第 34 卷 第 5 期　　应 用 数 学 学 报　　Vol. 34 No. 5
2011 年 9 月　　ACTA MATHEMATICAE APPLICATAE SINICA　　Sep., 2011

基于次梯度选取的非光滑优化强次可行方向法 *

唐春明　　简金宝 †

(广西大学数学与信息科学学院，南宁 530004)

(E-mail: cmtang@gxu.edu.cn; jianjb@gxu.edu.cn)

摘　要　本文结合次梯度选取技术及割平面法和强次可行方向法的思想，提出了一个求解目标函数非光滑约束优化问题的强次可行方向算法. 通过设计一个新的寻找搜索方向子问题和构造新型线搜索，算法不仅能接受不可行的初始点，而且能保持迭代点的强次可行性，同时避免在可行域外目标函数值的不适度增加. 算法具备全局收敛性，且初步的数值试验表明算法是稳定有效的.

关键词　次梯度选取；非光滑优化；强次可行方向法；全局收敛

MR(2000) 主题分类　90C30; 65K05

中图分类　O221.2

1　引言

考虑求解如下非线性不等式约束优化问题

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \quad f(x) \\ &\text{s.t.} \quad c_i(x)\le 0, \quad i\in I \triangleq \{1,\cdots,m\}, \end{aligned} \tag{1.1}$$

其中 $f:\mathbb{R}^n\to\mathbb{R}$ 为凸函数，但不一定可微 (即可能非光滑). 为简便考虑，我们假设 $c_i\ (i\in I):\ \mathbb{R}^n\to\mathbb{R}$ 是连续可微的凸函数，因为非光滑的约束可以通过适当的技术惩罚进入目标函数，而约束函数的凸性也并非本文算法的本质要求 (因为算法可以经过适当修改后处理非凸约束), 因此问题 (1.1) 具有一定的代表性. 记问题 (1.1) 的可行集为：$\mathcal{F}=\{x\in\mathbb{R}^n:\ c_i(x)\le 0,\ i\in I\}$.

本文 2010 年 10 月 4 日收到.

* 国家自然科学基金 (71061002), 广西自然科学基金 (0832052, 2010GXNSFB013047), 广西教育厅科研基金 (200911MS202) 和广西大学科研基金 (XBZ100086) 资助项目.

† 通讯作者.

对于光滑形式的问题 (1.1) (即函数 f 和 c_i 均至少一阶连续可微), 目前已经有了许多有效的求解方法，如罚函数法，信赖域法，可行方向法和内点法等．其中，可行方向法 [1−5] 能够产生可行迭代点和 (近似) 最优解，因此在工程设计、实时控制等领域得到了广泛应用 [6]. 然而可行方向法的一个缺陷是需要一个可行初始点，而计算这样一个点往往不易，有时计算量与求解原问题相当．为此， [7] 和 [8] 提出了一类强次可行方向法．该类方法能从不可行的初始点开始，并且能保持迭代点的强次可行性，即对于当前迭代点满足可行性的约束函数在下一个迭代点仍然满足，且能同时避免在可行域外目标函数值增加过快．此外，一旦某个迭代点落入可行域，即自动变为可行方向法．经过十余年的研究和完善, 强次可行方向法已经成功地与多类快速算法有机结合, 如序列二次规划 [9] 和序列二次约束二次规划 [10], 并有效地应用于求解斜拉桥索力优化问题 [11].

然而，现有的强次可行方向法只能求解光滑类型的优化问题．为了弥补这一缺陷，本文提出了一个求解非光滑优化问题 (1.1) 的强次可行方向法．首先，利用次梯度选取技术 [12] 和割平面法的思想 [13], 对非光滑的目标函数进行线性化近似．然后结合强次可行方向法思想, 构造了一个新的寻找搜索方向子问题, 并基于子问题的特点和收敛性分析的需要，提出了一个新型线搜索策略．算法不仅具备强次可行方向法的上述优点，而且是全局收敛的．初步的数值试验表明新算法是稳定有效的．符号 $\|\cdot\|$ 表示 Euclidean 向量范数，集合 $\partial f(x)$ 表示 f 在 x 处的次微分， $g\in\partial f(x)$ 称为一个次梯度， $\nabla c_i(x)$ 表示 c_i 在 x 处的梯度.

2 算法描述

首先作如下基本假设，并设其在全文中成立.

假设 2.1 (a) f 是凸函数但不一定可微， c_i $(i\in I)$ 是连续可微的凸函数； (b) 问题 (1.1) 满足 Slater 约束规格，即存在一个向量 $\widetilde{x}\in\mathbb{R}^n$ 使得 $c_i(\widetilde{x})<0,\ \forall i\in I$.

此外，跟以往的非光滑优化算法一样，我们假设有一个子程序能够计算 f 在任意一个点 $x\in\mathbb{R}^n$ 处的其中 (任意) 一个次梯度 $g(x)\in\partial f(x)$.

为导出本文算法，先考虑如下无约束凸极小化问题

$$\min\{f(x):\ x\in\mathbb{R}^n\}.$$

由于函数 f 的不可微性, 传统的求解光滑优化问题的梯度型算法不能用于求解该问题, 为此可以考虑对 $f(x)$ 进行适当地近似．考虑到凸函数可以精确表示为

$$f(x)=\sup\left\{f(y)+g_y^T(x-y):\ g_y\in\partial f(y),\ y\in\mathbb{R}^n\right\}.$$

但整个次微分 $\partial f(y)$ 一般难于计算，因此，对于当前迭代点 $x^k\in\mathbb{R}^n$, 割平面法 [14] 采用如下函数对 f 进行线性化近似

$$\widehat{f}^k(x)=\max\{f_j(x):\ j=1,\cdots,k\},$$

其中 $f_j(x) = f(y^j) + (g^j)^T(x - y^j)$, y^j, $j = 1, \cdots, k$ 为辅助迭代点, $g^j \in \partial f(y^j)$. 简单变形可得

$$f_j(x) = f_j^k + (g^j)^T(x - x^k), \qquad j = 1, \cdots, k,$$

其中 $f_j^k = f_j(x^k)$. 因此极小化 $\widehat{f}^k(x)$ 导致如下寻找搜索方向子问题

$$\begin{aligned} \min_{d \in \mathbb{R}^n,\, z \in \mathbb{R}} \quad & z + \frac{1}{2}\|d\|^2 \\ \text{s.t.} \quad & f_j^k + (g^j)^T d \le z, \ j = 1, \cdots, k, \end{aligned}$$

