第 40 卷第 6 期　　西南大学学报(自然科学版)　　2018 年 6 月
Vol. 40 No. 6　　Journal of Southwest University (Natural Science Edition)　　Jun. 2018

DOI: 10.13718/j.cnki.xdzk.2018.06.015

一类条件极值判定的拉格朗日方法①

吴燕春[1]，　胡　凯[2]

1. 重庆市南开中学，重庆 400030；2. 西南大学 数学与统计学院，重庆 400715

摘要：研究了多元函数条件极值的充分条件与判定方法，建立了可行集 $\mathscr{M}$ 与变分空间 $V(P_0)$ 上的点之间的联系. 利用二者间关键的误差估计，给出了条件极值判定定理的简化证明.

关　键　词：条件极值；拉格朗日乘数法；充分条件

中图分类号：O177.99　　**文献标志码**：A　　**文章编号**：1673-9868(2018)06-0097-04

条件极值问题是分析学和运筹学当中的重要内容，它主要是对一定约束条件下目标函数的求解，例如求解满足约束条件 $g_i(x_1, x_2, \cdots, x_n)=0\ (i=1,2,\cdots,k)$ 的目标函数 $\boldsymbol{z}=f(x_1, x_2, \cdots, x_n)$. 这里 f, g_i 都是二阶连续可微的函数，$1 \leqslant k < n$. 我们记向量值函数 $\boldsymbol{g}=(g_1, g_2, \cdots, g_k)^{\mathrm{T}}$，梯度算子 $D=\left(\dfrac{\partial}{\partial x_1}, \cdots, \dfrac{\partial}{\partial x_n}\right)$. 当 $D\boldsymbol{g}$ 行满秩时，约束条件 $\boldsymbol{g}=\boldsymbol{0}$ 确定了一个隐函数组 $\boldsymbol{y}=\boldsymbol{y}(\boldsymbol{x})$. 不妨设该隐函数组的自变量和因变量分别为 $\boldsymbol{x}=(x_1, x_2, \cdots, x_m)$，$\boldsymbol{y}=(x_{m+1}, x_{m+2}, \cdots, x_n)$，自然数 $m=n-k$. 于是原条件极值问题可记为

$$\boldsymbol{z}=f(\boldsymbol{x}, \boldsymbol{y}) \qquad \text{s.t.}\ \ \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})=\boldsymbol{0} \tag{1}$$

我们知道 $\boldsymbol{z}=f(\boldsymbol{x}, \boldsymbol{y})$ 的无约束极值可根据 Hessian 矩阵 $\boldsymbol{H}_f=\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}\right)_{n\times n}$ 来进行判断[1]. 然而条件极值问题(1)却没有完整的判定定理. 因为当拉格朗日函数

$$L(P, \lambda)=f(P)+\sum_i \lambda_i g_i(P)$$

关于 $P(\boldsymbol{x}, \boldsymbol{y})$ 的 Hessian 矩阵 $\boldsymbol{H}_L$ 不定时，无法确定 f 的极值情况. 文献[2]给出了相应的反例：

$$f(x_1, x_2, x_3)=x_1{}^2+x_2{}^2-x_3{}^2 \qquad \text{s.t.}\ \ g(x_1, x_2, x_3)=x_3=0$$

此时 H_L 在驻点 $(x_1, x_2, x_3, \lambda)=(0, 0, 0, 0)$ 处是不定的，然而上述问题的 f 却在点 $P(x_1, x_2, x_3)=(0, 0, 0)$ 处取得极小值. 这说明了条件极值问题在求解与判定方面都更加复杂和困难. 约束极值的研究价值不仅限于多元函数，它在泛函的临界点理论中也有着重要的应用，具体案例参见文献[3-5]. 本文将结合驻点的一阶可行变分空间，研究条件极值的二阶判定方法.

