数学年刊 A 辑
2017, 38(4):391–404
DOI: 10.16205/j.cnki.cama.2017.0032

非单调滤子曲率线搜索算法解无约束非凸优化*

顾　超[1]　朱德通[2]

提要　宇和濮在文 [Yu Z S, Pu D G. A new nonmonotone line search technique for unconstrained optimization [J]. *J Comput Appl Math*, 2008, 219:134–144] 中提出了一种非单调的线搜索算法解无约束优化问题. 和他们的工作不同, 当优化问题非凸时, 本文给出了一种非单调滤子曲率线搜索算法. 通过使用海森矩阵的负曲率信息, 算法产生的迭代序列被证明收敛于一个满足二阶充分性条件的点. 在不需要假设极限点存在的情况下, 证明了算法具有整体收敛性. 而且分析了该算法的收敛速率. 数值试验表明算法的有效性.

关键词　非凸优化, 非单调, 曲率线搜索, 滤子, 收敛性
MR (2000) 主题分类　90C30, 65K05
中图法分类　O221.2
文献标志码　A
文章编号　1000-8314(2017)04-0391-14

1　引　言

考虑如下无约束优化问题:

$$\min \quad f(x) \tag{1.1}$$

其中 $x \in \mathbb{R}^n$　$f(x) \in \mathbb{R}$ 是一个实值函数.

有很多方法可以解这类问题, 它们通常是下降方法, 通过产生一个迭代序列 $\{x_k\}$, 使得每个极限点 x_* 是稳定点, 即 $g(x_*) = 0$ 或者 $\liminf\limits_{k\to\infty} \|g(x_k)\| = 0$, 其中 $g(x)$ 是目标函数 $f(x)$ 的梯度. 然而, 当海森矩阵及其近似不是正定的, 线搜索方法会出现一些问题. 幸运的是, McCormick[1] 修改了 Armijo 准则并提出一种二阶 Armijo 步长准则, 在线搜索中包含了二阶导数信息. 运用负曲率作为搜索方向, 该方法能解决海森矩阵不正定的情况, 使得算法产生的迭代序列收敛到满足二阶条件的稳定点. 传统的二阶线搜索方法需要目标函数单调下降, 从而保持算法的整体收敛性[2−3].

最近的研究表明, 当迭代点到达狭长弧形谷底, 单调的线搜索技术会降低收敛速度. Grippo 等人[4]首先推广了 Armijo 准则并提出了一种非单调线搜索技术, 该技术允许函数值上升, 并且保持算法的整体收敛性. Dai[5]和 Mascarenhas[6] 证明了对于非凸无约束优化, 标准的 BFGS 方法结合传统的非单调线搜索无法保证算法的整体收敛性. 非单调技术[4]已经与曲率搜索和二阶线搜索结合, 用以解决非凸的无约束优化问题[7−11].

Fletcher 和 Leyffer[12]提出了滤子方法解非线性规划, 代替传统的罚函数方法来保证优化算法的整体收敛性. 其主要思想是: 当目标函数或约束违反度被改进时, 就接受

本文 2015 年 9 月 29 日收到, 2016 年 9 月 6 日收到修改稿.

[1]上海立信会计金融学院统计与数学学院, 上海 201620. E-mail: guchao@lixin.edu.cn

[2]上海师范大学数学系, 上海 200234. E-mail: dtzhu@shnu.edu.cn

*本文受到国家自然科学基金 (No. 11201304, No. 11371253), 上海市自然科学基金和上海立信会计金融学院统计学一级学科项目的资助.

试探点. 滤子方法已经被成功地应用到各种优化问题当中, 包括互补性和变分不等式问题[13−20].

基于以上的想法, 我们提出了非单调滤子曲率线搜索算法解无约束非凸优化问题. 我们的方法有以下几个好的性质:

(1) 发展了宇和濮的非单调技术[21]和负曲率方向解非凸问题.

(2) 运用滤子思想, 使得算法更容易接受试探点.

(3) 与以往的算法[8, 22]相比, 在不需要假设极限点存在的情况下, 新算法具有整体收敛性.

文章结构如下. 第 2 章叙述新的二阶线搜索算法; 第 3 章证明算法的整体收敛性; 第 4 章建立算法的局部超线性收敛速率; 第 5 章给出数值实验结果; 第 6 章总结本文的主要工作和结果.

2 非单调滤子曲率线搜索算法

考虑如下迭代格式:

$$x_{k+1} = x_k + \ {}_k^2 s_k + \ {}_k d_k$$

其中 s_k 和 d_k 是搜索方向, ${}_k$ 是步长.

定义 2.1 如果 $H(x)$ 至少有一个非负特征值, 那么点 x 是一个不定点. 进一步, 如果 x 是一个不定点, 那么, 当 $d^{\mathrm{T}}H(x)d \quad 0$ 时, d 是一个负曲率方向.

定义 2.2 如果 $s^{\mathrm{T}}g(x) \leqslant 0 \ \ d^{\mathrm{T}}g(x) \leqslant 0 \ \ d^{\mathrm{T}}H(x)d \quad 0$, 那么 $(s \ \ d)$ 在不定点 x 处被称为是一个下降对; 如果 x 不是一个不定点, 同时, $s^{\mathrm{T}}g(x) \quad 0 \ \ d^{\mathrm{T}}g(x) \leqslant 0 \ \ d^{\mathrm{T}}H(x)d = 0$, 那么 $(s \ \ d)$ 被称为零曲率方向上的下降对.