其中 $z \in \mathbb{R}$ 是一个辅助变量. 显然, 当 k 不断增大的时候, 这样的子问题会带来存储和计算上的困难. 为此, Kiwiel[15] 提出了一个基于次梯度选取技术的寻找搜索方向子问题, 即只需要部分次梯度产生线性化约束

$$\begin{aligned} \min_{d \in \mathbb{R}^n,\, z \in \mathbb{R}} \quad & z + \frac{1}{2}\|d\|^2 \\ \text{s.t.} \quad & f_j^k + (g^j)^T d \le z, \quad j \in J^k, \end{aligned} \tag{2.1}$$

其中 $J^k \subseteq \{1, \cdots, k\}$ 是通过适当方式产生的一个指标集 (参见下面的 (2.5) 及 (2.6) 式), 其元素个数有上界 $n+2$ (类似 (2.7) 式). 因此, 子问题 (2.1) 的计算量和存储量极大减少.

以上描述了如何利用割平面法思想和次梯度选取技术处理非光滑的目标函数, 接下来我们加入约束函数同时考虑, 并结合强次可行方向法思想提出了一个新的寻找搜索方向子问题. 为此, 先定义符号

$$I_-(x) = \{i \in I:\ c_i(x) \le 0\}, \quad I_+(x) = \{i \in I:\ c_i(x) > 0\}, \quad \varphi(x) = \max\{0,\ c_i(x),\ i \in I\}$$

及某种意义下的惩罚函数: $\delta(x):\ \mathbb{R}^n \to [0, +\infty)$ 是连续函数, 且 $\delta(x) = 0$ 当且仅当 $x \in \mathcal{F}$. 事实上, $\delta(x)$ 的一种最直接的取法就是 $\delta(x) = \varphi(x)$, 而这里给出一种更广泛的形式, 有利于改善数值效果. 对当前迭代点 $x^k \in \mathbb{R}^n$, 简记

$$I_-^k = I_-(x^k), \quad I_+^k = I_+(x^k), \quad \varphi^k = \varphi(x^k), \quad \delta^k = \delta(x^k).$$

考虑如下寻找搜索方向子问题

$$\begin{aligned} \min_{d \in \mathbb{R}^n,\, z \in \mathbb{R}} \quad & z + \frac{1}{2}\|d\|^2 \\ \text{s.t.} \quad & f_j^k - f(x^k) + (g^j)^T d \le z + \delta^k, \quad j \in J^k, \\ & c_i(x^k) + \nabla c_i(x^k)^T d \le z, \quad i \in I_-^k, \\ & c_i(x^k) + \nabla c_i(x^k)^T d \le z + \varphi^k, \quad i \in I_+^k, \end{aligned} \tag{2.2}$$

其中 $z \in \mathbb{R}$ 是一个辅助变量. 为初始化算法, 当 $k = 1$ 时取 $x^1 \in \mathbb{R}^n$, $y^1 = x^1$ 及

$$g^1 = g(x^1), \qquad f_1^1 = f(x^1), \qquad J^1 = \{1\}. \tag{2.3}$$

注 1 将约束函数的线性化分为两组 (I_-^k 和 I_+^k) 源自于强次可行方向法思想， φ^k 的引入用于保证子问题约束的相容性， δ^k 项的引入可防止在可行域外目标函数值不适当地增加. 此外，由 f 的凸性可知 $f_j^k - f(x^k) \le 0,\ j \in J^k$, 从而 $(d,z)=(0,0)$ 是子问题 (2.2) 的一个可行解.

设 (d^k, z^k) 是 (2.2) 的最优解. 显然 (2.2) 是一个带线性约束的凸规划，因此 (d^k, z^k) 也是 (2.2) 的一个 KKT 点，即存在乘子 $\lambda_j^k,\ j \in J^k,\ \mu_i^k,\ i \in I$ 使得

$$\begin{aligned}
& d^k + \sum_{j\in J^k} \lambda_j^k g^j + \sum_{i\in I} \mu_i^k \nabla c_i(x^k) = 0, \qquad \sum_{j\in J^k} \lambda_j^k + \sum_{i\in I} \mu_i^k = 1,\\
& 0 \le \lambda_j^k \perp \big(f_j^k - f(x^k) + (g^j)^T d^k - z^k - \delta^k\big) \le 0, \qquad j \in J^k,\\
& 0 \le \mu_i^k \perp \big(c_i(x^k) + \nabla c_i(x^k)^T d - z^k\big) \le 0, \qquad i \in I_-^k,\\
& 0 \le \mu_i^k \perp \big(c_i(x^k) + \nabla c_i(x^k)^T d - z^k - \varphi^k\big) \le 0, \qquad i \in I_+^k.
\end{aligned} \tag{2.4}$$

定义指标集

$$\widehat{J}^k = \{j \in J^k:\ \lambda_j^k > 0\}, \qquad \widehat{I}^k = \{i \in I:\ \mu_i^k > 0\}, \tag{2.5}$$

并更新指标集 J^k 如下

$$J^{k+1} = \widehat{J}^k \cup \{k+1\}. \tag{2.6}$$

推广 [12] 中引理 2.2.1 (该引理中不考虑约束) 的分析可知, 存在满足 (2.4) 的乘子 $\lambda_j^k,\ j \in J^k,\ \mu_i^k,\ i \in I$ 使得

$$|\widehat{J}^k| + |\widehat{I}^k| \le n+1. \tag{2.7}$$

[12] 还指出, 满足 (2.7) 的乘子可以通过求解一个线性规划获得, 并且绝大多数二次规划软件求解 (2.2) 所得到的乘子都自动满足 (2.7), 进而由 (2.6) 产生的指标集满足 $|J^{k+1}| \le n+2$. 这说明每一次迭代中，子问题 (2.2) 所含约束个数的上界是 $m+n+2$, 从而该子问题可以快速有效地求解. 这也正是本文要利用次梯度选取技术的目的所在.

在进一步分析子问题 (2.2) 解的性质之前，我们构造如下最优性衡量函数

$$H(y;x) = \max\big\{f(y) - f(x) - \delta(x);\ c_i(y),\ i \in I_-(x);\ c_i(y) - \varphi(x),\ i \in I_+(x)\big\},$$

其中 x 是固定的，而 $y \in \mathbb{R}^n$ 是自变量. 此函数是 [12] 中函数 $H(y;x) = \max\{f(y) - f(x); C(y)\}$ 的变形推广，其中 $C(x) = \max\{c_i(x),\ i \in I\}$. 以下引理给出了函数 H 的性质.

