运筹学与控制论

DOI:CNKI:50-1165/N.20121112.1642.002

# Armijo 型线搜索一个修正 LS 共轭梯度法的全局收敛性*

孟继东，杜学武
(重庆师范大学 数学学院，重庆 401331)

**摘要** 提出一个新的修正 Liu-Storey 共轭梯度(MLSCG)算法。在精确线搜索下 MLSCG 算法化归为标准的 Liu-Storey(LS)共轭梯度算法。MLSCG 算法产生的搜索方向不依赖于所使用的线搜索准则而具有充分下降性。本文证明了 MLSCG 算法在一个 Armijo 型线搜索下具有全局收敛性。数值试验表明,对于多数算例 MLSCG 算法比 PRP、HS、LS 等算法具有更好的计算结果。

**关键词** 共轭梯度法;Liu-Storey 共轭梯度法;Armijo 型线搜索;全局收敛性
**中图分类号**:O221.2　　**文献标志码**:A　　**文章编号**:1672-6693(2012)06-0006-03

本文考虑无约束优化问题

$$\min_{x\in\mathbf{R}^n} f(x) \tag{1}$$

其中 $f:\mathbf{R}^n\to\mathbf{R}$ 为连续可微函数,其梯度向量记为 $g(x)$。

共轭梯度法是求解问题(1),特别是大规模问题的主要方法之一。共轭梯度法的迭代格式为

$$x_{k+1}=x_k+\alpha_k d_k \tag{2}$$

$$d_k=\begin{cases}-g_k, k=1\\ -g_k+\beta_k d_{k-1}, k\geqslant 2\end{cases} \tag{3}$$

其中 $g_k=\nabla f(x_k)$,$d_k$ 为搜索方向,$\alpha_k$ 是用某种线搜索得到的步长,$\beta_k$ 是一个纯量,其取法满足:当 $f(x)$ 为二次凸函数且采用精确线搜索时,方法(2)~(3)化归为线性共轭梯度法。

常见的共轭梯度法包括 Fletcher-Reeves(FR)方法、Polak-Ribiere-Polyak(PRP)方法、Hestenes-Stiefel(HS)方法、共轭下降(CD)方法、Liu-Storey(LS)方法和 Dai-Yuan(DY)方法等。这些方法中的 $\beta_k$ 取值如下:

$$\beta_k^{\mathrm{FR}}=\frac{\|g_k\|^2}{\|g_{k-1}\|^2},\beta_k^{\mathrm{PRP}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{\|g_{k-1}\|^2},\beta_k^{\mathrm{HS}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{d_{k-1}^{\mathrm{T}}y_{k-1}}$$

$$\beta_k^{\mathrm{CD}}=\frac{\|g_k\|^2}{-g_{k-1}^{\mathrm{T}}d_{k-1}},\beta_k^{\mathrm{LS}}=\frac{g_k^{\mathrm{T}}y_{k-1}}{-g_{k-1}^{\mathrm{T}}d_{k-1}},\beta_k^{\mathrm{DY}}=\frac{\|g_k\|^2}{d_{k-1}^{\mathrm{T}}y_{k-1}}$$

其中 $y_{k-1}=g_k-g_{k-1}$。

FR 方法、CD 方法和 DY 方法具有较好的理论收敛性,而 PRP 方法、HS 方法和 LS 方法具有较好的数值计算效果[1-12]。

张丽[8]提出一个新的 Liu-Storey 型方法,并证明了该方法对一般非线性函数在 Grippo-Lucidi 线搜索下具有全局收敛性。张丽[8]还修正了她提出的 Liu-Storey 型方法,使得修正后的方法在强 Wolfe-Powell 线搜索下对非凸函数具有全局收敛性。

张丽[9]受拟牛顿方向迭代公式的启发,修正了 PRP 方法的搜索方向,取搜索方向 $d_k$ 为

$$d_k=-g_k+\beta_k^{\mathrm{PRP}}d_{k-1}+\theta_k y_{k-1}$$

其中 $\theta_k=\frac{g_k^{\mathrm{T}}d_{k-1}}{\|g_{k-1}\|^2}$,从而使新算法具有充分下降性和全局收敛性,并且获得了很好的数值试验结果。

本文受文献[9]的启发,对 Liu-Storey 方法进行了修正,得到一个新的修正 Liu-Storey(MLS)算法,并证明了新算法在一个 Armijo 型线搜索下的全局收敛性。所采用的 Armijo 型线搜索准则为:取步长 $\alpha_k=\max\{\rho^j; j=0,1,2,\cdots\}$,使得不等式

$$f(x_k+\alpha_k d_k)\leqslant f(x_k)+\delta_1\alpha_k g_k^{\mathrm{T}}d_k-\delta_2\alpha_k^2\|d_k\|^2 \tag{4}$$