定义 $y_k(\ ) = x_k + s_k 2^{-i} + d_k 2^{-\frac{i}{2}}$. MoCormick 的二阶 Armijo 步长准则试图从 $0 \ \ 1 \ \ \cdots$ 中找到最小非负整数 $(\ )$, 满足

$$f(y_k(\ )) \leqslant f(x_k) + \ \ 2^{-i}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k H_k d_k\Big) \tag{2.1}$$

其中 $0 \qquad 1$ 是一个预设的常数. 无论 H_k 是否不定, 我们可以在每一次迭代中利用准则 (2.1), 因为当 H_k 是正定时, 我们能够令 d_k 是一个零向量. 显而易见, 当 H_k 正定时, 二阶步长准则 (2.1) 就是经典的 Armijo 步长准则. 为了使有限整数 $(\ )$ 满足 (2.1), 以下情况是必然的: 当 $g_k \neq 0$ 时,

其中 $0 \quad {}_1 \quad {}_2 \quad 1 \quad \tau(0) = 0 \quad 0 \leqslant \tau(\ ) \leqslant \min\{\tau(\ -1)+1 \ M\}$　M 是一个整数, $\sum_{r=0}^{\tau(k)-1} \ {}_{kr} = 1 \quad {}_{kr} \in (0\ 1)$ 和 ${}_{kr} \geqslant \quad 0$. 宇和濮去掉条件

$$\|s_k\| \leqslant c_1 \|g_k\|$$

其中 $c_1 \quad 0$, 并建立了强收敛性结果

$$\lim_{k\to\infty} \|g_k\| = 0$$

进一步, 水平集是紧的和 $\nabla f(x)$ 的 Lipschitz 连续性这些条件都不需要.

将 (2.2) 推广到非单调曲率线搜索情形

$$f(x_k + \ {}_k^2 s_k + \ {}_k d_k) \leqslant \max\left[f(x_k) \ \sum_{r=0}^{\tau(k)-1} \ {}_{kr} f(x_{k-r})\right] + \ {}_k^2 \ {}_1 \left(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k H_k d_k\right) \tag{2.3}$$

其中 $(s_k \ d_k)$ 是一个下降对. 滤子方法的基本思想可以理解为将一个优化问题看成是一个最小化目标函数 $f(x)$ 和约束违反度 $\ (x)$ ($\ (x) = \|g(x)\| + |\min\{\tilde{\ } \ 0\}|$) 的双目标优化问题, 其中 $\tilde{\ }$ 是 D_k 的最小特征值, 可以由 (2.6) 得到. 在算法开始之初, 初始化滤子为

$$\mathcal{F}_k = \{(\ \ f) \in \mathbb{R}^2 : \ \ = \tilde{\ }\}$$

其中 $\tilde{\ } \in (\ {}_0 \ +\infty)$.

定义 2.3　当

$${}_1 \leqslant \ {}_2 \quad 和 \quad f_1 \leqslant f_2$$

时, 点 x_1 替换点 x_2.

定义 2.4　滤子 $\mathcal{F}$ 是形式为 $(\ {}_i \ f_i)$ 的一个序列, 满足

$${}_i \leqslant \ {}_j \quad 或 \quad f_i \leqslant f_j$$

对 $\ \neq \ $.

与传统的滤子思想不同, 一个试探点

$$x_k(\ {}_{k,l}) = x_k + \ {}_{k,l}^2 s_k + \ {}_{k,l} d_k$$

被非单调滤子接受, 当且仅当

$$(x_k(\ {}_{k,l})) \leqslant (1-\gamma_\theta) \max\left[\ (x_j) \ \sum_{r=0}^{\tau(k)-1} \ {}_{kr} \ (x_{j-r})\right] \tag{2.4a}$$

或

$$f(x_k(\ {}_{k,l})) \leqslant \max\left[f(x_j) \ \sum_{r=0}^{\tau(k)-1} \ {}_{kr} f(x_{j-r})\right] - \gamma_f \ (x_j) \tag{2.4b}$$

对所有的 $(h(x_j) \ f(x_j)) \in \mathcal{F}_k$. 如果试探点被 (2.4) 接受, 则滤子通过以下非单调校正公式进行增广

$$\mathcal{F}_{k+1} = \mathcal{F}_k \cup \{(\ \ f) \in \mathbb{R}^2 : \ \ \geqslant (1-\gamma_\theta)\overline{\ }_k \quad f \geqslant \overline{f}_k - \gamma_f \ (x_k)\} \tag{2.5}$$

其中

$$\bar{\ }_k = \max\Big[\ (x_k)\ \sum_{r=0}^{\tau(k)-1}\ {}_{kr}\ (x_{k-r})\Big] \quad \bar{f}_k = \max\Big[f(x_k)\ \sum_{r=0}^{\tau(k)-1}\ {}_{kr}f(x_{k-r})\Big]$$

如果 (2.3) 对步长成立, 则不增广滤子.

周和孙[11]引入了一种能够稳定分解一般对称矩阵的方法, 该方法最早由 Bunch 和 Parlett[23]提出. 这种方法将海森矩阵 H_k 分解为如下形式:

$$PH_kP = LD_kL^{\mathrm{T}}$$

其中 P 是一个置换阵, L 是一个单位下三角矩阵, D_k 是一个块对角矩阵, 包含 1×1 和 2×2 对角块. 如果 H_k 是正定的, 那么 D_k 是对角的. 这种分解有如下性质.

命题 2.1 D_k 和 H_k 有同样的惯性.

命题 2.2 存在正数 $a_1\ a_2\ a_3$ 和 a_4, 与 H_k 无关, 满足

$$a_1 \leqslant \|L\| \leqslant a_2 \quad a_3 \leqslant \|L^{-1}\| \leqslant a_4$$

命题 2.3 假设 H_k 不是正定的, 令 $\ _k$ 和 $\ _k$ 分别是 H_k 和 D_k 的最大的负特征值, 则

$$_k\|L\|^2 \leqslant \ _k \leqslant \frac{_k}{\|L^{-1}\|^2}$$

基于 Bunch-Parlett 分解, 可以得到 D_k 的谱分解:

$$D_k = \ \Lambda_k\ ^{\mathrm{T}} \tag{2.6}$$

其中 $\Lambda_k = \mathrm{diag}(\ ^k_1\ \ ^k_2\ \cdots\ \ ^k_n)$ 是一个正交矩阵. 那么令

$$\bar{\ }^k_j = \max\{|\ ^k_j|\ \ n \max_{1\leqslant i\leqslant n}|\ ^k_i|\ \ \} \quad \bar{\Lambda}_k = \mathrm{diag}(\bar{\ }^k_1\ \bar{\ }^k_2\ \cdots\ \bar{\ }^k_n)$$

