DOI:10.13656/j.cnki.gxkx.20180813.003
欧小梅,唐春明.带非欧氏范数的双稳定束方法[J].广西科学,2018,25(4):428-432.
OU X M,TANG C M.A doubly stabilized bundle method with non-Euclidean norm[J].Guangxi Sciences,2018,25(4):428-432.

带非欧氏范数的双稳定束方法*

A Doubly Stabilized Bundle Method with Non-Euclidean Norm

欧小梅,唐春明**
OU Xiaomei,TANG Chunming

(广西大学数学与信息科学学院,广西南宁 530004)
(College of Mathematics and Information Science,Guangxi University,Nanning,Guangxi,530004,China)

摘要:针对一类非光滑凸优化问题,提出一个带非欧氏范数的双稳定束方法.通过利用邻近函数代替传统的欧氏距离,形成更具广泛性的双稳定子问题,进而在计算上可充分利用可行集的几何结构,加快收敛速度、减少计算量.分析论证了算法的全局收敛性,当下降步有限时,最后一个稳定中心即为问题的最优解;当下降步无限时,稳定中心点列任意的聚点均为问题的最优解.该方法将传统邻近束方法和水平束方法的稳定性有机融合,从而具备更优越的理论性质和更稳定的数值效果.

关键词:双稳定束方法 邻近函数 非光滑优化 全局收敛
中图分类号:O29 **文献标识码**:A **文章编号**:1005-9164(2018)04-0428-05

Abstract:In this paper,a doubly stabilized bundle method with non-Euclidean norm was proposed to solve a class of non-smooth convex optimization problems.By using the proximity function to replace the traditional Euclidean distance,a more general doubly stabilized sub-problem was formed,and the geometric structure of the feasible set could be fully utilized in the calculation to speed up the convergence rate and reduce the computational cost.The global convergence of the algorithm was analyzed and demonstrated.When the number of descent steps was finite,the last stable center was the optimal solution of the problem.When the number of descent steps was infinite,each accumulation point of the sequence of stable centers was the optimal solution of the problem.This method well combined the stability of traditional proximal bundle method and level bundle method, so that it had more superior theoretical properties and more stable numerical effects.

Key words:doubly stabilized bundle method, proximal function, non-smooth optimization, global convergence

0 引言

考虑求解如下非光滑凸优化问题

$$f^* = \inf \{f(x) : x \in X\}, \tag{0.1}$$

其中,$f:R^n \to R$ 为非光滑凸函数,$X \subseteq R^n$ 为非空闭凸集.

非光滑优化是最优化研究[1-3]的重要分支,其在

收稿日期:2018-01-29
作者简介:欧小梅(1992-),女,在读研究生,主要从事最优化理论、方法及其应用研究。
*国家自然科学基金项目(11761013)和广西自然科学基金项目(2014GXNSFFA118001,2016GXNSFDA380019)资助。
**通信作者:唐春明(1979-),男,博士,教授,主要从事最优化理论、方法及其应用研究,E-mail:cmtang@gxu.edu.cn。

数据挖掘[4]、图像恢复[5]及机器学习[6]等领域有着广泛应用.束方法是求解非光滑优化最有效的方法之一[7-9].传统束方法包括邻近束方法[10]、水平束方法[11]和信赖域束方法[12]等.最近,De Oliveira 和 Solodov[13]结合邻近束方法与水平束方法的稳定性,提出求解问题(0.1)的一个双稳定束方法,其通过求解如下子问题产生新迭代点 x_{k+1}:

$$\min\left\{\check{f}_k(x)+\frac{1}{2\tau_k}\|x-\hat{x}_k\|^2:\check{f}_k(x)\leqslant l_k,x\in X\right\}, \tag{0.2}$$

其中,$\check{f}_k(x)$ 是在第 k 次迭代时产生的 $f(x)$ 的割平面模型,且满足 $\check{f}_k(x)\leqslant f(x)$,$\hat{x}_k$ 为当前稳定中心,$\tau_k>0$ 为邻近参数,$l_k\in R$ 为水平参数.

注意到子问题(0.2)的邻近项采用的是经典的欧氏距离,为充分利用可行集的几何结构,加快收敛速度,提升数值效果,许多方法引入了非欧氏距离.沈洁等[14]在文献[13]中引入矩阵范数,应用对偶思想进行求解,分析算法的全局收敛性.Kiwiel 等在邻近点算法中引入 Bregman 距离替代邻近项,获得算法的全局收敛性[15];在邻近束方法中引入 Bregman 距离替代邻近项,允许求解子问题时出现近似解,获得算法的全局收敛性[16].