成立,其中 $\rho\in(0,1)$,$\delta_1\in(0,1)$,$\delta_2>0$。

## 1 MLSCG 算法

本文给出的新算法迭代格式为(1)式和

$$d_k=\begin{cases}-g_k, k=1\\ -g_k+\beta_k^{\mathrm{LS}}d_{k-1}+\theta_k y_{k-1}, k\geqslant 2\end{cases} \tag{5}$$

其中

$$\theta_k=\frac{g_k^{\mathrm{T}}d_{k-1}}{g_{k-1}^{\mathrm{T}}d_{k-1}} \tag{6}$$

* 收稿日期:2012-03-22　　网络出版时间:2012-11-12 16:42:01
资助项目:国家自然科学基金(No. 10971241;No. 11171363);重庆师范大学自然科学基金(No. 08XLR022)
作者简介:孟继东,男,助教,硕士研究生,研究方向为最优化理论与算法;通讯作者:杜学武,E-mail:duxuewu@cqnu.edu.cn
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20121112.1642.201206.6_002.html

称(1)、(5)式为 MLS 方法。当线搜索精确时，新方法就得到标准的 LS 共轭梯度法。同时由(2)式和 $\beta_k^{LS}$ 知对 $k\geqslant 1$ 有

$$g_k^T d_k = -\|g_k\|^2 \tag{7}$$

算法　步骤 1，给出 $x_1\in\mathbf{R}^n$，$\varepsilon\geqslant 0$ 令 $d_1=-g_1=-\nabla f(x_1)$，$k:=1$，若 $\|g_1\|<\varepsilon$ 立即停止；步骤 2，找到步长 $\alpha_k$，满足(4)式；步骤 3，令 $x_{k+1}=x_k+\alpha_k d_k$，$g_{k+1}=g(x_{k+1})$，若 $\|g_{k+1}\|<\varepsilon$ 立即停止；步骤 4，通过(5)式求得 $d_{k+1}$；步骤 5，令 $k:=k+1$，返回步骤 2。

## 2 算法的收敛性

本文做如下假设：

(H1) $f(x)$ 在水平集 $\Omega=\{x\in\mathbf{R}^n \mid f(x)\leqslant f(x_1)\}$ 有下界，其中 $x_1$ 为初始点。(H2) $f$ 在水平集 $\Omega$ 的一个邻域 $N$ 内连续可微，且其梯度满足 Lipschitz 条件，即存在常数 $L>0$，使得

$$\|g(x)-g(y)\|\leqslant L\|x-y\|,\ \forall x,y\in N \tag{8}$$

显然，在上述假设下，存在某个 $\gamma_1>0$ 使得

$$\|g(x)\|\leqslant\gamma_1,\ \forall x\in\Omega \tag{9}$$

通过上述假设，由

$$f(x_k+\alpha_k d_k)\leqslant f(x_k)+\delta_1\alpha_k g_k^T d_k-\delta_2\alpha_k^2\|d_k\|^2$$

可知 $\sum\limits_{k\geqslant 1}^{\infty}(\delta_2\alpha_k^2\|d_k\|^2-\delta_1\alpha_k g_k^T d_k)<+\infty$，结合 $g_k^T d_k=-\|g_k\|^2$，可得到 $\sum\limits_{k\geqslant 1}^{\infty}\delta_2\alpha_k^2\|d_k\|^2<+\infty$ 和 $\sum\limits_{k\geqslant 1}^{\infty}\delta_1\alpha_k\|g_k\|^2<+\infty$。

特别地

$$\lim_{k\to\infty}a_k\|d_k\|=0,\ \lim_{k\to\infty}a_k\|g_k\|^2=0 \tag{10}$$

引理 1　如果存在常数 $\varepsilon>0$ 使得 $\|g_k\|\geqslant\varepsilon$，则对任意的 $k$，必存在常数 $M>0$ 满足 $\|d_k\|\leqslant M$。

证明　由(5)~(6)、(9)式及假设(H2)、Schwarz 不等式和引理的条件易知，对任意 $k\geqslant 2$ 有

$$\|d_k\|\leqslant\|g_k\|+\frac{2\|g_k\|\|y_{k-1}\|\|d_{k-1}\|}{\|g_{k-1}\|^2}\leqslant$$
$$\gamma_1+\frac{2\gamma_1 L\alpha_{k-1}\|d_{k-1}\|\|d_{k-1}\|}{\varepsilon^2}$$

由(10)式知，存在某个常数 $r\in(0,1)$ 和整数 $k_0$ 使得不等式 $\frac{2\gamma_1 L\alpha_{k-1}\|d_{k-1}\|}{\varepsilon^2}\leqslant r$ 对所有 $k\geqslant k_0$ 成立。从而对任意 $k\geqslant k_0$，有

$$\|d_k\|\leqslant\gamma_1+r\|d_{k-1}\|\leqslant\gamma_1(1+r+r^2+\cdots+r^{k-k_0-1})+$$
$$r^{k-k_0}\|d_{k_0}\|\leqslant\frac{\gamma_1}{1-r}+\|d_{k_0}\|$$

