文章编号:1000-0887(2012)02-0250-15

一种新的求解带约束的有限极大极小问题的精确罚函数*

马 骋[1], 李 迅[1], 姚家晖[1], 张连生[2]

(1. 香港理工大学 应用数学系, 九龙,香港;
2. 上海大学 数学系, 上海 200444)

摘要: 提出了一种新的精确光滑罚函数求解带约束的极大极小问题. 仅仅添加一个额外的变量,利用这个精确光滑罚函数,将带约束的极大极小问题转化为无约束优化问题. 证明了在合理的假设条件下,当罚参数充分大,罚问题的极小值点就是原问题的极小值点. 进一步,研究了局部精确性质. 数值结果表明这种罚函数算法是求解带约束有限极大极小问题的一种有效算法.

关 键 词: 带约束的极大极小问题; 约束优化问题; 罚函数

中图分类号: O221.2 **文献标志码:** A

DOI: 10.3879/j. issn. 1000-0887. 2012. 02. 010

引 言

极大极小规划问题有广泛的应用. 许多决策问题可以转化为极大极小问题. 大量的金融、工程管理、经济问题同样都是转化成极大极小规划模型来求解的. 有兴趣的读者可以在文献[1-2]中找到极大极小模型在金融、经济学中的实际应用的范例. 一些最优控制问题[3-5]也是转化成极大极小问题的.

本文中,我们考虑下面的带等式约束的极大极小问题:

$$(\bar{\mathrm{P}})\quad \begin{cases} \min \max\limits_{1\leqslant j\leqslant q} f_j(x) \\ \text{s. t. } F_j(x)=0, \qquad \forall j=q+1,\cdots,m, \\ \qquad x\in R^n, \end{cases}$$

其中, $f_j: R^n \to \mathbf{R}$, $j=1,2,\cdots,q$ 和 F_j: $R^n \to \mathbf{R}$, $j=q+1,\cdots,m$ 是连续可微的函数.

由于目标函数中包含极大项,因此它是连续但不是可微的,即使 $f_j(x)$, $j=1,2,\cdots,q$ 是连续可微函数. 因此,我们不能直接使用带导数的无约束优化算法来求解这个问题. 在已有的文献中,有几种求解极大极小问题的算法. 它们可以被分为下面两类:问题($\bar{\mathrm{P}}$)可以看作是带约束的非光滑优化问题, 因此,我们可以使用求解非光滑优化问题的一般算法,例如次梯度方法、捆绑法、切平面法(如文献[6-8]). 另一类方法是,考虑不可微的特定结构目的是利用某种光滑优化算法. 其中,包括利用光滑函数来正则极大项的技巧(见文献[9-13]). Ye 等[12]借助

* **收稿日期:** 2011-03-31; **修订日期:** 2011-11-23
基金项目: AMSS-PolyU 联合研究所资助项目
作者简介: 马骋(1983—),男,山东青岛人,博士生(联系人. E-mail: mc_0812@163.com).

于指数罚函数近似方法,提出光滑信赖域 Newton-CG 算法. DiPillo 等[9]首先将精确罚函数方法推广到求解带约束的极大极小问题中去. 同样,也存在一些其他方法是基于引入一个新的变量 $\theta \in \mathbf{R}$ 转化成下面等价的非线性规划问题来求解的:

$$(\mathrm{P})\quad \begin{cases} \min \theta \\ \text{s. t.} \quad f_j(x) \leqslant \theta, & \forall j = 1,2,\cdots,q, \\ \qquad F_j(x) = 0, & \forall j = q+1,\cdots,m, \\ \qquad (x,\theta) \in R^{n+1}. \end{cases}$$

很明显,求解带约束的有限极大极小规划问题($\bar{\mathrm{P}}$)等价于求解问题(P).

例如,Zhou 等[14]和 Zhu 等[15]基于问题(P)提出了求解极大极小规划问题的 SQP 算法,并且讨论了一些性质. Obasanjo 等[16]提出了求解极大极小规划问题的内点算法. 基于非线性规划问题(P)以及连续可微的精确障碍罚函数的构造,对于有限的罚参数值,罚问题的最小值点就是带约束的极大极小问题的最优值点.

尽管传统的罚函数方法是受欢迎的方法,但是也存在着缺陷. 如果罚函数是精确和光滑的,那么它就不是简单的;如果罚函数是简单和光滑的,它就不是精确的. 这里,"简单"的含义是指罚函数仅仅包括原问题的目标函数和约束函数. 为了说明这一点,我们举一个传统的约束优化问题例子.

$$\begin{cases} \min f(x) \\ \text{s. t.} \quad F_j(x) = 0, & \forall j \in E, \\ \qquad g_\ell(x) \leqslant 0, & \forall \ell \in I, \\ \qquad x \in R^n, \end{cases}$$

其中, $f: R^n \to \mathbf{R}$, $F_j: R^n \to \mathbf{R}$, $\forall j \in E$ 和 $g_\ell: R^n \to \mathbf{R}$, $\forall \ell \in I$ 是连续可微的函数, E, I 分别表示等式约束和不等式约束的指标集.

我们给出下面几种传统的罚函数:

$$f_\sigma(x) = f(x) + \sigma\left(\sum_{j\in E} | F_j(x) | + \sum_{\ell \in I} \max(0, g_\ell(x))\right), \tag{1}$$

$$f_\sigma(x) = f(x) + \sigma\left(\sum_{j\in E} F_j^2(x) + \sum_{\ell \in I} (\max(0, g_\ell(x)))^2\right) \tag{2}$$

和

$$f_\sigma(x) = f(x) + \boldsymbol{\lambda}^{\mathrm{T}}(x)\boldsymbol{F}(x) + \frac{\sigma}{2}\boldsymbol{F}^{\mathrm{T}}(x)\boldsymbol{F}(x), \tag{3}$$

其中,罚参数 $\sigma > 0$, $\boldsymbol{F}(x) := (\boldsymbol{F}_j(x), j \in E)$, $\boldsymbol{\lambda}(x) = (\nabla \boldsymbol{F}(x))^{+} \nabla f(x)$, $(\nabla \boldsymbol{F}(x))^{+}$ 是 $\nabla \boldsymbol{F}(x)$ 的广义逆矩阵.

对于上面的 3 类罚函数,我们知道:罚函数(1)是非光滑简单精确罚函数; 罚函数(2)是光滑简单罚函数,但不是精确的; 罚函数(3)是光滑精确罚函数,但不是简单的.

近来,在文献[17]中,对于等式约束优化问题, 通过引入一个新的变量,提出了一种新的精确罚函数来求解等式约束优化问题如下:

$$\min_{x \in S} f(x),\ T = \{x \in [u,v] : F_j(x) = 0,\ \forall j \in E\} \neq \varnothing, \tag{4}$$

其中, $[u,v] \in R^n$ 是具有非空内部的箱式约束, $[u,v] = \{x \in R^n : u \leqslant x \leqslant v\}$, $(\{-\infty\} \cup \mathbf{R})^n \leqslant u < v \leqslant (\{+\infty\} \cup \mathbf{R})^n$, $f: D \to \mathbf{R}$ 和 $F_j: D \to \mathbf{R}$, $\forall j \in E$ 都是定义在开集 D 上的连续可微函数,其中 D 包含 $[u,v]$. 固定 $w_j \in \mathbf{R}$, $\forall j \in E$ 考虑下面的等价问题:

$$\min_{(x,\varepsilon)\in S_{\varepsilon_0}} f(x),\ S_{\varepsilon_0}=\{(x,\varepsilon)\in[u,v]\times[0,\bar{\varepsilon}]: F_j(x)=\varepsilon w_j,\ \forall j\in E,\varepsilon=0\}=S\times\{0\}, \tag{5}$$