本文基于邻近函数,引入非欧氏距离,对子问题(0.2)进行改进,提出一个带非欧氏范数的双稳定束方法.该方法充分吸收了传统邻近束方法和水平束方法的优势,在产生新迭代点的二次规划子问题中引入邻近函数代替传统的欧氏距离,使得在计算上可充分利用可行集的几何结构,获得优良的数值效果.

本文符号说明.凸函数 f 在 x 处的次微 $\partial f(x)=\{g:f(y)\geqslant f(x)+\langle g,y-x\rangle,\forall y\}$,ε-次微分记为 $\partial_\varepsilon f(x)=\{g:f(y)\geqslant f(x)+\langle g,y-x\rangle-\varepsilon,\forall y\}$. $N_X(x)$ 表示 X 在点 x 处的法锥,即 $N_X(x)=\{y:\langle y,z-x\rangle\leqslant 0,\forall z\in X\}$. $i_X(x)$ 为指示函数,即:若 $x\in X$,则 $i_X(x)=0$;否则 $i_X(x)=+\infty$.

1 算法与性质

引入邻近函数[17]

$$D(x,y)=h(x)-h(y)-\langle\nabla h(y),x-y\rangle.$$

为叙述简便,假设 $h:X\to R$ 为一阶连续可微的强凸函数,满足

$$h(y)\geqslant h(x)+\langle\nabla h(x),y-x\rangle+\frac{\sigma_h}{2}\|y-x\|^2,\forall x,y\in X, \tag{1.1}$$

其中 $\sigma_h>0$ 为函数 h 的凸性参数.

将子问题(0.2)推广如下:

$$\min\left\{\check{f}_k(x)+\frac{1}{\tau_k}D(x,\hat{x}_k):\check{f}_k(x)\leqslant l_k,x\in X\right\}, \tag{1.2}$$

其中 $\check{f}_k(x)=\max_{j\in B_k}\{\bar{f}_j(x):=f(x_j)+\langle g_j,x-x_j\rangle\}$,指标集 $B_k\subseteq\{1,\cdots,k\}$,$x_j$ 为迭代点,$g_j\in\partial f(x_j)$.当 $h(\cdot)=\frac{1}{2}\|\cdot\|$ 时,子问题(1.2)变为子问题(0.2).

通过引入一个变量 t,问题(1.2)可写成:

$$\min_{(x,t)\in R^{n+1}}\{t+\frac{1}{\tau_k}D(x,\hat{x}_k):\check{f}_k(x)\leqslant t,t\leqslant l_k,x\in X\}, \tag{1.3}$$

易知问题(1.2)与问题(1.3)关于 x 部分的解是等价的.

以下引理给出子问题(1.2)的性质.

引理 1.1 若 $X_k=\{x\in X:\check{f}_k(x)\leqslant l_k\}\neq\varnothing$,则子问题(1.2)有唯一解 x_{k+1}.此外,若 X 为多面体或者 $riX\cap\{x\in R^n:\check{f}_k(x)\leqslant l_k\}\neq\varnothing$,则存在 $s_{k+1}\in\partial\check{f}_k(x_{k+1})$,$p_{k+1}\in N_X(x_{k+1})=\partial i_X(x_{k+1})$ 和拉格朗日乘子 $\mu_k\geqslant 1$,$\lambda_k\geqslant 0$ 使得

$$\nabla h(x_{k+1})=\nabla h(\hat{x}_k)-\tau_k\mu_k\hat{g}_k, \tag{1.4}$$

$$\hat{g}_k=s_{k+1}+\frac{1}{\mu_k}p_{k+1},\mu_k=\lambda_k+1, \tag{1.5}$$

$$\mu_k(\check{f}_k(x_{k+1})-t_{k+1})=0,\lambda_k(\check{f}_k(x_{k+1})-l_k)=0$$

成立.此外,对于任意的 $x\in X$,定义聚集线性化函数

$$\bar{f}_k^a(x):=\check{f}_k(x_{k+1})+\langle\hat{g}_k,x-x_{k+1}\rangle, \tag{1.6}$$

