Advances in Applied Mathematics 应用数学进展, 2013, 2, 48-54 Hans汉斯
http://dx.doi.org/10.12677/aam.2013.21007 Published Online February 2013 (http://www.hanspub.org/journal/aam.html)

A Modified LS Nonlinear Conjugate Gradient Algorithms

Hai Chen

College of Mathematics and Information Science, Guangxi University, Nanning
Email: hchengx@126.com

Received: Nov. 22nd, 2012; revised: Dec. 4th, 2012; accepted: Dec. 20th, 2012

Abstract: In this paper, a modified LS conjugate gradient formula is proposed. This formula can ensure that the scalar $\beta_k \geq 0$ holds and the search direction possesses the sufficiently descent property without any line search. The global convergence will be established for general functions under suitable conditions and numerical results are reported.

Keywords: Conjugate Gradient Method; Descent Property; Global Convergence

一个修改的 LS 非线性共轭梯度算法

陈 海

广西大学数学与信息科学学院，南宁
Email: hchengx@126.com

收稿日期：2012 年 11 月 22 日；修回日期：2012 年 12 月 4 日；录用日期：2012 年 12 月 20 日

摘 要：我们给出一个修改的 LS 共轭梯度公式，此公式能保证参数 β_k 非负且搜索方向在不需要任何线搜索下具有充分下降性。在适当条件下，证明该方法对一般函数具有全局收敛性，同时给出数值检验结果。

关键词：共轭梯度法；下降性；全局收敛性

1. 引言

求解下面无约束优化问题

$$\min_{x \in R^n} f(x) \tag{1.1}$$

其中 $f(x)$ 连续可微，此问题可来源于诸多实际，具有广泛的应用背景，是最优化问题的最基础形式，同时也是是其他类型优化问题的基础。一般情况下，首先探讨此问题的一般求解方法，然后再推广至其他类型问题。所以说，研究上述问题的求解方法既具有实际意义又具有理论根据。已存在众多方法求解此问题，如牛顿法、信赖域方法、拟牛顿方法和共轭梯度法等，其中非线性共轭梯度法拥有结构简单且效果明显的优点，是研究热点。此方法采用下述迭代公式：

$$x_{k+1} = x_k + \alpha_k d_k, k = 0,1,2,\cdots$$

x_k 是第 k 次迭代点，$\alpha_k > 0$ 是步长，d_k 是具有下面定义形式的搜索方向

$$d_k = \begin{cases} -g_k + \beta_k d_{k-1}, & k \ge 1, \\ -g_k, & k = 0, \end{cases} \tag{1.2}$$

其中 $\beta_k \in \Re$ 是一个参数，根据 β_k 的选取方式不同而称为相应的共轭梯度法。目前已有很多著名的共轭梯度公式：

$$\beta_k^{PRP} = \frac{g_{k+1}^T\left(g_{k+1}-g_k\right)}{\|g_k\|^2}\,[1,2],\quad \beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{\|g_k\|^2}\,[3],\quad \beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}\,[4],$$

$$\beta_k^{LS} = \frac{g_{k+1}^T\left(g_{k+1}-g_k\right)}{-d_k^T g_k}\,[5],\quad \beta_k^{HS} = \frac{g_{k+1}^T\left(g_{k+1}-g_k\right)}{\left(g_{k+1}-g_k\right)^T d_k}\,[6],\quad \beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{\left(g_{k+1}-g_k\right)^T d_k}\,[7],$$

$g_k = \nabla f\left(x_k\right)$ 和 $g_{k+1} = \nabla f\left(x_{k+1}\right)$，分别表示函数 $f\left(x\right)$ 在 x_k 和 x_{k+1} 的梯度值，$\|.\|$ 是欧氏向量范数。上述方法可归结为两类：一类是 PRP, LS 和 HS 方法，他们拥有数值表现优越但收敛性不理想的特点；另一类是 FR, CD 和 DY 方法，此类方法的收敛性好但数值表现却不优越。PRP 方法的数值表现非常，许多学者都以此方法为基础，寻找更为优越的方法，已取得众多成果(见[8-17]等)。Yuan[15]给出了下述 PRP 修改公式：

$$\beta_k^{MPRP} = \beta_k^{PRP} - \min\left\{\beta_k^{PRP}, \frac{\mu\|y_k\|^2}{\|g_k\|^4} g_{k+1}^T d_k\right\},$$