其中 $\bar{\varepsilon}>0$ 是固定数. 令

$$f_\sigma(x,\varepsilon)=\begin{cases} f(x), & \text{如果 } \varepsilon=0,\ x\in S, \\ f(x)+\dfrac{1}{2\varepsilon}\dfrac{\Delta(x,\varepsilon)}{1-q\Delta(x,\varepsilon)}+\sigma\beta(\varepsilon), & \text{如果 } 0<\varepsilon\leqslant\bar{\varepsilon}, \\ & \quad \Delta(x,\varepsilon)<q^{-1},q>0, \\ +\infty, & \text{否则.} \end{cases} \tag{6}$$

相应的罚问题可表示成

$$\min_{(x,\varepsilon)\in[u,v]\times[0,\bar{\varepsilon}]} f_\sigma(x,\varepsilon), \tag{7}$$

这里,约束违背测度表示为

$$\Delta(x,\varepsilon)=\sum_{j\in E}(F_j(x)-\varepsilon w_j)^2.$$

在较弱的条件下,文献[17]表明,非线性规划问题(4)的局部最优值点就是相应的罚问题的最优值点. 然而,文献[17]证明了,对于充分大的罚参数 $\sigma>0$, 当 $\varepsilon>0$ 时,罚问题(7)不存在 Karush-Kuhn-Tucker 点 (x,ε). 但是,我们设计算法得到的一般就是 Karush-Kuhn-Tucker 点,这是可以计算的,因此这是求解原问题(4)的一个实际问题.

受文献[17]的启发,本文中,针对极大极小问题,我们提出一种新的精确光滑罚函数,这就是本文的写作目的. 这种罚函数的主要特点是我们仅仅增加一个变量 ε 到约束中. 效益函数可以看作是 x 和 ε 的函数,即使不需要梯度和 Jacobi 矩阵的信息,也具有好的光滑性质和精确性质. 在 S 上,只要 $f(x)$ 是下有界的,罚函数就是下有界的,这一点是 l_1 精确罚函数所不具备的. 众所周知,罚参数过大会给计算结果带来不好的影响. 因此,本文中,对于算法,我们仅仅要求罚参数增加较小的常数目的是保持罚参数尽可能的小,避免变态的情况发生. 我们将会给出,如果罚问题的局部最优解处,满足线性独立约束规范,那么局部极小值点可以被表述成 $(x^*,\theta^*,0)$ 的形式. 另外,我们推出一个重要的结论,罚问题的极小值点 $(x^*,\theta^*,\varepsilon^*)$ 满足 $\varepsilon^*=0$ 当且仅当 (x^*,θ^*) 是原问题(P)的最优值点,其中 θ^* 是最优目标函数值. 这个性质表明引入的新变量 ε 可以被看作是原问题(P)的局部(全局)最小值点的指示变量. 除了具有上述的性质之外,我们指出这个罚问题还具有精确性,也就是说,存在一个有限数,当罚参数大于这个有限数时,罚问题的最小值点就是原问题的最小值点. 进一步,局部精确性质也将会被给出,这里目标函数和约束函数未必要求是光滑的.

本文结构安排如下:第 1 节,我们介绍求解带约束的极大极小规划问题的光滑精确罚函数;第 2 节,我们将会给出罚函数算法及其收敛性分析;第 3 节,讨论这个新的精确光滑罚函数的局部精确性质;第 4 节,数值结果;第 5 节,总结本文.

1 一种新的光滑精确罚函数

记

$$S=\{(x,\theta)\in R^{n+1}: f_j(x)-\theta\leqslant 0,\ \forall j=1,2,\cdots,q;\ F_j(x)=0,\ \forall j=q+1,\cdots,m\}. \tag{8}$$

正如在引言中所陈述的,通过引入变量 θ, 我们将规划问题(P)等价写作下面的优化问题:

$$\min_{(x,\theta,\varepsilon)\in S_0}\theta,\ S_0=\{(x,\theta,\varepsilon)\in R^{n+2}: f_j(x)-\theta\leqslant\varepsilon^\gamma w_j,\ \forall j=1,2,\cdots,q;$$

$$F_j(x)=\varepsilon^{\gamma}w_j,\ \forall j=q+1,\cdots,m,\varepsilon=0\},$$

其中，$w_j\in(0,1)$，$j=1,2,\cdots,m$. 类似地，我们记作

$$S_\varepsilon=\{(x,\theta,\varepsilon)\in R^{n+2}: f_j(x)-\theta\leqslant\varepsilon^{\gamma}w_j,\forall j=1,2,\cdots,q;$$
$$F_j(x)=\varepsilon^{\gamma}w_j,\forall j=q+1,\cdots,m\}.$$

我们做如下假设：

1）存在问题(P)的全局最优值点，这意味着 θ 在集合 S 上是下有界的；

2）如果 $(x^*,\theta^*)\in L(P)$，那么 $L_{(x^*,\theta^*)}=\{(x,\theta)\in L(P):\theta=\theta^*\}$ 是紧致集合，这里 $L(P)$ 表示问题(P)的局部最优值点集合.

我们建立求解极大极小问题的罚函数如下：

$$f_\sigma(x,\theta,\varepsilon)=\begin{cases}\theta, & \text{如果 } \varepsilon=0,(x,\theta)\in S,\\ \theta+\dfrac{\varepsilon^{-\alpha}\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\sigma\varepsilon^{\beta}, & \\ & \text{如果 } \varepsilon\neq 0,\ 0<1-2\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)<1,\\ +\infty, & \text{否则},\end{cases}\tag{9}$$

其中

$$\Delta(x,\theta,\varepsilon)=\sum_{j=1}^{q}(\max(f_j(x)-\theta-\varepsilon^{\gamma}w_j,0))^2+\sum_{j=q+1}^{m}(F_j(x)-\varepsilon^{\gamma}w_j)^2,$$

$\alpha,\beta,\gamma,\delta$ 是正偶数. 相应的罚问题(P_σ)是

$$(\mathrm{P}_\sigma)\quad \min_{(x,\theta,\varepsilon)\in R^{n+1}\times(-\bar\varepsilon,\bar\varepsilon)} f_\sigma(x,\theta,\varepsilon).$$

对于 $\varepsilon\neq 0,0<1-2\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)<1$，我们有

$$f_\sigma(x,\theta,\varepsilon)=\theta+\frac{\varepsilon^{-\alpha}\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\sigma\varepsilon^{\beta}\geqslant\theta,$$

当 $F_j(x)=0,\ \forall j=q+1,\cdots m$.

因此，$f_\sigma(x,\theta,\varepsilon)$ 在 $R^{n+1}\times[-\bar\varepsilon,\bar\varepsilon]$ 上是下有界的，只要 $f_j(x)$，$\forall j=1,2,\cdots,q$ 在集合

$$D'=\{x\in R^n\mid \|F(x)\|\leqslant\frac{\sqrt{2}}{2}\bar\varepsilon^{\delta}+\bar\varepsilon^{\gamma}\|w\|\}$$

上是下有界的. 这是合理的条件，因为当 $f_j(x)$，$\forall j=1,2,\cdots,q$ 在可行集合上是下有界的，$\bar\varepsilon$ 是充分小的. 下面，我们将刻画出，在较弱的条件下，$f_\sigma(x,\theta,\varepsilon)$ 是连续可微的并且在具有有限函数值的边界上具有连续的极限.