其中 是一个机器精度, 令

$$\bar{D}_k = \ \bar{\Lambda}_k\ ^{\mathrm{T}}$$

方向 s_k 通过以下方程得到:

$$(P^{\mathrm{T}}L\bar{D}_kL^{\mathrm{T}}P)s = -g_k$$

令 z_k 是 D_k 的最小特征值 $\tilde{\ }_k$ 的单位特征向量. 显而易见, 方向

$$t_k = |\min\{\tilde{\ }_k\ 0\}|^{\frac{1}{2}}P^{\mathrm{T}}L^{-\mathrm{T}}z_k$$

满足 $t_k^{\mathrm{T}}H_kt_k = \tilde{\ }_k|\min\{\tilde{\ }_k\ 0\}| \leqslant 0$. 这就说明 t_k 是一个非正曲率方向. 为了构造一个非升方向, 选择

$$d_k = \begin{cases} -t_k & \text{如果 } g_k^{\mathrm{T}}t_k\quad 0 \\ t_k & \text{否则.} \end{cases} \tag{2.7}$$

当且仅当 $\ _k$ 是非负的, 矩阵 H_k 是半正定的, 其原因是 Bunch-Parlett 分解的命题 2.1. 因此, 当 H_k 是半正定的, d_k 由 (2.7) 确定为是一个零向量.

我们正式叙述非单调滤子曲率线搜索算法解无约束非凸优化问题.

算法 2.1 给定: 初始点 x_0; 常数 $\tilde{\ } \in (\ _0\ +\infty)$; $\gamma_\theta\ \gamma_f \in (0\ 1)$; $0\quad\ _1\quad 1$; $0\quad \tau_1 \leqslant \tau_2\quad 1$; $\quad 0$; $M\quad 0$.

(1) 计算 $f(x_0)$　$g(x_0)$　$H(x_0)$, 迭代计数　$\leftarrow 0$.

(2) 如果停止准则成立, 算法终止.

(3) 计算下降对 $(s_k\ \ d_k)$.

(4) 令　$_{k,0} = 1$ 和　$\leftarrow 0$.

(5) 计算 $x_k(\ \ _{k,l}) = x_k + \ ^2_{k,l}s_k + \ \ _{k,l}d_k\ \ f(x_k(\ \ _{k,l}))\ \ g(x_k(\ \ _{k,l}))\ \ \ \ (x_k(\ \ _{k,l}))$.

(6) 如果 $x_k(\ \ _{k,l}) \in \mathcal{F}_k$, 转 (9). 否则, 转 (7).

(7) 如果 $x_k(\ \ _{k,l})$ 满足非单调二阶准则 (2.3), 转 (9).

(8) 选择　$_{k,l+1} \in [\tau_1\ \ _{k,l}\ \tau_2\ \ _{k,l}]$, 令　$\leftarrow\ \ + 1$, 转 (5).

(9) 接受 $x_{k+1} = x_k(\ \ _{k,l})$, 如果 x_{k+1} 被当前滤子接受, 用 (2.5) 增广滤子. 否则, 令 $\mathcal{F}_{k+1} = \mathcal{F}_k$. 计算 $f(x_{k+1})$　$g(x_{k+1})$　$H(x_{k+1})$, 令　$\leftarrow\ \ +1\ \ \ \ (\ \ ) = \min\{\ \ (\ \ -1)+1\ \ M\}$, 转步 (2).

3　整体收敛性

本节中, 将在不需要极限点存在的情况下, 证明算法的整体收敛性. 下面给出整体收敛性证明的必要假设.

假设 G　(G1) 函数 $f(x): \mathbb{R}^n \to \mathbb{R}$ 二次连续可微.

(G2) $f(x)$ 在开集 N 上下有界, 且 N 包含水平集 $L = \{x : f(x) \leqslant f(x_0)\}$, 其中 x_0 是迭代初始点.

引理 3.1 很容易由文 [11] 中的定理得到.

引理 3.1　假设 G 成立, 且下降对 $(s_k\ \ d_k)$ 由 (2.6) 和 (2.7) 得到, 那么

$$d_k^{\mathrm{T}} H_k d_k = \ \tilde{}_k|\min\{\tilde{}_k\ \ 0\}|$$

存在常数 $0\ \ \ \ c_1 \leqslant 1\ \ c_2\ \ \ \ 0$ 和 $c_3\ \ \ \ 0$, 满足

$$-\frac{s_k^{\mathrm{T}} g_k}{\|s_k\|\|g_k\|} \geqslant c_1 \quad \text{当 } g_k \neq 0 \text{ 时}$$

$$c_3\|g_k\| \geqslant \|s_k\| \geqslant c_2\|g_k\|$$

引理 3.2　非单调曲率线搜索过程无矛盾.

证　反证法. 假设内循环有限步没有终止. 在这种情况下, 当前滤子和下降条件将一直拒绝试探点, 导致　$_{k,l} \to 0$. 分两种情况.

情况 1　如果 $x_k(\ \ _{k,l})$ 被非单调滤子接受, 则 $x_{k+1} = x_k(\ \ _{k,l})$, 用 (2.5) 增广滤子. 引理的结论成立.

情况 2　如果 $x_k(\ \ _{k,l})$ 被非单调滤子拒绝, 那么 $x_k(\ \ _{k,l})$ 只能满足非单调条件 (2.3). 由于 $g_k^{\mathrm{T}} d_k \leqslant 0$　$g_k^{\mathrm{T}} s_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k$　0 和 0　　$_1$　1, 有

$$\begin{aligned}
&f(x_k + \ ^2_{k,l}s_k + \ \ _{k,l}d_k) \\
&= f(x_k) + g_k^{\mathrm{T}}(\ ^2_{k,l}s_k + \ \ _{k,l}d_k) + \frac{1}{2}(\ ^2_{k,l}s_k + \ \ _{k,l}d_k)^{\mathrm{T}} H_k(\ ^2_{k,l}s_k + \ \ _{k,l}d_k) + o(\ ^2_{k,l}) \\
&\leqslant f(x_k) + \ ^2_{k,l}g_k^{\mathrm{T}} s_k + \frac{1}{2}\ \ ^2_{k,l}d_k^{\mathrm{T}} H_k d_k + o(\ ^2_{k,l})
\end{aligned}$$

$$\leqslant \max\Big[f(x_k) \quad \sum_{r=0}^{\tau(k)-1} {}_{kr} f(x_{k-r})\Big] + {}^{2}_{k,l} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k H_k d_k\Big) + o({}^{2}_{k,l})$$

那么, 非单调曲率条件 (2.3) 当 充分大时成立. 引理的结论成立.

