若存在实数 λ 和非零向量 $X \in R^n$ 满足 $AX = \lambda X$，则称 λ 为矩阵 A 的特征值，X 为对应于 λ 的特征向量。

定义 2 数域 P 上的一个 n 元二次型是系数在 P 中的含 n 个变量的二次齐次多项式，其一般形式是

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

简写成 $f(x_1,x_2,\cdots,x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$，其中 $a_{ij} = a_{ji}\,(1 \le i,j \le n)$，令 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$，称 A 为二次型 $f(x_1,x_2,\cdots,x_n)$ 的矩阵，此时 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX$。

定义 3 设实二次型 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX$，$X = (x_1,x_2,\cdots,x_n)^{\mathrm{T}} \in R^n$，$A^{\mathrm{T}} = A$，则

1) 若对 $\forall X \ne 0$，有 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX > 0$，则称 f 为正定二次型，A 为正定矩阵。

2) 若对 $\forall X \ne 0$，有 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX < 0$，则称 f 为负定二次型，A 为负定矩阵。

3) 若对 $\forall X \ne 0$，有 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX \ge 0$，则称 f 为半正定二次型，A 为半正定矩阵。

4) 若对 $\forall X \ne 0$，有 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX \le 0$，则称 f 为半负定二次型，A 为半负定矩阵。

5) 其余情形，则称 f 为不定二次型，A 为不定矩阵。

定理 1 设实二次型 $f(x_1,x_2,\cdots,x_n) = X^{\mathrm{T}}AX$ 是正定(负定)二次型，其中 $X = (x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，$A = (a_{ij})_{n\times n}$，则 $a_{ii} > 0\ (a_{ii} < 0)$，$i = 1,2,\cdots,n$。

证明 这里以 f 是正定二次型为例给出证明，f 是负定二次型时，证明过程类似。

根据正定二次型的定义，取 $X_1 = (1,0,\cdots,0)^{\mathrm{T}} \ne 0$，有 $f = X_1^{\mathrm{T}}AX_1 = a_{11} > 0$。取 $X_2 = (0,1,\cdots,0)^{\mathrm{T}} \ne 0$，有 $f = X_2^{\mathrm{T}}AX_2 = a_{22} > 0$。依次下去，取 $X_n = (0,0,\cdots,1)^{\mathrm{T}} \ne 0$，有 $f = X_n^{\mathrm{T}}AX_n = a_{nn} > 0$。综上，结论成立。

定理 2 (极值的充分条件) p_0 是函数 $f(x_1,x_2,\cdots,x_n)$ 的驻点，$f(x_1,x_2,\cdots,x_n)$ 在 p_0 附近具有二阶连续偏导数，记 $H(x_1,x_2,\cdots,x_n) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$，则

1) 当 $H(p_0)$ 是正定矩阵时，p_0 是极小值点，$f(p_0)$ 是极小值；

2) 当 $H(p_0)$ 是负定矩阵时，p_0 是极大值点，$f(p_0)$ 是极大值；

3) 当 $H(p_0)$ 是不定矩阵时，p_0 不是极值点，$f(p_0)$ 不是极值；

4) 当 $H(p_0)$ 是半正定或半负定时，无法判断 p_0 是否为极值点；

分析 将函数 $f(x_1,x_2,\cdots,x_n)$ 在点 p_0 处 Taylor 展开，结合矩阵与二次型之间的关系即可得到结论，详

细证明过程可参看[3] [6]。

注 1 i) 若函数 $f\left(x_1,x_2,\cdots,x_n\right)$ 在点 p_0 处满足 $\left.\dfrac{\partial f}{\partial x_i}\right|_{p_0}=0$，则 p_0 是函数的驻点。

ii) 矩阵 $H\left(p_0\right)$ 是函数 $f\left(x_1,x_2,\cdots,x_n\right)$ 在点 p_0 处的 Hessian 矩阵。

iii) 对于条件极值而言，当 $H\left(p_0\right)$ 是正定(负定)矩阵时，p_0 是满足约束条件的极小(大)值点，$f\left(p_0\right)$ 是满足约束条件的极小(大)值；当 $H\left(p_0\right)$ 是不定矩阵时，无法判断 p_0 是否为满足约束条件的极值点。反例可参看本文的例 5、例 6。

