网络优先数字出版时间:2014-3-28
网络优先数字出版地址:http://www.cnki.net/kcms/doi/10.13656/j.cnki.gxkx.20140328.001.html
广西科学 Guangxi Sciences 2014,21(3):283～286

非光滑优化的强次可行方向邻近点束求解方法*

A Proximal Bundle Method of Strongly Sub-feasible Directions for Nonsmooth Optimization

唐春明[1],简金宝[1,2]
TANG Chun-ming[1],JIAN Jin-bao[1,2]

(1.广西大学数学与信息科学学院,广西南宁 530004;2.玉林师范学院数学与信息科学学院,广西玉林 537000)
(1.College of Mathematics and Information Science,Guangxi University,Nanning,Guangxi,530004,China;2.College of Mathematics and Information Science,Yulin Normal University,Yulin,Guangxi,537000,China)

摘要:通过引入新型邻近点参数修正策略及搜索方向子问题,提出一个求解非光滑优化的强次可行方向邻近点束方法.该方法稳定性好,能保证迭代点的强次可行性,且具备全局收敛性.
关键词:非光滑优化 邻近点 束方法 强次可行方向
中图分类号:O221.2 **文献标识码**:A **文章编号**:1005-9164(2014)03-0283-04
Abstract:By introducing a new updating strategy of proximal parameters and a new search direction finding subproblem,a proximal bundle method of strongly sub-feasible directions is proposed for solving nonsmooth optimization. The proposed method is stable,can preserve sub-feasibility of the iterations,and possesses global convergence.
Key words:nonsmooth optimization,proximal point,bundle method,strongly sub-feasible directions

考虑求解如下非线性不等式约束优化问题

$$\min_{x\in R^n} f(x)$$
$$\text{s.t. } c_i(x)\leqslant 0, i\in I\triangleq\{1,\cdots,m\}, \qquad (0.1)$$

其中 $f:R^n\to R$ 为凸函数,但不一定可微(即可能非光滑),$c_i(i\in I):R^n\to R$ 是连续可微的凸函数.

强次可行方向法[1]是求解光滑约束优化的一类有效方法.该类方法能接受不可行的初始迭代点,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,且同时避免在可行域外目标函数值增加过快,一旦某个迭代点落入可行域,即自动变为可行方向法.文献[2,3]结合次梯度聚集技术、次梯度选取技术及广义割平面法的思想[4],将强次可行方向法推广到求解非光滑类型的问题(0.1).本文将结合邻近点束方法的思想[5],对文献[2]的方法作进一步改进,旨在提升算法的稳定性和计算效率.文中提出一个带邻近点控制的产生搜索方向的二次规划子问题,并给出一个新的邻近点参数更新策略,最后证明算法的全局收敛性.

1 算法描述

假设 1.1 (i)设 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;

(ii)问题(0.1)满足 Slater 约束规格,即存在一个向量 $\tilde{x}\in R^n$ 使得 $c_i(\tilde{x})<0,\forall i\in I$.

记问题(0.1)的可行集为 $F=\{x\in R^n:c_i(x)\leqslant 0,i\in I\}$.定义符号

$I_-(x)=\{i\in I:c_i(x)\leqslant 0\}$,$I_+(x)=\{i\in I:c_i(x)>0\}$,$\varphi(x)=\max\{0,c_i(x),i\in I\}$,及惩罚函

收稿日期:2013-09-30
修回日期:2013-12-11
作者简介:唐春明(1979-),男,博士,教授,主要从事最优化理论与算法的研究。
*国家自然科学基金(11301095,11126341)和广西自然科学基金项目(2013GXNSFAA019013)资助。

数:$\delta(x):R^n \to [0,+\infty)$ 是连续函数,且 $\delta(x)=0$ 当且仅当 $x \in F$.

对 $\epsilon \geqslant 0$,凸函数 f 在 $x \in R^n$ 处的 ϵ-次微分定义为

$$\partial_\epsilon f(x)=\{g \in R^n : f(y) \geqslant f(x)+g^T(y-x)-\epsilon, \forall y \in R^n\}.$$

特别地,$\partial_0 f(x)=\partial f(x)$ 为通常的次微分,$g \in \partial f(x)$ 称为一个次梯度.