命题 1.1　令 $(x,\theta)\to(x^*,\theta^*)\in S,0\neq\varepsilon\to\varepsilon^*=0$. 假设

$$\begin{cases}\gamma>\delta>\alpha>0,\\ -\alpha-1+2\delta>0,\\ \beta>1,\end{cases}\tag{10}$$

那么

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} f_\sigma(x,\theta,\varepsilon)=f_\sigma(x^*,\theta^*,0)=\theta^*,\qquad \lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}}\nabla_x f_\sigma(x,\theta,\varepsilon)=0,$$

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}}\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\varepsilon}=0,\qquad \lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}}\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\theta}=1.$$

证明　由 $\varepsilon\neq0,0<1-2\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)<1$ 可知，我们有 $\Delta(x,\theta,\varepsilon)=O(\varepsilon^{2\delta})$，$|f_j(x)-\theta-\varepsilon^{\gamma}w_j|=O(\varepsilon^{\delta})$，$\forall j=1,2,\cdots,q$；$|F_j(x)-\varepsilon^{\gamma}w_j|=O(\varepsilon^{\delta})$，$\forall j=q+1,\cdots,m$ 和

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} 1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)=c^*\in[1/2,1].$$

从式(10),我们有 $2\delta>\alpha$ 和 $\beta>1$. 因此可知,

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} f_\sigma(x,\theta,\varepsilon)=\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} \theta+\frac{\varepsilon^{-\alpha}\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\sigma\varepsilon^{\beta}=\theta^*.$$

注意到, 当 (x,θ,ε) 满足 $\varepsilon\neq0, 0<1-2\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)<1$, $f_\sigma(x,\theta,\varepsilon)$ 是连续可微的. $f_\sigma(x,\theta,\varepsilon)$ 在 (x,θ,ε) 点处的梯度是

$$\nabla_{(x,\theta,\varepsilon)}f_\sigma(x,\theta,\varepsilon)=\left(\nabla_x f_\sigma(x,\theta,\varepsilon),\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\theta},\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\varepsilon}\right),$$

这里

$$\begin{aligned}
\nabla_x f_\sigma(x,\theta,\varepsilon)&=\varepsilon^{-\alpha}\left(\frac{\partial_x\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\Delta(x,\theta,\varepsilon)\frac{\varepsilon^{-2\delta}\partial_x\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}\right)=\\
&\varepsilon^{-\alpha}\frac{\partial_x\Delta(x,\theta,\varepsilon)(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))+\Delta(x,\theta,\varepsilon)\varepsilon^{-2\delta}\partial_x\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}=\\
&\varepsilon^{-\alpha}\frac{\partial_x\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}=\\
&\frac{2\varepsilon^{-\alpha}\left[\sum_{j=1}^{q}(\max(f_j(x)-\theta-\varepsilon^{\gamma}w_j,0)\nabla f_j(x)+\sum_{j=q+1}^{m}(F_j(x)-\varepsilon^{\gamma}w_j)\nabla F_j(x)\right]}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2},
\end{aligned}\tag{11}$$

$$\begin{aligned}
\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\theta}&=1+\varepsilon^{-\alpha}\left(\frac{\partial_\theta\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\Delta(x,\theta,\varepsilon)\frac{\varepsilon^{-2\delta}\partial_\theta\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}\right)=\\
&1+\varepsilon^{-\alpha}\frac{\partial_\theta\Delta(x,\theta,\varepsilon)(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))+\Delta(x,\theta,\varepsilon)\varepsilon^{-2\delta}\partial_\theta\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}=\\
&1+\varepsilon^{-\alpha}\frac{\partial_\theta\Delta(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}=\\
&1-\frac{2\varepsilon^{-\alpha}\sum_{j=1}^{q}\max(f_j(x)-\theta-\varepsilon^{\gamma}w_j,0)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}
\end{aligned}\tag{12}$$

和

$$\begin{aligned}
\frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\varepsilon}&=\\
&-\alpha\varepsilon^{-\alpha-1}\frac{\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\varepsilon^{-\alpha}\left(\frac{\partial_\varepsilon\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\right.\\
&\left.\frac{\Delta(x,\theta,\varepsilon)(-2\delta\varepsilon^{-2\delta-1}\Delta(x,\theta,\varepsilon)+\varepsilon^{-2\delta}\partial_\varepsilon\Delta(x,\theta,\varepsilon))}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}\right)+\beta\varepsilon^{\beta-1}\sigma=\\
&-\alpha\varepsilon^{-\alpha-1}\frac{\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)}+\\
&[\varepsilon^{-\alpha}(\partial_\varepsilon\Delta(x,\theta,\varepsilon)(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)))-2\delta\varepsilon^{-2\delta-1}\Delta^2(x,\theta,\varepsilon)+\\
&\varepsilon^{-2\delta}\partial_\varepsilon\Delta(x,\theta,\varepsilon)\Delta(x,\theta,\varepsilon)]/[(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2]+\beta\varepsilon^{\beta-1}\sigma=
\end{aligned}$$

$$- \alpha\varepsilon^{-\alpha-1}\frac{\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)} + \varepsilon^{-\alpha}\frac{\partial_\varepsilon\Delta(x,\theta,\varepsilon) - 2\delta\varepsilon^{-2\delta-1}\Delta^2(x,\theta,\varepsilon)}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2} +$$

$$\beta\varepsilon^{\beta-1}\sigma = \Big\{ -\alpha\varepsilon^{-\alpha-1}\Delta(x,\theta,\varepsilon)(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)) -$$

$$2\gamma\varepsilon^{-\alpha+\gamma-1}\Big[\sum_{j=1}^{q}\max(f_j(x)-\theta-\varepsilon^\gamma w_j,0)w_j + \sum_{j=q+1}^{m}(F_j(x)-\varepsilon^\gamma w_j)w_j\Big] -$$

$$2\delta\varepsilon^{-\alpha-2\delta-1}\Delta^2(x,\theta,\varepsilon)\Big\}/(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2 + \beta\varepsilon^{\beta-1}\sigma =$$

$$\frac{\varepsilon^{-\alpha-1}}{(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon))^2}\Big[-\alpha\Delta(x,\theta,\varepsilon)(1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)) -$$

$$2\gamma\varepsilon^{\gamma}\Big(\sum_{j=1}^{q}\max(f_j(x)-\theta-\varepsilon^\gamma w_j,0)w_j +$$

$$\sum_{j=q+1}^{m}(F_j(x)-\varepsilon^\gamma w_j)w_j\Big) - 2\delta\varepsilon^{-2\delta}\Delta^2(x,\theta,\varepsilon)\Big] + \beta\varepsilon^{\beta-1}\sigma. \tag{13}$$

结合式(10)、(11)、(12)和(13),对于任意罚参数 $\sigma > 0$, 我们可以推得

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} \nabla_x f_\sigma(x,\theta,\varepsilon) = 0, \qquad \lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} \frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\varepsilon} = 0,$$

$$\lim_{\substack{\varepsilon\to\varepsilon^*=0\\(x,\theta)\to(x^*,\theta^*)\in S}} \frac{\partial f_\sigma(x,\theta,\varepsilon)}{\partial\theta} = 1.$$ □

2 算　法

步 1　选择任意小的数 $\bar{\varepsilon},\tilde{\varepsilon} > 0,\eta > 0$. 另外,选择 $\sigma > 0,\rho > 0$ 和初始点 $(x_0,\theta_0,\varepsilon_0) \in R^{n+1} \times (-\bar{\varepsilon},\bar{\varepsilon}),\varepsilon_0 \neq 0$,设置 $k:=0$;

步 2　对于规划问题(P),我们建立下面的罚函数:

$$f_\sigma(x,\theta,\varepsilon) = \begin{cases} \theta, & \text{若 } \varepsilon = 0,(x,\theta) \in S, \\ \theta + \dfrac{\varepsilon^{-\alpha}\Delta(x,\theta,\varepsilon)}{1-\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon)} + \sigma\varepsilon^\beta, & \text{若 } \varepsilon \neq 0, \\ & \quad 0 < 1 - 2\varepsilon^{-2\delta}\Delta(x,\theta,\varepsilon) < 1, \\ +\infty, & \text{否则}, \end{cases}$$

其中

$$\Delta(x,\theta,\varepsilon) = \sum_{j=1}^{q}(\max(f_j(x)-\theta-\varepsilon^\gamma w_j),0)^2 + \sum_{j=q+1}^{m}(F_j(x)-\varepsilon^\gamma w_j)^2,$$
$$w_j \in (0,1),\ j = 1,2,\cdots m,$$

利用任何无约束优化算法来求解

$$\min_{(x,\theta,\varepsilon)\in R^{n+1}\times(-\bar{\varepsilon},\bar{\varepsilon})} f_\sigma(x,\theta,\varepsilon),$$

将问题 (P_{σ_k}) 的解记作 $(x_\sigma,\theta_\sigma,\varepsilon_\sigma)$;

步 3　如果 $|\varepsilon_\sigma| \leqslant \tilde{\varepsilon}$, $\|\nabla_{(x,\theta,\varepsilon)} f_\sigma(x,\theta,\varepsilon)\| \leqslant \eta$,那么停止. (x_σ,θ_σ) 是问题(P)的近似解. 否则, $\sigma:=\sigma+\rho$, 回到步 2.

