

# 解含多个复杂分量函数无约束minimax问题的积极集光滑化算法

周正勇, 秦丽娜

(山西师范大学数学与计算机科学学院, 山西 临汾 041004)

**摘要:** 本文利用分段三次多项式方程构造了一种积极集策略的二次连续可微的光滑化max函数, 给出积极集及稳定的光滑化max函数的计算方法. 基于该光滑化max函数, 结合Armijo线搜索, 负梯度和牛顿方向及光滑化参数的更新策略, 给出一种解含多个复杂分量函数无约束minimax问题的积极集光滑化算法. 初步的数值实验表明了该算法的有效性.

**关键词:** Minimax问题; 积极集; 光滑化max函数; 光滑化算法

**中图分类号:** O224

**AMS(2000)主题分类:** 90C47; 65D10; 11Y16

**文献标识码:** A

**文章编号:** 1001-9847(2020)03-0690-09

## 1. 引言

本文考虑如下无约束minimax问题:

$$\min_{x \in \mathbb{R}^n} \{F(x) = \max_{j \in Q} f_j(x)\}, \quad (1.1)$$

其中分量函数 $f_j: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j \in Q = \{1, \dots, q\}$ , 复杂且二次连续可微,  $q$ 较大. 问题(1.1)是一类典型的不可微优化问题, 在数据拟合<sup>[1]</sup>, 工程设计<sup>[2]</sup>, 结构优化<sup>[3]</sup>, 资源配置<sup>[4]</sup>, 选址问题<sup>[5]</sup>, 电路设计<sup>[6]</sup>等领域有广泛的应用, 此外, 曲线拟合问题,  $L_1$ 和 $L_\infty$ 逼近问题等均可等价转化为该类问题.

由于问题(1.1)的非光滑性, 光滑优化的理论与算法不能直接用于求解该类问题. 解该类问题的一类方法是将其等价转化为如下光滑约束优化问题:

$$\min_{x \in \mathbb{R}^n, v \in \mathbb{R}} \{v \mid f_j(x) - v \leq 0, j \in Q\}, \quad (1.2)$$

进而用求解一般光滑约束优化问题的算法对(1.2)进行求解. 文[7]指出该类方法有两个缺点, 首先, 该问题中Lagrange函数的海森矩阵通常奇异; 其次, 目标函数 $v$ 可由 $\alpha v$ 替代, 其中 $\alpha$ 是任意大于0的常数, 因而目标函数可被任意尺度化, 从而该问题病态. 另一类方法是非光滑优化方法, 如次梯度法, 割平面法, 捆集法等. 但这类方法大多是基于非光滑凸问题建立的, 其收敛性条件较强.

解问题(1.1)的第三类方法是光滑化方法, 该类方法通过构造max函数的光滑逼近函数(一般采用凝聚函数<sup>[8]</sup>— $F_p(x) = \ln(\sum_{j \in Q} \exp(pf_j(x)))/p$ , 其中 $p > 0$ 为光滑化参数), 将问题(1.1)转化为一族带光滑化参数的光滑无约束优化问题:

$$\min_{x \in \mathbb{R}^n} F_p(x), \quad (1.3)$$

\* 收稿日期: 2019-07-21

基金项目: 国家自然科学基金(11701350)

作者简介: 周正勇, 男, 汉族, 山东人, 副教授, 研究方向: 数值优化.

当 $p$ 趋于 $\infty$ 时, 这族带光滑化参数的优化问题的稳定点趋近于问题(1.1)的稳定点. 由于问题(1.3)为光滑无约束优化问题, 因此可采用许多解光滑无约束优化问题的算法对其进行求解.

直接计算, 凝聚函数的梯度及海森矩阵为:

$$\begin{aligned}\nabla F_p(x) &= \sum_{j \in Q} \mu_p^j(x) \nabla f_j(x), \\ \nabla^2 F_p(x) &= \sum_{j \in Q} \mu_p^j(x) \nabla^2 f_j(x) + p \sum_{j \in Q} \mu_p^j(x) \nabla f_j(x) \nabla f_j(x)^T \\ &\quad - p \sum_{j \in Q} \mu_p^j(x) \nabla f_j(x) \sum_{j \in Q} \mu_p^j(x) \nabla f_j(x)^T,\end{aligned}$$

其中

$$\mu_p^j(x) = \frac{\exp(pf_j(x))}{\sum_{j \in Q} \exp(pf_j(x))} \in (0, 1], \quad j \in Q, \quad \sum_{j \in Q} \mu_p^j(x) = 1.$$

由于凝聚函数的梯度与所有分量函数的梯度相关, 海森矩阵与所有分量函数的梯度及海森矩阵相关且表达式较为复杂, 因此, 当分量函数复杂且个数较多时, 凝聚函数的梯度及海森矩阵的计算量较大. 为了改善这一问题, 文[9]构造了如下方程:

$$\sum_{j \in Q} \phi_t(f_j(x) - \alpha) - 1 = 0, \quad (1.4)$$

其中 $t > 0$ 为光滑化参数. 对任意的 $t > 0$ ,  $\phi_t(z) : \mathbb{R} \rightarrow \mathbb{R}$ 二次连续可微且满足: 当 $z > -t$ 时,  $\phi_t(z)$ 严格单调递增, 当 $z \leq -t$ 时,  $\phi_t(z) = 0$ ;  $\phi_t(0) = 1$ . 根据 $\phi_t(z)$ 的定义, 可将其表示为

$$\phi_t(z) = \begin{cases} 0, & z \leq -t, \\ \bar{\phi}_t(z), & z > -t, \end{cases}$$