引理 2.1 设假设 2.1 成立，则以下三个命题等价

(a) $\overline{x}$ 是问题 (1.1) 的最优解;

(b) $\min\{H(y;\overline{x}):\ y \in \mathbb{R}^n\} = H(\overline{x};\overline{x}) = 0$;

(c) $0 \in \partial H(\overline{x};\overline{x})$.

证 (a)⇒(b). 设 $\overline{x}$ 是问题 (1.1) 的最优解，则有 $\delta(\overline{x}) = 0,\ I_-(\overline{x}) = I$ 及 $I_+(\overline{x}) = \emptyset$, 因此 $H(y;\overline{x}) = \max\big\{f(y) - f(\overline{x});\ c_i(y),\ i \in I\big\} \ge 0$ 且 $H(\overline{x};\overline{x}) = 0$, 故 (b) 成立.

(b)⇒(a). 设 (b) 成立，先证 $\overline{x}\in\mathcal{F}$. 反证法假设 $\overline{x}\notin\mathcal{F}$, 则有 $\delta(\overline{x})>0$. 由假设 2.1(b) 可知, 存在一个向量 $\widetilde{x}\in\mathbb{R}^n$ 满足 $c_i(\widetilde{x})<0,\ i\in I$. 定义 $y(\tau)=\overline{x}+\tau(\widetilde{x}-\overline{x})=(1-\tau)\overline{x}+\tau\widetilde{x}$, $\tau\in(0,1)$, 则由 f 的连续性可得

$$f(y(\tau))-f(\overline{x})-\delta(\overline{x})<0,$$

对充分小的 $\tau>0$ 成立．此外，由 c_i 的凸性，对 $\tau\in(0,1)$ 我们有

$$c_i(y(\tau))=c_i((1-\tau)\overline{x}+\tau\widetilde{x})\ \le(1-\tau)c_i(\overline{x})+\tau c_i(\widetilde{x})\ \le\tau c_i(\widetilde{x})<0,\qquad i\in I_-(\overline{x})$$

及

$$c_i(y(\tau))-\varphi(\overline{x})\le(1-\tau)c_i(\overline{x})+\tau c_i(\widetilde{x})-\varphi(\overline{x})\ \le\tau c_i(\widetilde{x})<0,\qquad i\in I_+(\overline{x}).$$

以上关系说明存在一个常数 $\overline{\tau}\in(0,1)$ 使得

$$H(y(\tau);\overline{x})<0,\qquad\forall\,\tau\in(0,\overline{\tau}),\tag{2.8}$$

这与 (b) 矛盾, 从而 $\overline{x}\in\mathcal{F}$. 进而有 $H(y;\overline{x})=\max\big\{f(y)-f(\overline{x});\ c_i(y),\ i\in I\big\}$. 如果 $\overline{x}$ 不是问题 (1.1) 的最优解，则存在一个向量 $\widehat{x}\in\mathcal{F}$ 使得 $f(\widehat{x})<f(\overline{x})$. 令 $y(\tau)=\widehat{x}+\tau(\widetilde{x}-\widehat{x})=(1-\tau)\widehat{x}+\tau\widetilde{x},\ \tau\in(0,1)$. 由 f 的凸性可知

$$f(y(\tau))-f(\overline{x})\le(1-\tau)f(\widehat{x})+\tau f(\widetilde{x})-f(\overline{x})=\tau(f(\widetilde{x})-f(\widehat{x}))+(f(\widehat{x})-f(\overline{x}))<0$$

对充分小的 $\tau\in(0,1)$ 成立．此外,

$$c_i(y(\tau))\le(1-\tau)c_i(\widehat{x})+\tau c_i(\widetilde{x})\le\tau c_i(\widetilde{x})<0,\qquad i\in I$$

对 $\tau\in(0,1)$ 成立．于是我们得到与 (2.8) 相同的矛盾 ($\overline{\tau}$ 不同), 故 (a) 成立.

最后，由于函数 $H(\cdot;\overline{x})$ 是凸函数， (b) 与 (c) 的等价性是显然的.

类似 [12], 计算聚集的次梯度及函数值如下：

$$(p^k,\widetilde{f}_p^k)=\frac{1}{\theta^k}\sum_{j\in J^k}\lambda_j^k(g^j,f_j^k),\tag{2.9}$$

其中 $\theta^k=\sum\limits_{j\in J^k}\lambda_j^k$. 显然此式隐含 $\theta^k\neq0$, 若 $\theta^k=0$, 则取 $p^k=0,\ \ \widetilde{f}_p^k=0$, 以下分析平凡地成立，故不再提及．以下引理给出了子问题 (2.2) 解的重要性质.

引理 2.2 设 (d^k,z^k) 是子问题 (2.2) 的最优解，则

(a) $z^k=-(\|d^k\|^2+\widetilde{\alpha}^k)$, 其中

$$\widetilde{\alpha}^k=\theta^k\big(f(x^k)-\widetilde{f}_p^k+\delta^k\big)-\sum_{i\in I_-^k}\mu_i^kc_i(x^k)-\sum_{i\in I_+^k}\mu_i^k\big(c_i(x^k)-\varphi^k\big);$$

(b) $-d^k$ 是函数 $H(\cdot;x^k)$ 在 x^k 处的 ε 次梯度，即 $-d^k\in\partial_\varepsilon H(x^k;x^k)$, 其中 $\varepsilon=\widetilde{\alpha}^k$;

(c) 如果 $z^k=0$, 则 $d^k=0$ 且 x^k 是问题 (1.1) 的最优解.

证　(a) 由 KKT 条件 (2.4) 得

$$z^k=-\|d^k\|^2+\sum_{j\in J^k}\lambda_j^k\big(f_j^k-f(x^k)-\delta^k\big)+\sum_{i\in I_-^k}\mu_i^kc_i(x^k)+\sum_{i\in I_+^k}\mu_i^k\big(c_i(x^k)-\varphi^k\big),$$

故结合 (2.9) 式可知 (a) 成立.