3. 经典解析

3.1. 多元函数 $f\left(x_1,x_2,\cdots,x_n\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\left(a_{ij}=a_{ji}\right)$ 的极值与矩阵 $A=\left(a_{ij}\right)_{n\times n}$ 的特征值之间的联系

例 1 (北京交通大学 2022 年数学分析考研真题)证明：n 元实二次型 $\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j\left(a_{ij}=a_{ji}\right)$ 在单位球面 $\sum_{i=1}^{n}x_i^2=1$ 上的最大值、最小值恰为对称矩阵 $A=\left(a_{ij}\right)_{n\times n}$ 的最大、最小特征值。

分析 即证多元函数 $f\left(x_1,x_2,\cdots,x_n\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$ 在条件 $\sum_{i=1}^{n}x_i^2=1$ 下的最大值、最小值分别是矩阵 $A=\left(a_{ij}\right)_{n\times n}$ 的最大、最小特征值。

解 首先求 $f\left(x_1,x_2,\cdots,x_n\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$ 在条件 $\sum_{i=1}^{n}x_i^2=1$ 下的最值，构造 Lagrange 函数

$$L=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-\lambda\left(\sum_{i=1}^{n}x_i^2-1\right),$$

L 关于 $x_1,x_2,\cdots,x_n,\lambda$ 求偏导，得

$$\begin{cases}L_{x_i}=2\sum_{j=1}^{n}a_{ij}x_j-2\lambda x_i=0,\quad i=1,2,\cdots,n\\ L_\lambda=\sum_{i=1}^{n}x_i^2-1=0.\end{cases}\tag{1}$$

将前 n 个方程分别乘以 $x_1,x_2,\cdots,x_n$ 后相加，并结合 $\sum_{i=1}^{n}x_i^2=1$ 可知

$$0=\sum_{i=1}^{n}x_iL_{x_i}=2\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j-2\lambda\sum_{i=1}^{n}x_i^2=2f\left(x_1,x_2,\cdots,x_n\right)-2\lambda\text{。}$$

即方程组(1)的解 $X=\left(x_1,x_2,\cdots,x_n\right)^{\mathrm{T}}$ 满足

$$f\left(x_1,x_2,\cdots,x_n\right)=\lambda\tag{2}$$

其次，易知方程组(1)中的前 n 个线性方程组的系数矩阵为 A，且满足 $AX=\lambda X$。即 λ 为矩阵 A 的特征值，而 $X=\left(x_1,x_2,\cdots,x_n\right)^{\mathrm{T}}$ 为对应于 λ 的特征向量。由于 f 在有界闭区域 $\sum_{i=1}^{n}x_i^2=1$ 上连续，从而存在最大值与最小值，所以由(2)式可知 f 的最大值与最小值分别为矩阵 A 的最大特征值与最小特征值。

3.2. 二次型理论在多元函数极值中的应用

例 2 求函数 $f(x,y)=x^4+y^4-x^2-2xy-y^2$ 的极值。

分析 先求可能的极值点，再进一步判断。

解 第一：求可能的极值点。

解方程 $\begin{cases} f_x=4x^3-2x-2y=0 \\ f_y=4y^3-2x-2y=0 \end{cases}$ 求得驻点： $p_0(0,0)$， $p_1(1,1)$， $p_2(-1,-1)$；