其中 $\bar{\phi}_t(z)$ 满足以下条件:

$$\bar{\phi}_t(0) = 1, \quad \bar{\phi}_t(-t) = 0, \quad \bar{\phi}_t'(-t) = 0, \quad \bar{\phi}_t''(-t) = 0.$$

满足以上条件的函数 $\bar{\phi}_t(z)$ 有多种类型, 如样条函数, 多项式函数等, 文[9]利用

$$\bar{\phi}_t(z) = \left(1 + \frac{z}{t}\right)^3, \quad z > -t, \quad (1.5)$$

构造了一种积极集策略的光滑化max函数 $\alpha_t(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ , 该函数仅与函数值较大的分量函数相关. 文[9-10]均采用迭代策略计算 $\alpha_t(x)$ 的函数值及与其相关的分量函数的指标集, 计算效率较低.

本文对文[9]中基于方程(1.4)及函数(1.5)构造的积极集策略的光滑化max函数, 给出了指标集的直接计算方法, 利用该指标集将方程(1.4)转化为一般的三次多项式方程, 根据该三次多项式方程根的性质给出了 $\alpha_t(x)$ 的一种稳定计算策略. 结合Armijo线搜索, 负梯度和牛顿方向及光滑化参数的更新策略, 给出了一种解含多个复杂分量函数无约束minimax问题的积极集光滑化算法, 初步的数值实验表明了该算法的有效性.

本文用到的假设和引理.

**假设1.1**  $f_j(x)$ ,  $j \in Q$ , 二次连续可微.

**引理1.1**<sup>[9]</sup> 对任意的 $t > 0$ , (1.5)式及方程(1.4)的解定义了一个二次连续可微的光滑化max函数 $\alpha_t(x)$ , 且 $\alpha_t(x)$ 满足

$$\max_{j \in Q} f_j(x) \leq \alpha_t(x) \leq \max_{j \in Q} f_j(x) + t - \frac{t}{\sqrt[3]{q}}.$$

**引理1.2**<sup>[9]</sup> 对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ ,  $\alpha_t(x)$ 仅与如下的指标集相关:

$$Q_t(x) = \{j \in Q \mid f_j(x) - \alpha_t(x) > -t\}, \quad (1.6)$$

且 $Q_t(x)$ 满足:

$$Q_t(x) \subseteq Q_t^0(x) = \{j \in Q | f_j(x) - \max_{j \in Q} f_j(x) > -t\}.$$

由 $Q_t(x)$ 的定义可知,  $\alpha_t(x)$ 的梯度和海森矩阵仅与函数值较大的分量函数有关. 令 $z_t^j(x) = f_j(x) - \alpha_t(x)$ , 则 $\nabla z_t^j(x) = \nabla f_j(x) - \nabla \alpha_t(x)$ , 直接计算可得

$$\begin{aligned} \nabla \alpha_t(x) &= \sum_{j \in Q_t(x)} \lambda_t^j(x) \nabla f_j(x), \\ \nabla^2 \alpha_t(x) &= \sum_{j \in Q_t(x)} \nabla \lambda_t^j(x) \nabla f_j(x)^T + \lambda_t^j(x) \nabla^2 f_j(x), \end{aligned}$$

其中

$$\begin{aligned} \lambda_t^j(x) &= \frac{(t+z_t^j(x))^2}{\sum_{j \in Q_t(x)} (t+z_t^j(x))^2} \in (0, 1] (j \in Q_t(x)), \quad \sum_{j \in Q_t(x)} \lambda_t^j(x) = 1, \\ \nabla \lambda_t^j(x) &= \frac{2(t+z_t^j(x)) \nabla z_t^j(x) - 2\lambda_t^j(x) \sum_{j \in Q_t(x)} (t+z_t^j(x)) \nabla z_t^j(x)}{\sum_{j \in Q_t(x)} (t+z_t^j(x))^2}. \end{aligned}$$

## 2. 指标集 $Q_t(x)$ 的直接计算方法

对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ , 令

$$\alpha_j^{xt} = f_j(x) + t, j \in Q, \quad (2.1)$$

定义方程(1.4)在 $\alpha \in \mathbb{R}$ 的残量为

$$r^{xt}(\alpha) = \sum_{j \in Q} \phi_t(f_j(x) - \alpha) - 1. \quad (2.2)$$

**定理2.1** 对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ , 及任意满足 $Q_t(x) \subseteq Q^0 \subseteq Q$ 的 $Q^0$ ,

$$Q_t(x) = \{j \in Q^0 | r^{xt}(\alpha_j^{xt}) < 0\}.$$

**证** 对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ ,  $j \in Q_t(x)$ , 由 $Q_t(x) \subseteq Q^0$ 可知 $j \in Q^0$ . 由(1.6)式可得

$$f_j(x) - \alpha_t(x) > -t,$$

结合(2.1)式可得

$$\alpha_j^{xt} > \alpha_t(x). \quad (2.3)$$

由 $f_j(x) + t \leq \max_{j \in Q} f_j(x) + t$ 可得

$$\alpha_j^{xt} \leq \max_{j \in Q} f_j(x) + t.$$

因此

$$\alpha_j^{xt}, \alpha_t(x) \in (-\infty, \max_{j \in Q} f_j(x) + t]. \quad (2.4)$$