(b) 由 f 的凸性可知

$$f(x)\ge f(x^k)+(g^j)^T(x-x^k)-\big(f(x^k)-f_j^k\big),\qquad j\in J^k.$$

上式两边对 $j\in J^k$ 求和得

$$f(x)\ge f(x^k)+(p^k)^T(x-x^k)-\big(f(x^k)-\widetilde{f}_p^k\big).$$

类似地，由 c_i 的凸性可得

$$\sum_{i\in I}\mu_i^kc_i(x)\ge\sum_{i\in I}\mu_i^k\big(c_i(x^k)+\nabla c_i(x^k)^T(x-x^k)\big).$$

因此

$$\begin{aligned}H(x;x^k)=&\max\big\{f(x)-f(x^k)-\delta^k;\ c_i(x),\ i\in I_-^k;\ c_i(x)-\varphi^k,\ i\in I_+^k\big\}\\ \ge&\theta^k\big(f(x)-f(x^k)-\delta^k\big)+\sum_{i\in I_-^k}\mu_i^kc_i(x)+\sum_{i\in I_+^k}\mu_i^k\big(c_i(x)-\varphi^k\big)\\ \ge&\theta^k\big((p^k)^T(x-x^k)-\big(f(x^k)-\widetilde{f}_p^k\big)-\delta^k\big)+\sum_{i\in I}\mu_i^kc_i(x^k)\\ &+\sum_{i\in I}\mu_i^k\nabla c_i(x^k)^T(x-x^k)-\sum_{i\in I_+^k}\mu_i^k\varphi^k\\ =&H(x^k;x^k)+(-d^k)^T(x-x^k)-\widetilde{\alpha}^k,\end{aligned}$$

其中最后一个等式用到 $H(x^k;x^k)=0$, 故 (c) 成立.

(c) 由 f 的凸性有 $f(x^k)-\widetilde{f}_p^k\ge0$, 进而可得 $\widetilde{\alpha}^k\ge0$, 故由 $z^k=0$ 可知 $d^k=0$ 且 $\widetilde{\alpha}^k=0$. 再结合结论 (b) 易知 $0\in\partial H(x^k,x^k)$, 从而根据引理 2.1 可得 x^k 是问题 (1.1) 的最优解.

基于以上分析，下面给出具体算法.

算法 2.1

步骤 0 (初始化) 选取初始点 $x^1\in\mathbb{R}^n$ 及参数 $\beta,\eta,\bar{t}\in(0,1)$. 令 $y^1=x^1$, 再根据 (2.3) 初始化算法. 置 $k=1,\ l=0$ 及 $k(0)=1$.

步骤 1 (计算方向) 求解子问题 (2.2) 得到最优解 (d^k,z^k) 及乘子 $\lambda_j^k,\ j\in J^k,\ \mu_i^k,\ i\in I$. 根据 (2.9) 计算 p^k 及 $\widetilde{f}_p^k$.

步骤 2 (终止准则) 令 $w^k=\frac{1}{2}\|d^k\|^2+\widetilde{\alpha}^k$. 如果 $w^k=0$, 终止；否则转步骤 3.

步骤 3 (线搜索)

3.1. 计算试探步长 t^k, 它是序列 $\{1,\beta,\beta^2,\cdots\}$ 中第一个满足以下不等式的 t 值

$$c_i(x^k+td^k)\le\varphi^k+\eta tz^k,\qquad i\in I_+^k,\tag{2.10}$$
$$c_i(x^k+td^k)\le 0,\qquad i\in I_-^k.\tag{2.11}$$

如果 $t^k\le\bar{t}$, 转步骤 3.2; 否则转步骤 3.3.

3.2. 如果

$$f(x^k+t^kd^k)\le f(x^k)+\eta t^kz^k+t^k\delta^k\tag{2.12}$$

成立，则令 $t_L^k=t^k$ (有效步) 及辅助步长 $t_R^k=t^k$, 置 $k(l+1)=k+1,\ \ l:=l+1$; 否则令 $t_L^k=0$ (无效步) 及 $t_R^k=t^k$.

3.3. 计算步长 γ^k, 它是序列 $\{1,\beta,\beta^2,\cdots\}\cap[\bar{t},t^k]$ 中第一个满足下式的 γ 值

$$f(x^k+\gamma d^k)\le f(x^k)+\eta\gamma z^k+\gamma\delta^k.\tag{2.13}$$

如果 γ^k 存在，则令 $t_L^k=\gamma^k$ (有效步) 及 $t_R^k=\gamma^k$, 置 $k(l+1)=k+1,\ \ l:=l+1$; 否则令 $t_L^k=0$ (无效步) 并找一个辅助步长 $t_R^k\in[\bar{t},t^k]$.

步骤 4 (更新) 令 $x^{k+1}=x^k+t_L^kd^k,\ \ y^{k+1}=x^k+t_R^kd^k$. 由 (2.6) 式产生 $\widehat{J}^k$, 并通过下式计算新的线性化函数值

$$f_j^{k+1}=f_j(x^{k+1})=f_j^k+(g^j)^T(x^{k+1}-x^k),\qquad j\in\widehat{J}^k.$$

计算新的次梯度 $g^{k+1}=g(y^{k+1})$, 及 $f_{k+1}^{k+1}=f(y^{k+1})+(g^{k+1})^T(x^{k+1}-y^{k+1})$. 再由 (2.6) 式产生 J^{k+1}, 令 $k:=k+1$, 返回步骤 1.

注 2 (a) 步骤 3 中给出了一个新的线搜索策略，该策略利用满足约束函数强次可行性的步长 t^k 作为目标函数搜索的初始试探步长，而以往的算法通常从 1 开始试探. 此做法保证了辅助步长 t_R^k 对 (2.10) 及 (2.11) 式总是满足的，故一旦算法产生无效步，则一定有目标函数线搜索不等式在 t_R^k 处违背，这在分析算法的全局收敛性时起到了重要的作用 (见引理 3.6). 此外，若 t^k 较小，则直接用其作为目标函数搜索的步长试探一次，不再减小 (步骤 3.2); 若 t^k 还较大，则可以进一步适当减小 (步骤 3.3).

(b) 线搜索 (2.10)–(2.12) 保证了迭代点的强次可行性. δ^k 的引入除了用于处理不可行的迭代点外，还能防止目标函数值在可行域外不适度地增加.

下面的引理给出了算法 2.1 的基本性质.

引理 2.3 若算法 2.1 有限步终止于第 k 个迭代点，即 $w^k=0$, 则 x^k 是问题 (1.1) 的一个最优解.

证 由于 $w^k=0$ 等价于 $z^k=0$, 故结合引理 2.2 立知结论成立.