第二：判断。求二阶偏导数， $f_{xx}=12x^2-2$， $f_{yy}=12y^2-2$， $f_{xy}=-2$。

$p_1(1,1)$ 处，对应的 Hessian 矩阵 $H(p_1)=\left.\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y} \\ \frac{\partial^2 f}{\partial y\partial x} & \frac{\partial^2 f}{\partial x^2} \end{pmatrix}\right|_{p_1}=\begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix}$，各阶顺序主子式：

$$|H_1|=10>0,\quad |H_2|=100-4=96>0,$$

即 $H(p_1)$ 是正定矩阵，故 $p_1(1,1)$ 是极小值点，极小值为 $f(1,1)=-2$。

$p_2(-1,-1)$ 处，对应的 Hessian 矩阵 $H(p_2)=\left.\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y} \\ \frac{\partial^2 f}{\partial y\partial x} & \frac{\partial^2 f}{\partial x^2} \end{pmatrix}\right|_{p_2}=\begin{pmatrix} 10 & -2 \\ -2 & 10 \end{pmatrix}$，各阶顺序主子式：$|H_1|>0$，$|H_2|>0$，即 $H(p_2)$ 是正定矩阵，故 $p_2(-1,-1)$ 是极小值点，极小值为 $f(-1,-1)=-2$。

$p_0(0,0)$ 处，对应的 Hessian 矩阵 $H(p_0)=\left.\begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y} \\ \frac{\partial^2 f}{\partial y\partial x} & \frac{\partial^2 f}{\partial x^2} \end{pmatrix}\right|_{p_0}=\begin{pmatrix} -2 & -2 \\ -2 & -2 \end{pmatrix}$，$|H|=0$。但注意到 $f(0,0)=0$，且

$$\Delta f=f(x,y)-f(0,0)=x^4+y^4-(x+y)^2,$$

在曲线上 $y=x, x\in(0,1)$， $\Delta f=2x^2(x^2-2)<0$，

在曲线上 $y=-x, x\in(0,1)$， $\Delta f=2x^4>0$，故 $(0,0)$ 不是极值点；

综上， $p_1(1,1)$、 $p_2(-1,-1)$ 是极小值点，相应的极小值为 $f(1,1)=f(-1,-1)=-2$。

3.3. 二次型理论在多元函数最值中的应用

例 3 (第十三届全国大学生数学竞赛)设 $B\subset R^n(n\geq 2)$ 是单位开球，函数 u，v 在 $\bar{B}$ 上连续，在 B 内二阶连续可导，满足

$$\begin{cases} -\Delta u-(1-u^2-v^2)u=0, & x\in B, \\ -\Delta v-(1-u^2-v^2)v=0, & x\in B, \\ u(x)=v(x)=0, & x\in\partial B, \end{cases} \tag{3}$$

其中 $x=(x_1,x_2,\cdots,x_n)$， $\Delta u=\frac{\partial^2 u}{\partial x_1^2}+\frac{\partial^2 u}{\partial x_2^2}+\cdots+\frac{\partial^2 u}{\partial x_n^2}$， ∂B 表示 B 的边界，证明： $u^2(x)+v^2(x)\leq 1(\forall x\in\bar{B})$。

分析 方程组(3)中的前两个方程均含有 u^2+v^2 的形式，且要证明的结论是 $u^2+v^2\le 1$，为便于讨论，利用形式统一法将 u^2+v^2 视为一个整体进行分析与探究。

证明 记 $w(x)=u^2(x)+v^2(x)$，则

$$\nabla w=2u\nabla u+2v\nabla v，\quad \Delta w=\operatorname{div}(\nabla w)=2|\nabla u|^2+2u\Delta u+2|\nabla v|^2+2v\Delta v。$$

即 w 满足以下方程

$$\begin{cases}-\Delta w-2(1-w)w=-2\left(|\nabla u|^2+|\nabla v|^2\right), & x\in B,\\ w(x)=0, & x\in\partial B,\end{cases}\tag{4}$$

