

基于积极集识别技术的半无限 minimax 问题非单调有限记忆 SQP 算法

杨永亮, 王福胜, 甄 娜
(太原师范学院数学系, 山西 晋中 030619)

摘 要: 本文研究了半无限 minimax 问题. 利用积极集识别技术结合非单调有限记忆序列二次规划 (SQP) 方法来求解半无限 minimax 问题. 在适当的条件下证明了算法的收敛性. 数值结果表明新算法在降低求解规模和迭代次数等方面均优于采用 Armijo 型线搜索的 SQP 方法.

关键词: 极大极小问题; 积极集; 离散化方法; SQP 算法; 非单调技术

MR(2010) 主题分类号: 90C30; 65K05

中图分类号: O224

文献标识码: A

文章编号: 0255-7797(2020)05-0577-08

1 引言

半无限 minimax 优化问题是一类非常重要的优化问题, 有着广泛的应用背景, 例如工程设计、最优控制、金融工程等领域的很多问题可以归结为求解这类优化问题, 很多学者对此进行了研究, 获得了丰富的研究成果 (如文献 [1–12]). 由于其特殊的结构, 大多数传统的方法不再适用, 离散化方法是求解这类型问题的重要数值方法之一 (如文献 [1–8]). 半无限 minimax 问题具有如下形式

$$\begin{cases} \min_{x \in R^n} \max_{y \in Y} f(x, y), \\ \text{s.t. } g(x, \omega) \leq 0, \omega \in \Omega, \end{cases} \quad (1.1)$$

其中指标集 $Y = [1, 2, \dots, m]$, $f: R^n \times Y \rightarrow R$, 关于 x, y 都连续可微关于 x 连续可微 $g: R^n \times [0, 1] \rightarrow R$. 为了方便起见, 记问题 (1.1) 的可行集 X 和水平集 $l(x^0, \Omega)$:

$$X = \{x \in R^n | g(x, \omega) \leq 0, \omega \in \Omega\}, l(x^0, \Omega) = \{x \in X | f(x, y) \leq f(x^0, y)\}.$$

再记: 对于任意的 $x \in R^n$, $F_Y(x) = \max_{y \in Y} f(x, y)$. 对于 $\hat{Y} \subset Y$, $F_{\hat{Y}}(x) = \max_{y \in \hat{Y}} f(x, y)$ 的方向导数

$$F'_{\hat{Y}}(x, d) = \max_{y \in \hat{Y}} (f(x, y) + \nabla_x f(x, y)^T d) - F_Y(x).$$

离散化方法的主要思想通过不断离散连续变量的连续区间来逼近约束函数, 将求解半无限规划问题转化为求解其离散后的一系列有限约束优化问题. 将 Ω 离散成有限

*收稿日期: 2019-10-15

接收日期: 2020-03-17

基金项目: 山西省回国留学人员科研项目 (2017-104); 太原师范学院研究生创新项目 (SYYSJSJC-1911).

作者简介: 杨永亮 (1993-), 男, 山西吕梁, 硕士, 主要研究方向: 最优化理论与方法

通讯作者: 王福胜

集: $\Omega_E = \left\{0, \frac{1}{q}, \frac{2}{q}, \dots, \frac{q-1}{q}, 1\right\}$, 其中 q 反映了离散水平, q 越大离散水平越好. 定义 Ω 和 Ω_E 之间的 Hausdorff 距离为 $\text{dist}(\Omega_E, \Omega) = \max_{\omega \in \Omega} \min_{\omega \in \Omega_E} \|\omega_E - \omega\|$, 其中集列 $\{\Omega_E^{q_i}\}_{i \in N^+}$ 满足条件

$$\Omega_E^{q_i} \subset \Omega, |\Omega_E^{q_i}| < \infty, \quad i \in N^+, \lim_{i \rightarrow \infty} \text{dist}(\Omega_E^{q_i}, \Omega) = 0. \quad (1.2)$$

基于离散化方法求解原问题 (1.1) 可归结为求解一系列具有如下形式的 minimax 离散化问题:

$$\min_{x \in R^n} \max_{y \in Y} f(x, y) \text{ s.t. } g(x, \omega) \leq 0, \quad \omega \in \Omega_E. \quad (1.3)$$