引理 3.3 令 $\{x_k\}$ 由非单调曲率条件 (2.3) 产生. 那么以下不等式成立:

$$f(x_k) \leqslant f(x_0) + {}_{1}\sum_{r=0}^{k-2} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) + {}_{1} \; {}^{2}_{k-1}\Big(s_{k-1}^{\mathrm{T}} g_{k-1} + \frac{1}{2} d_{k-1}^{\mathrm{T}} H_{k-1} d_{k-1}\Big)$$
$$f(x_0) + {}_{1}\sum_{r=0}^{k-1} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big)$$

证 数学归纳法. 如果 $= 1$, 有

$$f(x_1) \leqslant f(x_0) + {}^{2}_{0} \; {}_{1}\Big(s_0^{\mathrm{T}} g_0 + \frac{1}{2} d_0^{\mathrm{T}} H_0 d_0\Big) \quad f(x_0) + {}^{2}_{0} \; {}_{1}\Big(s_0^{\mathrm{T}} g_0 + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big)$$

假设结论在 时成立.

情况 a 如果 $\max\big[f(x_k) \quad \sum\limits_{r=0}^{\tau(k)-1} {}_{kr} f(x_{k-r})\big] = f(x_k)$, 易知

$$\begin{aligned} f(x_{k+1}) &\leqslant f(x_k) + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \\ &\leqslant f(x_0) + {}_{1}\sum_{r=0}^{k-1} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \\ &\quad f(x_0) + {}_{1}\sum_{r=0}^{k} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) \end{aligned}$$

情况 b 如果 $\max\big[f(x_k) \quad \sum\limits_{r=0}^{\tau(k)-1} {}_{kr} f(x_{k-r})\big] = \sum\limits_{r=0}^{\tau(k)-1} {}_{kr} f(x_{k-r})$, 令 $u = \tau(\quad) - 1$. 由 $\sum\limits_{t=0}^{u} {}_{kt} = 1 \quad \leqslant {}_{kt} \quad 1$, 可得

$$\begin{aligned} f(x_{k+1}) &\leqslant \sum_{t=0}^{u} {}_{kt} f(x_{k-t}) + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \\ &\leqslant \sum_{t=0}^{u} {}_{kt}\Big[f(x_0) + {}_{1}\sum_{r=0}^{k-t-2} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) + {}_{1} \; {}^{2}_{k-t-1}\Big(s_{k-t-1}^{\mathrm{T}} g_{k-t-1} \\ &\quad + \frac{1}{2} d_{k-t-1}^{\mathrm{T}} H_{k-t-1} d_{k-t-1}\Big)\Big] + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \\ &= f(x_0) + {}_{1}\sum_{r=0}^{k-t-2}\Big(\sum_{t=0}^{u} {}_{kt}\Big) {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) + {}_{1}\sum_{t=0}^{u} {}_{kt} \; {}^{2}_{k-t-1} \\ &\quad \cdot\Big(s_{k-t-1}^{\mathrm{T}} g_{k-t-1} + \frac{1}{2} d_{k-t-1}^{\mathrm{T}} H_{k-t-1} d_{k-t-1}\Big) + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \\ &\leqslant f(x_0) + {}_{1}\sum_{r=K}^{k-t-2} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) + {}_{1}\sum_{r=k-t-1}^{k-1} {}^{2}_{r}\Big(s_r^{\mathrm{T}} g_r + \frac{1}{2} d_r^{\mathrm{T}} H_r d_r\Big) \\ &\quad + {}^{2}_{k} \; {}_{1}\Big(s_k^{\mathrm{T}} g_k + \frac{1}{2} d_k^{\mathrm{T}} H_k d_k\Big) \end{aligned}$$

$$= f(x_0) + \quad {}_1\sum_{r=0}^{k-1} \ {}^2_r\Big(s_r^{\mathrm T} g_r + \frac{1}{2} d_r^{\mathrm T} H_r d_r\Big) + \ {}^2_k \ {}_1\Big(s_k^{\mathrm T} g_k + \frac{1}{2} d_k^{\mathrm T} H_k d_k\Big)$$

$$f(x_0) + \quad {}_1\sum_{r=0}^{k} \ {}^2_r\Big(s_r^{\mathrm T} g_r + \frac{1}{2} d_r^{\mathrm T} H_r d_r\Big)$$

定理 3.1　如果滤子增广有限次, 那么, 迭代序列的每一个极限点 x_* 都是一个二阶稳定点, 也就是说, $g(x_*) = 0$ 和海森矩阵 $H(x_*)$ 至少是半正定的.

证　由假设知, 存在 $K \in \mathbb{N}$, 对所有的迭代 $\quad \geqslant K$, 滤子不增广. 我们得到, 对所有的 $\ \geqslant K$, 迭代点 $\{x_k\}$ 由非单调曲率条件 (2.3) 产生. 与引理 3.3 的证明一样, 我们得到

$$f(x_{k+1}) \leqslant \overline{f}_K + \quad {}_1\sum_{r=K}^{k} \ {}^2_r\Big(s_r^{\mathrm T} g_r + \frac{1}{2} d_r^{\mathrm T} H_r d_r\Big)$$

从引理 3.1 知道

$$f(x_{k+1}) \leqslant \overline{f}_K + \quad {}_1\sum_{r=K}^{k} \ {}^2_r\Big(-c_1c_2\|g_r\|^2 + \frac{1}{2}\tilde{\ }_r|\min\{\tilde{\ }_r \ \ 0\}|\Big)$$

由于 $\{f_k\}$ 是下有界的, 当 $\quad \to \infty$ 时, 有

$$\sum_{r=K}^{k} \ {}^2_r\Big(-c_1c_2\|g_r\|^2 + \frac{1}{2}\tilde{\ }_r|\min\{\tilde{\ }_r \ \ 0\}|\Big) \qquad \infty$$

根据 $\ {}_r \quad 0 \quad -c_1c_2\|g_r\|^2 \leqslant 0$ 和 $\tilde{\ }_r|\min\{\tilde{\ }_r \ \ 0\}| \leqslant 0$, 可得

$$\lim_{k\to\infty} \ {}^2_k\|g_k\|^2 = 0 \qquad \lim_{k\to\infty} \ {}^2_k\tilde{\ }_k|\min\{\tilde{\ }_k \ \ 0\}| = 0$$

表明

$$\lim_{k\to\infty} \ {}_k\|g_k\| = 0 \qquad \lim_{k\to\infty} \ {}^2_k\tilde{\ }_k|\min\{\tilde{\ }_k \ \ 0\}| = 0$$

有两种情况需要考虑.

