文章编号:1008-1542(2018)02-0142-07 doi:10.7535/hbkd.2018yx02007

HS

黎 勇,王松华

(百色学院数学与统计学院,广西百色 533000)

摘 要: 为了提高大规模非光滑优化问题的求解效率,克服其他方法存储需求大、算法复杂等缺点,提出求解非光滑优化问题的一种修正 HS 共轭梯度算法。在经典 HS 三项共轭梯度法的基础上提出一种新的搜索方向,并利用 Moreau-Yosida 正则化技术和 Armijo-type 线搜索技术进行设计。新算法满足充分下降条件,搜索方向属于信赖域,在适当条件下证明了新算法全局收敛。初步的数值实验表明新算法在求解非光滑无约束优化问题方面比 LMBM 方法更有效。新算法不仅具有较好的收敛性质,而且数值表现良好,为更加高效地求解非光滑优化问题提供了新的方法。

关键词: 最优化;非光滑优化;共轭梯度法;充分下降条件;信赖域;全局收敛性

中图分类号: O224 **MSC(2010)主题分类:** 49J25 **文献标志码:** A

A modified three-term HS conjugate gradient method for solving nonsmooth minimizations

LI Yong, WANG Songhua

(School of Mathematics and Statistics, Baise University, Baise, Guangxi 533000, China)

Abstract: To improve the efficiency for large-scale nonsmooth optimization problems and overcome the large storage requirements and complex computation of other algorithms, a modified HS conjugate gradient algorithm for nonsmooth optimization problems is proposed. A new search direction based on the classical HS conjugate gradient method is given, then the Moreau-Yosida regularization technique and the Armijo-type line search technique are used to design the algorithm. The sufficient descent condition and the trust region are satisfied for this algorithm. Under suitable conditions, the global convergence of the new algorithm is proved. The preliminary numerical experiments show that the new algorithm is more efficient than the LMBM method for nonsmooth unconstrained optimization problems. The presented algorithm is efficiently for solving nonsmooth optimization problems since it has good convergence property and good numerical performance.

Keywords: optimization; nonsmooth optimization; conjugate gradient method; sufficient descent; trust region; global convergence

:2017-08-21; :2017-12-10; :

: (11661001,11661009); (YB2014389); (KY2016YB417)

: (1973—), , , , 。

E-mail:liyong3922@163.com

, . HS [J]. ,2018,39(2):142-148.

LI Yong, WANG Songhua. A modified three-term HS conjugate gradient method for solving nonsmooth minimizations[J]. Journal of Hebei University of Science and Technology, 2018, 39(2): 142-148.

，、。

，PRP[1-2]，HS[3]，LS[4]，FR[5]。，

，[6-10]。

，WEI [11] WYL ，，

。YAO [12] WYL ，HS ，：

$$\boldsymbol{\beta}_k^{MHS}=\frac{\boldsymbol{g}_k^{\mathrm{T}}(\boldsymbol{g}_k-\frac{\|\boldsymbol{g}_k\|}{\|\boldsymbol{g}_{k-1}\|}\boldsymbol{g}_{k-1})}{(\boldsymbol{g}_k-\boldsymbol{g}_{k-1})^{\mathrm{T}}\boldsymbol{d}_{k-1}}。$$

ZHANG [13] ，PRP ，

：

$$\boldsymbol{d}_k=\begin{cases}-\boldsymbol{g}_k, & \text{if } k=0,\\ -\boldsymbol{g}_k+\boldsymbol{\beta}_k^{PRP}\boldsymbol{d}_{k-1}-\boldsymbol{\theta}_k\boldsymbol{y}_{k-1}, & \text{if } k\geqslant 1,\end{cases}$$

$\boldsymbol{\theta}_k=\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_{k-1}}{\|\boldsymbol{g}_{k-1}\|^2}$，$\boldsymbol{y}_{k-1}=\boldsymbol{g}_k-\boldsymbol{g}_{k-1}$， $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+t_k\boldsymbol{d}_k$，$t_k$ 。

，。

，，、

、。，、、、

、，，

。[14] [15] Polak-Ribière-Polyak Hestenes-Stiefel

，Moreau-Yosida ，Armijo-type ，

。[16] Moreau-Yosida 、

LS WYL 。[12] [13] ，[14—16]