引理 2.1　如果 $(x_k,\theta_k,\varepsilon_k) \in L(P_{\sigma_k})$ 具有有限的目标函数值 $f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$ 和 $\varepsilon_k \neq 0$,那么 $(x_k,\theta_k,\varepsilon_k) \notin S_{\varepsilon_k}$.

证明　由 $(x_k,\theta_k,\varepsilon_k)\in L(P_{\sigma_k})$ 具有有限值 $f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$, $\varepsilon_k\neq 0$, 我们有

$$\frac{\partial f_\sigma(x_k,\theta_k,\varepsilon_k)}{\partial\varepsilon}=$$
$$\frac{\varepsilon_k^{-\alpha-1}}{(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2}\Big[-\alpha\Delta(x_k,\theta_k,\varepsilon_k)(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))-$$
$$2\gamma\varepsilon_k^{\gamma}\Big(\sum_{j=1}^{q}\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)w_j+\sum_{j=q+1}^{m}(F_j(x_k)-\varepsilon_k^{\gamma}w_j)w_j\Big)-$$
$$-2\delta\varepsilon_k^{-2\delta}\Delta^2(x_k,\theta_k,\varepsilon_k)\Big]+\beta\varepsilon_k^{\beta-1}\sigma_k=0.$$

如果 $(x_k,\theta_k,\varepsilon_k)\in S_{\varepsilon_k}$，那么上面式子的左边等于 $\sigma_k\beta\varepsilon_k^{\beta-1}\neq 0$，矛盾. 因此，$(x_k,\theta_k,\varepsilon_k)\notin S_{\varepsilon_k}$. □

注　我们有这样的结论：如果在点 $(x_k,\theta_k,\varepsilon_k)$ 处，满足 $\nabla_{(x,\theta,\varepsilon)}f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)=0$ 并且具有有限目标函数值 $f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$ 和 $\varepsilon_k\neq 0$，那么 $(x_k,\theta_k,\varepsilon_k)\notin S_{\varepsilon_k}$.

引理 2.2　如果 $(x_k,\theta_k,\varepsilon_k)\in L(P_{\sigma_k})$，具有有限的目标函数值 $f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$, $\varepsilon_k\neq 0$, $(x_k,\theta_k,\varepsilon_k)\xrightarrow{k}(x^*,\theta^*,\varepsilon^*)$ 并且 $\nabla F_j(x^*)$, $j=q+1,\cdots,m$ 是线性独立的，那么，$\varepsilon^*=0$, $(x^*,\theta^*)\in S$.

证明　首先证明 $\varepsilon^*=0$. 由 $\varepsilon_k\neq 0$ 和引理 2.1，我们有 $(x_k,\theta_k,\varepsilon_k)\notin S_\varepsilon$. 再由 $(x_k,\theta_k,\varepsilon_k)\in L(P_{\sigma_k})$，我们有

$$\nabla_x f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)=$$
$$2\varepsilon^{-\alpha}\Big[\sum_{j=1}^{q}(\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)\nabla f_j(x_k))+$$
$$\sum_{j=q+1}^{m}(F_j(x_k)-\varepsilon_k^{\gamma}w_j)\nabla F_j(x_k))\Big]\Big/(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2=0, \tag{14}$$

$$\frac{\partial f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)}{\partial\theta}=1-\frac{2\varepsilon_k^{-\alpha}\sum_{j=1}^{q}\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)}{(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2}=0 \tag{15}$$

和

$$\frac{\partial f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)}{\partial\varepsilon}=$$
$$\frac{\varepsilon_k^{-\alpha-1}}{(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2}\Big[-\alpha\Delta(x_k,\theta_k,\varepsilon_k)(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))-$$
$$2\gamma\varepsilon_k^{\gamma}\Big(\sum_{j=1}^{q}\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)w_j+$$
$$\sum_{j=q+1}^{m}(F_j(x_k)-\varepsilon_k^{\gamma}w_j)w_j\Big)-2\delta\varepsilon_k^{-2\delta}\Delta^2(x_k,\theta_k,\varepsilon_k)\Big]+\beta\varepsilon_k^{\beta-1}\sigma_k=0. \tag{16}$$

从式(16)，推得

$$-\alpha\Delta(x_k,\theta_k,\varepsilon_k)(\varepsilon_k^{2\delta}-\Delta(x_k,\theta_k,\varepsilon_k))-$$
$$2\gamma\varepsilon_k^{2\delta+\gamma}\Big(\sum_{j=1}^{q}\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)w_j+$$
$$\sum_{j=q+1}^{m}(F_j(x_k)-\varepsilon_k^{\gamma}w_j)w_j\Big)-2\delta\Delta^2(x_k,\theta_k,\varepsilon_k)+$$

$$\beta(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2\varepsilon_k^{\alpha+\beta+2\delta}\sigma_k=0. \tag{17}$$

令 $k\to+\infty$，由式(17)，第1项到第3项趋于有限值. 由罚函数 $f_\sigma(x,\theta,\varepsilon)$ 的结构，可知

$$\lim_{k\to+\infty}(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))\neq 0.$$

因此

$$\lim_{k\to+\infty}\varepsilon_k=\varepsilon^*=0. \tag{18}$$

下面我们证明 $(x^*,\theta^*)\in S$. 由式(15)，我们有

$$\varepsilon_k^{\alpha}(1-\varepsilon_k^{-2\delta}\Delta(x_k,\theta_k,\varepsilon_k))^2-2\sum_{j=1}^{q}\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)=0. \tag{19}$$

令 $k\to+\infty$，我们有

$$\sum_{j=1}^{q}\max(f_j(x^*)-\theta^*-(\varepsilon^*)^{\gamma}w_j,0)=0,$$

因而

$$f_j(x^*)-\theta^*\leqslant(\varepsilon^*)^{\gamma}w_j,\qquad\forall j=1,2,\cdots,q. \tag{20}$$

进一步，由式(14)，我们知道

$$\sum_{j=1}^{q}(\max(f_j(x_k)-\theta_k-\varepsilon_k^{\gamma}w_j,0)\nabla f_j(x_k))+\sum_{j=q+1}^{m}(F_j(x_k)-\varepsilon_k^{\gamma}w_j)\nabla F_j(x_k)=0.$$

令 $k\to+\infty$，根据式(20)，推得

$$\sum_{j=q+1}^{m}(F_j(x^*)-(\varepsilon^*)^{\gamma}w_j)\nabla F_j(x^*)=0.$$

根据引理3.2的假设，$\nabla F_j(x^*)$，$\forall j=q+1,\cdots,m$ 是线性独立的，可知

$$F_j(x^*)-(\varepsilon^*)^{\gamma}w_j=0,\qquad\forall j=q+1,\cdots,m. \tag{21}$$

因此

$$\Delta(x^*,\theta^*,\varepsilon^*)=\sum_{j=1}^{q}(\max(f_j(x^*)-\theta^*-(\varepsilon^*)^{\gamma}w_j,0))^2+\sum_{j=q+1}^{m}(F_j(x^*)-(\varepsilon^*)^{\gamma}w_j)^2=0.$$

我们有

$$f_j(x^*)-\theta^*\leqslant 0,\qquad\forall j=1,2,\cdots,q,$$
$$F_j(x^*)-(\varepsilon^*)^{\gamma}w_j=F_j(x^*)=0,\qquad\forall j=q+1,\cdots,m,$$

即 $(x^*,\theta^*)\in S$. □

定理2.1　假设 $(x_k,\theta_k,\varepsilon_k)\in L(P_{\sigma_k})$ 并且 $f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$ 是有限值，$\varepsilon_k\neq 0$. 对于任何聚点 $(x^*,\theta^*,\varepsilon^*)$，$\nabla F_j(x^*)$，$j=q+1,\cdots,m$ 满足线性独立约束规范，那么 (x^*,θ^*) 是(P)的局部最优解，其中 θ^* 是最优值.