取 $M=\max\left\{\|d_1\|,\|d_2\|,\cdots,\|d_{k_0}\|,\frac{\gamma_1}{1-r}+\|d_{k_0}\|\right\}$，则有 $\|d_k\|\leqslant M$ 对任意 $k\geqslant 1$ 成立。　证毕

定理 1　在假设(H1)、(H2)成立的条件下，设 $\{x_k\}$ 由采用 Armijo 型线搜索准则(4)的 MLSCG 算法产生，那么，或者 $\|g_k\|=0$ 对某个 $k$ 成立，或者有 $\liminf\limits_{k\to\infty}\|g_k\|=0$。

证明　反证法。假设定理的结论不成立，则存在常数 $\varepsilon>0$，使得 $\|g_k\|\geqslant\varepsilon$ 对任意 $k$ 成立。下面分 2 种情况进行讨论。

1) $\liminf\limits_{k\to\infty}\alpha_k>0$，则由(7)、(10)式容易推出 $\liminf\limits_{k\to\infty}\|g_k\|=0$，这与假设矛盾。

2) $\liminf\limits_{k\to\infty}\alpha_k=0$，则存在无限指标集 $K$ 使得 $\lim\limits_{k\in K,k\to\infty}\alpha_k=0$。

由算法的步骤 2 可知当 $k\in K$ 充分大时，$\rho^{-1}\alpha_k$ 不能满足

$$f(x_k+\alpha_k d_k)-f(x_k)\leqslant\delta_1\alpha_k g_k^T d_k-\delta_2\alpha_k^2\|d_k\|^2$$

即

$$f(x_k+\rho^{-1}\alpha_k d_k)-f(x_k)>$$
$$\delta_1\rho^{-1}\alpha_k g_k^T d_k-\delta_2\rho^{-2}\alpha_k^2\|d_k\|^2 \tag{11}$$

由中值定理和假设(H2)可知，存在 $h_k\in(0,1)$ 使得

$$f(x_k+\rho^{-1}\alpha_k d_k)-f(x_k)=$$
$$\rho^{-1}\alpha_k g(x_k+h_k\rho^{-1}\alpha_k d_k)^T d_k=\rho^{-1}\alpha_k g_k^T d_k+$$
$$\rho^{-1}\alpha_k(g(x_k+h_k\rho^{-1}\alpha_k d_k)-g_k)^T d_k\leqslant$$
$$\rho^{-1}\alpha_k g_k^T d_k+L\rho^{-2}\alpha_k^2\|d_k\|^2$$

再由(11)式，当 $k\in K$ 充分大时有

$$\delta_1\rho^{-1}\alpha_k g_k^T d_k-\delta_2\rho^{-2}\alpha_k^2\|d_k\|^2\leqslant$$
$$\rho^{-1}\alpha_k g_k^T d_k+L\rho^{-2}\alpha_k^2\|d_k\|^2$$
$$\delta_1 g_k^T d_k-\delta_2\rho^{-1}\alpha_k\|d_k\|^2\leqslant g_k^T d_k+L\rho^{-1}\alpha_k\|d_k\|^2$$
$$(\delta_1-1)g_k^T d_k\leqslant\delta_2\rho^{-1}\alpha_k\|d_k\|^2+L\rho^{-1}\alpha_k\|d_k\|^2$$
$$-g_k^T d_k\leqslant\frac{\delta_2\rho^{-1}\alpha_k\|d_k\|^2+L\rho^{-1}\alpha_k\|d_k\|^2}{1-\delta_1}$$

整理可得

$$\|g_k\|^2\leqslant\frac{\delta_2\rho^{-1}\alpha_k\|d_k\|^2+L\rho^{-1}\alpha_k\|d_k\|^2}{1-\delta_1}=$$
$$\frac{\delta_2\rho^{-1}+L\rho^{-1}}{1-\delta_1}\alpha_k\|d_k\|^2$$

由引理知，$\{d_k\}$ 有界，从而由上面不等式和 $\lim\limits_{k\in K,k\to\infty}\alpha_k=0$ 得知 $\liminf\limits_{k\to\infty}\|g_k\|=0$，这与假设矛盾。　证毕

## 3 数值试验

对于 PRP、HS、LS 共轭梯度算法和 MLSCG 算法，采用 Matlab 编程，对一些标准的测试函数进行了数值

试验[13]。数值试验结果见表 1。表中“ Problem ”表示测试函数的名称;“ *Dim* ”表示测试函数自变量的个数;“ *NI*,*NF*,*NG* ”依次表示为迭代次数、函数值计算次数、梯度值计算次数;– 表示迭代失败。