记可行集 $\mathscr{M}=\{(\boldsymbol{x}, \boldsymbol{y}): \boldsymbol{g}(\boldsymbol{x}, \boldsymbol{y})=\boldsymbol{0}\}$. 假设点 $P_0(\boldsymbol{x}, \boldsymbol{y}) \in \mathscr{M}$ 是 f 在 $\mathscr{M}$ 上的极值点，则存在拉格朗日乘子 $\lambda=(\lambda_1, \lambda_2, \cdots, \lambda_k)$，使得

$$DL(P_0, \lambda)=Df(P_0)+\sum_{i=1}^{k} \lambda_i Dg_i(P_0)=0 \tag{2}$$

---

① 收稿日期：2017-09-03

基金项目：国家自然科学基金项目(11426185).

作者简介：吴燕春(1978-)，男，高级教师，主要从事泛函分析与数学教学的研究.

记增量 $\Delta\boldsymbol{x}=(\Delta x_1,\cdots,\Delta x_m)$，$\Delta\boldsymbol{y}=(\Delta x_{m+1},\cdots,\Delta x_n)$. 若点 $P(\boldsymbol{x}+\Delta\boldsymbol{x},\boldsymbol{y}+\Delta\boldsymbol{y})\in\mathscr{M}$，则利用泰勒展开式可知：

$$f(P)-f(P_0)=Df(P_0)(\Delta\boldsymbol{x},\Delta\boldsymbol{y})^{\mathrm{T}}+\frac{1}{2}(\Delta\boldsymbol{x},\Delta\boldsymbol{y})\boldsymbol{H}_f(P_0)(\Delta\boldsymbol{x},\Delta\boldsymbol{y})^{\mathrm{T}}+ o(\|\Delta\boldsymbol{x}\|^2+\|\Delta\boldsymbol{y}\|^2) \tag{3}$$

$$0=g_i(P)-g_i(P_0)=Dg_i(P_0)(\Delta\boldsymbol{x},\Delta\boldsymbol{y})^{\mathrm{T}}+\frac{1}{2}(\Delta\boldsymbol{x},\Delta\boldsymbol{y})\boldsymbol{H}_{g_i}(P_0)(\Delta\boldsymbol{x},\Delta\boldsymbol{y})^{\mathrm{T}}+ o(\|\Delta\boldsymbol{x}\|^2+\|\Delta\boldsymbol{y}\|^2) \tag{4}$$

其中 $1\leqslant i\leqslant k$. 将(4)式乘以相应的 λ_i，再累加到(3)式上，结合(2)式可得

$$f(P)-f(P_0)=\frac{1}{2}(\Delta\boldsymbol{x},\Delta\boldsymbol{y})\boldsymbol{H}_L(P_0)(\Delta\boldsymbol{x},\Delta\boldsymbol{y})^{\mathrm{T}}+o(\|\Delta\boldsymbol{x}\|^2+\|\Delta\boldsymbol{y}\|^2) \tag{5}$$

由(5)式可以看出，仅凭借 $\boldsymbol{H}_L(P_0)$ 在 $\mathbb{R}^n$ 上的正定(负定)性来判断极值是十分粗略的，有些情形下甚至是失效的. 这是因为点 P 在可行集上变动时，对应的增量 $(\Delta\boldsymbol{x},\Delta\boldsymbol{y})$ 仅是 $\mathbb{R}^n$ 的一个子集. 我们记 $\mathscr{M}$ 上点 P_0 的一阶可行变分子空间为

$$V(P_0)=\{(\delta_x,\delta_y)\in\mathbb{R}^m\times\mathbb{R}^k: D\boldsymbol{g}\cdot(\delta_x,\delta_y)^{\mathrm{T}}=\boldsymbol{0}\}$$

进而得到如下结论.