证明　由定理2.1，存在序列 $\{(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\}_{\aleph}\subseteq\{(x_k,\theta_k,\varepsilon_k)\}$ 使得 $(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\xrightarrow{k}(x^*,\theta^*,\varepsilon^*)$. 再根据引理2.2可知 $\varepsilon^*=0$ 和 (x^*,θ^*) 是问题(P)的可行点，并且 $f_j(x^*)\leqslant\theta^*$，$\forall j=1,2,\cdots,q$. 因此，存在一个邻域 $o(x^*,\theta^*,0)$，对于邻域中的点 $(x,\theta,0)\in o(x^*,\theta^*,0)\cap(S\times\{0\})$，特别地，$\theta=\max_{j=1,2,\cdots,q}f_j(x)$. 那么，由 $(x_k,\theta_k,\varepsilon_k)$ 的定义，我们有，对于任何 $j=1,2,\cdots q$，

$$f_j(x^*)\leqslant\theta^*=f_\sigma(x^*,\theta^*,0)\leqslant f_\sigma(x,\theta,0)=\theta.$$

因此，(x^*,θ^*) 是问题(P)的局部最优值点并且 θ^* 是最优值. 事实上，$\theta^*=\max_{1\leqslant j\leqslant q}f_j(x^*)$.

证毕. □

推论 2.1　假设罚问题(P_σ)的每一个局部极小值点($x^*,\theta^*,\varepsilon^*$)并且$f_\sigma(x^*,\theta^*,\varepsilon^*)$是有限值,$\nabla F_j(x^*)$, $j=q+1,\cdots,m$满足线性独立约束规范, 那么(x^*,θ^*)是原问题(P)的局部极小值点当且仅当$\varepsilon^*=0$.

证明　如果(x^*,θ^*)是原问题(P)的局部极小值点,那么$f_j(x^*)\leqslant\theta^*$, $j=1,2,\cdots,q$和$F_j(x^*)=0$, $j=q+1,\cdots,m$. 利用反证法,由引理 2.2 可知,$\varepsilon^*=0$. 相应地,如果$\varepsilon^*=0$,考虑到罚函数$f_\sigma(x,\theta,\varepsilon)$的结构,那么$(x^*,\theta^*)\in S$,即$f_j(x^*)\leqslant\theta^*$和$F_j(x^*)=0$, $j=q+1,\cdots,m$. (x^*,θ^*)是问题(P)的可行点. 根据假设($x^*,\theta^*,0$)是问题(P_σ)的局部最优值点,那么(x^*,θ^*)是原问题(P)的局部极小值点. □

注　推论 2.1 表明这个罚函数的另外一个优点,就是ε可以被看作是局部极小值点的指示变量. 换句话说,在一般的条件下,对于($x^*,\theta^*,\varepsilon^*$),$\varepsilon^*=0$等价于($x^*,\theta^*$)是(P)的最优值点.

下一个定理探讨了罚函数$f_\sigma(x,\theta,\varepsilon)$的有限终止性质. 由这个结论,我们可知原问题(P)的最优解可以在有限步内得到.

定理 2.2　如果定理 2.1 的条件成立,并且$\alpha,\beta,\gamma,\delta$满足

$$-\alpha-\beta+2\delta\geqslant 0,\ \gamma>\delta, \tag{22}$$

那么,存在$k_0>0$,当$k\geqslant k_0$,我们有$\varepsilon_k=0,(x_k,\theta_k)\in L(P)$.

证明　反证法. 假设这个定理不成立,那么存在子列$\{(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\}_{\aleph}\subseteq\{(x_k,\theta_k,\varepsilon_k)\}$使得对任何$k_0>0$,当$n_k\geqslant k_0$,$(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\in L(P_{\sigma_{n_k}})$,并且具有有限值$f_{\sigma_{n_k}}(x_{n_k},\theta_{n_k},\varepsilon_{n_k})$,$\varepsilon_{n_k}\neq 0$并且定理 2.2 的条件在这个子序列上成立. 由引理 2.1,$(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\notin S_{\varepsilon_{n_k}}$成立. 因此,

$$\frac{\partial f_{\sigma_{n_k}}(x_{n_k},\theta_{n_k},\varepsilon_{n_k})}{\partial\varepsilon}=0,$$

即

$$\begin{aligned}&\frac{\varepsilon_{n_k}^{-\alpha-\beta}}{(1-\varepsilon_{n_k}^{-2\delta}\Delta(x_{n_k},\theta_{n_k},\varepsilon_{n_k}))^2}\Big[-\alpha\Delta(x_{n_k},\theta_{n_k},\varepsilon_{n_k})(1-\varepsilon_{n_k}^{-2\delta}\Delta(x_{n_k},\theta_{n_k},\varepsilon_{n_k}))-\\&\quad 2\gamma\varepsilon_{n_k}^{\gamma}\Big(\sum_{j=1}^{q}\max(f_j(x_{n_k})-\theta_{n_k}-\varepsilon_{n_k}^{\gamma}w_j,0)w_j+\sum_{j=q+1}^{m}(F_j(x_{n_k})-\varepsilon_{n_k}^{\gamma}w_j)w_j\Big)-\\&\quad 2\delta\varepsilon_{n_k}^{-2\delta}\Delta^2(x_{n_k},\theta_{n_k},\varepsilon_{n_k})\Big]+\beta\sigma_{n_k}=0.\end{aligned} \tag{23}$$

根据引理 2.2,我们推得$\varepsilon_{n_k}\to\varepsilon^*=0$, $(x_{n_k},\theta_{n_k})\xrightarrow{k}(x^*,\theta^*)\in S$. 结合$\varepsilon_{n_k}\neq 0$, $0<1-2\varepsilon_{n_k}^{-2\delta}\Delta(x_{n_k},\theta_{n_k},\varepsilon_{n_k})<1$, 我们有

$$\lim_{\substack{\varepsilon_{n_k}\to\varepsilon^*=0\\ x_{n_k}\to x^*\in S}} 1-\varepsilon_{n_k}^{-2\delta}\Delta(x_{n_k},\theta_{n_k},\varepsilon_{n_k})=c^*\in[1/2,1].$$

令$-\alpha-\beta+2\delta\geqslant 0$, $-\alpha-\beta+\delta+\gamma\geqslant 0$,式(23)的第 1 项趋于有限值,第 2 项趋于无穷,而式(23)的右边为0,这不可能. 因此,不存在这样的子列. 因此存在$k_0>0$,当$k\geqslant k_0$,我们有$\varepsilon_k=0,(x_k,\theta_k,0)\in L(P_{\sigma_k})$,这里$x_k,\theta_k$满足

$$f_j(x_k)\leqslant\theta_k,\qquad j=1,2,\cdots,q;$$

$$F_j(x_k)=0,\qquad j=q+1,\cdots,m.$$

因此,由$(x_k,\theta_k,0)\in L(P_{\sigma_k})$,在点$(x_k,\theta_k,0)$处存在一个邻域$o(x_k,\theta_k,0)$,对于所有的$(x,\theta,$

$0)\in o((x_k,\theta_k,0))\cap(S\times\{0\})$，特别地，$\theta=\max_{1\leqslant j\leqslant q}f_j(x)$，我们有

$$\theta_k=f_{\sigma_k}(x_k,\theta_k,0)\leqslant f_{\sigma_k}(x,\theta,0)=\theta,$$

因此，$(x_k,\theta_k)\in L(P)$. □

3　局部精确性质

在本节中，我们将证明，在一般较弱的条件下和一些附加的假设条件下，对于充分大的罚参数 σ，$(x^*,\theta^*,0)$ 是罚问题(P_σ)的局部最优值点，如果(x^*,θ^*)是原问题(P)的局部极小值点.