通过表 1 的数值结果,说明了新方法的可行性,优于其他方法。

表 1 计算结果

Tab. 1 Computation result

| Problem | *Dim* | PRP (*NI*/*NF*/*NG*) | HS (*NI*/*NF*/*NG*) | LS (*NI*/*NF*/*NG*) | MLS (*NI*/*NF*/*NG*) |
|---|---|---|---|---|---|
| rosenbrock | 2 | 151/546/152 | 172/846/173 | 155/633/156 | 138/528/139 |
| Freudenstein and Roth | 2 | 27/896/28 | 15/641/16 | 12/325/13 | 21/575/22 |
| Brown | 2 | – | 46/6 397/48 | 151/23 342/152 | 41/5 166/42 |
| Beale | 2 | 21/412/22 | 14/483/16 | 23/537/24 | 24/400/25 |
| Box | 3 | 4 612/9 396/4 613 | 8/212/10 | 5 363/10 781/5 364 | 13 135/44 296/13 136 |
| Gaussian | 3 | 11/452/12 | 9/58/10 | 10/442/11 | 6//47/7 |
| Powell Singular | 4 | 319/6 462/320 | 278/6 765/280 | 190/3 579/191 | 42/830/43 |
| Brown and Dennis | 4 | 9 790/20 057/9 791 | 18 906/38 370/18 907 | 9 810/20 056/9 811 | 9 760/19 980/9 761 |
| Wood | 4 | 204/6 080/205 | 239/6 930/240 | 197/5 676/198 | 190/5 243/191 |
| Extend rosenbrock | 8 | 28/472/29 | 18/321/19 | 28/829/29 | 26/398/27 |
| Extend rosenbrock | 10 | 38/536/34 | 27/342/32 | 36/936/33 | 32/543/28 |

## 参考文献:


[1] Braun W J,Wang L,Zhao Y Q. Properties of the geometric and related processes[J]. Naval Research Logistics,2005,52(7):607-616.

[2] Dai Y H ,Yuan Y X. A nonlinear conjugate gradient with a strong global convergence property[J]. SIAM Journal on Optimization,2000,10(1):177-182.

[3] Fletcher R,Reeves C. Function minimization by conjugate gradients[J]. Computer Journal,1964,7(2):149-154.

[4] Hestenes M R,Stiefel E L. Methods of conjugate gradients for solving linear systems[J]. J Res Nat Bur Standards Sect,1952,49(5):409-436.

[5] Liu Y,Storey C. Efficient generalized conjugate gradient algorithms[J]. Journal of Optimization Theory and Applications,1991,69(1):129-152.

[6] Polak E,Ribiere G. Note sur la convergence de directions conjugees[J]. Rev Francaise Informat Recherche Operatinelle,1969,3(1):35-43.

[7] Polyak B T. The conjugate gradient method in extremem problems[J]. USSR Comp Math and Math Phys,1969,9(4):94-112.

[8] Zhang L. A new Liu-Storey type nonlinear conjugate gradient method for unconstrained optimization problems[J]. Journal of Computational and Applied Mathematics,2009,225(1):146-157.

[9] Zhang L,Zhou W J,Li D H. A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence[J]. IMA Journal of Numerical Analysis,2006,26(4):629-640.

[10] 戴彧虹,袁亚湘. 非线性共轭梯度法[M]. 上海:上海科学技术出版社,2000.

[11] 王开荣,曹伟,王银河. Armijo 型线搜索下的谱 CD 共轭梯度法[J]. 山东大学学报:理学版,2010,45(11):104-108.

[12] 莫利柳. 一类修正 WYL 共轭梯度法的全局收敛性[J]. 重庆理工大学学报:自然科学版,2010,24(9):94-97.

[13] More J J,Garbow B S,Hillstrom K E. Testing unconstrained optimization software[J]. ACM Transactions on Mathematical Software,1981,7(1):17-41.


**Operations Research and Cybernetics**

## Global Convergence of a Modified LS Conjugate Gradient Method with an Armijo-Type Line Search


MENG Ji-dong, DU Xue-wu

(College of Mathematics, Chongqing Normal University, Chongqing 401331, China)



**Abstract**: This article presents a new modified Liu-Storey conjugate gradient (MLSCG) algorithm. MLSCG algorithm reduces to the Liu-Storey conjugate gradient method when the exact line search is used. MLSCG algorithm possesses the sufficient descent property without relying on the line search be used. The global convergence of MLSCG algorithm with an Armijo-type line search is proved. Preliminary numerical results show that MLSCG algorithm is efficient.



**Key words**: conjugate gradient method; Liu-Storey conjugate gradient method; Armijo-type line search; global convergence

(责任编辑 黄 颖)