其中 $\mu > \frac{1}{4}$ 是常数，$y_k = g_{k+1} - g_k$。在此公式的基础上，又将此思想进行了推广，得到了下面的修改的 LS：

$$\beta_k^{MLS} = \beta_k^{LS} - \min\left\{\beta_k^{LS}, \frac{\mu\|y_k\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k\right\}. \tag{1.3}$$

Wei[14]等给出了下述修改的 PRP 方法：

$$\beta_k^{WYL} = \frac{g_{k+1}^T y_k^m}{\|g_k\|^2}, \tag{1.4}$$

其中 $y_k^m = g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k$，容易得到 $\beta_k^{WYL} = \frac{g_{k+1}^T y_k^m}{\|g_k\|^2} \ge \frac{\|g_{k+1}\|\left(\|g_{k+1}\| - \frac{\|g_{k+1}\|}{\|g_k\|}\|g_k\|\right)}{\|g_k\|^2} = 0$ 成立。此方法能保证参数非负，克服了原 PRP 方法能为负值从而导致不收敛的缺点。对 LS 方法而言，它本身具有数值表现好的优点，但在理论上性质不是很理想：如本身没有充分下降性、不能保证参数非负、收敛性不理想等，为克服上述不足，我们结合(1.3)和(1.4)，给出修改的 LS 公式如下：

$$\beta_k^{MMLS*} = \beta_k^{MMLS} - \min\left\{\beta_k^{MMLS}, \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k\right\}. \tag{1.5}$$

其中 $\beta_k^{MMLS} = \frac{g_{k+1}^T y_k^m}{-d_k^T g_k}$，下面我们将分析(1.5)的充分下降性和全局收敛性。论文的创新点体现在如下几点：1) 相对于原 LS 方法而言，新方法能保证参数的非负性；2) 新方法在不需要任何条件基础上自动拥有充分下降性；3) 新方法的公式中采用了新的当前梯度与下一梯度差；4) 对一般函数具有全局收敛性。

本文结构安排如下：下一节给出算法及其步骤，第三节分析方法的充分下降性和全局收敛性，最后一节给出数值检验结果。

2. 算法

修改的 LS 算法步骤 0 给定 $x_0 \in \Re^n, \delta \in (0,1/2), \sigma \in (\delta,1)$ 和终止参数 $\varepsilon > 0$, 令 $d_0 = -g_0 = -\nabla f(x_0)$，置 $k := 0$。

步骤 1：若 $\|g_k\| \le \varepsilon$，停止。

步骤 2：利用下面的 WWP 线搜索技术寻找步长 α_k：

$$f(x_k + \alpha_k d_k) \le f_k + \delta \alpha_k g_k^T d_k \tag{2.1}$$

$$g(x_k + \alpha_k d_k)^T d_k \ge \sigma g_k^T d_k . \tag{2.2}$$

步骤 3：令 $x_{k+1} = x_k + \alpha_k d_k$。如果 $\|g_{k+1}\| \le \varepsilon$，停止。

步骤 4：利用下面公式计算搜索方向

$$d_{k+1} = -g_{k+1} + \beta_k^{MMLS*} d_k \tag{2.3}$$

步骤 5：置 $k := k+1$，转步骤 2。

3. 充分下降性和全局收敛性分析

下面的引理说明了修改的 LS 方向具有充分下降性。

引理 3.1 对 $k \ge 0$，修改的 LS 搜索方向对下式

$$d_k^T g_k \le -c\|g_k\|^2 \tag{3.1}$$

$$d_k^T y_k \ge c(1-\sigma)\|g_k\|^2 \tag{3.2}$$

满足，$c > 0$ 是常数。

证明：如果 $k = 0$，则 $g_0^T d_0 = -\|g_0\|^2$，则(3.1)成立。假设当 $k \ge 1$，(3.1)对修改的 LS(2.3)满足，对 $k+1$，利用(2.3)，获得

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \beta_k^{MMLS*} d_k^T g_{k+1} \\ &= -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} - \min\left\{ \frac{g_{k+1}^T y_k^m}{-d_k^T g_k}, \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k \right\} \right] d_k^T g_{k+1} \end{aligned} \tag{3.3}$$

对于(3.3)，有假设可知

$$-d_k^T g_k > 0. \tag{3.4}$$

取

$$u = \frac{\sqrt{-d_k^T g_k}}{\sqrt{2\mu}} g_{k+1}, v = \frac{\sqrt{2\mu} g_{k+1}^T d_k}{\sqrt{-d_k^T g_k}} y_k^m .$$

下面分两种情况讨论(3.4)。

情形 1) 若 $\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} < \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k$，立即得到 $g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2$。