则

$$\bar{f}_k^a(x)\leqslant\check{f}_k(x)\leqslant f(x),x\in X. \tag{1.7}$$

证明:由于问题(1.2)的目标函数是强凸函数且可行集非空,故存在唯一最优解 x_{k+1}.根据问题(1.3)的最优性条件,存在最优解 (x_{k+1},t_{k+1}) 及拉格朗日乘子 $\mu_k\geqslant 0$ 和 $\lambda_k\geqslant 0$ 使得

$$0\in\frac{1}{\tau_k}(\nabla h(x_{k+1})-\nabla h(\hat{x}_k))+\mu_k\partial\check{f}_k(x_{k+1})+N_X(x_{k+1}), \tag{1.8}$$

$$0=1-\mu_k+\lambda_k,$$

$$\mu_k(\check{f}_k(x_{k+1})-t_{k+1})=0,\quad\lambda_k(t_{k+1}-l_k)=0.$$

由 $\mu_k=1+\lambda_k\geqslant 1$,可知 $\check{f}_k(x_{k+1})=t_{k+1}$.由公式(1.8)可知,存在 $s_{k+1}\in\partial\check{f}_k(x_{k+1})$,$p_{k+1}\in N_X(x_{k+1})$ 使得

$$\nabla h(x_{k+1})=\nabla h(\hat{x}_k)-\tau_k(\mu_k s_{k+1}+p_{k+1})=\nabla h(\hat{x}_k)-\tau_k\mu_k\hat{g}_k.$$

对于任意的 $x\in X$,结合公式(1.5),(1.6)及 $p_{k+1}\in N_X(x_{k+1})$,$s_{k+1}\in\partial\check{f}_k(x_{k+1})$,有

$$\bar{f}_k^a(x)=\check{f}_k(x_{k+1})+<s_{k+1},x-x_{k+1}>+\frac{1}{\mu_k}<p_{k+1},$$

$x - x_{k+1} > \leqslant \check{f}_k(x_{k+1}) + < s_{k+1}, x - x_{k+1} > \leqslant \check{f}_k(x) \leqslant f(x)$.

基于文献[13]中的算法1及子问题(1.3),下面给出带非欧氏范数的双稳定束方法.

算法1(带非欧氏范数的双稳定束方法)

步骤0 (初始化)取参数$\beta, m_l \in (0,1)$,终止参数$Tol_\Delta, Tol_e, Tol_g > 0$. 取$x_1 \in X$,令$\hat{x}_1 = x_1$. 若$f^*$的一个下界$f_1^{low}$可获得,令$v_1^l = (1-m_l)(f(\hat{x}_1) - f_1^{low})$;否则,令$f_1^{low} = -\infty$,并取$v_1^l > 0$. 选取$\tau_{min} > 0$,$\tau_1 \geqslant \tau_{min}$,令$k := 1$.

步骤1 (第一终止测试)令$\Delta_k = f(\hat{x}_k) - f_k^{low}$. 若$\Delta_k \leqslant Tol_\Delta$,终止算法.

步骤2 (可行性检测)计算水平参数$l_k = f(\hat{x}_k) - v_k{}^l$. 若水平集$X_k$为空,令$f_k^{low} = l_k$,$v_k^l = (1-m_l)(f(\hat{x}_k) - f_k^{low})$,返回步骤1.

步骤3 (寻找新迭代点)解问题(1.3)得(x_{k+1}, t_{k+1})及相应于水平参数$t \leqslant l_k$的乘子λ_k,且有$\mu_k = \lambda_k + 1$,$v_k^\tau = f(\hat{x}_k) - t_{k+1}$,$\hat{g}_k = \frac{1}{\tau_k \mu_k}(\nabla h(\hat{x}_k) - \nabla h(x_{k+1}))$,$\hat{e}_k = v_k^\tau - < \hat{g}_k, \hat{x}_k - x_{k+1} >$.

步骤4 (第二终止测试)若$\hat{e}_k \leqslant Tol_e$且$\| \hat{g}_k \| \leqslant Tol_g$,终止算法.

步骤5 (下降测试)取$f_{k+1}^{low} \in [f_k^{low}, f^*]$,若

$$f(x_{k+1}) \leqslant f(\hat{x}_k) - \beta v_k^\tau \tag{1.9}$$

成立,转至步骤5.1(下降步);否则转至步骤5.2(无效步).

步骤5.1 (下降步)令$\hat{x}_{k+1} = x_{k+1}$,$\tau_{k+1} = \tau_k \mu_k$和$v_{k+1}^l = \min\{v_k^l, (1-m_l)(f(\hat{x}_{k+1}) - f_{k+1}^{low})\}$.

选取模型$\check{f}_{k+1}$满足$\check{f}_{k+1}(\cdot) \leqslant f(\cdot)$.