在一定的条件下, 当 $\text{dist}(\Omega_E, \Omega) \rightarrow 0$ 时式 (1.3) 的最优解趋向于原问题 (1.1) 的最优解. 当 q 非常大的时候, 问题 (1.3) 的约束个数非常多, 求解的成本也会很高, 如何设计高效的算法求解问题 (1.3) 是解决半无限 minimax 问题的一个关键. 文献 [5] 提出了一种求解半无限规划问题的超线性收敛的模松弛 SQP 算法, 每次迭代只需要求解一个 QP 子问题就可以获得搜索方向, 遗憾的是上述算法要求初始点可行, 而通常求解可行点的计算量很大. 为了克服这一问题, 文献 [6] 提出了一种初始点任意的模松弛强次可行 SQP 算法. 另外, 在模松弛强次可行方向法中常利用线搜索来确定步长, 传统的线搜索方法都要求目标函数值严格下降, 其缺点是当迭代点“陷入很窄的峡谷时”, 常常会导致小步长或出现锯齿现象, 而采用非单调线搜索技术可以克服这些缺点 (如文献 [13–15]). 文献 [9] 提出一种约束积极集识别技术, 可以对积极集进行精确识别和估计来减少计算量. 文献 [16] 提出了一种求解无约束优化问题的有限记忆 BFGS 修正规则 (简称 L-BFGS), L-BFGS 修正方法无需存贮近似海森矩阵 H_k , 从而大大减少了计算机存储量, 提高运行效率, 对大规模的优化问题更有效. 受文献 [5–9] 启发, 本文结合离散化方法和模松弛强次可行 SQP 方法, 基于新型约束积极集识别技术, 采用非单调搜索和有限记忆 L-BFGS 修正方法更新 H_k , 提出了一种新的求解半无限 minimax 问题的非单调 SQP 算法.

2 算法描述

定义 2.1 点 x 称为 QP 子问题 (2.1) 的稳定点 (KKT 点), 如果存在乘子向量 $\lambda_\omega, (\omega \in \Omega_E); \gamma_y, (y \in Y)$ 使得

$$\sum_{y \in Y} \gamma_y \nabla_x f(x, y) + \sum_{\omega \in \Omega_E} \lambda_\omega \nabla_x g(x, \omega) = 0, \sum_{y \in Y} \gamma_y = 1, \quad (2.1)$$

$$\begin{cases} 0 \leq \gamma_y \perp (f(x, y) - F_Y(x)) \leq 0, \forall y \in Y, \\ 0 \leq \lambda_\omega \perp g(x, \omega) \leq 0, \forall \omega \in \Omega_E. \end{cases} \quad (2.2)$$

以下给出了式 (1.3) 的拉格朗日函数 $L(x, \lambda, \gamma)$, 可行集 X_E 和有效集 $Y(x)$ 和 $\Omega_E(x)$:

$$\begin{aligned} L(x, \lambda, \gamma) &= \sum_{y \in Y} \gamma_y f(x, y) + \sum_{\omega \in \Omega_E} \lambda_\omega g(x, \omega), \\ X_E &= \{x \in R^n, g(x, \omega) \leq 0, \forall \omega \in \Omega_E\}, \\ Y(x) &= \{y \in Y : f(x, y) = F(x)\}, \\ \Omega_E(x) &= \{\omega \in \Omega_E : g(x, \omega) = 0\}. \end{aligned}$$

由式 (2.1), (2.2), 定义如下的约束积极集识别函数

$$\phi(x, \lambda, \gamma) = \begin{pmatrix} \nabla_x L(x, \lambda, \gamma) \\ \sum_{y \in Y} \gamma_y^k - 1 \\ \min(F_Y(x)e - f(x, y), \gamma) \\ \min(-g(x, \omega), \lambda) \end{pmatrix}, \quad (2.3)$$

$$\rho(x, \lambda, \gamma) = \|\phi(x, \lambda, \gamma)\|^\delta.$$

其中 $e = (1, 1, \dots, 1)^T \in R^n$, $0 < \delta < 1$, 由文 [5] 可得到问题 (1.3) 的两个积极约束识别集

$$\begin{cases} Y(x, \lambda, \gamma) = \{y \in Y : f(x, y) - F_Y(x) + \rho(x, \lambda, \gamma) \geq 0\}, \\ \Omega_E(x, \lambda, \gamma) = \{\omega \in \Omega_E : g(x, \omega) + \rho(x, \lambda, \gamma) \geq 0\}. \end{cases} \quad (2.4)$$