情形 2) 若 $\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} \ge \frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k$，则(3.4)可写成

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\|g_{k+1}\|^2 + \left(\frac{g_{k+1}^T y_k^m}{-d_k^T g_k} - \frac{\mu \|y_k^m\|^2}{\left(-d_k^T g_k\right)^2} g_{k+1}^T d_k \right) d_k^T g_{k+1} \\ &= \frac{d_k^T g_{k+1} g_{k+1}^T y_k^m - \|g_{k+1}\|^2 \left(-d_k^T g_k\right) - \frac{\mu \|y_k^m\|^2}{-d_k^T g_k} \left(g_{k+1}^T d_k\right)^2}{-d_k^T g_k} \\ &= \frac{u^T v - \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)}{-d_k^T g_k} + \frac{-\left(1-\frac{1}{4\mu}\right)\|g_{k+1}\|^2 d_k^T y_k^n}{-d_k^T g_k} \le -\left(1-\frac{1}{4\mu}\right)\|g_{k+1}\|^2, \end{aligned}$$

最后一个不等式利用了 $u^T v \le \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)$。令 $c = 1 - \frac{1}{4\mu}$，则(3.1)成立。结合线搜索技术(2.2)可得(3.2)成立。所以此命题成立，证毕。

为了证明建议算法的收敛性，采用反证法来得到全局收敛性，即假设存在常数 $\varepsilon_0 > 0$ 满足

$$\|g_k\| \ge \varepsilon_0, \forall k \ge 0 \tag{3.5}$$

导出矛盾，证明结论。

假设条件：

(A): 1) 水平集 $\Omega = \left\{x \in \Re^n : f(x) \le f(x_0)\right\}$ 有界。2) f 在 Ω 上有下界且连续可微，它的梯度 g 满足 Lipschitz 条件，即存在常数 $L > 0$ 满足

$$\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in \Omega. \tag{3.6}$$

下述引理 3.2 和 3.3 可在众多共轭梯度算法文献中找到，本文也给出其证明过程。

引理 3.2 设假设(A)满足，序列 $\{g_k\}$ 和 $\{d_k\}$ 由修改的 LS 算法产生。如果不等式(3.5)成立，则 $d_k \ne 0$ 且 $\sum_{k=0}^{\infty} \|u_{k+1} - u_k\|^2 < \infty$，其中 $u_k = \frac{d_k}{\|d_k\|}$。

证明：利用(2.2)和 Lipschitz 条件(3.6)，得到 $-(1-\sigma) g_k^T d_k \le \left(g_{k+1} - g_k\right)^T d_k \le \alpha_k L \|d_k\|^2$，联立(3.1)，有 $\alpha_k \ge \frac{1-\sigma}{L} \frac{\left|g_k^T d_k\right|}{\|d_k\|^2}$，代入(2.1)且由假设(A)中函数 f 有下界，则

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty \tag{3.7}$$

成立。式(3.5)和引理 3.1 隐含着 $d_k \ne 0$，否则 $g_k = 0$ 成立，全局收敛性得证。则可令 $u_k = \frac{d_k}{\|d_k\|}$ 和 $\delta_k = \beta_k^{MMLS*} \frac{\|d_k\|}{\|d_{k+1}\|}, r_{k+1} = -\frac{g_{k+1}}{d_{k+1}}$。由(2.3)，当 $k \ge 1$，得 $u_{k+1} = r_{k+1} + \delta_k u_k$，利用 $\|u_{k+1}\| = \|u_k\| = 1$，则

$$\|r_{k+1}\| = \|u_{k+1} - \delta_k u_k\| = \|\delta_k u_{k+1} - u_k\|. \tag{3.8}$$

利用 $\beta_k^{MMLS*} \ge 0$，得 $\delta_k \ge 0$。利用上述等式和三角不等式，有

$$\|u_{k+1} - u_k\| \le \left\|\left(1+\delta_k\right) u_{k+1} - \left(1+\delta_k\right) u_k\right\| \le \|u_{k+1} - \delta_k u_k\| + \|\delta_k u_{k+1} - u_k\| = 2\|r_{k+1}\|. \tag{3.9}$$

再利用(3.1)和(3.7)，得 $\sum_{k\ge1} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} = \sum_{k\ge1} \|r_{k+1}\|^2 \|g_{k+1}\|^2 < \infty$。由(3.5)，则 $\sum_{k\ge1} \|r_{k+1}\|^2 < \infty$。