情况 1　如果 $\{\ _k\}$ 是一致下有界, 则

$$\lim_{k\to\infty} \|g_k\| = 0 \qquad \lim_{k\to\infty} \tilde{\ }_k|\min\{\tilde{\ }_k \ \ 0\}| = 0 \tag{3.1}$$

这意味着, 当 $\quad \to \infty$ 时, $\tilde{\ }_k \to 0$, 或者对充分大的 $\ $, 有 $\tilde{\ }_k \quad 0$. 下面用反证法. 假设存在一个无限子序列 $\{\ _i\}$, 对某个 $\qquad 0$, 满足

$$\tilde{\ }_{k_i} \leqslant -$$

因而

$$\lim_{i\to\infty} \tilde{\ }_{k_i}|\min\{\tilde{\ }_{k_i} \ \ 0\}| = \lim_{i\to\infty} -\tilde{\ }^2_{k_i} \leqslant - \ ^2$$

与 (3.1) 矛盾. 根据命题 2.1, 可得

$$\lim_{k\to\infty} \ {}_{\min}(H_k) \geqslant 0 \tag{3.2}$$

因此, $H(x_*)$ 至少是半正定的.

情况 2　如果 $\{\ _k\}$ 不是一致下有界, 存在一个子序列 $\{\ _{k_i}\}$, 满足

$$\lim_{i\to\infty} \ {}_{k_i} = 0$$

因为 ${}_{k_i}$ 的定义, 最后被拒绝的步长

$$ {}_{k_i,l_i} \in \Big[\frac{{}_{k_i}}{\tau_2}\ \ \frac{{}_{k_i}}{\tau_1}\Big] $$

满足

$$ \begin{aligned} f(x_{k_i} + {}^2_{k_i,l_i} s_{k_i} + {}_{k_i,l_i} d_{k_i}) \quad & \max\Big[f(x_{k_i}) \ \sum_{r=0}^{\tau(k_i)-1} {}_{kr} f(x_{k_i - r})\Big] \\ & + {}^2_{k_i,l_i}\ {}_1\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i} H_{k_i} d_{k_i}\Big) \\ & \geqslant f(x_{k_i}) + {}^2_{k_i,l_i}\ {}_1\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i} H_{k_i} d_{k_i}\Big) \end{aligned} \tag{3.3} $$

展开 (3.3) 的左边, 有

$$ \begin{aligned} & f(x_{k_i} + {}^2_{k_i,l_i} s_{k_i} + {}_{k_i,l_i} d_{k_i}) \\ = {} & f(x_{k_i}) + {}^2_{k_i,l_i} s_{k_i}^{\mathrm{T}} g_{k_i} + {}_{k_i,l_i} d_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2}({}^2_{k_i,l_i} s_{k_i} + {}_{k_i,l_i} d_{k_i})^{\mathrm{T}} \\ & \cdot H_{k_i}({}^2_{k_i,l_i} s_{k_i} + {}_{k_i,l_i} d_{k_i}) + o({}^2_{k_i,l_i}) \\ = {} & f(x_{k_i}) + {}^2_{k_i,l_i} s_{k_i}^{\mathrm{T}} g_{k_i} + {}_{k_i,l_i} d_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2}\ {}^2_{k_i,l_i} d_{k_i}^{\mathrm{T}} H_{k_i} d_{k_i} + o({}^2_{k_i,l_i}) \\ \leqslant {} & f(x_{k_i}) + {}^2_{k_i,l_i}\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i}^{\mathrm{T}} H_{k_i} d_{k_i}\Big) + o({}^2_{k_i,l_i}) \end{aligned} \tag{3.4} $$

结合 (3.3)–(3.4), 可得

$$ \begin{aligned} & {}^2_{k_i,l_i}\ {}_1\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i} H_{k_i} d_{k_i}\Big) \leqslant {}^2_{k_i,l_i}\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i}^{\mathrm{T}} H_{k_i} d_{k_i}\Big) + o({}^2_{k_i,l_i}) \\ & {}^2_{k_i,l_i}({}_1 - 1)\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i} H_{k_i} d_{k_i}\Big) \leqslant o({}^2_{k_i,l_i}) \\ & ({}_1 - 1)\Big(s_{k_i}^{\mathrm{T}} g_{k_i} + \frac{1}{2} d_{k_i} H_{k_i} d_{k_i}\Big) \leqslant \frac{o({}^2_{k_i,l_i})}{{}^2_{k_i,l_i}} \\ & ({}_1 - 1)\Big(- c_1 c_2 \|g_{k_i}\|^2 + \frac{1}{2}\tilde{}_{k_i} |\min\{\tilde{}_{k_i}\ \ 0\}|\Big) \leqslant \frac{o({}^2_{k_i,l_i})}{{}^2_{k_i,l_i}} \end{aligned} $$

当 $\ \to \infty$ 时, 取极限满足

$$ \lim_{i\to\infty} \|g_{k_i}\| = 0 \qquad \lim_{i\to\infty} \tilde{}_{k_i} |\min\{\tilde{}_{k_i}\ \ 0\}| = 0 $$

与 (3.2) 的证明类似, 得到海森矩阵 $H(x_*)$ 至少是半正定的.

定理 3.2 如果滤子增广无限次, 那么, 迭代序列至少存在一个极限点是一个二阶稳定点, 也就是说, $g(x_*) = 0$ 和海森矩阵 $H(x_*)$ 至少是半正定的.

证 有两种情况.