设 $x^k \in R^n$ 是当前迭代点,$y^j, j \in J^k \subseteq \{1,\cdots,k\}$ 为辅助迭代点,对应次梯度 $g^j \in \partial f(y^j), j \in J^k$. 定义 f 在 y^j 处的线性化函数为

$$f_j(x)=f(y^j)+(g^j)^T(x-y^j), \forall x \in R^n, j \in J^k.$$

进而得到 f 的分片线性(下)近似函数

$$\hat{f}^k(x)=\max\{f_j(x): j \in J^k\}.$$

令 $f_j^k=f_j(x^k)$ 得

$$f_j(x)=f_j^k+(g^j)^T(x-x^k), j \in J^k.$$

为了平衡目标函数的下降性和约束函数的可行性,对于任意给定的 $y \in R^n$,引入改进函数[2]:

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

以下引理给出了函数 H 的重要性质.

引理 1.1[2] 若假设 1.1 成立,则以下 3 个命题等价:(i) $\bar{x}$ 是问题(0.1)的最优解;(ii) $\min\{H_{\bar{x}}(x): x \in R^n\}=H_{\bar{x}}(\bar{x})=0$;(iii) $0 \in \partial H_{\bar{x}}(\bar{x})$.

由引理 1.1 可知,求解原问题(0.1)等价于求改进函数 H 的"稳定点".为此,在当前迭代点 x^k 处,定义函数 $H_{x^k}(x)$ 的分片线性(下)近似函数:

$$\hat{H}_{x^k}(x)=\max\{\hat{f}^k(x)-f(x^k)-\delta^k; \hat{c}_i^k(x), i \in I_-^k; \hat{c}_i^k(x)-\varphi^k, i \in I_+^k\},$$

其中 $\hat{c}_i^k(x)=c_i(x^k)+\nabla c_i(x^k)^T(x-x^k)$, $I_-^k=I_-(x^k)$, $I_+^k=I_+(x^k)$, $\varphi^k=\varphi(x^k)$, $\delta^k=\delta(x^k)$.

结合邻近点束方法思想[5],可选取如下新试探点

$$y^{k+1}=\arg\min\{\hat{H}_{x^k}(x)+\frac{1}{2}\gamma^k\|x-x^k\|^2 : x \in R^n\},$$

其中 $\gamma^k>0$ 是邻近点参数.进一步引入次梯度聚集技术[4],并令 $d=x-x^k$,则求解 y^{k+1} 可转化为求解如下寻找搜索方向子问题:

$$\begin{aligned}
&\min_{d \in R^n, z \in R} \ z+\frac{1}{2}\gamma^k\|d\|^2 \\
&\text{s.t. } f_j^k-f(x^k)+(g^j)^T d \leqslant z+\delta^k, j \in J^k, \\
&\qquad f_p^k-f(x^k)+(p^{k-1})^T d \leqslant z+\delta^k, \\
&\qquad c_i(x^k)+\nabla c_i(x^k)^T d \leqslant z, i \in I_-^k, \\
&\qquad c_i(x^k)+\nabla c_i(x^k)^T d \leqslant z+\varphi^k, i \in I_+^k.
\end{aligned} \tag{1.1}$$

其中 $z \in R$ 是一个辅助变量,f_p^k 和 p^{k-1} 分别是前 $k-1$ 个函数值 $f_j^k, j=1,\cdots,k-1$ 和次梯度 $g^j, j=1,\cdots,k-1$ 的聚集(非负线性组合,具体更新见下面算法).

由 f 和 c_i 的凸性可知,$(d,z)=(0,0)$ 是子问题(1.1)的一个可行解.设 (d^k,z^k) 是问题(1.1)的最优解,其亦为问题(1.1)的 KKT 点,即存在乘子 $\lambda_j^k, j \in J^k, \lambda_p^k, \mu_i^k, i \in I$ 使得