为了叙述方便, 对任给的 $x^k \in R^n, \hat{Y} \subset Y(x, \lambda, \gamma), \Omega_k \subset \Omega_E(x, \lambda, \gamma)$, 定义记号如下

$$\begin{aligned} \phi(x) &= \max \left\{ 0, \max_{\omega \in \Omega_E} g(x, \omega) \right\}, \\ \varphi_k &= \varphi(x^k) = \max \left\{ 0, \max_{\omega \in \Omega_k} g(x^k, \omega) \right\} = \max \{0, g(x^k, \omega), \omega \in \Omega_k\}, \\ \Omega_k^- &= \Omega^-(x^k) = \{\omega \in \Omega_k | g(x^k, \omega) \leq 0\}, \Omega_k^+ = \Omega^+(x^k) = \{\omega \in \Omega_k | g(x^k, \omega) > 0\}. \end{aligned}$$

构造如下的 QP 子问题

$$\begin{cases} \min r_0 z + \frac{1}{2} d^T H_k d, \\ \text{s.t. } F'_{\hat{Y}}(x^k, d) \leq r_0 z + r \varphi_k^\theta, \\ g(x^k, \omega) + \nabla_x g(x^k, \omega)^T d \leq r_\omega \sigma_k z, \forall \omega \in \Omega_k^-, \\ g(x^k, \omega) + \nabla_x g(x^k, \omega)^T d \leq r_\omega \sigma_k z - \varepsilon_k z^2 + \phi_k, \forall \omega \in \Omega_k^+, \end{cases} \quad (2.5)$$

其中参数 $r_0, r, r_\omega (\omega \in \Omega_E), \theta$ 都是正的常数, σ_k 是随迭代调整的正参数, H_k 是 $\nabla_{xx}^2 f(x^k, y)$ 的近似矩阵. 我们引入一个重要的项 $-\varepsilon_k z^2$ 是为了确保当 $d^k \rightarrow 0$ 有 $z_k \rightarrow 0$. 以下引理说明 QP 子问题具有良好的性质.

假设 2.1 对于任意的 $y \in Y, \omega \in \Omega$, 函数 $f(\cdot, y)$ 和 $g(\cdot, \omega)$ 在可行集上至少一阶连续可微.

假设 2.2 对于任意的 $x \in X_E$, 弱 MFCQ 成立, 即存在 $d \in R^n$ 使得 $\nabla_x g(x, \omega)^T d < 0$, 对于所有的 $\omega \in \Omega(x)$ 成立.

引理 2.1 设假设 2.1 和假设 2.2 成立, $x^k \in X, H_k$ 是对称正定, 若 (z_k, d^k) 是子问题 (2.5) 的最优解, 则

- (1) $r_0 z_k + \frac{1}{2} (d^k)^T H_k d^k \leq 0, z_k \leq 0$;
- (2) $z_k = 0 \Leftrightarrow d^k = 0$;
- (3) $z_k = 0 \Rightarrow x^k$ 是问题 (1.3) 的一个稳定点;
- (4) 如果 $x^k \notin X$ 则 $z_k < 0$;
- (5) 若 $d^k \neq 0$, 则 $z_k < 0, d^k$ 为问题 (1.3) 在 x^k 处的可行下降方向.

证 (1), (2) 的证明类似于文献 [9](引理 2.2,2.3) 其余证明可参考文献 [7](引理 2.1).

在文献 [13] 中 Zhang 和 Hager 提出了新的非单调搜索技术, 受文献 [13] 启发, 本文对其进行了改进, 具体形式如下

$$F_{\hat{Y}}(x^k + \alpha_k d^k) \leq C_k + \delta \alpha_k \left[\nabla_x f(x, y)^T d^k + \frac{1}{2} \alpha_k \gamma L_k \|d^k\|^2 \right], \quad (2.6)$$

$$C_k = \begin{cases} F_{\hat{Y}}(x^k), & k = 1; \\ \eta C_{k-1} + (1 - \eta) F_{\hat{Y}}(x^k), & k \geq 2. \end{cases} \quad (2.7)$$

其中 $\gamma \in [0, 2)$, $\delta \in (0, \frac{1}{2})$, $s_k = \frac{-\nabla_x f(x, y)^T d^k}{L_k \|d^k\|^2}$, $t \in (0, 1)$, α_k 为 $\{s_k, ts_k, t^2 s_k, \dots\}$ 中满足上式的最大者, L_k 的调整规则参见文献 [15].