情况 1 如果存在某个 $K \in \mathbb{N}$, 使得滤子对所有的迭代 $\ \geqslant K$ 都要校正, 我们将证明, 序列的每一个极限点 x_* 是一个二阶稳定点. 证明用到反证法. 假设存在一个无限子序列 $\{\ _i\}$, 满足

$$ {}_{k_i} = \|g_{k_i}\| + |\min\{\tilde{}_{k_i}\ \ 0\}| \geqslant \tag{3.5} $$

对某个 $\quad 0$. 每一次迭代 ${}_i$ $(\ _{k_i}\ \ f_{k_i})$ 被加入到滤子, 这意味着没有其他的 $(\ \ f)$ 被添加

到滤子且也不在以下正方形中:

$$[\bar{\ }_{k_i} - \gamma_\theta \quad \bar{\ }_{k_i}] \times [\bar{f}_{k_i} - \gamma_f \quad \bar{f}_{k_i}]$$

这些正方形的面积至少是 $\gamma_\theta^{\mathrm{non}}\gamma_f^{\mathrm{non}\ \ 2}$. 假设对所有的 ,

$$f^{\min} \leqslant f_k \leqslant f^{\max} \quad 和 \quad 0 \leqslant \ _k \leqslant \ ^{\max}$$

其中 $f^{\min}$ $f^{\max}$ 和 $^{\max}$ 是常数. 因此, 滤子中的 (f) 对被限制在

$$\mathcal{B} = [0 \quad ^{\max}] \times [f^{\min} \quad f^{\max}]$$

易见 $\mathcal{B}$ 的面积是有限的. 因而, $\mathcal{B}$ 会被有限个这样的正方形完全覆盖, 与满足 (3.5) 的无限序列矛盾.

情况 2 如果对所有的 $K \in \mathbb{N}$, 存在 $_1$ $_2 \geqslant K$, 使得 x_{k_1} 由非单调曲率搜索条件 (2.3) 产生和 x_{k_2} 由非单调滤子 (2.4) 产生. 与情况 1 相似, 可以证明, 迭代序列至少存在一个极限点 x_* 是二阶稳定点.

4 局部收敛性

为了分析算法的局部收敛性, 需要更多的假设条件.

假设 H (H1) 序列 $\{x_k\}$ 收敛到点 x_* 满足 $g_* = 0$ 和 H_* 正定, 其中 $g_* = g(x_*)$ 和 $H_* = \nabla^2 f(x_*)$.

(H2) H_k 和 s_k 由算法 2.1 产生, 满足以下条件

$$\lim_{k\to\infty} \frac{\|(H_k - \nabla^2 f(x_*))s_k\|}{\|s_k\|} = 0 \tag{4.1}$$

引理 4.1 假设 H 成立. 那么, 存在 $K \in \mathbb{N}$, 对所有的 $\geqslant K$, 满足 $_k = 1$.

证 通过假设知 $\nabla^2 f(x_k)$ 对充分大的 是正定的, 这就表明, $\tilde{\ }$ 0 $d_k = 0$ 和 $(x_k) = \|g_k\|$. 从滤子的结构可以看出 ((x_k) $f(x_k)) \in \mathcal{F}_k$, 也就是说,

$$(x_k) \leqslant (1 - \gamma_\theta) \max\Big[\ (x_j) \quad \sum_{r=0}^{\tau(j)-1} \ _{jr} \ (x_{j-r}) \Big]$$

或

$$f(x_k) \leqslant \max\Big[f(x_j) \quad \sum_{r=0}^{\tau(j)-1} \ _{jr} f(x_{j-r}) \Big] - \gamma_f \ (x_j)$$

对所有的 ((x_j) $f(x_j)) \in \mathcal{F}_k$.

情况 1 如果 $f(x_k) \leqslant \max\big[f(x_j) \ \sum_{r=0}^{\tau(j)-1} \ _{jr} f(x_{j-r})\big] - \gamma_f \ (x_j)$, 可得

$$\begin{aligned}
& f(x_k + s_k) - \max\Big[f(x_j) \quad \sum_{r=0}^{\tau(j)-1} \ _{kr} f(x_{j-r}) \Big] + \gamma_f \ (x_j) \\
&= g_k^{\mathrm{T}} s_k + \frac{1}{2} s_k^{\mathrm{T}} \nabla^2 f(x_k) s_k + o(\|s_k\|^2) + f(x_k) - \max\Big[f(x_j) \quad \sum_{r=0}^{\tau(k)-1} \ _{kr} f(x_{j-r}) \Big] + \gamma_f \ (x_j) \\
&\leqslant -s_k^{\mathrm{T}} (P^{\mathrm{T}} L \overline{D}_k L^{\mathrm{T}} P) s_k + \frac{1}{2} s_k^{\mathrm{T}} \nabla^2 f(x_k) s_k + o(\|s_k\|^2)
\end{aligned}$$

$$= -\frac{1}{2}s_k^{\mathrm{T}}\nabla^2 f(x_k)s_k - s_k^{\mathrm{T}}[(P^{\mathrm{T}}L\overline{D}_kL^{\mathrm{T}}P) - \nabla^2 f(x_k)]s_k + o(\|s_k\|^2)$$
$$\leqslant -\frac{1}{2}\ \ \|s_k\|^2 + o(\|s_k\|^2)$$

其中 是一个正数. 对充分大的 , (2.4b) 成立.

情况 2 如果 $(x_k) \leqslant (1-\gamma_\theta)\max\big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\big]$, 可得

$$(x_k+s_k) - (1-\gamma_\theta)\max\Big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\Big]$$
$$= \|g_k + \nabla^2 f(x_k)s_k + O(\|s_k\|^2)\| - (1-\gamma_\theta)\max\Big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\Big]$$
$$= \|g_k + H_ks_k + (\nabla^2 f(x_k) - H_k)s_k + O(\|s_k\|^2)\| - (1-\gamma_\theta)\max\Big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\Big]$$
$$= \|(\nabla^2 f(x_k) - H_k)s_k + O(\|s_k\|^2)\| - (1-\gamma_\theta)\max\Big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\Big]$$
$$= o(\|s_k\|) - (1-\gamma_\theta)\max\Big[\ (x_j)\ \sum_{r=0}^{\tau(j)-1}\ {}_{jr}\ (x_{j-r})\Big]$$

因此 (2.4a) 对充分大的 成立. 这意味着 x_{k+1} 对充分大的 将被滤子接受.