$$\begin{aligned}
&\gamma^k d^k+\sum_{j \in J^k}\lambda_j^k g^j+\lambda_p^k p^{k-1}+\sum_{i \in I}\mu_i^k \nabla c_i(x^k)=0, \\
&\sum_{j \in J^k}\lambda_j^k+\lambda_p^k+\sum_{i \in I}\mu_i^k=1, \\
&0 \leqslant \lambda_j^k \perp (f_j^k-f(x^k)+(g^j)^T d^k-z^k-\delta^k) \leqslant 0, j \in J^k, \\
&0 \leqslant \lambda_p^k \perp (f_p^k-f(x^k)+(p^{k-1})^T d^k-z^k-\delta^k) \leqslant 0, \\
&0 \leqslant \mu_i^k \perp (c_i(x^k)+\nabla c_i(x^k)^T d-z^k) \leqslant 0, i \in I_-^k, \\
&0 \leqslant \mu_i^k \perp (c_i(x^k)+\nabla c_i(x^k)^T d-z^k-\varphi^k) \leqslant 0, i \in I_+^k.
\end{aligned} \tag{1.2}$$

计算满足如下关系的聚集次梯度 p^k 和聚集函数值 $\tilde{f}_p^k$:

$$\theta^k(p^k,\tilde{f}_p^k)=\sum_{j \in J^k}\lambda_j^k(g^j,f_j^k)+\lambda_p^k(p^{k-1},f_p^k), \tag{1.3}$$

其中 $\theta^k=\sum_{j \in J^k}\lambda_j^k+\lambda_p^k$.

以下引理给出了子问题(1.1)解的重要性质,其证明类似于文献[2]的引理 2.2.

引理 1.2 设 (d^k,z^k) 是子问题(1.1)的最优解,则

(i) $z^k=-(\gamma^k\|d^k\|^2+\tilde{\alpha}^k) \leqslant 0$,其中

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

(ii) $-\gamma^k d^k$ 是函数 $H_{x^k}(\cdot)$ 在 x^k 处的 ϵ 次梯度,即 $-\gamma^k d^k \in \partial_\epsilon H_{x^k}(x^k)$,其中 $\epsilon=\tilde{\alpha}^k$;

(iii) 如果 $z^k=0$,则 $d^k=0$ 且 x^k 是问题(0.1)的最优解.

算法 1.1

步骤 0 (初始化).选取初始点 $x^1 \in R^n$,次梯度 $g^1 \in \partial f(x^1)$,及参数 $\beta \in (0,1)$, $\eta \in (0.5,1)$, $\gamma_{\min}>0$, $\gamma^1>0$.令 $y^1=x^1$, $p^0=g^1$, $f_p^1=f_1^1=f(x^1)$, $J^1=\{1\}$.置 $k=1, l=0$ 及 $k(0)=1$.

步骤 1 (计算方向).求解子问题(1.1)得到最优解 (d^k,z^k) 及乘子 $\lambda_j^k, j \in J^k, \lambda_p^k, \mu_i^k, i \in I$.根据(1.3)式计算 p^k 及 $\tilde{f}_p^k$.

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

步骤 3 （线搜索）. 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：

$$c_i(x^k+td^k)\leqslant\varphi^k+\eta tz^k,i\in I_+^k,\tag{1.4}$$

$$c_i(x^k+td^k)\leqslant\eta tz^k,i\in I_-^k.\tag{1.5}$$

如果

$$f(x^k+t^kd^k)\leqslant f(x^k)+\eta t^kz^k+t^k\delta^k\tag{1.6}$$

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 $t_R^k=t^k$.

步骤 4 （更新）. 令 $x^{k+1}=x^k+t_L^kd^k,y^{k+1}=x^k+t_R^kd^k$. 选取 $\hat{J}^k\subseteq J^k$，并计算新的线性化函数值

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

$$f_p^{k+1}=\tilde{f}_p^k+(p^k)^T(x^{k+1}-x^k).$$

计算 $g^{k+1}\in\partial f(y^{k+1})$，以及 $f_{k+1}^{k+1}=f(y^{k+1})+(g^{k+1})^T(x^{k+1}-y^{k+1})$. 令 $J^{k+1}=\hat{J}^k\cup\{k+1\}$，

步骤 5 （邻近点参数修正）. 令

$$\gamma^{k+1}=\begin{cases}\max\{u^{k+1},\frac{\gamma^k}{10},\gamma_{\min}\},\text{如果 } x^{k+1}\neq x^k;\\ \gamma^k,\text{否则},\end{cases}$$

其中 $u^{k+1}=2\gamma^k(1-\frac{H_{x^k}(x^{k+1})}{t^kz^k})$.