我们考虑非光滑的情况. 假设$f_j(x)$，$\forall j=1,2,\cdots,q$ 和 $F_j(x)$，$\forall j=q+1,\cdots,m$ 是非光滑函数. 为了正则化f_j，$j=1,2,\cdots,q$ 和 F_j，$j=q+1,\cdots,m$，我们通过引入上述的变量 ε，将$f_j(x)$，$\forall j=1,2,\cdots,q$ 和 $F_j(x)$，$j=q+1,\cdots,m$ 光滑化为$f_j(x,\varepsilon)$，$F_j(x,\varepsilon)$. 因此，引入的变量 ε 对求解问题(P)起到了重要的作用，它不仅在约束系统上起到了扰动的作用也在非光滑情况下起到了正则的作用. 正则化之后，函数$f_j(x,\varepsilon)$，$\forall j=1,2,\cdots,q$ 和 $F_j(x,\varepsilon)$，$\forall j=q+1,\cdots,m$ 在(x,ε)上都是连续可微的，$\varepsilon\neq 0$ 并且满足

$$f_j(x)=f_j(x,0)=\lim_{\varepsilon\to 0}f_j(x,\varepsilon),\qquad \forall j=1,2,\cdots,q,$$
$$F_j(x)=F_j(x,0)=\lim_{\varepsilon\to 0}F_j(x,\varepsilon),\qquad \forall j=q+1,\cdots,m.$$

我们考虑下面的系统：

$$(\mathrm{P}_\varepsilon)\quad\begin{cases}\min\limits_{\theta\in\mathbf{R}}\theta\\ \text{s.t. } f_j(x,\varepsilon)\leqslant\theta, & \forall j=1,2,\cdots,q,\\ F_j(x,\varepsilon)=0, & \forall j=q+1,\cdots,m.\end{cases}\tag{24}$$

令 $h_j(x,\theta,\varepsilon)=f_j(x,\varepsilon)-\theta$，$\forall j=1,2,\cdots,q$，那么$(P_\varepsilon)$可以被转化为

$$\begin{cases}\min\limits_{\theta\in\mathbf{R}}\theta\\ \text{s.t. } h_j(x,\theta,\varepsilon)\leqslant 0, & \forall j=1,2,\cdots,q,\\ F_j(x,\varepsilon)=0, & \forall j=q+1,\cdots,m.\end{cases}\tag{25}$$

现在我们介绍误差界的定义.

定义 3.1　我们将满足下面的系统的 $x\in R^n$ 组成的集合记为 S

$$\begin{cases}F(x)=0,\\ g(x)\leqslant 0,\end{cases}$$

这个系统在 x^* 处满足误差界条件，对于所有的 $x\in x^*+\delta\mathbb{B}$，如果存在正常数 $k>0$ 和 $\delta>0$ 使得

$$\mathrm{dist}(x\mid S)\leqslant k(\|F(x)\|+\|g(x)^+\|)$$

成立，这里 $\mathbb{B}$ 是 R^n 上的一个闭球.

接下来，我们讨论问题(P)的误差界成立的条件. 我们做以下假设：

(A_1) 对于(P_ε)，Managasarian-Fromovitz 约束规范在$(x^*,\theta^*,0)$处成立，也就是说，$\nabla F_j(x^*,0)$，$j=q+1,\cdots,m$ 是线性独立的，并且存在向量 $\boldsymbol{s}\in R^n$ 使得

$$(\boldsymbol{s},1)^{\mathrm{T}}\begin{pmatrix}\nabla_x h_j(x^*,\theta^*,0)\\ \dfrac{\partial h_j(x^*,\theta^*,0)}{\partial\theta}\end{pmatrix}>0,\qquad\forall j\in I(x^*,\theta^*);$$

$$s^{\mathrm{T}} \nabla F_j(x^*,0) = 0, \qquad \forall j = q+1,\cdots,m,$$

其中，$I(x^*,\theta^*) = \{j = 1,2,\cdots,q \mid h_j(x^*,\theta^*,0) = 0\}$。根据文献[18]，我们知道假设($A_1$)能够保证误差界的条件成立。进一步结合文献[18]中推论 2.3.1 和推论 2.4.1 或者文献[19]中的定理 3.1，我们有以下结论：

引理 3.1 如果假设(A_1)成立，那么在点(x^*,θ^*)处，存在一个邻域N_0，和一个常数$\tau > 0$使得对于$(x,\theta) \in N_0$，

$$\theta \geqslant \theta^* - \tau \left(\sum_{j=1}^{q} \| h_j(x,\theta,0)^+ \| + \sum_{j=q+1}^{m} \| F_j(x,0) \| \right)$$

成立。

下面我们给出局部精确性质的重要理论结果。在证明之前，我们首先需要做如下假设：

(H_1) δ,β,γ 是正偶整数并且满足 $\delta \geqslant \beta$ 和 $\gamma \geqslant \beta$；

(H_2) 对于充分小的$0 < \varepsilon' \ll 1$，

$$\| h_j(x,\theta,\varepsilon) - h_j(x,\theta,0) \| \leqslant K\varepsilon^{\beta},$$
$$\forall j = 1,2,\cdots,q,\ \varepsilon \in [-\varepsilon',0) \cup (0,\varepsilon'],$$
$$\| F_j(x,\varepsilon) - F_j(x,0) \| \leqslant K\varepsilon^{\beta}, \qquad \forall j = q+1,\cdots,m,\ \varepsilon \in [-\varepsilon',0) \cup (0,\varepsilon'].$$

基于上面的假设，我们给出这一节的主要结果：

定理 3.1 假设(A_1)、(H_1)和(H_2)都成立，对于充分大的罚参数σ，在(x^*,θ^*)处存在一个邻域$N \subseteq N_0$以及充分小的$0 < \varepsilon' \ll 1$使得

$$f_{\sigma}(x,\theta,\varepsilon) > f_{\sigma}(x^*,\theta^*,0) = \theta^*,$$
$$\text{对所有}(x,\theta,\varepsilon) \in N \times [-\varepsilon',0) \cup (0,\varepsilon'].$$

特别地，$(x^*,\theta^*,0)$是$f_{\sigma}(x,\theta,\varepsilon)$的局部极小值点。

证明 假设罚参数

$$\sigma \geqslant m\tau(K+2).$$

我们分成两种情况进行分析。对于$(x,\theta) \in N, \varepsilon \in [-\varepsilon',0) \cup (0,\varepsilon']$：

(i) $\Delta(x,\theta,\varepsilon) \geqslant \varepsilon^{2\delta}$；

(ii) $\Delta(x,\theta,\varepsilon) < \varepsilon^{2\delta}$。

情况(i) 由罚函数的构造，$f_{\sigma}(x,\theta,\varepsilon) = +\infty$。因此，$f_{\sigma}(x,\theta,\varepsilon) > f_{\sigma}(x^*,\theta^*,0)$。

我们进一步分析情况(ii)。

对于情况(ii)，$\Delta(x,\theta,\varepsilon) < \varepsilon^{2\delta}$，即

$$\sum_{j=1}^{q} (\max(h_j(x,\theta,\varepsilon) - \varepsilon^{\gamma} w_j,0))^2 + \sum_{q+1}^{m} (F_j(x,\varepsilon) - \varepsilon^{\gamma} w_j)^2 < \varepsilon^{2\delta},$$