由于采用离散化技术后, 问题 (1.3) 的约束个数较多, 导致算法求解的计算量增加, 效率降低, 如何降低计算量成为本文的关键. 在 SQP 算法中通常用 BFGS 公式更新 H_k , 文 [16] 提出一种新的有限记忆 L-BFGS 更新规则, 它最显著的特点是不需要存储 H_k , 对给定的 H_k 和非负整数 m , 利用前 m 步的信息对 H_0 进行修正 m 次得到 H_k , 仅存贮 $m+1$ 个向量组就能计算 H_{k+1} , 从而降低了算法对计算量和存储量的要求, 本文将应用到算法设计中更新 H_k .

算法 A (求解离散化问题 (1.3))

步 1 初始化. 取适当大正整数 q , 将区间 $[0, 1]$ 离散成有限集 Ω_k 选取参数 $r_0, r, r_\omega, \theta \in (0, 1)$, $\sigma_k, \eta \in (0, 1)$. 选取初始可行点 $x^0 \in X_E$, 初始对称正定矩阵 H_0 , $(\lambda^0, \gamma^0)^T = (1, 1, \dots, 1)^T \in R^{m+q+1}$, 并令 $k := 0$.

步 2 由 (2.3) 式计算 $\rho(x^k, \lambda^k, \gamma^k)$, 由 (2.4) 式生成积极约束识别集 $Y(x^k, \lambda^k, \gamma^k)$ 和 $\Omega_E(x^k, \lambda^k, \gamma^k)$. 如果 $\rho(x^k, \lambda^k, \gamma^k) = 0$, 则 x^k 是问题 (1.3) 的一个稳定点, 停止. 否则进入步 3.

步 3 计算搜索方向. 对当前迭代点 x^k 求解 QP 子问题 (2.5) 得到一个 KKT 点对 (z_k, d^k) . 如果 $d^k = 0$, 则 x^k 是问题 (1.3) 的一个稳定点, 停止. 否则进入步 4.

步 4 求搜索步长. 由新的非单调搜索 (2.6) 获得步长 α_k .

步 5 令 $x^{k+1} = x^k + \alpha_k d^k$, 对称正定矩阵 k_{k+1} 可按文献 [16] 中 L-BFGS 更新规则得到, σ_{k+1} 由 $\sigma_{k+1} = \min \{\bar{\sigma}, \sigma_1 \|d^k\|^\xi\}$, $\forall k \geq 1$ 获得, 其中 $\bar{\sigma}, \sigma_1, \xi$ 均为正常数. 置 $k = k+1$, 返回步 2.

对于半无限 minimax 问题的离散化方法, 由文献 [7] 可知, 若下面的假设 2.3 成立, 且存在 $x^0 \in X$, 使得 $l(x^0, \Omega_E^{q_0})$ 紧, 那么离散化问题 (1.3) 解序列的某个子列 $\{\tilde{x}_{q_i}\}_{i \in N^k}$, $N^k \subseteq N, k \in R$ 收敛于原问题 (1.1) 的解 $\tilde{x}$.

假设 2.3 集列 $\{\Omega_E^{q_i}\}_{i \in N^+}$ 满足条件 (1.2), 且 $\Omega_E^{q_i} \subseteq \Omega_E^{q_{i+1}}, i \in N^+$ 成立.

下面给出求解原问题 (1.1) 的算法

算法 B

初始步 给定参数: $\{\pi_i\}_{i \in N^+}$ 满足 $0 < \varepsilon < \pi_{i+1} < \pi_i, \forall i \in N^+$ 和 $\lim_{i \rightarrow \infty} \pi_i = 0$, $q_0 \in N^+ \setminus \{0\}, \varepsilon = 10^{-4}$ 以及算法 A 的初始化参数.

步 1 选取 $x^0 \in X$ 离散集 $\Omega_E^{q_0} \subset \Omega$, 满足 $|\Omega_E^{q_0}| < \infty$. 令 $i = 0$.

步 2 利用算法 A 求解离散化问题 (1.3), 其最优解记为 $\tilde{x}_{q_i}$.