显然，$w(x)\in C^2(B)\cap C(\bar{B})$。所以 $w(x)$ 必然在 $\bar{B}$ 上达到最大值，设最大值点为 p_0。

若 $p_0\in B$，即 p_0 是函数 $w(x)$ 的内部极值点，则 $\nabla w(p_0)=0$。记 $H(p_0)$ 是函数 $w(x)$ 在点 p_0 处的 Hessian 矩阵，则 $H(p_0)$ 是负定矩阵，由定理 1 知，$\dfrac{\partial^2 w(p_0)}{\partial x_i^2}<0$，$i=1,2,\cdots,n$。从而 $-\Delta w(p_0)>0$。于是由(4)得到，在 p_0 处，

$$0<-\Delta w=2(1-w)w-2\left(|\nabla u|^2+|\nabla v|^2\right)\le 2(1-w)w，$$

而 $w(p_0)\ge 0$，故上式表明 $w(p_0)\le 1$。

若 $p_0\in\partial B$，则由(4)，$w(p_0)=0$。

综上，恒有 $0\le w(x)\le 1$，$x\in\bar{B}$；即 $u^2(x)+v^2(x)\le 1\left(\forall x\in\bar{B}\right)$。

3.4. 二次型理论在多元函数条件极值中的应用及其存在的局限性

3.4.1. 二次型理论在多元函数条件极值中的应用

例 4 求函数 $f(x_1,x_2,\cdots,x_n)=\sum_{i=1}^{n}a_ix_i^2\,(a_i>0,i=1,2,\cdots,n)$ 在条件 $\sum_{i=1}^{n}x_i=c\,(x_i>0)$ 下的极值。

解 第一：构造 Lagrange 函数 $L(x_1,x_2,\cdots,x_n,\lambda)=\sum_{i=1}^{n}a_ix_i^2+\lambda\left(\sum_{i=1}^{n}x_i-c\right)$。

第二：求 Lagrange 函数的驻点。解方程

$$\begin{cases}L_{x_1}=2a_1x_1+\lambda=0\\ L_{x_2}=2a_2x_2+\lambda=0\\ \qquad\vdots\\ L_{x_n}=2a_nx_n+\lambda=0\\ L_\lambda=\sum_{i=1}^{n}x_i-c=0\end{cases}\tag{5}$$

得 $\begin{cases}x_i=-\dfrac{\lambda}{2a_i},i=1,2,\cdots,n\\ \lambda=-\dfrac{2c}{\sum_{i=1}^{n}\dfrac{1}{a_i}}\end{cases}$，记 $M_0\left(-\dfrac{\lambda}{2a_1},-\dfrac{\lambda}{2a_2},\cdots,-\dfrac{\lambda}{2a_n}\right)$。

函数 L 在点 M_0 处的 Hessian 矩阵

$$H\left(M_0\right)=\left.\begin{pmatrix} L_{x_1x_1} & L_{x_1x_2} & \cdots & L_{x_1x_n} \\ L_{x_2x_2} & L_{x_2x_1} & \cdots & L_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ L_{x_nx_n} & L_{x_nx_1} & \cdots & L_{x_nx_n} \end{pmatrix}\right|_{M_0}=\left.\begin{pmatrix} 2a_1 & 0 & \cdots & 0 \\ 0 & 2a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2a_n \end{pmatrix}\right|_{M_0},$$

因 $a_i>0,i=1,2,\cdots,n$，故 $H\left(M_0\right)$ 是正定矩阵。即 $f\left(x_1,x_2,\cdots,x_n\right)$ 有唯一的极小值 $f\left(M_0\right)=\dfrac{c^2}{\sum\limits_{i=1}^{n}\dfrac{1}{a_i}}$，并且也是最小值。故 $f\left(x_1,x_2,\cdots,x_n\right)=\sum\limits_{i=1}^{n}a_ix_i^2\left(a_i>0,i=1,2,\cdots,n\right)$ 在条件 $\sum\limits_{i=1}^{n}x_i=c\left(x_i>0\right)$ 下的极小值为 $\dfrac{c^2}{\sum\limits_{i=1}^{n}\dfrac{1}{a_i}}$。