由 $\phi_t(z)$ 在 $\mathbb{R}$ 上单调递增及在 $[-t, +\infty)$ 上严格单调递增可知,  $r^{xt}(\alpha)$ 在 $(-\infty, \max_{j \in Q} f_j(x) + t]$ 上严格单调递减, 进而由(2.3), (2.4)式及 $r^{xt}(\alpha_t(x)) = 0$ 可得 $r^{xt}(\alpha_j^{xt}) < 0$ , 因此 $Q_t(x) \subseteq \{j \in Q^0 | r^{xt}(\alpha_j^{xt}) < 0\}$ .

对任意满足 $r^{xt}(\alpha_j^{xt}) < 0$ 的 $j \in Q^0$ , 由 $r^{xt}(\alpha)$ 在 $\mathbb{R}$ 上单调递减及 $r^{xt}(\alpha_t(x)) = 0$ 可知 $\alpha_j^{xt} > \alpha_t(x)$ , 结合(2.1)式可得

$$f_j(x) - \alpha_t(x) > -t,$$

因此 $\{j \in Q^0 | r^{xt}(\alpha_j^{xt}) < 0\} \subseteq Q_t(x)$ .

综上所述, 对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ ,  $Q_t(x) = \{j \in Q^0 | r^{xt}(\alpha_j^{xt}) < 0\}$ .

对任意的 $x \in \mathbb{R}^n$ ,  $t > 0$ ,  $i \in Q$ , 定义 $Q_i^+(x) = \{j \in Q | f_j(x) - f_i(x) > 0\}$ , 由(2.1)式可知 $Q_i^+(x) = \{j \in Q | f_j(x) - \alpha_i^{xt} > -t\}$ , 则由 $\phi_t(z)$ 及 $\alpha_i^{xt}$ 的定义可知

$$r^{xt}(\alpha_i^{xt}) = \sum_{j \in Q} \phi_t(f_j(x) - \alpha_i^{xt}) - 1$$

$$\begin{aligned}
 &= \sum_{j \in Q_i^+(x)} \bar{\phi}_t(f_j(x) - \alpha_i^{xt}) - 1 = \sum_{j \in Q_i^+(x)} \left(1 + \frac{f_j(x) - \alpha_i^{xt}}{t}\right)^3 - 1 \\
 &= \frac{\sum_{j \in Q_i^+(x)} (f_j(x) - f_i(x))^3 - t^3}{t^3}.
 \end{aligned} \tag{2.5}$$

**定理2.2** 对任意的  $x \in \mathbb{R}^n, t > 0$ , 及任意满足  $Q_t(x) \subseteq Q^0 \subseteq Q$  的  $Q^0$ ,

$$Q_t(x) = \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}, \tag{2.6}$$

其中

$$(f_j(x) - f_i(x))_+ = \begin{cases} f_j(x) - f_i(x), & f_j(x) - f_i(x) > 0, \\ 0, & f_j(x) - f_i(x) \leq 0. \end{cases}$$

**证** 对任意的  $x \in \mathbb{R}^n, t > 0, i \in Q_t(x)$ , 由定理2.1可知  $r^{xt}(\alpha_i^{xt}) < 0, i \in Q^0$ , 由  $r^{xt}(\alpha)$  在  $\mathbb{R}$  上单调递减及  $r^{xt}(\alpha_t(x)) = 0$  可得  $\alpha_i^{xt} > \alpha_t(x)$ , 由  $Q_i^+(x)$  及  $Q_t(x)$  的定义可得  $Q_i^+(x) \subseteq Q_t(x)$ , 再由  $Q_t(x) \subseteq Q^0 \subseteq Q$  可得  $Q_i^+(x) \subseteq Q_t(x) \subseteq Q^0 \subseteq Q$ , 则(2.5)式可等价变形为

$$r^{xt}(\alpha_i^{xt}) = \frac{\sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3}{t^3}.$$

由  $t > 0$  及  $r^{xt}(\alpha_i^{xt}) < 0$  可得  $\sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0$ , 因此  $Q_t(x) \subseteq \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}$ .