步 3 如果 $\text{dist}(\Omega_E^{q_i}, \Omega) \leq \varepsilon$, 则算法停止, 否则, 取更为精细的离散集 $\Omega_E^{q_{i+1}} \subset \Omega$, 使得 $\{\omega_{q_i}\} \cup \Omega_E^{q_i} \subset \Omega_E^{q_{i+1}}$, 其中 $q_i \in \Omega_E^{q_{i+1}}, |\Omega_E^{q_{i+1}}| < \infty$ 和 $\text{dist}(\Omega_E^{q_{i+1}}, \Omega) \leq \text{dist}(\Omega_E^{q_i}, \Omega) \leq \pi_{i+1}$, 且满足 $g(\tilde{x}_{q_i}, \omega) = \max_{\omega \in \Omega_{q_{i+1}}} g(\tilde{x}_{q_i}, \omega)$.

步 4 令 $x^0 = \tilde{x}_{q_i}, i = i + 1$, 转步 2.

3 收敛性分析

本节对算法 B 进行收敛性分析, 首先给出如下的记号

$$\Phi(x) = \max \left\{ 0, \max_{\omega \in \Omega} g(x, \omega) \right\}, \varphi_{q_i}(x) = \max \left\{ 0, \max_{\omega \in \Omega_E^{q_i}} g(x, \omega) \right\}, X = \{x \in R^n | g(x, \omega) \leq 0, \omega \in \Omega\},$$

$$\Omega(x) = \{\omega \in \Omega | g(x, \omega) = 0\}, \Omega_E^{q_i}(x) = \{\omega \in \Omega_E^{q_i} | g(x, \omega) = 0\}.$$

对于算法 B, 定义如下的水平集

$$l(x^0, \Omega_E^{q_0}) = \{x \in R^n | \varphi_{q_0}(x) \leq \Phi(x^0)\}, l(x^0, \Omega_E^{q_i}) = \{x \in R^n | \varphi_{q_i}(x) \leq \Phi(x^0)\}. \quad (3.1)$$

假设 3.1 水平集 $l(x^0, \Omega_E^{q_0})$ 有界, 且存在 Lipschitz 常数 l_{g_x} 和常数 l_{g_w} 使得下式成立.

$$\begin{aligned} |g(x^1, \omega) - g(x^2, \omega)| &\leq l_{g_x} \|x^1 - x^2\|, \omega \in \Omega, x^1, x^2 \in l(x^0, \Omega_E^{q_0}), \\ |g(x, \omega_1) - g(x, \omega_2)| &\leq l_{g_w} \|\omega_1 - \omega_2\|, \omega_1, \omega_2 \in \Omega, x \in l(x^0, \Omega_E^{q_0}). \end{aligned} \quad (3.2)$$

假设 3.2 问题 (1.1) 在点 $\tilde{x} \in X$ 时有 $\{\nabla_x g(\tilde{x}, \omega), \omega \in \Omega(\tilde{x})\}$ 线性无关.

引理 3.1 [7] 若 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 为算法 B 生成的迭代点列, 则序列 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 存在聚点 $\tilde{x}$.

引理 3.2 令 $\{\tilde{x}_{q_i}\}_{N^k}, N^k \subset N^+, |N^k| = \infty$ 是算法 B 生成的序列 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 存在聚点 $\tilde{x}$ 的一个收敛于 $\tilde{x}$ 的子列, 则 $\{\varphi_{q_i}(\tilde{x}_{q_i})\}_{i \in N^k}$ 收敛, 且 $\lim_{i \rightarrow \infty, i \in N^k} \varphi_{q_i}(\tilde{x}_{q_i}) = \Phi(\tilde{x})$.

证 先证 $0 \leq \Phi(x) - \varphi_{q_i}(x) \leq l_{g_w} \text{dist}(\Omega_E^{q_i}, \Omega), \forall x \in l(x^0, \Omega_E^{q_0}), i \in N^+$.