注 2 由例 4，可得到如下结论

函数 $f\left(x_1,x_2,\cdots,x_n\right)=\sum\limits_{i=1}^{n}x_i^2$ 在条件 $\sum\limits_{i=1}^{n}x_i=c\left(x_i>0\right)$ 下的最小值为 $\dfrac{c^2}{n}$，即

$$x_1^2+x_2^2+\cdots+x_n^2\geq\frac{\left(x_1+x_2+\cdots+x_n\right)^2}{n}。$$

故利用 Lagrange 函数求多元函数的条件极值时，还可以用来证明不等式。

例 5 求函数 $f\left(x,y,z\right)=xy+2xz+2yz$ 在条件 $xyz=4$ 下的极值。

解 方法 1 转化为无条件极值(略)；

方法 2 Lagrange 乘数法

第一：构造 Lagrange 函数 $L\left(x,y,z,\lambda\right)=xy+2xz+2yz+\lambda\left(xyz-4\right)$；

第二：求 Lagrange 函数的驻点，解方程

$$\begin{cases} L_x=y+2z+\lambda yz=0 \\ L_y=x+2z+\lambda xz=0 \\ L_z=2x+2y+\lambda xy=0 \\ L_\lambda=xyz-4=0 \end{cases}\tag{6}$$

得 $x=2$，$y=2$，$z=1$，$\lambda=-2$，记 $M_0\left(2,2,1\right)$；当 $z=-\dfrac{1}{\lambda}$ 时，将其代入(6)得，$z=0$ (舍)。

第三：判断。

构造函数 $\bar{L}\left(x,y,z\right)=L\left(x,y,z,\lambda_0\right)=xy+2xz+2yz+\lambda_0\left(xyz-4\right)$；计算函数 $\bar{L}$ 在点 M_0 处的 Hessian 矩阵

$H\left(M_0\right)=\left.\begin{pmatrix} \bar{L}_{xx} & \bar{L}_{xy} & \bar{L}_{xz} \\ \bar{L}_{yx} & \bar{L}_{yy} & \bar{L}_{yz} \\ \bar{L}_{zx} & \bar{L}_{zy} & \bar{L}_{zz} \end{pmatrix}\right|_{M_0}=\begin{pmatrix} 0 & -1 & -2 \\ -1 & 0 & -2 \\ -2 & -2 & 0 \end{pmatrix}$，则 $H\left(M_0\right)$ 的各阶顺序主子式：

$$\left|H_1\right|=0，\left|H_2\right|=\begin{vmatrix} 0 & -1 \\ -1 & 0 \end{vmatrix}=-1<0，\left|H_3\right|=\begin{vmatrix} 0 & -1 & -2 \\ -1 & 0 & -2 \\ -2 & -2 & 0 \end{vmatrix}=8>0，$$

即 $H\left(M_0\right)$ 是不定矩阵。此时无法判断 M_0 是否为满足约束条件的极值点。

接下来使用二阶微分法判断。

$$\mathrm{d}\bar{L}\left(x,y,z\right)=y\mathrm{d}x+x\mathrm{d}y+2z\mathrm{d}x+2x\mathrm{d}z++2y\mathrm{d}z+2z\mathrm{d}y+\lambda_0\left(yz\mathrm{d}x+xz\mathrm{d}y+xy\mathrm{d}z\right)，$$

$$\mathrm{d}^2\bar{L}(x,y,z)=(2+2\lambda_0 z)\mathrm{d}x\mathrm{d}y+(4+2\lambda_0 y)\mathrm{d}x\mathrm{d}z+(4+2\lambda_0 x)\mathrm{d}z\mathrm{d}y\text{；}$$