Hestenes-Stiefel ，。

1 预备知识

，。

：

$$\min\{f(x)\mid x\in\mathbf{R}^n\},\tag{1}$$

$f:\mathbf{R}^n\rightarrow\mathbf{R}$ 。

f Moreau-Yosida $F:\mathbf{R}^n\rightarrow\mathbf{R}$，

$$F(x)=\min_{z\in\mathbf{R}^n}\{f(z)+\frac{1}{2\mu}\|z-x\|^2\},\tag{2}$$

$\|\cdot\|$ ，$\mu>0$。$Q(z,x)=f(z)+\frac{1}{2\mu}\|z-x\|^2$，$h(x)=\arg\min_z Q(z,x)$，

$Q(z,x)$ x ，$h(x)$ 。$F(x)=f(h(x))+\frac{1}{2\mu}\|h(x)-x\|^2$。

[17—19] $F(x)$ ：

i)$F(x)$ ，。：

$$g(x)=\nabla F(x)=\frac{x-h(x)}{\mu},\tag{3}$$

$g(x)$ Lipschitz ，

$$\|g(x)-g(y)\|\leqslant\frac{1}{\mu}\|x-y\|,\quad\forall x,y\in\mathbf{R}^n。\tag{4}$$

ii)x (1) ，$\nabla F(x)=0$ ，$h(x)=x$。

，$\arg\min_z Q(z,x)$ ，$F(x)$ $g(x)$ 。

$\arg\min_z Q(z,x)=h(x)$ ，。

$h(x)$ $F(x)$ $g(x)$。 [20] 。

$\forall x\in\mathbf{R}^n,\varepsilon>0$， $\boldsymbol{h}^\alpha(x,\varepsilon)\in\mathbf{R}^n$， ：

$$f(\boldsymbol{h}^\alpha(x,\varepsilon))+\frac{1}{2\mu}\|\boldsymbol{h}^\alpha(x,\varepsilon)-x\|^2\leqslant F(x)+\varepsilon, \tag{5}$$

ε $\boldsymbol{h}^\alpha(x,\varepsilon)$ $F(x)$ $g(x)$ ：

$$\boldsymbol{F}^\alpha(x,\varepsilon)=f(\boldsymbol{h}^\alpha(x,\varepsilon))+\frac{1}{2\mu}\|\boldsymbol{h}^\alpha(x,\varepsilon)-x\|^2, \tag{6}$$

$$\boldsymbol{g}^\alpha(x,\varepsilon)=\frac{x-\boldsymbol{h}^\alpha(x,\varepsilon)}{\mu}。 \tag{7}$$

[20] ，$\boldsymbol{F}^\alpha(x,\varepsilon)$ $\boldsymbol{g}^\alpha(x,\varepsilon)$ 。

命题 1 (5)— (7) ， ：

$$F(x)\leqslant\boldsymbol{F}^\alpha(x,\varepsilon)\leqslant F(x)+\varepsilon, \tag{8}$$

$$\|\boldsymbol{h}^\alpha(x,\varepsilon)-h(x)\|\leqslant\sqrt{2\mu\varepsilon}, \tag{9}$$

$$\|\boldsymbol{g}^\alpha(x,\varepsilon)-g(x)\|\leqslant\sqrt{2\varepsilon/\mu}。 \tag{10}$$

， ε ，$\boldsymbol{F}^\alpha(x,\varepsilon)$ $\boldsymbol{g}^\alpha(x,\varepsilon)$ $F(x)$ $g(x)$。 [20]。

2 修正 HS 三项共轭梯度法

[12—15] ， (1) HS ，

：

$$\boldsymbol{d}_{k+1}=\begin{cases}-\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})+\dfrac{\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})^{\mathrm{T}}\boldsymbol{y}_k^*\boldsymbol{d}_k-\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\boldsymbol{y}_k^*}{\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\}}, & \text{if } k\geqslant 1,\\ -\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1}), & \text{if } k=0,\end{cases} \tag{11}$$

：$\boldsymbol{d}_k$ ， $\boldsymbol{y}_k^*=\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})-\dfrac{\|\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|}{\|\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)\|}\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)$，$\boldsymbol{y}_k=\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})-\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)$， $c>0$。 。

1

Step1 $x_0\in\mathbf{R}^n$， $\sigma\in(0,1),c>0,s>0,\mu>0,\boldsymbol{d}_0=-\boldsymbol{g}^\alpha(x_0,\varepsilon_0),\gamma\in[0,1]$， $k=0$。