若 $\Phi(x) = 0$, 则以上的关系式显然成立. 若 $\Phi(x) > 0$, 取 $\omega \in X$, 则存在 $\omega_{q_i} \in \Omega_E^{q_i}$, 满足 $\|\omega - \omega_{q_i}\| \leq l_{g_o} \text{dist}(\Omega_E^{q_i}, \Omega)$, 因而有

$$0 \leq \Phi(x) - \varphi_{q_i}(x) \leq g(x, \omega) - g(x, \omega_{q_i}) \leq l_{g_w} \|\omega - \omega_{q_i}\| \leq l_{g_w} \text{dist}(\Omega_E^{q_i}, \Omega). \quad (3.3)$$

再证 $\{\varphi_{q_i}(\tilde{x}_{q_i})\}_{i \in N^k}$ 收敛. 由式 (3.3) 可知 $\varphi_{q_i}(\tilde{x}_{q_i}) \leq \Phi(\tilde{x}_{q_i})$ 且有

$$\varphi(\tilde{x}_{q_i}) = \varphi(\tilde{x}_{q_i}) - \Phi(\tilde{x}_{q_i}) + \Phi(\tilde{x}_{q_i}) \geq -l_{g_w} \text{dist}(\Omega_{q_i}, \Omega) + \Phi(\tilde{x}_{q_i}) \geq -l_{g_w} \pi_i + \Phi(\tilde{x}_{q_i}).$$

由算法 B 可知 $\lim_{\pi_i \rightarrow 0} \pi_i = 0$. 根据式 (3.3) 可知 $\lim_{i \rightarrow \infty, i \in N^k} \varphi_{q_i}(\tilde{x}_{q_i}) = \Phi(\tilde{x}_{q_i})$.

引理 3.3 令 $\{\tilde{x}_{q_i}\}_{i \in N^k}, N^k \subseteq N^+, |N^k| = \infty$ 是算法 B 生成的序列 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 的一个收敛于 $\tilde{x}$ 的子列, 则对于 $\tilde{\omega} \in \Omega(\tilde{x})$, 存在点列 $\{\omega_{q_i}\}_{i \in N^k}$ 使得 $q_i \in \Omega_E^{q_i}(\tilde{x}_{q_i})$ 对于充分大的满足 $\omega_{q_i} \rightarrow \tilde{\omega}$.

证明可参考文献 [7] 的引理 6.1.4.

定理 3.1 若假设 3.1, 3.2 成立, 如果 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 是由算法 B 生成的迭代点列, 则存在 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 的一个聚点 $\tilde{x}$ 是半无限 minimax 问题 (1.1) 的 KKT 点, 即算法 B 是收敛的.

证 类似于文献 [8] 中定理 2.1 的证明过程, 由引理 3.1 可知序列 $\{\tilde{x}_{q_i}\}_{i \in N^+}$ 存在聚点 $\tilde{x}$, 取其收敛于 $\tilde{x}$ 的子列, $\{\tilde{x}_{q_i}\}, N^k \subseteq N^+, |N^k| = \infty$. 由算法 B 可知 $\tilde{x}_{q_i}$ 是问题 (1.3) 的 KKT 点. 因此, 由 Caratheodory 定理, 对于 $m = n + 1, i \in N^k$ 和 $\omega_{q_i} \in \Omega_E^{q_i}(\tilde{x}_{q_i})$, 有

$$\nabla f(\tilde{x}_{q_i}, y) + \sum_{j=1}^m \xi_i^j \nabla_x g(\tilde{x}_{q_i}, \omega_{q_i}^j) = 0, \quad (3.4)$$

$$\xi_i^j \cdot g(\tilde{x}_{q_i}, \omega_{q_i}^j) = 0, j = 1, 2, \dots, m, \quad (3.5)$$