若  $Q_t(x) \subset \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}$ , 则存在  $i^* \in \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}$ , 但  $i^* \notin Q_t(x)$ . 由  $i^* \notin Q_t(x)$  可知, 对任意的  $j \in Q_t(x) \subseteq Q^0, f_j(x) > f_{i^*}(x), f_{i^*}(x) \leq \alpha_t(x) - t$ , 即  $\alpha_{i^*}^{xt} \leq \alpha_t(x)$ . 结合  $\phi_t(z)$  在  $\mathbb{R}$  上的单调递增性可得,

$$\begin{aligned}
 \frac{\sum_{j \in Q^0} ((f_j(x) - f_{i^*}(x))_+)^3 - t^3}{t^3} &\geq \frac{\sum_{j \in Q_t(x)} ((f_j(x) - f_{i^*}(x))_+)^3 - t^3}{t^3} \\
 &= \sum_{j \in Q_t(x)} \phi_t(f_j(x) - \alpha_{i^*}^{xt}) - 1 \\
 &\geq \sum_{j \in Q_t(x)} \phi_t(f_j(x) - \alpha_t(x)) - 1 = 0,
 \end{aligned}$$

与  $i^* \in \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}$  矛盾.

综上所述, 对任意的  $x \in \mathbb{R}^n, t > 0, Q_t(x) = \{i \in Q^0 \mid \sum_{j \in Q^0} ((f_j(x) - f_i(x))_+)^3 - t^3 < 0\}$ .

### 3. 积极集光滑化max函数的稳定计算策略

对任意的  $x \in \mathbb{R}^n, t > 0$ , 由  $Q_t(x), \phi_t(z)$  的定义及(1.5)式可将方程(1.4)等价转化为一般三次多项式方程:

$$\sum_{j \in Q_t(x)} \left(1 + \frac{f_j(x) - \alpha}{t}\right)^3 - 1 = a\alpha^3 + b\alpha^2 + c\alpha + d = 0, \tag{3.1}$$

其中  $\bar{q} = \#(Q_t(x))$ ,

$$\begin{cases} a = -\frac{\bar{q}}{t^3}, \\ b = \sum_{j \in Q_t(x)} \frac{3f_j(x)}{t^3} + \frac{3\bar{q}}{t^2}, \\ c = -\sum_{j \in Q_t(x)} \frac{3f_j(x)^2}{t^3} - \sum_{j \in Q_t(x)} \frac{6f_j(x)}{t^2} - \frac{3\bar{q}}{t}, \\ d = \sum_{j \in Q_t(x)} \frac{f_j(x)^3}{t^3} + \sum_{j \in Q_t(x)} \frac{3f_j(x)^2}{t^2} + \sum_{j \in Q_t(x)} \frac{3f_j(x)}{t} + \bar{q} - 1. \end{cases}$$

令  $r(\alpha) = \sum_{j \in Q_t(x)} (1 + \frac{f_j(x) - \alpha}{t})^3 - 1$ , 由  $r'(\alpha) = -\frac{3}{t} \sum_{j \in Q_t(x)} (1 + \frac{f_j(x) - \alpha}{t})^2$  可知  $r(\alpha)$  在  $\mathbb{R}$  上严格单调递减, 因此方程(3.1)在  $\mathbb{R}$  上仅有一个实根  $\alpha_t(x)$ , 且为单根.

文[10]给出了一种解方程(3.1)的稳定计算策略, 本节结合三次多项式根的性质及[10]的结果, 给出了另外一种解方程(3.1)的计算策略. 数值结果表明本文给出的计算策略在数值计算中更加稳定.

令  $p = \frac{-b^2 + 3ac}{3a^2}$ ,  $q = \frac{2b^3 - 9abc + 27a^2d}{27a^3}$ ,  $\Delta = (\frac{q}{2})^2 + (\frac{p}{3})^3$ . 由三次多项式方程根的性质, 及方程(3.1)在  $\mathbb{R}$  上仅有一个实根  $\alpha_t(x)$  且为单根, 可知  $\Delta > 0$  且

$$\alpha_t(x) = -\frac{b}{3a} + \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}.$$

为了减小计算中的舍入误差, 可将其等价转化为

$$\alpha_t(x) = \begin{cases} -\frac{b}{3a} - \frac{p}{3\sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\Delta}}, & q > 0, \\ -\frac{b}{3a} - \frac{p}{3\sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}}} + \sqrt[3]{-\frac{q}{2} + \sqrt{\Delta}}, & q < 0. \end{cases}$$

为了进一步减小计算中的舍入误差, 令  $\bar{f} = \frac{\sum_{j \in Q_t(x)} f_j(x)}{\bar{q}}$ ,  $f_j^*(x) = f_j(x) - \bar{f}$ , 将  $\alpha$  进行平移变换可得

$$\sum_{j \in Q_t(x)} (1 + \frac{f_j^*(x) - \alpha}{t})^3 - 1 = a_1 \alpha^3 + b_1 \alpha^2 + c_1 \alpha + d_1,$$

其相应的系数为:

$$\begin{cases} a_1 = -\frac{\bar{q}}{t^3}, \\ b_1 = \frac{3\bar{q}}{t^2}, \\ c_1 = -\sum_{j \in Q_t(x)} \frac{3f_j^*(x)^2}{t^3} - \frac{3\bar{q}}{t}, \\ d_1 = \sum_{j \in Q_t(x)} \frac{f_j^*(x)^3}{t^3} + \sum_{j \in Q_t(x)} \frac{3f_j^*(x)^2}{t^2} + \bar{q} - 1, \end{cases} \quad (3.2)$$