Step2 $\|\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)\|<\gamma$， ， 。

Step3 ε_{k+1}， $0<\varepsilon_{k+1}<\varepsilon_k$， Armijo-type [21] t_k：

$$\boldsymbol{F}^\alpha(\boldsymbol{x}_k+t_k\boldsymbol{d}_k,\varepsilon_{k+1})-\boldsymbol{F}^\alpha(\boldsymbol{x}_k,\varepsilon_k)\leqslant\sigma t_k\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k, \tag{12}$$

$t_k=s2^{-i_k},i_k\in\{0,1,2,\cdots\}$。

Step4 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+t_k\boldsymbol{d}_k$， $\|\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|<\gamma$， ； ， 。

Step5 (11) $\boldsymbol{d}_{k+1}$。

Step6 $k:=k+1$， Step2。

引理 1 $k\in\mathbf{N}\cup\{0\}$， ：

$$\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k=-\|\boldsymbol{g}^\alpha(\boldsymbol{x}_k,\varepsilon_k)\|^2。 \tag{13}$$

证明

$k=0$ ，$\boldsymbol{d}_0=-\boldsymbol{g}^\alpha(x_0,\varepsilon_0)$， 。

$k\geqslant1$ ， (11) $\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})$ ：

$$\begin{aligned}\boldsymbol{d}_{k+1}^{\mathrm{T}}\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})=&-\|\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|^2+\\&\left[\frac{\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})^{\mathrm{T}}\boldsymbol{y}_k^*\boldsymbol{d}_k-\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\boldsymbol{y}_k^*}{\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\}}\right]^{\mathrm{T}}\boldsymbol{g}^\alpha(\boldsymbol{x}_{k+1},\varepsilon_{k+1})=\\&-\|\boldsymbol{g}^\alpha(x_{k+1},\varepsilon_{k+1})\|^2。\end{aligned}$$

。

， $k\in\mathbf{N}\cup\{0\}$ 。

引理 2 $\forall k \in \mathbf{N} \cup \{0\}$， ：

$$\|\boldsymbol{d}_k\| \leqslant \left(1+\frac{1}{c}\right)\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\|。\tag{14}$$

证明 $\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\} \geqslant 2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|$， (11) ：

$$\begin{aligned}
\|\boldsymbol{d}_{k+1}\| = & \left\| -\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1}) + \frac{\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})^{\mathrm{T}}\boldsymbol{y}_k^*\boldsymbol{d}_k - \boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\boldsymbol{y}_k^*}{\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\}} \right\| \leqslant \\
& \|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\| + \left\| \frac{\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})^{\mathrm{T}}\boldsymbol{y}_k^*\boldsymbol{d}_k - \boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\boldsymbol{y}_k^*}{\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\}} \right\| \leqslant \\
& \|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\| + \frac{\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|\|\boldsymbol{y}_k^*\|\|\boldsymbol{d}_k\| + \|\boldsymbol{d}_k\|\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|\|\boldsymbol{y}_k^*\|}{\max\{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|,|\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{y}_k|\}} \leqslant \\
& \|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\| + \frac{\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|\|\boldsymbol{y}_k^*\|\|\boldsymbol{d}_k\| + \|\boldsymbol{d}_k\|\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|\|\boldsymbol{y}_k^*\|}{2c\|\boldsymbol{d}_k\|\|\boldsymbol{y}_k^*\|} \leqslant \\
& \left(1+\frac{1}{c}\right)\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})\|。
\end{aligned}$$

1 。 2

， 。

3 全局收敛性

1 。

A i) $\forall \xi_k \in [\boldsymbol{x}_k,\boldsymbol{x}_{k+1}]$ $\forall k \in \mathbf{N}$， ρ， ：

$$\|\nabla^2 F(\xi_k)\| \leqslant \rho,\tag{15}$$

F f Moreau-Yosida 。

ii) F 。

iii) $\{\varepsilon_k\}$ 0。

引理 3 $\{\boldsymbol{x}_k\}$ 1 ， A ， $\varepsilon_k = o(t_k^2\|\boldsymbol{d}_k\|^2)$， k， l，

$$t_k \geqslant l。\tag{16}$$

证明 。

， $t'_k = \frac{t_k}{2}$， (12) ， ：

$$\boldsymbol{F}^{\alpha}(\boldsymbol{x}_k + t'_k\boldsymbol{d}_k,\varepsilon_{k+1}) - \boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k) > \sigma t'_k\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k。$$