步骤5.2 (无效步)令$\hat{x}_{k+1} = \hat{x}_k$,取$\tau_{k+1} \in [\tau_{min}, \tau_k]$. 若$\mu_k > 1$,令$v_{k+1}^l = m_l v_k^l$;否则令$v_{k+1}^l = v_k^l$.

选取模型$\check{f}_{k+1}$满足$\max\{\bar{f}_{k+1}(\cdot), \bar{f}_k^a(\cdot)\} \leqslant \check{f}_{k+1}(\cdot) \leqslant f(\cdot)$.

步骤6 (循环)令$k =: k+1$,返回步骤1.

以下引理给出算法的近似最优性条件,其证明类似文献[13].

引理1.2 聚集线性化误差$\hat{e}_k$,满足

$$\hat{e}_k \geqslant 0, \hat{e}_k + < \hat{g}_k, \hat{x}_k - x_{k+1} > = v_k^\tau \geqslant v_k^l. \tag{1.10}$$

若$\mu_k > 1$,则$v_k^\tau = v_k^l$. 此外,对于任意的$x \in X$,有

$$f(\hat{x}_k) + < \hat{g}_k, x - \hat{x}_k > - \hat{e}_k \leqslant f(x). \tag{1.11}$$

2 收敛性分析

根据算法1的步骤可知在迭代过程中必定会出现以下3种情况之一:水平集$X_k = \emptyset$出现无限次;水平集$X_k = \emptyset$出现有限次时,算法1产生有限多下降步;水平集$X_k = \emptyset$出现有限次时,算法1产生无限多下降步.

首先讨论水平集$X_k = \emptyset$出现无限次的情形.

定理2.1[13] 假设$X_k = \emptyset$出现无限次,则$\Delta_k \to 0$,$\{f(\hat{x}_k)\} \to f^*$,且序列$\{\hat{x}_k\}$的每个聚点(若存在)都是问题(0.1)的解;或者当$\{\hat{x}_k\}$有限时,最后一个迭代点$\hat{x}_k$是问题(0.1)的解.

下面讨论$X_k = \emptyset$出现有限次的情形. 不失一般性,假设$X_k \neq \emptyset$, $\forall k \geqslant 1$.

首先讨论算法1产生有限多下降步的情形.

对于$x \in X$,定义函数

$$\bar{\phi}_k(x) := \bar{f}_k^a(x) + \frac{1}{\tau_k} D(x, \hat{x}_k), \tag{2.1}$$

$$\check{\phi}_k(x) := \check{f}_k(x) + \frac{1}{\tau_k} D(x, \hat{x}_k). \tag{2.2}$$

引理2.3 若$\mu_k = 1$,则

$$D(x, x_{k+1}) + D(x_{k+1}, \hat{x}_k) - D(x, \hat{x}_k) = < \nabla h(\hat{x}_k) - \nabla h(x_{k+1}), x - x_{k+1} > = \tau_k < \hat{g}_k, x - x_{k+1} >, \tag{2.3}$$

$$\bar{\phi}_k(x) = \bar{\phi}_k(x_{k+1}) + \frac{1}{\tau_k} D(x, x_{k+1}). \tag{2.4}$$

证明:由公式(1.4)有

$D(x, x_{k+1}) + D(x_{k+1}, \hat{x}_k) - D(x, \hat{x}_k) = < \nabla h(\hat{x}_k) - \nabla h(x_{k+1}), x - x_{k+1} > = \tau_k < \hat{g}_k, x - x_{k+1} >$.

根据公式(1.6),(2.1)与(2.3)有

$\bar{\phi}_k(x) = \check{f}_k(x_{k+1}) + < \hat{g}_k, x - x_{k+1} > + \frac{1}{\tau_k} D(x, \hat{x}_k) = \bar{f}_k^a(x_{k+1}) + < \hat{g}_k, x - x_{k+1} > + \frac{1}{\tau_k} D(x, \hat{x}_k) = \bar{\phi}_k(x_{k+1}) - \frac{1}{\tau_k} D(x_{k+1}, \hat{x}_k) + \frac{1}{\tau_k}(D(x, x_{k+1}) + D(x_{k+1}, \hat{x}_k) - D(x, \hat{x}_k)) + \frac{1}{\tau_k} D(x, \hat{x}_k) = \bar{\phi}_k(x_{k+1}) + \frac{1}{\tau_k} D(x, x_{k+1})$.