定理 4.1 假设 H 成立, 存在 x^* 的一个邻域 $N(x^*\ \ ) = \{x \in \mathbb{R}^n \mid \|x - x^*\|\ \ \}$, 满足 $\nabla^2 f(x)$ 在 $N(x^*\ \ )$ 上是 Lipschitz 连续的, 也就是说, 存在一个常数 $L\ \ 0$, 满足

$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \leqslant L\|x - y\| \qquad \forall x\ \ y \in N(x^*\ \ )$$

那么 $\{x_k\}$ 二次收敛于 x^*.

证 由于 $\nabla^2 f(x_k)$ 对充分大的 是正定的, 有 $P^{\mathrm{T}}L\overline{D}_kL^{\mathrm{T}}P_k = \nabla^2 f(x_k)$ 和 $d_k = 0$. 这就是说 s_k 是一个牛顿步. 定理的结论成立.

5 数值实验

本节中, 我们将给出非单调滤子曲率线搜索算法的数值结果, 采用 HP (i5 CPU) 个人电脑, 内存为 2G. 参数选择: ${}_1 = 0\ 3$; $\tau_1 = \tau_2 = 0\ 5$; $\ = 10^{-6}$; $M = 8$. 在 (2.3) 中选取 ${}_{kr} = \frac{1}{\tau(k)}$, 其中 $r = 0\ \cdots\ \ \tau(\ \ ) - 1$. 例如, 如果 $\tau(\ \ ) = 3$, 有 ${}_{k0} = {}_{k1} = {}_{k2} = \frac{1}{3}$. 当终止条件 $\|g_k\| \leqslant$ 成立, 计算停止.

Fail 代表算法失败. NIT, NF 和 NFI 分别代表迭代次数, 函数计算次数和滤子增广次数. 为了进行比较, 我们包含了非单调线搜索算法 (2.2), 不使用负曲率和滤子 (用 NLA 表示). NLA 的参数选择与算法 2.1 一样. 试验结果显示在表 1 中. 为了比较算法 2.1 和 NLA 的数值表现, 使用 Dolan 和 Moré 在文 [24] 中提出的图表法. 我们的图表基于迭代次数和函数计算次数. 检测结果通过运行软件计算一组测试题, 记录有关的信息, 比如函数计算次数. 令 $\mathcal{S}$ 是一组需要比较的软件. 对每个问题 p 和每个算法 s, 定义

$$t_{p,s} = \text{算法 } s \text{ 计算问题 } p \text{ 所需要的函数计算次数}$$

相应地, 可以令 $t_{p,s}$ 是迭代次数. 算法间的比较通过实验比率来定义

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in \mathcal{S}\}}$$

如果定义

$$_s(\tau) = \frac{1}{n_p}\mathrm{size}\{p \in \mathcal{P} : r_{p,s} \leqslant \tau\}$$

那么 $_s(\tau)$ 是算法 $s \in \mathcal{S}$ 成功的概率. $_s(1)$ 的值是算法胜过其他算法的概率.

表 1　数值结果

	算法 2.1			NLA	
Function	NIT	NF	NFI	NIT	NF
Rosenbrock	51	62	13	58	81
Freudenstein and Roth	10	17	9	10	22
Brown badly scaled	16	29	15	18	60
Beale	18	22	10	18	25
Helical vally	36	48	15	42	72
Bard	12	17	8	18	26
Bard 2	5	6	4	4	7
Meyer	85	173	15	Fail	
Gulf research and development	1	2	1	6	13
Box three-dimensional	26	36	11	30	39
Powell singular	38	49	13	38	59
Wood	99	125	23	115	174
Kowalik and Osborne	33	34	8	29	32
Brown and Dennis	40	68	19	42	95
Biggs EXP 6	14	17	10	14	19
Watson (n=12)	36	47	16	31	52
Watson (n=6)	36	47	16	31	52
Broyden banded	32	56	14	57	101
Raydan 2	7	8	7	7	8
Raydan 1	138	166	23	140	202
Extended Penalty	43	55	18	38	60
Perturbed Quadratic	19	30	8	20	39
Diagonal 2	94	95	24	94	95

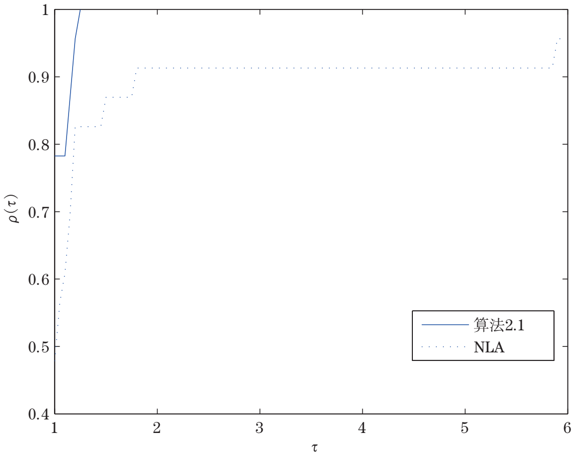

图 1 迭代次数比较

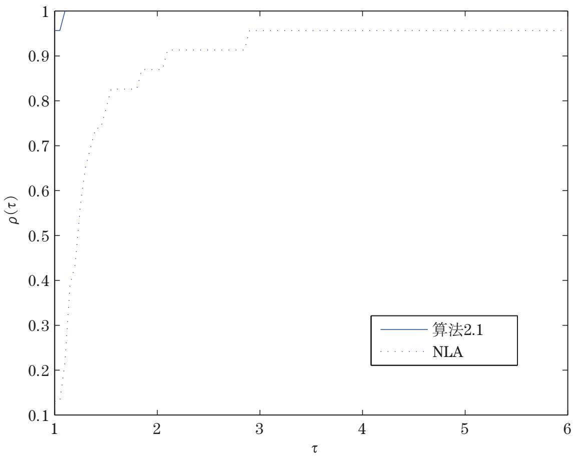

图 2 函数计算次数比较

从图 1–2 可以看出, 算法 2.1 在给定的测试题中胜过 NLA, 主要源于滤子和负曲率方向的运用. 由于滤子使得算法更容易接受试探点, 算法 2.1 比 NLA 更有效. 当不定的 H_k 在计算中出现, 算法 2.1 能够构造负曲率方向, 继续迭代. 但是, 算法 NLA 会失败. 比如, 算法 2.1 成功地求解测试题 Function Meyer, 其中不定的 H_k 出现了 13 次. 然而算法 NLA 失败. 因此新算法表现更好, 更有吸引力.