则

$$\alpha_t(x) = \begin{cases} -\frac{b_1}{3a_1} - \frac{p_1}{3\sqrt[3]{-\frac{q_1}{2} - \sqrt{\Delta_1}}} + \sqrt[3]{-\frac{q_1}{2} - \sqrt{\Delta_1}} + \bar{f}, & q_1 > 0, \\ -\frac{b_1}{3a_1} - \frac{p_1}{3\sqrt[3]{-\frac{q_1}{2} + \sqrt{\Delta_1}}} + \sqrt[3]{-\frac{q_1}{2} + \sqrt{\Delta_1}} + \bar{f}, & q_1 \leq 0, \end{cases}$$

其中  $p_1 = \frac{-b_1^2 + 3a_1c_1}{3a_1^2}$ ,  $q_1 = \frac{2b_1^3 - 9a_1b_1c_1 + 27a_1^2d_1}{27a_1^3}$ ,  $\Delta_1 = (\frac{q_1}{2})^2 + (\frac{p_1}{3})^3$ . 由于系数(3.2)中有多项以光滑化参数  $t$  做分母, 当  $t$  较小时, 计算中容易产生溢出, 因此将系数同时乘以  $t^3$  得到

$$\begin{cases} a_2 = -\bar{q}, \\ b_2 = 3\bar{q} \cdot t, \\ c_2 = -\sum_{j \in Q_t(x)} 3f_j^*(x)^2 - 3\bar{q} \cdot t^2, \\ d_2 = \sum_{j \in Q_t(x)} f_j^*(x)^3 + \sum_{j \in Q_t(x)} 3f_j^*(x)^2 \cdot t + \bar{q} \cdot t^3 - t^3, \end{cases}$$

则

$$\alpha_t(x) = \begin{cases} -\frac{b_2}{3a_2} - \frac{p_2}{3\sqrt[3]{-\frac{q_2}{2} - \sqrt{\Delta_2}}} + \sqrt[3]{-\frac{q_2}{2} - \sqrt{\Delta_2}} + \bar{f}, & q_2 > 0, \\ -\frac{b_2}{3a_2} - \frac{p_2}{3\sqrt[3]{-\frac{q_2}{2} + \sqrt{\Delta_2}}} + \sqrt[3]{-\frac{q_2}{2} + \sqrt{\Delta_2}} + \bar{f}, & q_2 \leq 0, \end{cases} \quad (3.3)$$

其中  $p_2 = \frac{-b_2^2 + 3a_2c_2}{3a_2^2}$ ,  $q_2 = \frac{2b_2^3 - 9a_2b_2c_2 + 27a_2^2d_2}{27a_2^3}$ ,  $\Delta_2 = (\frac{q_2}{2})^2 + (\frac{p_2}{3})^3$ .

**定理3.1** 对任意的  $x^1, x^2 \in \mathbb{R}^n$ ,  $t > 0$ , 若下述两个条件成立,

$$\min_{j \in Q_t(x^1)} f_j(x^2) - \alpha_t^{x^1}(x^2) + t > 0, \quad (3.4)$$

$$\max_{j \in Q \setminus Q_t(x^1)} f_j(x^2) - \alpha_t^{x^1}(x^2) + t \leq 0, \tag{3.5}$$

其中 $\alpha_t^{x^1}(x^2)$ 为方程  $\sum_{j \in Q_t(x^1)} (1 + \frac{f_j(x^2) - \alpha}{t})^3 - 1 = 0$ 的根, 则

$$\alpha_t(x^2) = \alpha_t^{x^1}(x^2).$$

证 由(3.4)及(3.5)式可得

$$\sum_{j \in Q} \phi_t(f_j(x^2) - \alpha_t^{x^1}(x^2)) - 1 = \sum_{j \in Q_t(x^1)} (1 + \frac{f_j(x^2) - \alpha_t^{x^1}(x^2)}{t})^3 - 1,$$

再由  $\sum_{j \in Q_t(x^1)} (1 + \frac{f_j(x^2) - \alpha_t^{x^1}(x^2)}{t})^3 - 1 = 0$ 可知,

$$\sum_{j \in Q} \phi_t(f_j(x^2) - \alpha_t^{x^1}(x^2)) - 1 = 0,$$

即 $\alpha_t^{x^1}(x^2)$ 是方程  $\sum_{j \in Q} \phi_t(f_j(x^2) - \alpha) - 1 = 0$ 的根, 由于该方程在 $\mathbb{R}$ 上仅有一个实根 $\alpha_t(x^2)$ , 因此 $\alpha_t^{x^1}(x^2) = \alpha_t(x^2)$ .

#### 4. 积极集策略的光滑化算法及其收敛性

本节基于前两节给出的积极集策略的光滑化max函数 $\alpha_t(x)$ , 结合文[11]中的Newton-Armijo算法, 给出一种解含多个复杂分量函数无约束minimax问题的积极集策略的光滑化算法.

##### 算法4.1

步0 (参数)  $\alpha, \beta, \omega \in (0, 1), \bar{l} > 0, \theta \in \mathbb{N}_+, \kappa \gg 1$ , 满足  $\lim_{t \downarrow 0} \tau(t) = 0$ 的 $\tau(t) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ .

$k = 0, x^0 \in \mathbb{R}^n, t_0 \in \mathbb{R}_+, h_{t_0}^0(x^0) = 1$ .

步1 (计算搜索方向) 若 $\nabla^2 \alpha_{t_k}(x^k)$ 正定且条件数 $c(\nabla^2 \alpha_{t_k}(x^k)) < \kappa$ , 选择牛顿方向(4.1); 否则, 选择负梯度方向(4.2).

$$d_{t_k}(x^k) = -\nabla^2 \alpha_{t_k}(x^k) \setminus \nabla \alpha_{t_k}(x^k), \tag{4.1}$$

$$d_{t_k}(x^k) = -\nabla \alpha_{t_k}(x^k), \tag{4.2}$$

转步2.

步2 (计算步长) 计算 $h_{t_k}(x^k) = h_{t_k}^0(x^k)\beta^l$ , 其中 $l$ 为满足下述条件的最小非负整数

$$\alpha_{t_k}(x^k + h_{t_k}^0(x^k)\beta^l d_{t_k}(x^k)) - \alpha_{t_k}(x^k) \leq \alpha h_{t_k}^0(x^k)\beta^l \langle \nabla \alpha_{t_k}(x^k), d_{t_k}(x^k) \rangle.$$

若 $l > \bar{l}$ , 终止算法; 否则, 转步3.

步3 (更新迭代点) 令 $x^{k+1} = x^k + h_{t_k}(x^k)d_{t_k}(x^k)$ , 转步4.

步4 若 $\|\nabla \alpha_{t_k}(x^k)\| < \tau(t_k)$ , 令 $h_{t_k}^0(x^k) = 1$ , 转步5; 否则, 令 $t_{k+1} = t_k, h_{t_{k+1}}^0(x^{k+1}) = \min\{1, \frac{h_{t_k}^0(x^k)}{\beta^\theta}\}$ ,  $k = k + 1$ , 转步1.

步5 (调整光滑化参数) 令 $t_{k+1} = \omega \cdot t_k, k = k + 1$ , 转步4.

**注4.1** 当光滑化参数 $t$ 较小, 步2中满足线搜索条件的步长一般较小, 在步4中令初始迭代搜索步长 $h_{t_k}^0(x^k) = \min\{1, \frac{h_{t_k}^0(x^k)}{\beta^\theta}\}$ , 此设置可减少步长的搜索次数, 提高算法的计算效率.

**注4.2** 在算法中, 若相邻的迭代点满足定理3.1中的条件, 则这两个迭代点相对应的指标集相等, 无需计算第二个迭代点对应的指标集, 否则利用(2.6)式计算相应的指标集, 该策略可减少指标集的计算次数, 提高算法的计算效率.

类似于文[9]中定理3.2的证明, 可以得到算法4.1的收敛性定理.

**定理4.1** 若假设1.1成立, 设 $\{t_k\}_{k=0}^\infty$ 和 $\{x^k\}_{k=0}^\infty$ 是由算法4.1产生的有界序列, 则存在无穷子序列 $N' \subseteq \mathbb{N}$ 和 $x^* \in \mathbb{R}^n$ , 当 $k \rightarrow^{N'} \infty$ 时,  $x^k \rightarrow x^*, 0 \in \partial F(x^*)$ , 即 $x^*$ 是问题(1.1)的稳定点.

## 5. 数值实验

本节将文[11]中基于凝聚函数提出的第二种解无约束minimax问题的算法记为PRW2, 其光滑化参数采用与算法4.1对应的更新策略 $p_{k+1} = p_k/\omega$ . 本节选取了两个梯度及海森矩阵较为复杂的半无限minimax问题, 将其等距离离散化后得到两类含多个复杂分量函数的无约束minimax问题, 将算法4.1同算法PRW2进行了数值比较. 算法在MATLAB R2014a中编程实现, 在配置为INTEL(R) Core(TM) i5-6200U 2.30GHz, 内存为4.00GB的笔记本上运行. 数值结果表明, 对含多个复杂分量函数的无约束minimax问题, 算法4.1的稳定性与计算效率均高于算法PRW2.

### 例5.1<sup>[12]</sup>

$$F(x) = \max_{y \in Y} f(x, y),$$

$$f(x, y) = x_1^2 \exp(-x_2 y) \cos^2(x_3 y + x_4) - \cos(y) \\ + x_2^2 x_3^2 \exp(-x_1 y) \sin^2(x_2 y) + \exp((1 - x_6)^2 y + x_5^2),$$

$n = 6, Y = [0, 10], F(x^*) = 2, x^0 = (1, 1, 1, 1, 1, 1)$ .