引理2.4 假设存在$k_1 \geqslant 1$,对任意$k \geqslant k_1$,下降测试公式(1.9)不成立,即产生的都是无效步. 若$\mu_k = 1$,则

$$f(x_{k+1}) - \check{f}_k(x_{k+1}) \to 0, \tag{2.5}$$

此外,当$\{\tau_k\} \to \tau_\infty > 0$时,有

$$x_{k+1} \to x_\infty := \operatorname{Argmin}\{f(x) + \frac{1}{\tau_\infty} D(x, \hat{x}_{k_1}), x \in X\}. \tag{2.6}$$

证明:对任意的$k \geqslant k_1$,由公式(1.7),(2.1),(2.2)和(2.4)可得

$f(\hat{x}_{k_1}) \geqslant \check{f}_{k+1}(\hat{x}_{k_1}) = \check{\phi}_{k+1}(\hat{x}_{k_1}) \geqslant \check{\phi}_{k+1}(x_{k+2}) =$

$\bar{\phi}_{k+1}(x_{k+2}) \geqslant \bar{\phi}_k(x_{k+2}) \geqslant \bar{\phi}_k(x_{k+1}) + \frac{1}{\tau_k}D(x_{k+2},x_{k+1}) \geqslant \bar{\phi}_k(x_{k+1})$.

因为 $\hat{x}_{k_1}$ 是固定的，由单调有界定理可知序列 $\{\bar{\phi}_k(x_{k+1})\}$ 收敛，再结合 $\tau_{\min} \leqslant \tau_{k+1} \leqslant \tau_k$ 可得 $\{D(x_{k+2},x_{k+1})\} \to 0$. 进而由公式(1.1)有 $x_{k+2} - x_{k+1} \to 0$. 在公式(2.4)中固定 x，结合 $\{\bar{\phi}_k(x_{k+1})\}$ 收敛，可知 $\{x_{k+1}\}$ 有界. 由算法1步骤5.2，有

$f(x_{k+2}) - f(x_{k+1}) \geqslant \check{f}_{k+1}(x_{k+2}) - f(x_{k+1}) \geqslant < g_{k+1}, x_{k+2} - x_{k+1} >$，其中 $g_{k+1} \in \partial f(x_{k+1})$. 结合 $\{x_{k+1}\}$ 有界和 g_{k+1} 的有界性有 $\check{f}_{k+1}(x_{k+2}) - f(x_{k+1}) \to 0$，即 $\check{f}_k(x_{k+1}) - f(x_{k+1}) \to 0$.

设 $\bar{x}$ 为 $\{x_{k+1}\}$ 的任一聚点，则存在无限指标集 K'，使得 $x_{k+1} \to \bar{x}, k \in K'$. 由公式(1.7)，(2.1)和(2.4)有

$$f(x) + \frac{1}{\tau_k}D(x,\hat{x}_{k_1}) \geqslant \bar{f}_k^a(x) + \frac{1}{\tau_k}D(x,\hat{x}_{k_1}) = \bar{\phi}_k(x) \geqslant \bar{\phi}_k(x_{k+1}) = \bar{f}_k^a(x_{k+1}) + \frac{1}{\tau_k}D(x_{k+1},\hat{x}_{k_1}) = f(x_{k+1}) + (\check{f}_k(x_{k+1}) - f(x_{k+1})) + \frac{1}{\tau_k}(h(x_{k+1}) - h(\hat{x}_{k_1}) - < \nabla h(\hat{x}_{k_1}), x_{k+1} - \hat{x}_{k_1} >), x \in X.$$

上式对 $k \in K', k \to \infty$ 取极限有

$$f(x) + \frac{1}{\tau_\infty}D(x,\hat{x}_{k_1}) \geqslant f(\bar{x}) + \frac{1}{\tau_\infty}D(\bar{x},\hat{x}_{k_1}), x \in X.$$

由问题(2.6)解的唯一性知 $\bar{x} = x_\infty$，因此结合 $\{x_{k+1}\}$ 的有界性及 $\bar{x}$ 的任意性可得 $\{x_{k+1}\} \to x_\infty$.

定理2.2 设存在 $k_1 \geqslant 1$，当 $k \geqslant k_1$ 时，下降测试公式(1.9)不成立，则集合 $K := \{k : \mu_k > 1\}$ 为无限集，且 $\hat{x}_{k_1}$ 为问题(0.1)的最优解.

证明：根据算法1知序列 $\{v_k^l\}$ 是非增的，由步骤5.1有 $v_{k+1}^l = m_l v_k^l, \forall k \in K$.