(8) A， ， ：

$$\begin{aligned}
\sigma t'_k\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k < & \boldsymbol{F}^{\alpha}(\boldsymbol{x}_k + t'_k\boldsymbol{d}_k,\varepsilon_{k+1}) - \boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k) \leqslant \\
& F(\boldsymbol{x}_k + t'_k\boldsymbol{d}_k) + \varepsilon_{k+1} - \boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k) \leqslant \\
& F(\boldsymbol{x}_k + t'_k\boldsymbol{d}_k) - F(\boldsymbol{x}_k) + \varepsilon_{k+1} = \\
& F(\boldsymbol{x}_k) + t'_k\boldsymbol{d}_k^{\mathrm{T}}g(\boldsymbol{x}_k) + \frac{\nabla F^2(\xi_k)(t'_k\boldsymbol{d}_k)^2}{2} - F(\boldsymbol{x}_k) + \varepsilon_{k+1} = \\
& t'_k\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}_k) + \frac{(t'_k)^2\boldsymbol{d}_k^{\mathrm{T}}\ \nabla F^2(\xi_k)\boldsymbol{d}_k}{2} + \varepsilon_{k+1} \leqslant \\
& t'_k\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}_k) + \frac{\rho}{2}(t'_k)^2\|\boldsymbol{d}_k\|^2 + \varepsilon_{k+1},
\end{aligned}\tag{17}$$

$\xi_k = \boldsymbol{x}_k + at'_k\boldsymbol{d}_k$, $a \in (0,1)$。 (17) ， (10)、 (13) (14) $\varepsilon_{k+1} \leqslant \varepsilon_k$， $\varepsilon_k = o(t_k^2\|\boldsymbol{d}_k\|^2)$， ：

$$\frac{t_k}{2}=t'_k \geqslant \frac{\sigma \boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k-\boldsymbol{d}_k^{\mathrm{T}}g(\boldsymbol{x}_k)-\frac{\varepsilon_{k+1}}{t'_k}}{\frac{\rho}{2}\|\boldsymbol{d}_k\|^2} >$$

$$\left[\frac{(\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)-g(\boldsymbol{x}_k))^{\mathrm{T}}\boldsymbol{d}_k-(1-\sigma)\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k-\frac{\varepsilon_{k+1}}{t'_k}}{\|\boldsymbol{d}_k\|^2}\right]\frac{2}{\rho} \geqslant$$

$$\left[\frac{(1-\sigma)\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\|^2-\sqrt{\frac{2\varepsilon_k}{\mu}}\|\boldsymbol{d}_k\|-\frac{\varepsilon_k}{t'_k}}{\|\boldsymbol{d}_k\|^2}\right]\frac{2}{\rho}=$$

$$\left[\frac{(1-\sigma)\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\|^2}{\|\boldsymbol{d}_k\|^2}-\frac{o(t_k)}{\sqrt{\mu}}-o(t_k)\right]\frac{2}{\rho},$$

t_k, :

$$\frac{1}{2} \geqslant \lim_{k\to\infty}\left(\frac{2(1-\sigma)}{\rho(1+\frac{1}{c})^2}-o(t_k)\right)\frac{1}{t_k} \geqslant +\infty,$$

, 。

定理 1 $\{x_k\},\{t_k\}$ 1 , A , $\varepsilon_k=o(t_k^2\|\boldsymbol{d}_k\|^2)$, $\lim\limits_{k\to\infty}\|\boldsymbol{g}(\boldsymbol{x}_k)\|=0$, $\{\boldsymbol{x}_k\}$ (1) 。

证明 $\lim\limits_{k\to\infty}\|g(\boldsymbol{x}_k)\|=0$,

$$\lim_{k\to\infty}\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\|=0 \tag{18}$$

, 。 (18) , η_0 k_0, $\forall k>k_0$, :

$$\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\| \geqslant \eta_0。 \tag{19}$$

(12) , (13)、(16)、(19), :

$$\boldsymbol{F}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})-\boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k) \leqslant \sigma t_k \boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)^{\mathrm{T}}\boldsymbol{d}_k=-\sigma t_k\|\boldsymbol{g}^{\alpha}(x_k,\varepsilon_k)\|^2 \leqslant -\sigma l\eta_0,\forall k>k_0,$$

: $\sum\limits_{k>k_0}(\boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)-\boldsymbol{F}^{\alpha}(\boldsymbol{x}_{k+1},\varepsilon_{k+1})) \geqslant \sum\limits_{k>k_0}\sigma l\eta_0$, , $k\to\infty$, $\boldsymbol{F}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)\to\infty$, A ii) 。 (18) 。