### 例5.2<sup>[12]</sup>

$$F(x) = \max_{y \in Y} |f(x, y)|,$$

$$f(x, y) = \frac{x_3}{x_2} \exp(-y x_1) \sin(y x_2) + x_1 \exp(-x_2 y) \cos(x_3 y + x_4) + x_5 \exp(-x_6 y) \\ - \left( \frac{3}{20} \exp(-y) + \frac{1}{52} \exp(-5y) - \frac{1}{65} \exp(-2y) (3 \sin(2y) + 11 \cos(2y)) + \frac{1}{2} \exp(-y) \right. \\ \left. - \exp(-2y) + \frac{1}{2} \exp(-3y) + \frac{3}{2} \exp(-\frac{3}{2}y) \sin(7y) + \exp(-\frac{5}{2}y) \sin(5y) \right),$$

$n = 6, Y = [0, 1], F(x^*) = 0.162825, x^0 = (2, 2, 7, 0, -2, 1)$ .

算法4.1的参数设置如下:

$$\alpha = 0.5, \beta = 0.8, \omega = 0.1, \bar{l} = 60, \theta = 4, \kappa = 10^{10}, \tau(t_k) = 10^{-3}, t_0 = 1.$$

终止准则为

$$\|\nabla \alpha_{t_k}(x^k)\| < 10^{-3}, t_k < 10^{-3}.$$

相对应的将算法PRW2的终止准则设为

$$\|\nabla F_{p_k}(x^k)\| < 10^{-3}, p_k > 10^3.$$

例5.1及例5.2等距离离散化后分量函数的个数分别取 $2.5k_1 \times 10^5 (k_1 = 1, \dots, 40)$ 与 $2k_2 \times 10^3 (k_2 = 2, \dots, 40)$ . 在算法PRW2中, 当计算机内存溢出时, 算法停止.

图5.1给出了例5.1及例5.2在不同离散程度下算法4.1与算法PRW2的总迭代时间. 由图可知, 两者的总迭代时间基本随着分量函数个数的增加呈上升趋势, 且算法4.1的总迭代时间均比算法PRW2的总迭代时间少. 随着分量函数个数的增加, 算法4.1的总迭代时间增长稳定性较好, 而算法PRW2稳定性较差, 例如在例5.2中, 当分量函数个数为 $1.2 \times 10^4, 3 \times 10^4$ 和 $3.2 \times 10^4$ 时, 算法PRW2的迭代次数分别为86, 68和60, 总迭代时间分别为3.80秒, 6.37秒和5.53秒, 而其它离散化问题的迭代次数介于4650到5489之间, 总时间介于176.02秒到1029.10秒之间.

图5.2给出了例5.1及例5.2在不同离散程度下算法4.1与算法PRW2的平均迭代时间. 由图可知, 算法4.1的平均迭代时间均比算法PRW2的平均迭代时间少. 算法4.1在例5.1中平均时间较高的离散问题有8个, 对应分量函数的个数为 $6.5 \times 10^6, 6.75 \times 10^6, 7 \times 10^6, 7.25 \times 10^6, 7.5 \times 10^6, 7.75 \times 10^6, 8 \times 10^6, 8.25 \times 10^6$ , 分量函数赋值次数分别为39, 40, 41, 43, 56, 48, 57, 54, 梯度赋值次数分别为13, 13, 13, 13, 14, 15, 18, 21, 海森矩阵赋值次数及迭代次数分别为9, 9,

9, 9, 10, 11, 14, 17, 平均迭代时间分别为10.58秒, 11.67秒, 12.09秒, 12.87秒, 13.40秒, 11.71秒, 10.84秒, 9.87秒, 而其它离散化问题的分量函数赋值次数为41或42, 梯度赋值次数为23或24, 海森矩阵赋值次数及迭代次数为19, 20或21, 平均迭代时间介于3.10秒到7.04秒之间. 此外, 由图5.1及5.2可知, 当例5.1及例5.2中分量函数的个数分别大于 $4 \times 10^6$ 和 $4 \times 10^3$ 时, 由于计算机内存溢出, 算法PRW2失败.