步骤 6 令 $k:=k+1$，返回步骤 1.

注 1 步骤 5 给出了一个新的邻近点参数修正策略，其用意在于：当 $x^{k+1}\neq x^k$（有效步）时，当前的分片线性化近似模型较好，从而在下次迭代可以适当加大步长，因此需要减小邻近点参数 γ^k（见下面引理 2.1）；反之，模型近似程度较差，可保持邻近点参数不变或增大参数（此处，为简便描述，我们保持邻近点参数不变，若适当增加，也不会影响下面的理论分析[5]）.

下面的引理给出了算法 1.1 的基本性质.

引理 1.3 (i) 算法 1.1 是适定的，即线搜索 (1.4) 和 (1.5) 能在有限次计算后终止；

(ii) 算法 1.1 必定出现以下两种情形之一：(a) 存在指标 k_0 使得 $\varphi^{k_0}=0$，从而 $\varphi^k\equiv0,\delta^k\equiv0$ 及 $f(x^{k+1})\leqslant f(x^k)$，对所有 $k\geqslant k_0$ 成立；(b) $\varphi^k>0,\varphi^{k+1}\leqslant\varphi^k,I_-^k\subseteq I_-^{k+1},k=1,2,3,\cdots$.

证明 (i) 根据引理 1.2，若算法进入步骤 3，则有 $z^k<0$. 由 Taylor 展开，及 (1.2) 可得

$c_i(x^k+td^k)-\varphi^k-\eta tz^k=c_i(x^k)-\varphi^k-\eta tz^k+t\nabla c_i(x^k)^Td^k+o(t)\leqslant(1-t)(c_i(x^k)-\varphi^k)+(1-\eta)tz^k+o(t)\leqslant(1-\eta)tz^k+o(t),i\in I_+^k$.

再结合 $z^k<0,\eta\in(0.5,1)$ 可知 (1.4) 式对 $t>0$ 充分小成立. 另一方面，根据

$c_i(x^k+td^k)-\eta tz^k=c_i(x^k)+t\nabla c_i(x^k)^Td^k-\eta tz^k+o(t)\leqslant(1-t)c_i(x^k)+(1-\eta)tz^k+o(t)\leqslant(1-\eta)tz^k+o(t),i\in I_-^k$,

知 (1.5) 式对 $t>0$ 充分小成立，故 (i) 得证.

由算法 1.1 的步骤 3 易知 (ii) 成立.

2 算法全局收敛性

若算法 1.1 有限步终止于 x^k 点，则由步骤 2 可知 $w^k=0$，进而有 $z^k=0$，故根据引理 1.2(iii) 可得 x^k 是问题 (0.1) 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 (0.1) 的最优解.

引理 2.1 邻近点参数序列 $\{\gamma^k\}$ 单调不增，且有正的下界.

证明 当 $x^{k+1}\neq x^k$ 时，由线搜索 (1.4)～(1.6) 可知 $H_{x^k}(x^{k+1})\leqslant\eta t^kz^k<0$，故结合 $\eta\in(0.5,1)$ 得

$$u^{k+1}=2\gamma^k(1-\frac{H_{x^k}(x^{k+1})}{t^kz^k})\leqslant2\gamma^k(1-\eta)<\gamma^k.$$

再结合算法 1.1 的步骤 5 可知引理 2.1 成立.

引理 2.2 设有无限指标集 $K\subseteq\{1,2,\cdots\}$ 及 $\bar{x}\in R^n$ 满足 $x^k\xrightarrow{K}\bar{x}$ 及 $w^k\xrightarrow{K}0$，则 $\bar{x}$ 是问题 (0.1) 的最优解.

证明 由 $w^k\xrightarrow{K}0$ 及引理 2.1 可得 $\gamma^kd^k\rightarrow0$，$\tilde{\alpha}^k\rightarrow0,k\in K$. 故由 $x^k\xrightarrow{K}\bar{x}$ 及引理 1.2(ii) 知 $0\in\partial H_{\bar{x}}(\bar{x})$，故再由引理 1.1 可得引理 2.2 成立.