推得

$$\| F_j(x,\varepsilon) \| \leqslant \varepsilon^{\gamma} | w_j | + \| F_j(x,\varepsilon) - \varepsilon^{\gamma} w_j \| < \varepsilon^{\gamma} | w_j | + \varepsilon^{\delta},$$
$$\forall j = q+1,\cdots,m;$$
$$\| h_j(x,\theta,\varepsilon) \| \leqslant \varepsilon^{\gamma} | w_j | + \| h_j(x,\theta,\varepsilon) - \varepsilon^{\gamma} w_j \| < \varepsilon^{\gamma} | w_j | + \varepsilon^{\delta},$$
$$\forall j \in J^+(x,\theta,\varepsilon),$$

其中$J^+(x,\theta,\varepsilon) = \{j = 1,2,\cdots,q \mid h_j(x,\theta,\varepsilon) \geqslant \varepsilon^{\gamma} w_j\}$。进一步结合引理 3.1 和假设($H_1$)、($H_2$)，我们有

$$\theta^* \leqslant \theta + \tau \left(\sum_{j=1}^{q} \| h_j(x,\theta,0)^+ \| + \sum_{j=q+1}^{m} \| F_j(x,0) \| \right) =$$

$$
\begin{aligned}
&\theta+\tau\left(\sum_{j\in J^+(x,\theta,0)}\|h_j(x,\theta,0)\|+\sum_{j=q+1}^{m}\|F_j(x,0)\|\right)\leqslant\\
&\theta+\tau\left(\sum_{j\in J^+(x,\theta,0)}\|h_j(x,\theta,\varepsilon)\|+\sum_{j\in J^+(x,\theta,0)}K\varepsilon^{\beta}+\sum_{j=q+1}^{m}\|F_j(x,\varepsilon)\|+\sum_{j=q+1}^{m}K\varepsilon^{\beta}\right)<\\
&\theta+\tau\left(\sum_{j=1}^{q}\varepsilon^{\gamma}|w_j|+q\varepsilon^{\delta}+Kq\varepsilon^{\beta}+\sum_{j=q+1}^{m}\varepsilon^{\gamma}|w_j|+(m-q)\varepsilon^{\delta}+K(m-q)\varepsilon^{\beta}\right)\leqslant\\
&\theta+m\tau(K+2)\varepsilon^{\beta}\leqslant\\
&\theta+\sigma\varepsilon^{\beta}.
\end{aligned}
$$

第 2 个不等式和第 4 个不等式分别源自于假设(H_2)和(H_1). 因此,

$$f_\sigma(x^*,\theta^*,0)=\theta^*<\theta+\sigma\varepsilon^{\beta}\leqslant f_\sigma(x,\theta,\varepsilon),$$

对于所有$(x,\theta,\varepsilon)\in N\times[-\varepsilon',0)\cup(0,\varepsilon']$.

$(x^*,\theta^*,0)$ 是 $f_\sigma(x,\theta,\varepsilon)$ 的局部极小值点. □

4 数值实验

为了检验算法的可实施性和有效性,我们进行如下数值实验. 该实验利用 Matlab 7.8.0 软件,在 Intel Core 2CPU 2.39 GHz, 1.99 GB memory 机器上运行. 我们使用 $\|\nabla_{(x,\theta,\varepsilon)}f_\sigma(x,\theta,\varepsilon)\|\leqslant 10^{-6}$ 作为停机准则. 表格 1 ~ 5 中 σ_k 表示罚参数, $x_k,\theta_k,\varepsilon_k$ 表示最终迭代点, 以及 $f(x_k)$ 表示函数 f 在 x_k 上的函数值.

例 1 令

$$f_1(x)=10(x_2-x_1^2),\ f_2(x)=-10(x_2-x_1^2),\ f_3(x)=1-x_1,$$
$$f_4(x)=1+x_1,\ F_1(x)=100x_1^2+x_2^2-101,\ F_2(x)=80x_1^2-x_2^2-79.$$

为了使得 $\alpha,\beta,\gamma,\delta$ 满足式(10)和(22),这个例子中,算法参数我们设置为

$$\alpha=6,\ \beta=8,\gamma=10,\delta=8.$$

最优解和最优值分别为 $x^*=(1,1)$ 和 $f(x^*)=0$. 初始点为

$$x_0=(2,-1),\theta_0=-5,\varepsilon_0=2.$$

表 1 例 1 的数值结果

Table 1 Numerical results of example 1

k	σ_k	x_k	θ_k	ε_k	$f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$
1	2	(1.000 0, 1.000 0)	0.000 0	0.226 6	1.755 9E-005
2	4	(1.000 0, 1.000 0)	0.000 0	0.299 1	2.877 9E-004
3	6	(1.000 0, 1.000 0)	0.000 0	0.265 5	1.500 5E-004
4	8	(1.000 0, 1.000 0)	0.000 0	0.304 1	6.092 7E-004
5	10	(1.000 0, 1.000 0)	0.000 4	0.051 5	5.790 5E-004

例 2 令

$$f_1(x)=\frac{1}{2}\left(x_1+\frac{10x_1}{x_1+0.1}+2x_2^2\right),\ f_2(x)=\frac{1}{2}\left(-x_1+\frac{10x_1}{x_1+0.1}+2x_2^2\right),$$
$$f_3(x)=\frac{1}{2}\left(x_1-\frac{10x_1}{x_1+0.1}-2x_2^2\right),\ F_1(x)=x_1^2+x_2^2+x_1x_2,\ F_2(x)=-x_1+x_2^2.$$

这里,算法参数设置为 $\alpha=6,\beta=8,\gamma=10,\delta=8$. 最优解和最优值分别为 $x^*=(0,0)$, $f(x^*)=0$. 初始点是 $x_0=(2,1),\theta_0=-1,\varepsilon_0=2$.

表 2 例 2 的数值结果

Table 2 Numerical results of example 2

k	σ_k	x_k	θ_k	ε_k	$f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$
1	2	(−0.000 0, 0.000 3)	0.000 0	0.193 8	5.292 2E−006
2	4	(−0.000 0, 0.000 0)	0.000 1	0.221 0	9.822 9E−005
3	6	(−0.000 2, −0.014 9)	0.008 2	0.412 5	0.015 8
4	8	(0.000 0, 0.002 2)	0.000 2	0.225 9	2.729 8E−004
5	10	(0.000 0, 0.002 0)	0.000 2	0.237 7	3.781 3E−004

例 3 令

$$f_1(x)=x_1^2+x_2^2,\ f_2(x)=(2-x_1)^2+(2-x_2)^2,\ f_3(x)=2\exp(x_2-x_1),$$
$$F_1(x)=x_1+x_2-2,\ F_2(x)=-x_1^2-x_2^2+2.25.$$

算法参数设置为 $\alpha=6,\beta=8,\gamma=10,\delta=8$. 最优解和最优值是 $x^*=(1.353\,55,\ 0.646\,45)$ 和 $f(x^*)=2.25$. 初始点是 $x_0=(0,0),\theta_0=1,\varepsilon_0=2$.

表 3 例 3 的数值结果

Table 3 Numerical results of example 3

k	σ_k	x_k	θ_k	ε_k	$f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$
1	2	(1.007 1, 1.077 1)	2.134 7	0.788 6	2.522 7
2	4	(1.353 6, 0.646 4)	2.250 0	0.180 3	2.250 0
3	6	(1.353 6, 0.646 4)	2.250 0	0.206 4	2.250 0
4	8	(1.353 6, 0.646 4)	2.250 0	0.160 3	2.250 0
5	10	(1.353 5, 0.646 5)	2.250 0	0.240 1	2.250 1

例 4 令

$$f_1(x)=\exp\left(\frac{x_1^2}{1\,000}+(x_2-1)^2\right),\ f_2(x)=\exp\left(\frac{x_1^2}{1\,000}+(x_2+1)^2\right),$$
$$F_1(x)=\frac{x_1^2}{1\,000}+x_2^2+x_1x_2,\ F_2(x)=-x_1+x_2^2.$$

算法参数设置为

$$\alpha=6,\ \beta=8,\ \gamma=10,\ \delta=8.$$

最优解和最优值是 $x^*=(0,0)$ 和 $f(x^*)=2.718\,28$. 初始点为 $x_0=(0,0),\theta_0=1,\varepsilon_0=2$.