当 $\lambda_0=-2$ 时，$\mathrm{d}^2\bar{L}\big|_{M_0}=-2\mathrm{d}x\mathrm{d}y-4\mathrm{d}x\mathrm{d}z-4\mathrm{d}y\mathrm{d}z$，此时无法判断 $\mathrm{d}^2\bar{L}\big|_{M_0}$ 的符号；结合条件 $xyz=4$，等式两边同时微分得，$yz\mathrm{d}x+xz\mathrm{d}y+xy\mathrm{d}z=0$，在点 $M_0(2,2,1)$ 处，$2\mathrm{d}z=-(\mathrm{d}x+\mathrm{d}y)$，将其代入 $\mathrm{d}^2\bar{L}\big|_{M_0}$ 得，

$$\mathrm{d}^2\bar{L}\big|_{M_0}=-2\mathrm{d}x\mathrm{d}y-4(\mathrm{d}x+\mathrm{d}y)\cdot\frac{\mathrm{d}x+\mathrm{d}y}{-2}=\mathrm{d}x^2+\mathrm{d}y^2+4\mathrm{d}z^2>0\text{，}\ (\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)\neq(0,0,0)\text{；}$$

故 $M_0(2,2,1)$ 是极小值点，极小值为 $f(2,2,1)=12$。

例 6 求函数 $f(x,y,z)=x^2-y^2-z^2$ 在条件 $y+z=0$ 下的极值。

解 方法 1 转化为无条件极值(略)；

方法 2 Lagrange 乘数法

第一：构造 Lagrange 函数 $L(x,y,z,\lambda)=x^2-y^2-z^2+\lambda(y+z)$；

第二：求 Lagrange 函数的驻点，解方程

$$\begin{cases}L_x=2x=0\\L_y=-2y+\lambda=0\\L_z=-2z+\lambda=0\\L_\lambda=y+z=0\end{cases}\tag{7}$$

解得 $x=y=z=\lambda=0$，记 $M_0(0,0,0)$。

第三：判断。

构造函数 $\bar{L}(x,y,z)=L(x,y,z,\lambda_0)=x^2-y^2-z^2$；函数 $\bar{L}$ 在点 M_0 处的 Hessian 矩阵

$H(M_0)=\left.\begin{pmatrix}\bar{L}_{xx}&\bar{L}_{xy}&\bar{L}_{xz}\\\bar{L}_{yx}&\bar{L}_{yy}&\bar{L}_{yz}\\\bar{L}_{zx}&\bar{L}_{zy}&\bar{L}_{zz}\end{pmatrix}\right|_{M_0}=\begin{pmatrix}2&0&0\\0&-2&0\\0&0&-2\end{pmatrix}$，显然 $H(M_0)$ 是不定矩阵。但在约束条件 $y+z=0$ 下 $f(x,y,z)=x^2-y^2-z^2=x^2-2y^2$，取 $M_1\left(\frac{\sqrt{3}}{8},\frac{1}{8},-\frac{1}{8}\right)$，$M_2\left(\frac{1}{8},\frac{1}{8},-\frac{1}{8}\right)$，则 $f(M_1)=\frac{1}{64}>f(0,0,0)=0$，$f(M_2)=-\frac{1}{64}<f(0,0,0)=0$，可知 $M_0(0,0,0)$ 不是满足约束条件的极值点。

3.4.2. 二次型理论在求解多元函数条件极值时所存在的局限性

由例 4 可知，利用二次型的正定性来判断多元函数的条件极值是一个非常便捷的方法，但也存在一定的局限性。因为二次型的正定、负定仅是 M_0 成为满足约束条件极值点的充分条件而非必要条件。比如例 5，很多学生会计算函数 $\bar{L}$ 在点 M_0 处的 Hessian 矩阵，但此时判别法失效。这里通过二阶微分法有效地判断出 M_0 是条件极值的极小值点。对于例 6，判别法也失效，这里根据目标函数本身的结构特征进行判断。

虽然二次型理论在多元函数极值和条件极值的应用上存在一定的局限性，但有必要进一步的探讨和研究。在判断多元函数的极值和条件极值时，为了克服这些不足，我们可根据具体问题情况选择恰当的方法进行分析和求解。