以下分两种情形证明算法的全局收敛性. 首先分析算法产生无限多个有效步的情形. 类似文献[2]引理 3.3 的分析，可得如下引理.

引理 2.3 设有无限指标集 $L\subseteq\{1,2,\cdots\}$ 和向量 $\bar{x}\in R^n$ 使得 $x^{k(l)}\rightarrow\bar{x},l\rightarrow\infty,l\in L$，则 $\bar{x}$ 是问题 (0.1) 的一个最优解.

下面分析有限个有效步的情形，即存在指标 $\tilde{k}$ 使得 $x^k=x^{\tilde{k}},k\geqslant\tilde{k}$. 我们将证明 $x^{\tilde{k}}$ 是问题 (0.1) 的最优解. 根据算法 1.1 步骤 5 可知，存在常数 $\gamma>0$，使得 $\gamma^k\equiv\gamma,\forall k\geqslant\tilde{k}$. 不失一般性，在以下分析中我们将假设 $k\geqslant\tilde{k}$.

引理 2.4 w^k 是如下二次规划问题的最优值

$\min\frac{1}{2}\|\sum_{j\in J^k}\lambda_jg^j+\lambda_pp^{k-1}+\sum_{i\in I}\mu_i\nabla c_i(x^k)\|^2-\gamma^k\sum_{j\in J^k}\lambda_j(f_j^k-f(x^k)-\delta^k)-\gamma^k\lambda_p(f_p^k-f(x^k)-\delta^k)-\gamma^k\sum_{i\in I_-}\mu_ic_i(x^k)-\gamma^k\sum_{i\in I_+}\mu_i(c_i(x^k)-\varphi^k)$

$$\text{s.t.}\ \lambda_j \geqslant 0, j \in J^k, \lambda_p \geqslant 0, \mu_i \geqslant 0, i \in I,$$

$$\sum_{j\in J^k}\lambda_j + \lambda_p + \sum_{i\in I}\mu_i = 1. \tag{2.1}$$

证明 由于问题(2.1)是子问题(1.1)的对偶问题,因此(1.1)的KKT乘子是问题(2.1)的最优解,再结合 w^k 的定义可得结论成立.

接下来,我们推广文献[4]的引理2.4.10的结果,得到如下引理.

引理2.5 设 n 维向量 d,g,及数 $\eta \in (0.5,1)$, $\gamma > 0, C, w, z, \tilde{\alpha}$ 和 α 满足

$$w = \frac{1}{2}\|\gamma d\|^2 + \gamma\tilde{\alpha}, z = -(\gamma\|d\|^2 + \tilde{\alpha}), -\alpha + g^T d \geqslant \eta z, C \geqslant \max\{\|\gamma d\|, \|g\|, \tilde{\alpha}, 1\}.$$

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

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

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

引理2.6 假设对某个 $k > 1$ 有 $t_L^{k-1} = 0$,令 $\alpha_k^k = f(x^k) - f_k^k$,则

$$\text{(i)}\ -(\alpha_k^k + \delta^{k-1}) + (g^k)^T d^{k-1} \geqslant \eta z^{k-1}; \tag{2.2}$$

$$\text{(ii)}\ w^k \leqslant \phi_{C^k}(w^{k-1}), \tag{2.3}$$

其中 C^k 满足 $C^k \geqslant \max\{\|\gamma d^{k-1}\|, \|g^k\|, \tilde{\alpha}^{k-1}, 1\}$.

证明 (i) 根据线搜索(1.6),该部分的证明与文献[2]的引理3.6(i)完全平行.