引理 2.4 算法 2.1 是适定的，即线搜索 (2.10) 和 (2.11) 能在有限次计算后终止.

证 首先易知若算法进入步骤 3, 则 $z^k<0$. 由 Taylor 展开及 (2.4) 可得

$$\begin{aligned}c_i(x^k+td^k)-\varphi^k-\eta tz^k=&c_i(x^k)-\varphi^k-\eta tz^k+t\nabla c_i(x^k)^Td^k+o(t)\\ \le&c_i(x^k)-\varphi^k-\eta tz^k+t\big(-c_i(x^k)+z^k+\varphi^k\big)+o(t)\end{aligned}$$

$$
\begin{aligned}
&=(1-t)\big(c_i(x^k)-\varphi^k\big)+(1-\eta)tz^k+o(t)\\
&\leq(1-\eta)tz^k+o(t),\qquad i\in I_+^k.
\end{aligned}\tag{2.14}
$$

上面关系结合 $\eta\in(0,1)$ 及 $z^k<0$ 可知 (2.10) 对 $t>0$ 充分小成立．类似地，我们有

$$
\begin{aligned}
c_i(x^k+td^k)&=c_i(x^k)+t\nabla c_i(x^k)^Td^k+o(t)\leq c_i(x^k)+t\big(-c_i(x^k)+z^k\big)+o(t)\\
&=(1-t)c_i(x^k)+tz^k+o(t)\leq tz^k+o(t)<0,\qquad i\in I_-^k,
\end{aligned}\tag{2.15}
$$

对 $t>0$ 充分小成立．从而引理得证．

引理 2.5 算法 2.1 必定出现以下两种情形之一．

情形 I 存在一个指标 k_0 使得 $\varphi^{k_0}=0$, 从而 $\varphi^k\equiv 0$, $\delta^k\equiv 0$ 及 $f(x^{k+1})\leq f(x^k)$, 对所有 $k\geq k_0$ 成立；

情形 II $\varphi^k>0$, $\varphi^{k+1}\leq\varphi^k$, $I_-^k\subseteq I_-^{k+1}$, $k=1,2,3,\cdots$.

由于 I_-^k 及 I_+^k 均单调变化，且为 I 的子集，易得以下引理．

引理 2.6 当 k 充分大时， I_-^k 及 I_+^k 能固定下来，即存在固定指标集 I_- 和 I_+ 使得 $I_-^k\equiv I_-$ 及 $I_+^k\equiv I_+$.

3　全局收敛性

本节我们分析算法 2.1 的全局收敛性．一旦算法有限步终止于 x^k 点，则由引理 2.3 可知 x^k 是问题 (1.1) 的一个最优解．现假设算法产生一个无限迭代序列 $\{x^k\}$, 以下将证明其任意一个聚点都是问题 (1.1) 的最优解．

引理 3.1 假设存在一个无限指标集 $K\subseteq\{1,2,\cdots\}$ 及点 $\overline{x}\in\mathbb{R}^n$ 满足 $x^k\xrightarrow{K}\overline{x}$ 及 $w^k\xrightarrow{K}0$, 则 $\overline{x}$ 是问题 (1.1) 的最优解．

证　由 $w^k\xrightarrow{K}0$ 可得 $d^k\to 0$, $\widetilde{\alpha}^k\to 0$, $k\in K$. 故由 $x^k\xrightarrow{K}\overline{x}$ 及引理 2.2 (b) 知 $0\in\partial H(\overline{x};\overline{x})$. 这结合引理 2.1 可得结论成立．

引理 3.2 假设以下条件之一成立： (a) 引理 2.5 的情形 I 出现，及序列 $\{f(x^k)\}$ 有下界； (b) 引理 2.5 的情形 II 出现，及序列 $\{\varphi(x^k)\}$ 有下界，则

$$
\sum_{k=1}^{\infty}\big(t_L^k\|d^k\|^2+t_L^k\widetilde{\alpha}^k\big)<+\infty.\tag{3.1}
$$

证　如果条件 (a) 成立，则由 (2.12) 及 (2.13) 可知

$$
f(x^{k_0})-f(x^k)=f(x^{k_0})-f(x^{k_0+1})+\cdots+f(x^{k-1})-f(x^k)\geq\eta\sum_{i=k_0}^{k-1}t_L^i(-z^i).
$$

此关系结合 $\{f(x^k)\}$ 的有界性说明 (3.1) 成立．类似地，由 (2.10) 及 $\{\varphi(x^k)\}$ 的有界性，亦可得在条件 (b) 下 (3.1) 成立．

由算法 2.1 的步骤 3 可知，序列 $\{k(l)\}$ 用于记录那些由有效步得到的迭代点的指标，故而序列 $\{x^k\}$ 满足

$$x^k = x^{k(l)}, \qquad \forall\, k = k(l), k(l)+1, \cdots, k(l+1)-1. \tag{3.2}$$

若只有有限个有效步，即存在 l_0, 使得 $x^k = x^{k(l_0)}$, $\forall\, k \ge k(l_0)$, 则在上式中视 $k(l_0+1)$ 为无穷大.

下面分成两种情况证明： (a) 算法产生无限多个有效步； (b) 算法产生有限个有效步. 以下引理首先考虑情况 (a), 且根据引理 2.6, 在下面的分析中可假设 $I_-^k \equiv I_-$ 及 $I_+^k \equiv I_+$.

引理 3.3 假设存在一个无限指标集 $L \subseteq \{1, 2, \cdots\}$ 和向量 $\overline{x} \in \mathbb{R}^n$ 使得 $x^{k(l)} \to \overline{x}$, $l \to \infty$, $l \in L$, 则 $\overline{x}$ 是问题 (1.1) 的一个最优解.

证 记 $K = \{k(l+1)-1 : l \in L\}$, 则由 (3.2) 得 $x^k \to \overline{x}$, $k \in K$.