为巩固读者对二次型理论在多元函数极值和条件极值中应用的理解，提供以下题目供其练习。

1) 设实二次型 $f(x_1,x_2,\cdots,x_n)=X^{\mathrm{T}}AX$ 的矩阵 A 的最大特征值为 $\lambda_{\max}$，最小特征值为 $\lambda_{\min}$，证明：对于任意实数 $x_1,x_2,\cdots,x_n$，都有 $\lambda_{\min}X^{\mathrm{T}}X\le X^{\mathrm{T}}AX\le\lambda_{\max}X^{\mathrm{T}}X$。

2) 求函数 $f(x,y)=x^2y^2+x\ln x$ 的极值。

3) 证明 $\dfrac{x^n+y^n}{2}\ge\left(\dfrac{x+y}{2}\right)^n(x,y\ge 0)$。

4) 求函数 $f(x,y,z)=x-2y+2z$ 在条件 $x^2+y^2+z^2=1$ 下的极值。

4. 结束语

二次型理论在多元函数极值和条件极值的计算及解析过程中有广泛的应用，是最基础也是最重要的工具之一。掌握二次型理论，有助于深入理解相关领域的知识，提高分析和计算能力，达成对实际问题的准确把握。

但在实际应用中发现二次型理论也存在一些不足。例如，二次型理论没有给出多元函数极值和条件极值的具体计算公式，这在多元函数极值和条件极值中是十分重要的。又如对于高阶甚至非二次的函数，二次型理论就很难适用。因此，在处理复杂问题时需要借助更加高级的技巧或者通过计算机进行数值模拟分析。另外二次型理论通常只能处理实域函数，在特定领域中无法完全满足处理比如虚函数等复杂问题的需求。

综上，虽然二次型理论在多元函数极值和条件极值中是一种非常有效的工具，但在实际运用时，有必要将二次型理论应用到实际问题中去，同时还需根据特殊情况和需求来选择合适的数学模型和相应的工具以达到更好的效果。

基金项目

信息工程大学教育教学研究课题(JXYJ2022C010)，强军新工科一般课题：强联[2022] 2 号。

参考文献

[1] 米少田. 二次型理论在多元极值问题上应用的剖析[J]. 天津商学院学报, 1997(4): 26-32.

[2] 梁伟秋, 尹小艳. 二次型矩阵在多元可微函数极值问题的应用[J]. 高师理科学刊, 2017, 37(8): 28-30.

[3] 董丽华. 利用正定矩阵法解决多元函数的极值问题[J]. 连云港职业技术学院学报(综合版), 2016, 19(1): 59-60.

[4] 徐阳栋. 二次型在多元函数极值问题上的应用[J]. 教育教学论坛, 2015(28): 180-181.

[5] 卢里举. 基于正定矩阵的多元函数极值判定及应用[J]. 现代职业教育, 2018(27): 148-149.

[6] 刘爱德, 吕蕴霞, 刘华巧. 用矩阵讨论多元函数的极值问题[J]. 烟台师范学院学报(自然科学版), 2002, 18(4): 306-309+313.

[7] 陈荣群, 黄勇. 二次型在求条件极值中的应用[J]. 福建教育学院学报, 2018, 9(10): 98-101.

[8] 黄朝军. 二次型矩阵与函数的极值[J]. 凯里学院学报, 2011, 29(6): 7-9.

[9] 杨廷鸿, 林琮, 但琦, 付诗禄. 条件极值与矩阵特征值的结合[J]. 高等数学研究, 2012, 15(4): 30-33.

[10] 王萼芳, 石生明. 高等代数[M]. 第 5 版. 北京: 高等教育出版社, 2019.

[11] 陈纪修, 於崇华, 金路. 数学分析下册[M]. 第 2 版. 北京: 高等教育出版社, 2004.