定理1 设函数 $f,g_i(i=1,\cdots,k)$ 均在点 $P_0\in\mathbb{R}^n$ 的某邻域内二阶连续可微，且矩阵 $D\boldsymbol{g}$ 的秩为 k. 若点 $(P_0,\boldsymbol{\lambda})$ 是拉格朗日函数 $L(P,\lambda)=f(P)+\sum\limits_i\lambda_i g_i(P)$ 的驻点，则：

(a) 当 $\boldsymbol{H}_L(P_0)$ 在空间 $V(P_0)$ 上正定时，函数 f 在点 P_0 取条件极小值；

(b) 当 $\boldsymbol{H}_L(P_0)$ 在空间 $V(P_0)$ 上负定时，函数 f 在点 P_0 取条件极大值；

(c) 当 $\boldsymbol{H}_L(P_0)$ 在空间 $V(P_0)$ 上不定时，函数 f 在点 P_0 一定不取极值.

注 文献[6-7]提出了结论(a),(b)的一种证明方法. 但是文献[6-7]中的方法较为繁琐，不易推广，无法解决 $\boldsymbol{H}_L(P_0)$ 不定时的判定. 于是本文建立了一个新的估计式，它既给出了结论(a),(b)的简洁证明，同时得到更丰富的结果.

我们记多元函数的偏微分算子：

$$D_x=\left(\frac{\partial}{\partial x_1},\cdots,\frac{\partial}{\partial x_m}\right)$$

$$D_y=\left(\frac{\partial}{\partial x_{m+1}},\cdots,\frac{\partial}{\partial x_n}\right)$$

则向量值函数 $\boldsymbol{g}$ 的偏微分记为：

$$D_x\boldsymbol{g}=(D_xg_1,\cdots,D_xg_k)^{\mathrm{T}} \qquad D_y\boldsymbol{g}=(D_yg_1,\cdots,D_yg_k)^{\mathrm{T}}$$

一般情况下，点 P 在 P_0 的邻域 $N(P_0)\cap\mathscr{M}$ 内变动时，自变量的增量 $(\Delta\boldsymbol{x},\Delta\boldsymbol{y})$ 与 $(\delta_x,\delta_y)\in V(P_0)$ 是不同的两个量. 我们需要先建立估计式，找到集合 $\mathscr{M}$ 与空间 $V(P_0)$ 的联系.

命题1 任何 $(\delta_x,\delta_y)\in V(P_0)$ 和实数 t，当 $\Delta\boldsymbol{x}=t\delta_x$ 时，增量 $\Delta\boldsymbol{y}=\boldsymbol{y}(\boldsymbol{x}+t\delta_x)-\boldsymbol{y}(\boldsymbol{x})$ 满足 $\Delta\boldsymbol{y}=t\delta_y+o(t)\cdot\boldsymbol{I}_{1\times n}$，其中 $\boldsymbol{I}_{1\times n}$ 是 $1\times n$ 阶的全1矩阵.

证 对 $\boldsymbol{g}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0}$ 两端关于 $\boldsymbol{x}$ 求偏微分，得

$$D_x\boldsymbol{g}+D_y\boldsymbol{g}\cdot D_x\boldsymbol{y}^{\mathrm{T}}=\boldsymbol{0}$$

于是

$$D_x\boldsymbol{y}^{\mathrm{T}}=-(D_y\boldsymbol{g})^{-1}\cdot D_x\boldsymbol{g}$$

由 $(\delta_x,\delta_y)\in V(P_0)$ 知

$$D_x\boldsymbol{g}\cdot\delta_x^{\mathrm{T}}+D_y\boldsymbol{g}\cdot\delta_y^{\mathrm{T}}=\boldsymbol{0}$$

于是

$$\delta_y^{\mathrm{T}} = -(D_y \boldsymbol{g})^{-1} \cdot D_x \boldsymbol{g} \cdot \delta_x^{\mathrm{T}} = D_x \boldsymbol{y}^{\mathrm{T}} \cdot \delta_x^{\mathrm{T}}$$