$$\xi_i^j \geq 0, \quad (3.6)$$

$$\varphi_{q_i}(\tilde{x}_{q_i}) = 0. \quad (3.7)$$

由文献 [7] 中引理 2.3.2, 3.3.3 可知, 乘子序列 $\{\xi_i^j\}_{i \in N^k}, j = 1, 2, \dots, m$ 有界, 故存在子列收敛于 $\tilde{\xi}^j, j = 1, 2, \dots, m$, 不妨记为原数列. 再记 $J(\tilde{x})$ 为 $\Omega_E^{q_i}(\tilde{x})$ 的指标集, 由引理 3.3 知, 对于 $\tilde{\omega}^j \in \Omega(\tilde{x}), j \in J(\tilde{x})$, 存在点列 $\{\omega_{q_i}^j\}_{i \in N^k}$ 使得 $\omega_{q_i}^j \in \Omega_E^{q_i}(\tilde{x}_{q_i})$. 对充分大的 $i \in N^k$ 满足, 即若 $i \in N^k$ 充分大, 则通过整理可使得 $K(= 1, 2, \dots, m)$ 中前 $l(= |J(\tilde{x})|)$ 个对应的 $\Omega_E^{q_i}(\tilde{x}_{q_i})$ 中的指标 $\omega_{q_i}^j \rightarrow \tilde{\omega}^j, j = l + 1, \dots, m$. 此外由 $|K| < \infty$ 和 Ω 是紧集可知, 序列 $\{\omega_{q_i}^j\}_{i \in N^k}, j = l + 1, \dots, m$ 收敛于 $\tilde{\omega}^j \in \Omega, j = l + 1, \dots, m$. 在式 (3.4)–(3.7) 中令 $i \rightarrow \infty, i \in N^k$, 再结合引理 3.2 可得

$$c \nabla f(\tilde{x}, y) + \sum_{j=1}^m \tilde{\xi}^j \nabla_x g(\tilde{x}, \tilde{\omega}^j) = 0, \quad (3.8)$$

$$\tilde{\xi}^j \cdot g(\tilde{x}, \tilde{\omega}^j) = 0, j = 1, 2, \dots, m, \quad (3.9)$$

$$\tilde{\xi}^j \geq 0, \quad (3.10)$$

$$\Phi(\tilde{x}) = 0. \quad (3.11)$$

由式 (3.11) 可知 $\tilde{x} \in X$. 结合文 [7] 中定义 6.1.2, 并由式 (3.8)–(3.11) 可知, $\tilde{x}$ 是半无限 minimax 问题 (1.1) 的 KKT 点.

4 数值实验

本节将对算法 B 在 MATLAB2016a 上编程进行数值实验. 其中 P1, P2, P3 三个算例来源于文献 [4], 三个问题如下

$$\min \max\{f(x), 0\},$$

$$\text{s.t. } g(x, \omega) \leq 0, \omega \in \Omega.$$

$$\text{P1: } f(x) = 1.21 \exp(x_1) + \exp(x_2) + x_3,$$

$$\text{s.t. } g(x, \omega) = \omega - \exp(x_1 + x_2 + x_3) \leq 0, \forall \omega \in [0, 1], \text{ 初始点 } (-1, -1.5, 2);$$

$$\text{P2: } f(x) = x_1^2 + x_2^2,$$

$$\text{s.t. } g(x, \omega) = x_1 + x_2 \exp(\omega) + \exp(2\omega) - 2 \sin(4\omega) \leq 0, \forall \omega \in [0, 1], \text{ 初始点 } (-1, -2);$$

$$\text{P3: } f(x) = 1.21 \exp(x_1) + \exp(x_2) + x_3,$$

$$\text{s.t. } g(x, \omega) = \omega - \exp(x_1 + x_2 + x_3) \leq 0, \forall \omega \in [0, 1], \text{ 初始点 } (1, 1, 1).$$

将算法与文献 [7] 中的算法 C (文献 [7] 第二章 2.6 节的算法) 进行对比. 参数选取 $q = 100, r_0 = 5, r = 1, r_\omega = 0.01, \theta = 0.1, \gamma = 1, \delta = 0.38, L_k = 1, t = 0.87, L_k = 1$,

$\bar{\sigma} = 0.01, \sigma_1 = 100, \zeta = 0.5, H_0 = I$. 算法的终止准则为 $|z_k| \leq 10^{-4}$, 其中 Ni 为迭代次数, x^0 为初始点, x^* 为算法终止时近似的最优解, $F(x^*)$ 为相应的最优值. P 为所有迭代求解问题使用的约束个数, 数值结果见表 1.

表 1 数值结果

算例	x^0	算法	P	Ni	x^*	$F(x^*)$
P1	(-1,-1.5,2)	B	50	7	(-0.205594,-1.356656,1.859369)	5.33452113
	(-1,-1.5,2)	C	53	13	(-0.203265,-1.361444,1.853371)	5.33468779
P2	(-1,-2)	B	95	8	(-0.807574,-0.618937)	0.19474636
	(-1,-2)	C	97	17	(-0.750250,-0.618031)	0.19445626
P3	(1,1,1)	B	52	10	(-0.203781,-1.339076,1.86535)	5.37027732
	(1,1,1)	C	54	15	(-0.207218,-1.374695,1.83673)	5.33485403

数值实验表明, 算法 B 的迭代次数较算法 C 有明显减少, 近似最优值略有差距, 求解问题所使用的约束个数也有所减少. 在精度要求不太高的情况下, 算法 B 采用的积极约束集识别技术和非单调技术可以降低求解计算量和迭代次数, 因而本文所提出的算法是有效的.