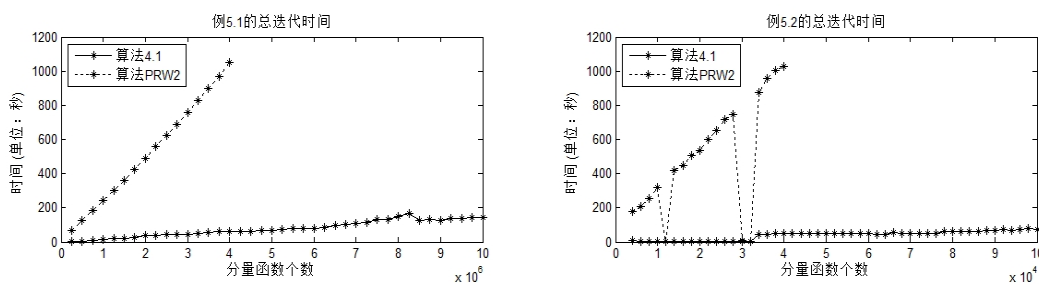


图5.1 算法4.1与算法PRW2总迭代时间

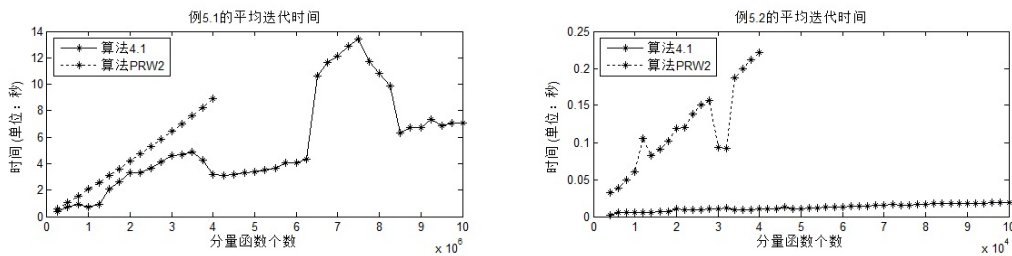


图5.2 算法4.1与算法PRW2平均迭代时间

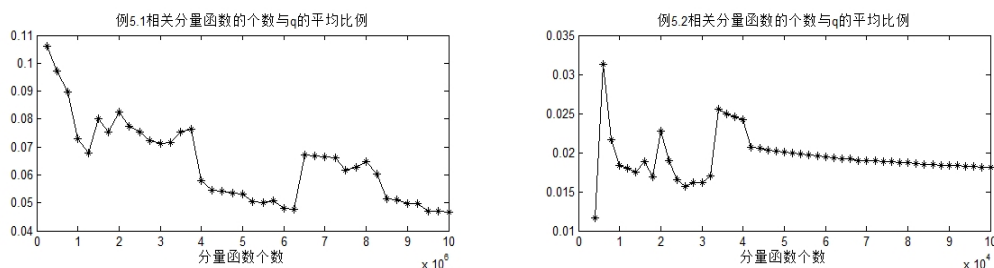


图5.3 算法4.1中相关分量函数的个数与q的平均比例

图5.3给出了例5.1及例5.2在不同离散程度下算法4.1中相关分量函数的个数与 $q$ 的平均比例. 由图可知, 随着分量函数个数的增多, 相关分量函数的个数与 $q$ 的平均比例总体呈递减趋势, 例5.1及例5.2在算法4.1的迭代中相关分量函数的个数平均比例不超过11%和3.5%. 因此, 当分量函数较多且复杂时, 算法4.1仅需计算少部分分量函数的梯度及海森矩阵, 而算法PRW2在每次迭代中均需计算所有分量函数的梯度及海森矩阵, 因此, 算法4.1的计算效率优于算法PRW2.

参考文献:

[1] Al-SUBAIHI I, WATSON G A. Fitting parametric curves and surfaces by  $l_\infty$  distance regression[J]. BIT Numerical Mathematics, 2005, 45(3): 443-461.



- [2] POLAK E. On the mathematical foundations of nondifferentiable optimization in engineering design[J]. Siam Review, 1987, 29(1): 21-89.
- [3] BANICHUK N V. Minimax approach to structural optimization problems[J]. Journal of Optimization Theory Applications, 1976, 20(1): 111-127.
- [4] LUSS H. Minimax resource allocation problems: Optimization and parametric analysis[J]. European Journal of Operational Research, 1992, 60(1):76-86.
- [5] HANSEN P, PEETERS D, RICHARD D, et al. The minisum and minimax location problems revisited[J]. Operations Research, 1985, 33(6): 1251-1265.
- [6] FORTS S, STEPHEN E. Approximate direct-search minimax circuit optimization[J]. International Journal for Numerical Methods in Engineering, 1989, 28(2): 359-368.
- [7] POLAK E, WOMERSLEY R S, YIN H X. An algorithm based on active sets and smoothing for discretized semi-infinite minimax problems[J]. Journal of Optimization Theory and Applications, 2008, 138(2): 311-328.
- [8] LI X S. An aggregate function method for nonlinear programming[J]. Science in China, Ser. A, 1991, 34(12): 1467-1473.
- [9] 贾智媛. 解有限维无约束极大极小问题的积极集光滑化牛顿法[D]. 山西师范大学, 2016.
- [10] 李群. 解无约束minimax的积极集光滑化牛顿法[D]. 山西师范大学, 2016.
- [11] POLAK E, ROYSET J O, WOMERSLEY R S. Algorithms with adaptive smoothing for finite minimax problems[J]. Journal of Optimization Theory and Applications, 2003(3): 459-484.
- [12] 肖瑜. 截断凝聚光滑化算法[D]. 大连理工大学, 2010.

## An Active Set Smoothing Method for Solving Unconstrained Minimax Problem with Multiple Complex Component Functions

ZHOU Zhengyong, QIN Lina

(School of Mathematics and Computer Science, Shanxi Normal University,  
Linfen 041004, China)

**Abstract:** In this paper, a twice continuously differentiable active set smoothing maximum function is proposed, the smoothing maximum function is constructed by a piecewise cubic polynomial equation, and the stable strategies for computing the active set and the smoothing maximum function are given. Combining the active set smoothing maximum function, the Armijo line search, the steep descent direction and the Newton direction, and an update strategy for the smoothing parameters, an active set smoothing method is proposed for solving unconstrained minimax problems with multiple component functions. Preliminary numerical experiments show the efficiency of the proposed method.

**Key words:** Minimax problem; Active set; Smoothing max function; Smoothing method