利用多元函数的泰勒展开式得：

$$y_i = y_i(\boldsymbol{x} + t\delta_x) - y_i(\boldsymbol{x}) = D_x y_i \cdot t\delta_x^{\mathrm{T}} + \frac{1}{2} t\delta_x \cdot \boldsymbol{H}_{y_i}(\xi_i) \cdot t\delta_x^{\mathrm{T}} = D_x y_i \cdot t\delta_x^{\mathrm{T}} + o(t)$$

$$\Delta \boldsymbol{y} = (D_x \boldsymbol{y}^{\mathrm{T}} \cdot t\delta_x^{\mathrm{T}})^{\mathrm{T}} + o(t) \cdot \boldsymbol{I}_{1\times n} = t\delta_y + o(t) \cdot \boldsymbol{I}_{1\times n}$$

其中 ξ_i 是 $\boldsymbol{x}$ 与 $\boldsymbol{x} + t\delta_x$ 之间的中值点，$i = 1, \cdots, k$. 证毕.

定理 1 的证明

(a) 已知 $\boldsymbol{H}_L(P_0)$ 在空间 $V(P_0)$ 上正定. 于是在驻点 $P_0(\boldsymbol{x}, \boldsymbol{y})$ 的 ε -邻域 $N_\varepsilon(P_0)$ 内任取一点 $P(\boldsymbol{x} + \Delta \boldsymbol{x}, \boldsymbol{y} + \Delta \boldsymbol{y}) \in \mathscr{M}$，令：

$$\bar{\delta}_x = \Delta \boldsymbol{x}$$

$$\bar{\delta}_y = -[(D_y \boldsymbol{g})^{-1} \cdot D_x \boldsymbol{g} \cdot \Delta \boldsymbol{x}^{\mathrm{T}}]^{\mathrm{T}}$$

可知

$$(\bar{\delta}_x, \bar{\delta}_y) \in V(P_0)$$

记

$$t = \sqrt{\|\bar{\delta}_x\|^2 + \|\bar{\delta}_y\|^2}$$

显然当 $\varepsilon \to 0$ 时，$\|\Delta \boldsymbol{x}\|$ 和 t 都收敛到 0. 令：

$$\delta_x = t^{-1} \bar{\delta}_x \qquad \delta_y = t^{-1} \bar{\delta}_y \qquad \boldsymbol{Q}(\delta_x, \delta_y) = (\delta_x, \delta_y) \boldsymbol{H}_L(P_0) (\delta_x, \delta_y)^{\mathrm{T}}$$

运用(5) 式与命题 1 可得

$$f(P) - f(P_0) = \frac{1}{2}(t\delta_x, t\delta_y) \boldsymbol{H}_L(P_0)(t\delta_x, t\delta_y)^{\mathrm{T}} + o(t^2) + o(\|t\delta_x\|^2 + \|t\delta_y\|^2) =$$

$$\frac{1}{2} t^2 \boldsymbol{Q}(\delta_x, \delta_y) + o(t^2) \tag{6}$$

因为二次型 $\boldsymbol{Q}(\delta_x, \delta_y)$ 在单位闭球上一定能取到最小值，所以存在正数 $\boldsymbol{Q}_m > 0$，使得 $\boldsymbol{Q}(\delta_x, \delta_y) \geqslant \boldsymbol{Q}_m > 0$. 注意到(6) 式中 $o(t^2)$ 关于 ε 的一致性. 我们推出：只要 ε 充分小，则任何 $P \in N_\varepsilon(P_0)$ 都满足

$$f(P) - f(P_0) > \frac{1}{4} t^2 \boldsymbol{Q}_m > 0$$

即 f 在 P_0 取到极小值.

(b) 证明与(a) 同理.