下面将证明 $w^k \to 0$, $k \in K$. 由于 $\{f(x^k)\}_{k \ge k_0}$ (若情形 I 发生) 及 $\{\varphi(x^k)\}$ 为非增序列，且 $f(x)$, $\varphi(x)$ 连续，故有 $f(x^k) \downarrow f(\overline{x})$, $k \ge k_0$ 及 $\varphi(x^k) \downarrow \varphi(\overline{x})$. 此结合引理 3.2 可知

$$\sum_{k=1}^{\infty} \left(t_L^k \|d^k\|^2 + t_L^k \widetilde{\alpha}^k \right) < +\infty. \tag{3.3}$$

反证法假设 $w^k \nrightarrow 0$, $k \in K$, 则存在一个无限指标集 $K' \subseteq K$ 和常数 $\overline{w} > 0$ 使得

$$-z^k \ge w^k \ge \overline{w} > 0, \qquad k \in K'. \tag{3.4}$$

因此，由 (2.14) 及 (2.15), 对于 $k \in K'$, 我们有

$$c_i(x^k + td^k) - \varphi^k - \eta t z^k \le (1-\eta) t z^k + o(t) \le -(1-\eta) t \overline{w} + o(t), \qquad i \in I_+$$

及

$$c_i(x^k + td^k) \le t z^k + o(t) \le -t\overline{w} + o(t), \qquad i \in I_-.$$

以上关系说明存在常数 $\widetilde{t} > 0$ 使得 $t^k \ge \widetilde{t} > 0$, $k \in K'$ 成立. 取 $\widehat{t} = \min\{\widetilde{t}, \overline{t}\}$, 可得 $t_L^k \ge \widehat{t}$, $k \in K'$. 因此由 (3.3) 可知

$$\widehat{t} \sum_{k \in K'} w^k \le \sum_{k \in K'} \widehat{t}(-z^k) \le \sum_{k \in K'} t_L^k(-z^k) \le \sum_{k=1}^{\infty} \left(t_L^k \|d^k\|^2 + t_L^k \widetilde{\alpha}^k \right) < +\infty,$$

这与 (3.4) 矛盾，从而 $w^k \to 0$, $k \in K$. 因此由引理 3.1 可得结论成立.

以下分析有限个有效步的情形，即存在指标 $\widetilde{k}$ 使得 $x^k = x^{\widetilde{k}}$, $k \ge \widetilde{k}$. 我们将证明 $x^{\widetilde{k}}$ 是问题 (2.2) 的最优解.

引理 3.4 w^k 是以下问题的最优值，该问题为子问题 (2.2) 的对偶问题

$$\begin{aligned}\min \quad & \frac{1}{2}\Big\|\sum_{j\in J^k}\lambda_j g^j+\sum_{i\in I}\mu_i\nabla c_i(x^k)\Big\|^2-\sum_{j\in J^k}\lambda_j\big(f_j^k-f(x^k)-\delta^k\big)\\ & -\sum_{i\in I_-}\mu_i c_i(x^k)-\sum_{i\in I_+}\mu_i\big(c_i(x^k)-\varphi^k\big) \qquad (3.5)\\ \text{s.t.} \quad & \lambda_j\ge 0,\ j\in J^k,\ \mu_i\ge 0,\ i\in I,\ \sum_{j\in J^k}\lambda_j+\sum_{i\in I}\mu_i=1.\end{aligned}$$

证 由于 (2.4) 的乘子是问题 (3.5) 的最优解，故由 w^k 的定义可得结论成立.

下面的证明需要用到 [12] 中如下的一个结果 (引理 2.4.10).

引理 3.5[12] 设 n 维向量 d, g 及数 $\eta\in(0,1)$, C, w, z, $\widetilde{\alpha}$ 和 α 满足

$$w=\frac{1}{2}\|d\|^2+\widetilde{\alpha},\quad z=-(\|d\|^2+\widetilde{\alpha}),\quad -\alpha+g^Td\ge\eta z,\quad C\ge\max\{\|d\|,\|g\|,\widetilde{\alpha},1\}.$$

令 $\overline{w}=\min\{Q(v):\ v\in[0,1]\}$, 其中

$$Q(v)=\frac{1}{2}\|(1-v)(-d)+vg\|^2+(1-v)\widetilde{\alpha}+v\alpha.$$

则 $\overline{w}\le\phi_C(w)$, 其中 $\phi_C(t)=t-(1-\eta)^2t^2/(8C^2)$.

引理 3.6 假设对某个 $k>1$ 有 $t_L^{k-1}=0$, 令 $\alpha_k^k=f(x^k)-f_k^k$, 则

(a)

$$-(\alpha_k^k+\delta^{k-1})+(g^k)^Td^{k-1}\ge\eta z^{k-1}; \qquad (3.6)$$

(b)

$$w^k\le\phi_{C^k}(w^{k-1}), \qquad (3.7)$$

其中 C^k 满足 $C^k\ge\max\{\|d^{k-1}\|,\|g^k\|,\widetilde{\alpha}^{k-1},1\}$.

证 (a) 若 $t_L^{k-1}=0$, 则由算法步骤 3 及 f, c_i 的凸性可知 $x^k=x^{k-1}$, $y^k=x^{k-1}+t_R^{k-1}d^{k-1}$ 及

$$f(y^k)-f(x^{k-1})>\eta t_R^{k-1}z^{k-1}+t_R^{k-1}\delta^{k-1}.$$

因此

$$\begin{aligned}&-(\alpha_k^k+\delta^{k-1})+(g^k)^Td^{k-1}=-(f(x^k)-f_k^k)-\delta^{k-1}+(g^k)^Td^{k-1}\\ =&-(f(x^{k-1})-f(y^k)-(g^k)^T(x^{k-1}-y^k))-\delta^{k-1}+(g^k)^Td^{k-1}\\ =&f(y^k)-f(x^{k-1})+(1-t_R^{k-1})(g^k)^Td^{k-1}-\delta^{k-1}\\ \ge&\eta t_R^{k-1}z^{k-1}+t_R^{k-1}\delta^{k-1}+(1-t_R^{k-1})(g^k)^Td^{k-1}-\delta^{k-1}. \qquad (3.8)\end{aligned}$$

由 f 的凸性及 (3.8) 的第二个等式，我们有 $\alpha_k^k=f(x^{k-1})-f(y^k)+t_R^{k-1}(g^k)^Td^{k-1}\ge 0$, 这说明

$$(g^k)^Td^{k-1}\ge(f(y^k)-f(x^{k-1}))/t_R^{k-1}\ge(\eta t_R^{k-1}z^{k-1}+t_R^{k-1}\delta^{k-1})/t_R^{k-1}=\eta z^{k-1}+\delta^{k-1}.$$

上式结合 (3.8) 可知

$$-(\alpha_k^k+\delta^{k-1})+(g^k)^T d^{k-1} \ge \eta t_R^{k-1} z^{k-1}+t_R^{k-1}\delta^{k-1}+(1-t_R^{k-1})(\eta z^{k-1}+\delta^{k-1})-\delta^{k-1}=\eta z^{k-1},$$

故 (a) 成立.