2 。

(10) :

$$\|\boldsymbol{g}^{\alpha}(\boldsymbol{x}_k,\varepsilon_k)-g(\boldsymbol{x}_k)\| \leqslant \sqrt{\frac{2\varepsilon_k}{\mu}},$$

A iii), :

$$\lim_{k\to\infty}\|g(\boldsymbol{x}_k)\|=0。 \tag{20}$$

x^* $\{\boldsymbol{x}_k\}$, $\{\boldsymbol{x}_k\}_K$:

$$\lim_{k\in K,k\to\infty}\boldsymbol{x}_k=\boldsymbol{x}^*。 \tag{21}$$

(3) $F(x)$ $g(\boldsymbol{x}_k)=\frac{\boldsymbol{x}_k-h(x_k)}{\mu}$, (20) (21) $\boldsymbol{x}^*=h(\boldsymbol{x}^*)$, $\boldsymbol{x}^*$ (1) 。

4 数值结果

。

Windows7+Fortran90, 2.0 GB, $s=\mu=1,\sigma=0.8,\varepsilon_k=1/(k+2)^2$; $\|\boldsymbol{g}^{\alpha}(\boldsymbol{x},\varepsilon)\| \leqslant 10^{-15}$ $N_i>10^4$。 [12], Fortran [12] , [22] LMBM(limited memory bundle method) N_i,N_f $f(x)$ 。 N_i ,N_f

, $f(x)$ 。

1 ,2 2 , Problem , x_0 ,Dim 。

表 1 测试问题

Tab.1 Test problems

No.	Problem	x_0
1	Generalization of MAXQ	$(1,2,\cdots,n/2,\cdots-(n/2+1),\cdots,-n)$
2	Generalization of MXHILB	$(1,1,\cdots)$
3	Chained LQ	$(-0.5,-0.5,\cdots)$
4	Number of active faces	$(1,1,\cdots)$
5	Nonsmooth generalization of Brown function 2	$(1,0,\cdots)$
6	Chained Mifflin2	$(-1,-1,\cdots)$

表 2 数值结果

Tab.2 Numerical results

No.	Dim	MHS $N_i/N_f/f(x)$	LMBM $N_i/N_f/f(x)$
1	1 000	133/2 735/2.760×10^{-7}	21 492/22 259/6.710×10^{-6}
	3 000	149/2 990/1.485×10^{-6}	94 144/96 680/1.134×10^{-5}
	5 000	167/3 381/3.107×10^{-7}	191 470/196 034/3.450×10^{-5}
	10 000	152/3 060/ 2.021×10^{-5}	512 415/523 351/5.835×10^{-5}
2	1 000	67/1 086/4.280×10^{-2}	441/861/6.166×10^{-3}
	3 000	88/1 392/3.013 5×10^{-7}	209/579/5.872×10^{-2}
	5 000	104/1 564/1.238×10^{-11}	1 258/2 487/3.521×10^{-2}
	10 000	107/1 681/1 .797×10^{-12}	7 027/7 810/5.116×10^{-2}
3	1 000	4/30/1.129×10^{-8}	300/1 824/−1.413×10^{5}
	3 000	7/51/−1.096×10^{-2}	275/1 373/−4.241×10^{5}
	5 000	2/11/1.201×10^{-4}	365/2 198/−7.070×10^{5}
	10 000	3/152/ −9.655×10^{-3}	376/2 281/−1.414×10^{6}
4	1 000	35/583/2.458×10^{-4}	523/569/1.377×10^{-14}
	3 000	51/778/4.076×10^{-8}	1 576/1 577/1.555×10^{-10}
	5 000	50/782/1.599×10^{-3}	2 585/2 586/1.213×10^{-10}
	10 000	59/945/9 .789×10^{-9}	5 069/5 073/5.384×10^{-10}
5	1 000	4/30/2.258×10^{-8}	467/3 873/4.058×10^{-9}
	3 000	7/40/2.118×10^{-7}	542/5 245/6.227×10^{-8}
	5 000	2/12/5.273×10^{-9}	453/4 073/1.080×10^{-8}
	10 000	3/20/1 .083×10^{-10}	736/7 453/2.522×10^{-8}
6	1 000	4/30/−2.498×10^{4}	1 254/7 355/−7.065×10^{4}
	3 000	7/51/−7.498×10^{4}	411/2 353/−2.121×10^{5}
	5 000	2/11/−1.250×10^{5}	219 /7 382/−3.535×10^{5}
	10 000	3/125/ −2.500×10^{5}	267/743/ −7.070×10^{5}

6 4 , 2 , LMBM 。 , , , 。

5 结 论

, , ,HS 。 , HS , Moreau-Yosida [21] Armijo-type , HS 。 , 。 , 。 。

，，，

，、、、、

，。

参考文献/References：

[1] POLAK E，RIBIERE G. Note surla convergence de method de directions conjugees[J]. Rev Francaise informat Recherche Opèrationelle，3e Annèe，1969，16(16)：35-43.