参 考 文 献

- [1] 周广路, 王长钰, 孙清滢. 半无限极大极小问题的全局收敛方法 [J]. 运筹学报, 1998, 2(2): 42–51.
- [2] Zhang L P, Fang S C, Wu S Y. An entropy based central cutting plane algorithm for convex min-max semi-infinite programming problems[J]. Sci. China. Math., 2013, 56(1): 198–208.
- [3] Polak E, Womersley R S, Yin H X. An Algorithm Based on Active Sets and Smoothing for Discretized Semi-Infinite minimax Problems[J]. J. Optim. Theory Appl., 2008, 138(2): 311–328.
- [4] Ling C, Qi L, Zhou G, Wu S Y. Global convergence of a robust smoothing SQP method for semi-infinite programming[J]. J. Optim. Theory. Appl., 2006, 129(1): 147–164.
- [5] Jian J B, Xu Q J, Han D L. A norm relaxed method of feasible directions for finely discretized problems form semi-infinite programming[J]. Eur. J. Oper. Res., 2008, 186(1): 41–62.
- [6] 徐庆娟, 简金宝. 半无限规划离散化问题一个两阶段序列二次规划算法 [J]. 数学杂志, 2014, 34(6): 1155–1162.
- [7] 徐庆娟. 半无限规划的有效数值算法研究 [D]. 上海: 上海大学, 2014.
- [8] 徐庆娟, 简金宝. 半无限规划基于离散化方法和局部约化的两个算法框架 (英文)[J]. 数学杂志, 2018, 38(5): 851–860.
- [9] Jian J B, Hu Q J, Tang C M. Superlinearly convergent norm Relaxed SQP method based on active set identification and new line search for constrained minimax problems[J]. J. Optim. Theory Appl., 2014, 163(3): 859–883.
- [10] Wang F S. A hybrid algorithm for linearly constrained minimax problems[J]. Ann. Oper. Res., 2013, 206(1): 501–525 .
- [11] He S X, Liu X F, Wang C M. A nonlinear lagrange algorithm for minimax problem with general constraints[J]. Numer. Funct. Anal. Optim., 2016, 37(6): 680–698.
- [12] Marco L, Georg S. Semi-infinite programming[J]. Eur. J. Operat. Res., 2006, 180(2): 491–518.

- [13] Zhang H C, Hager W W. A nonmonotone line search techniques and its application to unconstrained optimization[J]. SIAM J. Optim., 2004, 14(4): 1043–1056.
- [14] Gu N Z, Mo J T. Incorporating nonmonotone strategies into the trust region method for unconstrained optimization [J]. Comput. Math. Appl., 2008, 55(9): 2158–2172 .
- [15] Shi Z J, Shen J. New inexact line search method for unconstrained optimization[J]. J. Optim. Theory Appl., 2005, 127(2): 425–446.
- [16] 李廷锋. 求解大规模无约束优化问题的修正 L-BFGS 方法 [D]. 开封: 河南大学, 2008.

A NON-MONOTONE SQP ALGORITHM WITH L-BFGS UPDATE FOR SOLVING SEMI-INFINITE MINIMAX PROBLEM BASED ON ACTIVE SET TECHNIQUE

YANG Yong-liang, WANG Fu-sheng, ZHEN Na

(Department of Mathematics, Taiyuan Normal University, Jinzhong 030619, China)

Abstract: This paper studies the semi-infinite minimax problem. By using active set recognition technology combined with non-monotonic finite memory sequential quadratic programming (SQP) method to solve the semi-infinite minimax problem. The convergence of the algorithm is proved under appropriate conditions. Numerical results show that the new algorithm is superior to the SQP method using Armijo-type line search in terms of reducing the solution scale and the number of iterations.

Keywords: minimax problem; constraint active set; discretization method; SQP algorithm; non-monotone technique

2010 MR Subject Classification: 90C30; 65K05