表 4 例 4 的数值结果

Table 4 Numerical results of example 4

k	σ_k	x_k	θ_k	ε_k	$f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$
1	2	(−0.000 0, −0.000 0)	2.718 3	0.177 3	2.718 3
2	4	(0.000 0, 0.000 0)	2.718 3	0.232 5	2.718 3
3	6	(0.000 0, 0.000 0)	2.718 3	0.226 0	2.718 4

例 5 令

$$f_1(x)=x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4,$$
$$f_2(x)=f_1(x)+10F_1(x),\ f_3(x)=f_1(x)+10F_2(x),\ f_4(x)=f_1(x)+10F_3(x),$$
$$F_1(x)=x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4-8,$$
$$F_2(x)=x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4-9,$$
$$F_3(x)=x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4-5.$$

算法参数设置为

$$\alpha=6,\ \beta=8,\ \gamma=10,\ \delta=8.$$

最优解和最优值是 $x^*=(0,0,2,-1)$ 和 $f(x^*)=-44$. 初始点是 $x_0=(-1,1.5,1.8,-1)$, $\theta_0=-50$, $\varepsilon_0=2$.

表 5　例 5 的数值结果

Table 5　Numerical results of example 5

k	σ_k	x_k	θ_k	ε_k	$f_{\sigma_k}(x_k,\theta_k,\varepsilon_k)$
1	2	(−0.006 8, 1.000 8, 2.004 6, −0.994 2)	−43.999 8	0.379 1	−43.998 8
2	10	(0.043 4, 0.994 5, 1.969 9, −1.036 0)	−43.985 4	0.396 7	−43.978 8
3	18	(0.030 9, 0.996 1, 1.978 7, −1.025 8)	−43.993 0	0.007 3	−43.978 6

数值结果表明,通过适当的参数选取,我们不要求很大的罚参数就可以获得问题的最优值点,这说明本文设计的罚函数是合理的,并且运算效果良好.

注　在本文设计的算法的执行过程中, 与其他的罚函数方法有所不同. 不同点在于初始点的选取, 也就是说, 我们在执行该算法前, 应该选择初始点满足 $0<1-2\varepsilon_0^{-2\delta}\Delta(x_0,\theta_0,\varepsilon_0)<1$. 这个可以通过求解方程 $1-2\varepsilon_0^{-2\delta}\Delta(x_0,\theta_0,\varepsilon_0)=c$ 来实现,这里 c 是属于区间(0,1)中的任意实数.

5　结　　论

在本文中,针对带约束的极大极小规划问题,我们设计一种新的精确光滑罚函数. 我们的目的是提供一种新的求解非光滑的极大极小规划问题的光滑算法. 在合理的假设条件下,当罚参数充分大,罚问题的极小值点就是原问题的极小值点. 进一步,我们也研究了精确罚的逆定理,即我们讨论了在一般的假设条件下,原问题的解 x^* 也是罚问题的解 $(x^*,0)$. 数值结果表明这种新的罚函数算法是求解带约束有限极大极小问题的一种有效算法.

致谢　本文第三作者感谢香港理工大学 RGC 项目(5365/09E)对本文的资助.

参考文献(References):

[1]　Rustem B, Howe M A. *Algorithms for Worst-Case Design With Applications to Risk Management*[M]. Princeton: Princeton University Press, 2001.

[2]　Zhu S S, Fukushima M. Worst-case conditional value-at-risk with application to robust portfolio management[J]. *Operations Research*, 2009, **57**(5): 1155-1168.

[3]　Lemarechal C. Nondifferentiable optimization. Nemhauser G L, Rinnooy Kan A H G, Todd M J. *Handbooks in Operations Research and Management Science. Vol* 1: *Optimization*[M]. Amsterdam, Netherlands: North-Holland, 1989.

[4]　Rockafellar R T. Linear-quadric programming and optimal control[J]. *SIAM Journal on Control and Optimization*, 1987, **25**(3): 781-814.

[5]　Rockafellar R T. Computational schemes for large-scale problems in extended linear-quadratic programming[J]. *Mathematical Programming*, 1990, **48**(1/3): 447-474.

[6]　Gaudioso M, Monaco M F. A bundle type approach to the unconstrained minimization of convex nonsmooth functions[J]. *Mathematical Programming*, 1982, **23**(1): 216-226.

[7]　Polak E, Mayne D H, Higgins J E. Superlinearly convergent algorithm for min-max problems[J]. *Journal of Optimization Theory and Applications*, 1991, **69**(3): 407-439.

[8]　Zowe J. Nondifferentiable optimization: a motivation and a short introduction into the subgradient and the bundle concept[C]//Schittkowski K. *Computational Mathematical Programming, NATO SAI Series*. New York:Springer, 1985.

[9]　DiPillo G, Grippo L, Lucidi S. A smooth method for the finite minimax problem[J]. *Mathe-*

matical Programming, 1993, **60**(1/3): 187-214.

[10] Gigola C, Gomez S. A regularization method for solving the finite convex min-max problems [J]. *SIAM Journal on Numerical Analysis*, 1990, **27**(6): 1621-1634.

[11] Li X S, Pan S. Solving the finite min-max problem via an exponential penalty method[J]. *Comput Technol*, 2003, **8**(2): 3-15.

[12] Ye F, Liu H, Zhou S, Liu S. A smoothing trust-region Newton-CG method for minimax problem[J]. *Applied Mathematics and Computation*, 2008, **199**(2): 581-589.

[13] Zang I. A smoothing technique for min-max optimization[J]. *Mathematical Programming*, 1980, **19**(1): 61-77.

[14] Zhou J L, Tits A L. An SQP algorithm for finely discretized continuous minimax problems and other minimax problems with many objective functions[J]. *SIAM J Optimization*, 1996, **6**(2): 461-487.

[15] Zhu Z, Cai X, Jian J. An improved SQP algorithm for solving minimax problems[J]. *Applied Mathematics Letters*, 2009, **22**(4): 464-469.

[16] Obasanjo E, Tzallas-Regas G, Rustem B. An interior-point algorithm for nonlinear minimax problems[J]. *Journal of Optimization Theory and Applications*, 2010, **144**(2): 291-318.

[17] Huyer W, Neumaier A. A new exact penalty function[J]. *SIAM Journal on Optimization*, 2003, **13**(4): 1141-1158.

[18] Burke J V. An exact penalization viewpoint of constrained optimization[J]. *SIAM J Control and Optimization*, 1991, **29**(4): 968-998.

[19] Pang J S. Error bound in mathematical programming[J]. *Mathematical Programming*, 1997, **79**(1/3): 299-332.

A New Exact Penalty Function for Solving Constrained Finite Min-Max Problems

MA Cheng[1], LI Xun[1], YIU Ka-Fai Cedric[1], ZHANG Lian-sheng[2]

(1. *Department of Applied Mathematics*, *Hong Kong Polytechnic University*, *Kowloon*, *Hong Kong*, *P. R. China*;

2. *Department of Mathematics*, *Shanghai University*, *Shanghai* 200444, *P. R. Chian*)

Abstract: A new exact yet smooth penalty function to tackle constrained min-max problems was introduced. Using this new penalty function and adding just one extra variable, a constrained min-max problem was transformed into an unconstrained optimization one. It was proved that, under certain reasonable assumptions and when the penalty parameter was sufficiently large, the minimizer of this unconstrained optimization problem was equivalent to the minimizer of the original constrained one. Moreover, the local exactness property was also studied. The numerical results demonstrate that this penalty function method is an effective and promising approach for solving constrained finite min-max problems.

Key words: min-max problem; constrained optimization; penalty function