联立上述不等式和(3.9)，此引理结论成立。证毕。

下面的性质(P)由 Gilbert 和 Nocedal[18]给出，具体内容为：
性质(P)：如果

$$0<\gamma_1\le\|g_k\|\le\gamma_2. \tag{3.10}$$

若对所有 k，存在常数 $b>1$ 和 $\lambda>0$ 满足 $|\beta_k|\le b$ 和 $\|s_k\|\le\lambda$ 得到 $|\beta_k|\le\dfrac{1}{2b}$，我们说该方法满足性质(P)。

引理 3.3 设假设(A)满足，序列 $\{g_k\}$ 和 $\{d_k\}$ 由修改的 LS 算法产生，如果存在常数 $M>0$ 满足 $\|d_k\|\le M$，则 β_k^{MMLS*} 满足性质(P)。

证明：关于修改的 LS 方法，如果 $\dfrac{g_{k+1}^T y_k^m}{-d_k^T g_k}\le\dfrac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2}g_{k+1}^T d_k$ 成立，结论显然得证。否则，利用假设(A)-1)，则存在常数 $M_1>0$ 使得

$$\|s_k\|\le M_1 \tag{3.11}$$

成立。利用 $\|d_k\|\le M$，(3.1)，(3.2)，(3.6)，(3.10)和(3.11)可得

$$\begin{aligned}
\left|\beta_k^{MMLS*}\right| &\le\left|\frac{g_{k+1}^T y_k^m}{d_k^T g_k}\right|+\frac{\mu\|y_k^m\|^2}{\left(-d_k^T g_k\right)^2}\left|g_{k+1}^T d_k\right| \\
&\le\frac{\|g_{k+1}\|\left\|\dfrac{g_{k+1}\left(\|g_k\|-\|g_{k+1}\|\right)+\|g_{k+1}\|\left(g_{k+1}-g_k\right)}{\|g_k\|}\right\|}{c\|g_k\|^2} \\
&\quad+\frac{\mu\|g_{k+1}\|\|d_k\|\left\|\dfrac{g_{k+1}\left(\|g_k\|-\|g_{k+1}\|\right)+\|g_{k+1}\|\left(g_{k+1}-g_k\right)}{\|g_k\|}\right\|^2}{c^2\|g_k\|^4} \\
&\le\frac{2L\gamma_2^2\|s_k\|}{c\gamma_1^3}+\frac{2\mu\gamma_2^2ML^2M_1\|s_k\|}{c^2\gamma_1^5}=\left(\frac{2cL\gamma_2^2\gamma_1^2+2\mu\gamma_2^2ML^2M_1}{c^2\gamma_1^5}\right)\|s_k\|,
\end{aligned} \tag{3.12}$$

取 $b=\max\left\{2,\left(\dfrac{2cL\gamma_2^2\gamma_1^2+2\mu\gamma_2^2ML^2M_1}{c^2\gamma_1^5}\right)\right\}$ 和 $\lambda=\dfrac{c^2\gamma_1^5}{b\left(2cL\gamma_2^2\gamma_1^2+2\mu\gamma_2^2ML^2M_1\right)}$。

结合(3.12)和上述 b 和 λ 的取法以及 $b>1$，则 $\left|\beta_k^{MMLS*}\right|\le b$ 和

$$\left|\beta_k^{MMLS*}\right|\le\left(\frac{2cL\gamma_2^2\gamma_1^2+2\mu\gamma_2^2ML^2M_1}{c^2\gamma_1^5}\right)\|s_k\|\le\left(\frac{2cL\gamma_2^2\gamma_1^2+2\mu\gamma_2^2ML^2M_1}{c^2\gamma_1^5}\right)\lambda=\frac{1}{2b}.$$

所以修改的 LS 方法都具有性质(P)。证毕。

利用假设(A)，引理 3.1~3.3，与文献[9]中的定理 3.2 的证明类似，可得到修改的 LS 算法的全局收敛性定理，本文只给出此定理不再证明。

定理 3.1 假设(A)满足，序列 $\{d_k,x_k,g_k\}$ 由修改的 LS 算法产生，则 $\lim_{k\to\infty}\inf\|g_k=0\|$ 成立。

4. 数值结果

为了验证所给算法的有效性，下面给出数值检验结果，检验函数是在工程领域经常用到的 Benchmark 问题[19]，列举如下。

1) Sphere function $f_{Sph}(x)=\sum_{i=1}^{n}x_i^2, x_i\in[-5.12,5.12], x^*=(0,0,\cdots,0), f_{Sph}(x^*)=0$

2) Schwefel's function $f_{SchDS}(x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_j\right)^2, x_i\in[-65.536,65.536], x^*=(0,0,\cdots,0), f_{SchDS}(x^*)=0$