(c) 已知 $\boldsymbol{H}_L(P_0)$ 在空间 $V(P_0)$ 上不定，则存在两个线性无关的向量(δ_x^+, δ_y^+)，$(\delta_x^-, \delta_y^-) \in V(P_0)$，使得

$$\boldsymbol{Q}(\delta_x^+, \delta_y^+) > 0 \qquad \boldsymbol{Q}(\delta_x^-, \delta_y^-) < 0$$

对任意的 $t > 0$，记：

$$\boldsymbol{x}_t = \boldsymbol{x} + t\delta_x^+ \qquad \boldsymbol{y}_t = \boldsymbol{y}(\boldsymbol{x} + t\delta_x^+)$$

则增量 $\Delta \boldsymbol{x} = t\delta_x^+$，$\Delta \boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x} + t\delta_x^+) - \boldsymbol{y}(\boldsymbol{x})$. 取点 $P_t(\boldsymbol{x}_t, \boldsymbol{y}_t) \in \mathscr{M}$，利用命题 1 知，当 t 充分小时，有

$$f(P_t) - f(P_0) = \frac{1}{2}(\Delta \boldsymbol{x}, \Delta \boldsymbol{y}) \boldsymbol{H}_L(P_0)\ (\Delta \boldsymbol{x}, \Delta \boldsymbol{y})^{\mathrm{T}} + o(\|\Delta \boldsymbol{x}\|^2 + \|\Delta \boldsymbol{y}\|^2) =$$

$$\frac{1}{2}(t\delta_x^+, t\delta_y^+) \boldsymbol{H}_L(P_0)\ (t\delta_x^+, t\delta_y^+)^{\mathrm{T}} + o(t^2) + o(\|t\delta_x^+\|^2 + \|t\delta_y^+\|^2) =$$

$$\frac{1}{2} t^2 \boldsymbol{Q}(\delta_x^+, \delta_y^+) + o(t^2) > 0$$

因此在点 P_0 的任何邻域内都存在点 $P_t \in \mathscr{M}$，使得 $f(P_t) > f(P_0)$. 同理利用 $\boldsymbol{Q}(\delta_x^-, \delta_y^-) < 0$ 可知，存在点 $P'_t \in \mathscr{M}$，使得 $f(P'_t) < f(P_0)$. 综上所述，函数 f 在点 P_0 一定不取条件极值.

参考文献:

[1] 华东师范大学数学系. 数学分析 [M]. 4版. 北京:高等教育出版社,2010.

[2] 张筑生. 数学分析新讲 [M]. 北京:北京大学出版社,1990.

[3] HU K,TANG C L. Existence of Positive Solutions of Shroinger Equations with Potentials in Nonlinearity [J]. 西南大学学报(自然科学版),2010,32(6):129-134.

[4] 刘 磊,吴行平,唐春雷. 一类带临界指数的半线性椭圆方程多重变号解的存在性 [J]. 西南师范大学学报(自然科学版),2013,38(5):5-9.

[5] 吕 颖,唐春雷. 一类次二次二阶 Hamilton 系统的同宿轨 [J]. 西南大学学报(自然科学版),2013,35(4):84-87.

[6] AVRIEL M. Nonlinear Programming:Analysis and Methods [M]. New Jersey:Prentice-Hall,1976.

[7] BERTSEKAS D P. Nonlinear Programming [M]. 2th ed. New Hampshire:Athena Scientific,1999.

The Lagrangian Method to a Class of Conditional Extremum Problems

WU Yan-chun[1], HU Kai[2]

1. Chongqing Nankai Middle School, Chongqing 400030, China;
2. School of Mathematics and Statistics, Southwest University, Chongqing 400715, China

Abstract: In this paper, the sufficient conditions and determination method of the conditional extremum of multivariate functions are studies, and a connection is established between the points on the feasible set $\mathscr{M}$ and the variational set $V(P_0)$. By a critical error estimate of the two sets, the proof of the conditional extremum theorem is remarkably simplified.

Key words: conditional extremum; Lagrange multiplier method; sufficient condition

责任编辑 廖 坤