反证法，假设 K 是有限集. 由引理1.2，存在 $k_2 \geqslant k_1$ 使得

$\mu_k = 1, \lambda_k = 0, v_k^l = v_{k_2}^l > 0, \forall k \geqslant k_2$.

结合引理2.4，得 $\{f(x_{k+1})\} \to f(x_\infty)$. 若下降测试公式(1.9)不成立，则有

$f(x_{k+1}) - \check{f}_k(x_{k+1}) > f(\hat{x}_{k_1}) - \beta v_k^\tau - \check{f}_k(x_{k+1}) = (1-\beta)v_k^\tau, \forall k \geqslant k_1$.

由 $\beta \in (0,1)$，结合公式(2.5)，可知 $v_k^\tau \to 0$，这与 $v_k^\tau \geqslant v_k^l = v_{k_2}^l > 0, \forall k \geqslant k_2$ 矛盾. 因此，K 为无限集.

由 K 为无限集，根据算法1步骤5.1有 $v_{k+1}^l = m_l v_k^l, \forall\ k \in K$，因为 $m_l \in (0,1)$ 且 $\{v_k^l\}$ 单调非增，所以 $v_k^l \to 0, k \to \infty$. 由引理1.2知，$v_k^\tau = v_k^l, k \in K$. 因此 $\{v_k^\tau\} \to 0, k \in K$. 由 h 的强凸性，根据文献[18]定理2.1.9有

$$< \nabla h(\hat{x}_{k_1}) - \nabla h(x_{k+1}), \hat{x}_{k_1} - x_{k+1} > \geqslant \sigma_h \| \hat{x}_{k_1} - x_{k+1} \|^2, \tag{2.7}$$

结合公式(1.4)，(1.10)和(2.7)，有

$$v_k^\tau = \hat{e}_k + < \hat{g}_k, \hat{x}_{k_1} - x_{k+1} > = \hat{e}_k + \frac{1}{\tau_k \mu_k} < \nabla h(\hat{x}_{k_1}) - \nabla h(x_{k+1}), \hat{x}_{k_1} - x_{k+1} > \geqslant \hat{e}_k + \frac{\sigma_h}{\tau_k \mu_k} \| \hat{x}_{k_1} - x_{k+1} \|^2 \geqslant 0, k \in K.$$

所以 $\hat{e}_k \to 0, \hat{x}_{k_1} - x_{k+1} \to 0, k \in K$. 由公式(1.4)，$\nabla h(x)$ 连续可知 $\hat{g}_k \to 0, k \in K$，结合公式(1.11)可知 $\hat{x}_{k_1}$ 是问题(0.1)的解.

接下来讨论算法1产生无限多下降步的情形.

定理2.3 假设算法1产生无限多下降步，则 $\{f(\hat{x}_k)\} \to f^*$ 且序列 $\{\hat{x}_k\}$ 的任意一个聚点都是问题(0.1)的解.

证明：记 $\{\hat{x}_{k(j)}\}$ 为 $\{\hat{x}_k\}$ 的一个子列，$k(j)$ 表示第 j 次下降步的指标，定义 $i(j) = k(j+1) - 1, \forall j \geqslant 1$. 由公式(1.4)，(1.11)和引理1.2有

$$\nabla h(\hat{x}_{k(j+1)}) = \nabla h(\hat{x}_{k(j)}) - \tau_{i(j)} \mu_{i(j)} \hat{g}_{i(j)}, \tag{2.8}$$

$$\hat{g}_{i(j)} \in \partial_{\hat{e}_{k(j)}}(f + i_X)(\hat{x}_{k(j)}), \tag{2.9}$$

$$\tau_{i(j)} \mu_{i(j)} \geqslant \tau_{min},$$

$$f(\hat{x}_{k(j)}) - f(\hat{x}_{k(j+1)}) \geqslant \beta v_{i(j)}^\tau. \tag{2.10}$$

由算法1知 $\{f(\hat{x}_{k(j)})\}$ 是下降序列. 若 $\{f(\hat{x}_{k(j)})\} \to -\infty$，证毕，以下不妨假设 $\{f(\hat{x}_{k(j)})\}$ 有界.