(b) 定义乘子

$$\lambda_k(v)=v,\quad \lambda_j(v)=(1-v)\theta^{k-1}\lambda_j^{k-1},\quad j\in \widehat{J}^{k-1},\quad \mu_i(v)=(1-v)\mu_i^{k-1},\quad i\in I,\quad v\in[0,1]. \tag{3.9}$$

易知 (3.9) 为问题 (3.5) 的一个可行解. 此外, 我们有

$$\begin{aligned}\sum_{j\in J^k}\lambda_j(v)g^j+\sum_{i\in I}\mu_i(v)\nabla c_i(x^k)=&vg^k+(1-v)\theta^{k-1}p^{k-1}+(1-v)\sum_{i\in I}\mu_i^{k-1}\nabla c_i(x^{k-1})\\ =&vg^k+(1-v)(-d^{k-1})\end{aligned}$$

及

$$\begin{aligned}&-\sum_{j\in J^k}\lambda_j\big(f_j^k-f(x^k)-\delta^k\big)-\sum_{i\in I_-}\mu_i c_i(x^k)-\sum_{i\in I_+}\mu_i\big(c_i(x^k)-\varphi^k\big)\\ =&-v\big(f_k^k-f(x^k)-\delta^k\big)-(1-v)\theta^{k-1}\big(\widetilde{f}_p^{k-1}-f(x^k)-\delta^k\big)\\ &-(1-v)\sum_{i\in I_-}\mu_i^{k-1}c_i(x^k)-(1-v)\sum_{i\in I_+}\mu_i^{k-1}\big(c_i(x^k)-\varphi^k\big)\\ =&v(\alpha_k^k+\delta^{k-1})+(1-v)\widetilde{\alpha}^{k-1}.\end{aligned}$$

考虑问题

$$\min_{v\in[0,1]}\frac{1}{2}\big\|(1-v)(-d^{k-1})+vg^k\big\|^2+(1-v)\widetilde{\alpha}^{k-1}+v(\alpha_k^k+\delta^{k-1}). \tag{3.10}$$

设 $\overline{w}$ 为问题 (3.10) 的最优值, 则由 (3.9) 及引理 3.4 可知 $w^k\le\overline{w}$. 在引理 3.5 中取 $d=d^{k-1}$, $\widetilde{\alpha}=\widetilde{\alpha}^{k-1}$, $w=w^{k-1}$, $z=z^{k-1}$, $g=g^k$, $\alpha=\alpha_k^k+\delta^{k-1}$, 再结合 (3.6) 式, 我们有 $w^k\le\overline{w}\le\phi_{C^k}(w^{k-1})$, 故结论 (b) 成立.

引理 3.7 若算法 2.1 只产生有限个有效步, 即存在指标 $\widetilde{k}$ 使得 $x^k=x^{\widetilde{k}}$, $\forall k\ge\widetilde{k}$, 则 $x^{\widetilde{k}}$ 是问题 (1.1) 的最优解.

证 设引理条件成立, 推广 [12] 中相应分析可知, 存在一个常数 $C>0$ 使得 $C\ge\max\big\{\|d^{k-1}\|,\|g^k\|,\widetilde{\alpha}^{k-1},1\big\}$, $\forall k\ge\widetilde{k}$. 这结合 (3.7) 说明 $w^k\le\phi_C(w^{k-1})=w^{k-1}-(1-\eta)^2(w^{k-1})^2/(8C^2)$, 进而有 $w^k\to 0$, $k\to\infty$. 因此, 由引理 3.1 可知 $x^{\widetilde{k}}$ 是问题 (1.1) 的一个最优解.

基于上述分析, 我们可以得到如下全局收敛性定理.

定理 3.1 算法 2.1 或有限步终止于问题 (1.1) 的一个最优解, 或产生一个无限迭代序列 $\{x^k\}$, 使得其任意的聚点都是问题 (1.1) 的最优解.

证 由引理 2.3, 引理 3.3 及引理 3.7 可知定理成立.

4　数值试验

本节我们用 MATLAB 编程测试算法 2.1 的实际有效性，其中 4 个算例如下.

P1. Rosen-Suzuki 问题 [14].

$$\begin{aligned}\min\quad & f(x)=\max\{f_j(x):j=1,\cdots,4\}\\ \text{s.t.}\quad & c_1(x)=x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4-8\le 0,\\ & c_2(x)=x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4-10\le 0,\\ & c_3(x)=x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4-5\le 0,\end{aligned}$$

其中 $f_1(x)=x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4$，$f_2(x)=f_1(x)+10c_1(x)$，$f_3(x)=f_1(x)+10c_2(x)$，$f_4(x)=f_1(x)+10c_3(x)$.

P2. 增加约束后的 CB3 II 问题 [15], 其中 $n=10$.

$$\begin{aligned}\min\quad & f(x)=\max\{f_j(x):j=1,2,3\}\\ \text{s.t.}\quad & c_i(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-1\le 0,\qquad i=1,\cdots,n-1,\end{aligned}$$

$f_1(x)=\sum\limits_{t=1}^{n-1}(x_t^4+x_{t+1}^2)$，$f_2(x)=\sum\limits_{t=1}^{n-1}((2-x_t)^2+(2-x_{t+1})^2)$，$f_3(x)=\sum\limits_{t=1}^{n-1}2e^{(-x_t+x_{t+1})}$.

P3. 增加约束后的 Mifflin1 问题 [16].

$$\begin{aligned}\min\quad & f(x)=-x_1+20\max\{x_1^2+x_2^2-1,0\}\\ \text{s.t.}\quad & c_1(x)=x_1+2x_2-500\le 0,\ c_2(x)=x_1^2+x_2^2-4000\le 0.\end{aligned}$$

P4. 增加约束后的 Maxl 问题 [16], 其中 $n=20$.

$$\begin{aligned}\min\quad & f(x)=\max_{1\le i\le n}|x_i|\\ \text{s.t.}\quad & c_i(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-1\le 0,\qquad i=1,\cdots,n-1.\end{aligned}$$

算法参数选取为 $\beta=0.5$，$\eta=\bar{t}=0.01$, 并取 $\delta^k=2\varphi^k$, 终止准则为 $w^k\le 10^{-6}$. 数值试验结果列示在表 1 中，其列中符号含义如下：`ni`: 算法迭代次数 (该列的元素为两个数之和，前面一个表示可行域外的迭代次数，后面一个表示可行域内的迭代次数，因此两个数之和即为总的迭代次数); `nf`: 目标函数计算次数； `nfg`: 目标函数次梯度计算次数； `nc`: 约束函数计算次数； `ncg`: 约束函数梯度计算次数； `ObjValue`: 近似最优目标函数值.