[2] POLYAK B T. The conjugate gradient method in extremal problems[J]. Ussr Computational Mathematics and Mathematical Physics，1969，9(4)：94-112.

[3] HESTENES M R，STIEFEL E. Method of conjugate gradient for solving linear systems[J]. Journal of Research of the National Bureau of Standards，1952，49(6)：409-436.

[4] LIU Y，STOREY C. Efficient generalized conjugate gradient algorithms，part 1：Theory[J]. Journal of Optimization Theory and Applications，1991，69(1)：129-137.

[5] FLETCHER R，REEVES C. Function minimization by conjugate gradients[J]. Computer Journal，1964，7(2)：149-154.

[6] GILBERT J C，NOCEDAL J. Global convergence properties of conjugate gradient methods for optimization[J]. Siam Journal on Optimization，1992，2(1)：21-42.

[7] HU Y F，STOREY C. Global convergence result for conjugate gradient methods[J]. Journal of Optimization Theory & Applications，1991，71(2)：399-405.

[8] ZENG X W，GUO Y L，LI Q Q. Global convergence of the Polak-Ribière-Polyak conjugate gradient methods with inexact line search for nonconvex unconstrained optimization problems[J]. Mathematics of Computation，2008，77(264)：2173-2193.

[9] TOUATI-AHMED D，STOREY C. Efficient hybrid conjugate gradient techniques[J]. Journal of Optimization Theory & Applications，1990，64 (2)：379-397.

[10] ALBAALI M. Descent property and global convergence of the Flecher-Reves method with inexact line search[J]. Ima Journal of Numerical Analysis，1985，5(1)：121-124.

[11] WEI Z X，YAO S W，LIU L Y. The convergence properties of some new conjugate gradient methods[J]. Applied Mathematics and Computation，2006，183(2)：1341-1350.

[12] YAO S W，WEI Z X，HUANG H. A note about WYL's conjugate gradient method and its applications[J]. Applied Mathematics and Computation，2007，191(2)：381-388.

[13] ZHANG L，ZHOU W J，LI D H. A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence[J]. Ima Journal of Numerical Analysis，2006，26(4)：629-640.

[14] YUAN G L，WEI Z X，LI G Y. A modified Polak-Ribière-Polyak conjugate gradient algorithm for nonsmooth covex programs[J]. Journal of Computational and Applied Mathematics，2014，255(1)：86-96.

[15] YUAN G L，MENG Z H，LI Y. A modified Hestenes and Stiefel conjugate gradient algorithm for large-scale nonsmooth minimizations and nonlinear equations[J]. Journal of Optimization Theory and Applications，2016，168(1)：129-152.

[16] ． [D]． ： ,2014.
HU Yaping. Numerical Methods for Nonlinear Monotone Equations and Nonsmooth Optimization Problems[D]. Shanghai：East China University of Science and Technology，2014.

[17] BONNAN J F，GILBERT J C，LEMARECHAL C，et al. A family of variable metric proximal methods[J]. Mathematical Programming，1995，68(1)：15-47.

[18] CORREA R，LEMARÉCHAL C. Convergence of some algorithms for convex minimization[J]. Mathematical Programming，1993，62(1/2/3)：261-273.

[19] HIRIART-URRUTY J B，LEMARECHAL C. Convex Analysis and Minimization Algorithms Ⅱ[M]. Berlin：Springer，1993.

[20] FUKUSHIMA M，QI L. A globally and superlinearly convergent algorithm for nonsmooth convex minimization[J]. Siam Journal on Optimization，1996，6(4)：1106-1120.

[21] ZHANG H C，HARGER W W. A nonmonotone line search technique and its application to unconstrained optimization[J]. Siam Journal on Optimization，2006，14(4)：1043-1056.

[22] HAARALA M，MIETTINEN K，MÄKELÄ M M. New limited memory bundle method for large-scale nonsmooth optimization[J]. Optimization Methods & Software，2004，19(6)：673-692.