由 $\tau_{i(j)} \mu_{i(j)} \geqslant \tau_{min} > 0, \forall j \geqslant 1$，有

$$\sum_{j=1}^{\infty} \tau_{i(j)} u_{i(j)} = +\infty. \tag{2.11}$$

结合 $\beta \in (0,1)$ 和 $\{f(\hat{x}_{k(j)})\}$ 是单调有界序列，由公式(2.10)可得到 $v_{i(j)}^\tau \to 0, j \to \infty$. 再由公式(1.10)，(2.7)和(2.8)有

$$v_{i(j)}^\tau = \hat{e}_{i(j)} + < \hat{g}_{i(j)}, \hat{x}_{k(j)} - \hat{x}_{k(j+1)} > = \hat{e}_{i(j)} + \frac{1}{\tau_{i(j)} u_{i(j)}} < \nabla h(\hat{x}_{k(j)}) - \nabla h(\hat{x}_{k(j+1)}), \hat{x}_{k(j)} - \hat{x}_{k(j+1)} > \geqslant \hat{e}_{i(j)} + \frac{\sigma_h}{\tau_{i(j)} u_{i(j)}} \| \hat{x}_{k(j)} - \hat{x}_{k(j+1)} \|^2 \geqslant 0,$$

所以

$$\hat{e}_{i(j)} \to 0, \hat{x}_{k(j)} - \hat{x}_{k(j+1)} \to 0, j \to \infty. \tag{2.12}$$

因为下降序列 $\{f(\hat{x}_{k(j)})\}$ 有界，不妨假设它收敛到 $\bar{f}$. 要证 $\{f(\hat{x}_{k(j)})\} \to f^*$，因为 $\bar{f} \geqslant f^*$，需证 $\bar{f} \leqslant f^*$. 反证法，假设 $\bar{f} > f^*$，则存在 $\zeta > 0$ 和 $z \in X$，使得 $f(z) + \zeta < f(\hat{x}_{k(j)})$. 结合公式(1.4)，(2.3)，(2.8)和(2.9)有

$$D(z,\hat{x}_{k(j+1)}) = D(z,\hat{x}_{k(j)}) - D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) + < \nabla h(\hat{x}_{k(j)}) - \nabla h(\hat{x}_{k(j+1)}), z - \hat{x}_{k(j+1)} > = D(z,\hat{x}_{k(j)}) - D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) + \tau_{i(j)} \mu_{i(j)} < \hat{g}_{i(j)}, z - \hat{x}_{k(j+1)} > \leqslant$$

$D(z,\hat{x}_{k(j)}) - D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) + \tau_{i(j)}\mu_{i(j)}(f(z) - f(\hat{x}_{k(j)}) + \hat{e}_{i(j)}) + \tau_{i(j)}\mu_{i(j)} < \hat{g}_{i(j)},\hat{x}_{k(j)} - \hat{x}_{k(j+1)} >= D(z,\hat{x}_{k(j)}) - D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) + \tau_{i(j)}\mu_{i(j)}(f(z) - f(\hat{x}_{k(j)}) + \hat{e}_{i(j)}) + < \nabla h(\hat{x}_{k(j)}) - \nabla h(\hat{x}_{k(j+1)}),\hat{x}_{k(j)} - \hat{x}_{k(j+1)} > \leqslant D(z,\hat{x}_{k(j)}) + \tau_{i(j)}\mu_{i(j)}(\frac{1}{\tau_{i(j)}\mu_{i(j)}} < \nabla h(\hat{x}_{k(j)}) - \nabla h(\hat{x}_{k(j+1)}),\hat{x}_{k(j)} - \hat{x}_{k(j+1)} > - \frac{1}{\tau_{i(j)}\mu_{i(j)}} D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) - \zeta + \hat{e}_{i(j)})$.

由公式(2.12),可得

$$\frac{1}{\tau_{i(j)}\mu_{i(j)}} < \nabla h(\hat{x}_{k(j)}) - \nabla h(\hat{x}_{k(j+1)}),\hat{x}_{k(j)} - \hat{x}_{k(j+1)} > - \frac{1}{\tau_{i(j)}\mu_{i(j)}} D(\hat{x}_{k(j+1)},\hat{x}_{k(j)}) + \hat{e}_{i(j)} \to 0.$$

因此存在正整数 q 使得

$$D(z,\hat{x}_{k(j+1)}) \leqslant D(z,\hat{x}_{k(j)}) - \frac{\zeta}{2}\tau_{i(j)}\mu_{i(j)}, \forall j \geqslant q.$$

对上式求和有 $0 \leqslant D(z,\hat{x}_{k(j+1)}) \leqslant D(z,\hat{x}_{k(q)}) - \frac{\zeta}{2}\sum_{p=q}^{j}\tau_{i(p)}\mu_{i(p)}, \forall j \geqslant q$. 令 $j \to +\infty$,有

$$\sum_{p=q}^{+\infty}\tau_{i(p)}\mu_{i(p)} \leqslant \frac{2D(z,\hat{x}_{k(q)})}{\zeta} < +\infty,$$

与公式(2.11)矛盾,所以 $\{f(\hat{x}_{k(j)})\} \to f^*$,即 $\{f(\hat{x}_k)\} \to f^*$.