对于每个问题我们分别测试了三个不同的初始点. 基于本文算法的特点，这些初始点绝大部分选为不可行的点. 从表 1 中可以看出，对给出的 4 个测试问题及不同的初始点，算法 2.1 均能够快速、有效地求解.

表 1 算法 2.1 的数值试验结果

算例	x^1	$\varphi(x^1)$	ni	nf	nfg	nc	ncg	ObjValue
P1	$(1,\cdots,1)^T$	0	0+58	117	59	177	177	−4.399999690e+01
	$(10,\cdots,10)^T$	570	15+61	154	77	234	231	−4.399999700e+01
	$(-5,\cdots,-5)^T$	150	5+63	137	69	209	207	−4.399999684e+01
P2	$(1,\cdots,1)^T$	2	87+0	181	88	792	783	3.643067617e+01
	$(5,\cdots,5)^T$	74	85+0	171	86	774	774	3.643067366e+01
	$(10,\cdots,10)^T$	299	78+0	157	79	711	711	3.643066832e+01
P3	$(1,\cdots,1)^T$	0	0+34	185	35	128	68	−1.000000000e+00
	$(50,\cdots,50)^T$	1000	3+72	230	76	196	150	−1.000000000e+00
	$(-50,\cdots,-50)^T$	1000	2+127	280	130	270	258	−9.999998415e-01
P4	$(1,\cdots,1)^T$	2	2+28	156	31	627	304	1.084202172e-18
	$(50,\cdots,50)^T$	7499	13+28	185	42	798	475	1.734723476e-17
	$(-50,\cdots,-50)^T$	7499	13+28	185	42	798	475	1.734723476e-17

参 考 文 献

[1] Zoutendijk G. Methods of Feasible Directions. Amsterdam: Elsevier, 1960

[2] Topkis D M, Veinott Jr A F. On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming. *SIAM J. Control*, 1967, 5: 268–279

[3] Panier E R, Tits A L. A Superlinearly Convergent Feasible Method for the Solution of Inequality Constrained Optimization Problems. *SIAM Journal on Control and Optimization*, 1987, 25: 934–950

[4] Lawarence C T, Tits A L. A Computationally Efficient Feasible Sequential Quadratic Programming Algorithm. *SIAM Journal on Optimization*, 2001, 11: 1092–1118

[5] Jian J B. New Sequential Quadratically Constrained Quadratic Programming Norm-relaxed Method of Feasible Directions. *Journal of Optimization Theory and Applications*, 2006, 129: 109–130

[6] Kostreva M M. Applications of MFD in Engineering Design Optimization. http://www.math.clemson.edu /flstgla/mfdapp.ps, Department of Mathematical Sciences, Clemson University

[7] 简金宝．非线性约束最优化超线性与二次收敛算法的研究．博士论文，西安交通大学， 2000 (Jian J B. Researches on Superlinearly and Quadratically Convergent Algorithms for Nonlinearly Constrained Optimization. Ph. D. Thesis, Xi'an Jiaotong University, 2000)

[8] 简金宝．光滑约束优化快速算法 – 理论分析与数值试验．北京：科学出版社， 2010 (Jian J B. Fast Algorithms for Smooth Constrained Optimization—Theoretical Analysis and Numerical Experiments. Beijing: Science Press, 2010)

[9] Jian J B, Tang C M, Hu Q J, Zheng H Y. A New Superlinearly Convergent Strongly Sub-feasible Sequential Quadratic Programming Algorithm for Inequality-constrained Optimization. *Numerical Functional Analysis and Optimization*, 2008, 29: 376–409

[10] Jian J B, Tang C M, Zheng H Y. Sequential Quadratically Constrained Quadratic Programming Algorithm of Strongly Sub-feasible Directions. *European Journal of Operational Research*, 2010, 200:

645–657

[11] 陶海，沈祥福．斜拉桥索力优化的强次可行序列二次规划法．力学学报， 2006, 38: 381–384
(Tao H, Shen X F. Strongly Subfeasible Sequential Quadratic Programming Method of Cable Tension Optimization for Cable-stayed Bridges. *Chinese Journal of Theoretical and Applied Mechanics*, 2006, 38: 381–384)

[12] Kiwiel K C. Methods of Descent for Nondifferentiable Optimization. Lecture Notes in Mathematics 1133, Springer-Verlag, 1985

[13] Kelley J E. The Cutting Plane Method for Solving Convex Programs. *Journal of the SIAM*, 1960, 8: 703–712

[14] Rustem B, Nguyen Q. An Algorithm for the Inequality-constrained Discrete Minimax Problem. *SIAM Journal on Optimization*, 1998, 8: 265–283

[15] Karmitsa N. Test Problems for Large-scale Nonsmooth Minimization. Reports of the Department of Mathematical Information Technology, Series B, Scientific computing, No. B 4/2007, University of Jyväskylä, Jyväskylä, 2007

[16] Lukšan L, Vlček J. Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization. Technical Report 798, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2000

Strongly Sub-feasible Direction Method with Subgradient Selection for Nonsmooth Optimization

TANG Chunming JIAN Jinbao

(*College of Mathematics and Information Science, Guangxi University, Nanning* 530004)

(E-mail: cmtang@gxu.edu.cn; jianjb@gxu.edu.cn)

Abstract In this paper, by combining subgradient selection technique with the ideas of cutting plane method and strongly sub-feasible direction method, a strongly sub-feasible direction algorithm is proposed for the solution of constrained optimization problems with nonsmooth objective function. By introducing a new search direction finding subproblem and constructing a new line search, the algorithm can not only accept infeasible starting points, but also preserve the sub-feasibility of the iterations, and meanwhile prevent the objective value from increasing unduly. The algorithm possesses global convergence, and some preliminary numerical results show that the proposed algorithm is stable and efficient.

Key words subgradient selection; nonsmooth optimization; strongly sub-feasible direction method; global convergence

MR(2000) Subject Classification 90C30; 65K05

Chinese Library Classification O221.2