6 总 结

本文中, 我们发展了非单调滤子技术和负曲率方向求解无约束非凸优化问题; 在合理假设下, 建立了整体收敛性和局部收敛性理论; 数值试验结果表明新算法是稳定的和有效的.

参 考 文 献

[1] McCormick G P. A modification Armijo's step-size rule for negative curvature [J]. *Math Program,* 1977, 13:111–115.

[2] Goldfarb D. Cuvilinear path steplength algorithms for minimization which use direction of negative curvature [J]. *Math Program*, 1980, 18:31–40.

[3] Moré J J, Sorensen D C. On the use of direction of negative curvature in a modified Newton method [J]. *Math Program*, 1979, 16:1–20.

[4] Grippo L, Lampariello F, Ludidi S. A nonmonotone line search technique for Newton's method [J]. *SIAM J Numer Anal*, 1986, 23:707–716.

[5] Dai Y H. Convergence properties of the BFGS algorithm [J]. *SIAM J Optim*, 2002, 13:693–701.

[6] Mascarenhas W F. The BFGS method with exact line searches fails for non-convex objective functions [J]. *Math Program*, 2004, 99(1):49–61.

[7] Ferris M C, Lucidi S, Roma M. Nonmonotone curvilinear line search methods for unconstrained optimization [J]. *Comput Optim Appl*, 1996, 6:117–136.

[8] Sun W, Zhou Q Y. An unconstrained optimization method using nonmonotone second order Goldstein's line search [J]. *Sci China Ser A: Math*, 2007, 50:1389–1400.

[9] Xiao Y, Zhou F I. Nonmonotone trust region method with curvilinear path in unconstrained minimization [J]. *Computing*, 1992, 48:303–317.

[10] Zhou Q Y, Sun W. Adaptive nonmonotone trust region method with curvilinear searches [J]. *J Comput Math*, 2006, 24:761–770.

[11] Zhou Q Y, Sun W. A nonmonotone second-order steplength method for unconstrained minimization [J]. *J Comput Math*, 2007, 24:104–112.

[12] Fletcher R, Leyffer S. Nonlinear programming without a penalty function [J]. *Math Program*, 2002, 91:239–269.

[13] Gu C, Zhu D. A dwindling filter algorithm with modified subproblem for nonlinear inequality constrained optimization [J]. *Chin Ann Math Ser B*, 2014, 35:209–224.

[14] Gu C, Zhu D. A secant algorithm with line search filter method for nonlinear optimization [J]. *Appl Math Model*, 2011, 35:879–894.

[15] Li C, Sun W. On filter-successive linearization methods for nonlinear semidefinite programming [J]. *Sci China Ser A*, 2009, 52:2341–2361.

[16] Nie P Y. Sequential penalty quadratic programming filter methods for nonlinear programming [J]. *Nonlinear Analysis: Real World Applications*, 2007, 8:118–129.

[17] Su K. A globally and superlinearly convergent modified SQP-filter method [J]. *Journal of Global Optimization*, 2008, 41:203–217.

[18] Shen C G, Xue W J, Pu D G. A filter SQP algorithm without a feasibility restoration phase [J]. *Comp Appl Math*, 2009, 28:167–194.

[19] Wächter A, Biegler L T. Line search filter methods for nonlinear programming: Local convergence [J]. *SIAM J Optim*, 2005, 6:32–48.

[20] Wächter A, Biegler L T. Line search filter methods for nonlinear programming: Motivation and global convergence [J]. *SIAM J Optim*, 2005, 16:1–31.

[21] Yu Z S, Pu D G. A new nonmonotone line search technique for unconstrained optimization [J]. *J Comput Appl Math*, 2008, 219:134–144.

[22] Han X, Sun W, Dang C Y. Nonmonotone second-order Wolfe's line search method for unconstrained optimization problems [J]. *Comput Math Appl*, 2010, 60:2517–2525.

[23] Bunch J R, Parlett B N. Direct methods for solving symmetric indefinite systems of linear equations [J]. *SIAM J Numer Anal*, 1971, 8:639–655.

[24] Dolan E D, Moré J J. Benchmarking optimization software with performance profiles [J]. *Math Program*, 2002, 91:201–213.

A Nonmonotone Filter Curvilinear Line Search Algorithm for Unconstrained Nonconvex Optimization

GU Chao[1] ZHU Detong[2]

[1]School of Statistics and Mathematics, Shanghai Lixin University of Accounting and Finance, Shanghai 201620, China. E-mail: guchao@lixin.edu.cn

[2]Department of Mathematics, Shanghai Normal University, Shanghai 200234, China. E-mail: dtzhu@shnu.edu.cn

Abstract Yu and Pu [Yu Z S, Pu D G. A new nonmonotone line search technique for unconstrained optimization [J]. *J Comput Appl Math*, 2008, 219:134–144] introduced a nonmonotone line search algorithm for unconstrained optimization. Different from their work, the authors propose a nonmonotone filter curvilinear line search algorithm when a problem may be nonconvex. By using the negative curvature information of the Hessian, the generated sequence is shown to converge to stationary points that satisfy second-order optimality conditions. Global convergence is established even without requiring a priori the existence of a limit point. Moreover, the authors analyze the convergence rate of the new algorithm. The numerical experiments are reported to show the effectiveness of the proposed algorithm.

Keywords Nonconvex optimization, Nonmonotone, Curvilinear line search, Filter, Convergence

2000 MR Subject Classification 90C30, 65K05

The English translation of this paper will be published in
Chinese Journal of Contemporary Mathematics, Vol. 38 No. 4, 2017
by ALLERTON PRESS, INC., USA