设 x^* 是 $\{\hat{x}_k\}$ 的任一聚点,由 X 是非空闭凸集可知 $x^* \in X$,因此 x^* 是问题(0.1)的解.

参考文献:

[1] 袁功林,韦增欣.一个新的BFGS信赖域算法[J].广西科学,2004,11(3):195-196.
YUAN G L,WEI Z X. A new BFGS trust region algorithm[J]. Guangxi Sciences,2004,11(3):195-196.

[2] 张秀军,徐安农.一种新的非线性共轭梯度法的全局收敛性[J].广西科学,2005,12(4):282-283.
ZHANG X J,XU A N. Global convergence properties of a new class of nonlinear conjugate gradient methods[J]. Guangxi Sciences,2005,12(4):282-283.

[3] 韦增欣,谢品杰,顾能柱.一类拟牛顿算法的收敛性[J].广西科学,2006,13(4):282-287.
WEI Z X,XIE P J,GU N Z. Convergence properties of a class of quasi-Newton algorithm[J]. Guangxi Sciences,2006,13(4):282-287.

[4] BAGIROV A M,UGON J,MIRZAYEVA H G. Nonsmooth optimization algorithm for solving clusterwise linear regression problems[J]. Journal of Optimization Theory and Applications,2015,164(3):755-780.

[5] JUNG M,KANG M. Efficient nonsmooth nonconvex optimization for image restoration and segmentation[J]. Journal of Scientific Computing,2015,62(2):336-370.

[6] YU J,VISHWANATHAN S V N,GÜNTER S,et al. A quasi-newton approach to nonsmooth convex pptimization problems in machine learning[J]. Journal of Machine Learning Research,2010,11(2):1145-1200.

[7] MÄKELÄ M. Survey of bundle methods for nonsmooth optimization[J]. Optimization Methods and Software,2002,17(1):1-29.

[8] 唐春明,简金宝.非光滑优化的强次可行方向邻近点束求解方法[J].广西科学,2014,21(3):283-286.
TANG C M,JIAN J B. A Proximal bundle method of strongly sub-feasible directions for nonsmooth optimization[J]. Guangxi Sciences,2014,21(3):283-286.

[9] 唐春明,律金曼.基于非精确数据的非光滑优化强次可行方向法[J].广西科学,2016,23(5):404-408.
TANG C M,LU J M. Strongly sub-feasible direction method with inexact data for nonsmooth optimization[J]. Guangxi Sciences,2016,23(5):404-408.

[10] KIWIEL K C. A proximal bundle method with approximate subgradient linearizations[J]. Siam Journal on Optimization,2006,16(4):1007-1023.

[11] LAN G. Bundle-level type methods uniformly optimal for smooth and nonsmooth convex optimization[J]. Mathematical Programming:Ser A,2015(149):1-45.

[12] KIWIEL K C. An ellipsoid trust region bundle method for nonsmooth convex minimization[J]. Siam Journal on Control and Optimization,1989,27(4):737-757.

[13] DE OLIVEIRA W,SOLODOV M. A doubly stabilized bundle method for nonsmooth convex optimization[J]. Mathematical Programming,2016,156(1/2):125-159.

[14] 沈洁,李娜,田佳茜.双稳定束方法以及收敛性分析[J].沈阳师范大学学报:自然科学版,2015,33(2):177-180.
SHEN J,LI N,TIAN J Q. Doubly stabilized bundle method and its convergence[J]. Journal of Shenyang Normal University:Natural Science Edition,2015,33(2):177-180.

[15] KIWIEL K C. Proximal minimization methods with generalized Bregman functions[J]. SIAM Journal on Control and Optimization,1997,35(4):1142-1168.

[16] KIWIEL K C. A bundle Bregman proximal method for convex nondifferentiable minimization[J]. Mathematical Programming,1999,85(2):241-258.

[17] CENSOR Y,LENT A. An iterative row-action method for interval convex programming[J]. Journal of Optimization Theory and Applications,1981,34(3):321-353.

[18] NESTEROV Y. Introductory lectures on convex optimization:A basic course[M]. New York:Springer US,2004.

(责任编辑:米慧芝)