(ii) 对任意 $v \in [0,1]$,定义乘子

$$\lambda_k(v) = v, \lambda_j(v) = 0, j \in J^k\backslash\{k\}, \lambda_p(v) = (1-v)\theta^{k-1}, \mu_i(v) = (1-v)\mu_i^{k-1}, i \in I, \tag{2.4}$$

其为问题(2.1)的一个可行解.此外,我们有

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

$$-\sum_{j\in J^k}\lambda_j(f_j^k - f(x^k) - \delta^k) - \lambda_p(f_p^k - f(x^k) - \delta^k) - \sum_{i\in I_-}\mu_i c_i(x^k) - \sum_{i\in I_+}\mu_i(c_i(x^k) - \varphi^k) = -v(f_k^k - f(x^k) - \delta^k) - (1-v)\theta^{k-1}(f_p^k - f(x^k) - \delta^k) - (1-v)\sum_{i\in I_-}\mu_i^{k-1}c_i(x^k) - (1-v)\sum_{i\in I_+}\mu_i^{k-1}(c_i(x^k) - \varphi^k) = v(\alpha_k^k + \delta^{k-1}) + (1-v)\tilde{\alpha}^{k-1}.$$

设 $\bar{w}$ 为如下问题的最优值

$$\min \frac{1}{2}\|(1-v)(-\gamma d^{k-1}) + vg^k\|^2 + \gamma[(1-v)\tilde{\alpha}^{k-1} + v(\alpha_k^k + \delta^{k-1})]$$

$$\text{s.t.}\ v \in [0,1]. \tag{2.5}$$

则由引理2.4可知 $w^k \leqslant \bar{w}$.在引理2.5中令 $d = d^{k-1}$, $\tilde{\alpha} = \tilde{\alpha}^{k-1}, w = w^{k-1}, z = z^{k-1}, g = g^k, \alpha = \alpha_k^k + \delta^{k-1}$,$\gamma$ 取为邻近点参数,并结合(2.2),可得 $w^k \leqslant \bar{w} \leqslant \phi_{C^k}(w^{k-1})$,故(ii)得证.

基于上述分析,可得全局收敛性定理.

定理2.1 算法1.1或有限步终止于问题(0.1)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(0.1)的最优解.

证明 若算法1.1有限步终止于 x^k 点,则根据引理1.2(iii)可得 x^k 是问题(0.1)的一个最优解.现假设算法1.1产生一个无限迭代序列 $\{x^k\}$,且 x^* 为任一给定的聚点.下面分两种情形证明.

情形Ⅰ 有无限多个有效步.此时,必存在无限指标集 $L \subseteq \{1,2,\cdots\}$ 使得 $x^{k(l)} \to x^*, l \to \infty, l \in L$.因此,由引理2.3可知 x^* 是问题(0.1)的最优解.

情形Ⅱ 有限多个有效步.则存在指标 $\tilde{k}$ 使得 $x^k \equiv x^{\tilde{k}} = x^*, \gamma^k \equiv \gamma, \forall k \geqslant \tilde{k}$.此时,存在一个常数 $C > 0$ 使得

$$C \geqslant \max\{\|\gamma d^{k-1}\|, \|g^k\|, \tilde{\alpha}^{k-1}, 1\}, \forall k \geqslant \tilde{k}.$$

结合(2.3)式,说明

$$w^k \leqslant \phi_C(w^{k-1}) = w^{k-1} - (1-\eta)^2(w^{k-1})^2/(8C^2),$$

进而有 $w^k \to 0, k \to \infty$.因此,由引理2.2可知 x^* 是问题(0.1)的一个最优解.

参考文献:

[1] 简金宝.光滑约束优化快速算法——理论分析与数值试验[M].北京:科学出版社,2010.
Jian J B. Fast algorithms for smooth constrained optimization: theoretical analysis and numerical experiments [M]. Beijing: Science Press, 2010.

[2] Tang C M, Jian J B. Strongly sub-feasible direction method for constrained optimization problems with nonsmooth objective functions[J]. European Journal of Operational Research, 2012, 218: 28-37.

[3] 唐春明,简金宝.基于次梯度选取的非光滑优化强次可行方向法[J].应用数学学报,2011,34(5):924-937.
Tang C M, Jian J B. Strongly sub-feasible direction method with subgradient selection for nonsmooth optimization [J]. Acta Mathematicae Applicatae Sinica, 2011, 34(5): 924-937.

[4] Kiwiel K C. Methods of descent for nondifferentiable optimization[M]. [S. L.]: Lecture Notes in Mathematics 1133, Springer-Verlag, 1985.

[5] Kiwiel K C. Proximity control in bundle methods for convex nondifferentiable minimization[J]. Mathematical Programming, 1990, 46: 105-122.

(责任编辑:尹 闯)