3) Rastrigin function $f_{Ras}(x)=10n+\sum_{i=1}^{n}\left(x_i^2-10\cos(2\pi x_i)\right), x_i\in[-5.12,5.12], x^*=(0,0,\cdots,0), f_{Ras}(x^*)=0$

4) Griewank function $f_{Gri}(x)=1+\sum_{i=1}^{n}\frac{x_i^2}{4000}-\prod_{i=1}^{n}\cos\frac{x_i}{i}, x_i\in[-600,600],\ x^*=(0,0,\cdots,0),\ f_{Gri}(x^*)=0$

在我们实际计算中，参数选取如下：$\varepsilon=10^{-5},\delta=0.1,\sigma=0.9$，停止准则采用 Himmeblau 准则：如果 $|f(x_k)|>e_1$, 令 $\mathrm{stop1}=\frac{|f(x_k)-f(x_{k+1})|}{|f(x_k)|}$；否则令 $\mathrm{stop1}=|f(x_k)-f(x_{k+1})|$。如果 $\|g(x)\|<\varepsilon$ 或者 $\mathrm{stop1}<e_2$ 满足，算法终止，$e_1=e_2=10^{-5}$。如果总的迭代次数超过 1000 次，我们也终止算法。为验证新方法的有效性，我们也给出通常 LS 方法的数值结果，并将两者进行比较。一般情况下，方法有效性的体现在算法执行时需要函数值次数的多少、梯度值次数的多少与所需要的 CPU 时间，我们也将从这几个方面进行比较，下面两表中各代码含义是：

x_0：初始点；Dim：问题的维数；NI：迭代次数；NFG：函数值与梯度值迭代次数和；Time：以秒为单位的计算机 CPU 时间；$f(\bar{x})$：算法终止时的函数值。

Table 1. The numerical results
表 1. 数值结果

		修改的 LS 算法		通常的 LS 算法	
		NI/NFG/ $f(\bar{x})$/Time	NI/NFG/ $f(\bar{x})$/Time	NI/NFG/ $f(\bar{x})$/Time	NI/NFG/ $f(\bar{x})$/Time
Sphere	x_0	$(-4,-4,\cdots,-4)$	$(3,3,\cdots,3)$	$(-4,-4,\cdots,-4)$	$(3,3,\cdots,3)$
	10	2/6/7.888609e − 030/0	3/7/1.972152e − 030/0	2/6/7.888609e − 030/0.047	3/7/1.972152e − 030/0
Dim	100	2/6/1.262177e − 027/0	2/6/1.597443e − 027/0	2/6/1.262177e − 027/0	2/6/1.597443e − 027/0
	300	2/6/2.863565e − 026/0	2/6/2.366583e − 026/0	2/6/2.863565e − 026/0	2/6/2.366583e − 026/0
Schwefel's	x_0	$(-0.001,\cdots,-0.001)$	$(0.0001,\cdots,-0.0001)$	$(-0.001,\cdots,-0.001)$	$(0.0001,\cdots,-0.0001)$
	10	3/8/3.539977e − 7/0	2/5/4.378593e − 8/0	3/8/3.539977e − 7/0.031	2/5/4.378593e − 8/0
Dim	100	5/14/1.889258e − 5/0.11	3/8/3.711002e − 6/0.047	5/14/1.889258e − 5/0.078	3/8/3.711002e − 6/0.0624
	300	8/23/3.264067e − 5/3.59	4/11/1.871432e − 5/1.825	8/23/3.263579e − 5/3.78	4/11/1.871432e − 5/1.763
Rastrigin	x_0	$(0.01,0.01,\cdots,0.01)$	$(0.001,\cdots,0.001)$	$(0.01,0.01,\cdots,0.01)$	$(0.001,\cdots,0.001)$
	10	3/10/0.000000e + 000/0	3/9/9.237056e − 012/0	3/11/0.000000e + 0/0	3/10/0.000000e + 0/0
Dim	100	3/9/2.273737e − 013/0	3/8/4.646120e − 008/0	3/10/0.000000e + 0/0	3/9/0.000000e + 0/0
	300	4/11/4.547474e − 013/0	3/8/1.477929e − 010/0	3/9/0.000000e + 0/0	3/9/0.000000e + 0/0.0312
Griewank	x_0	$(-100,-100,\cdots,-100)$	$(30,30,\cdots,30)$	$(-100,-100,\cdots,-100)$	$(30,30,\cdots,30)$
	10	6/32/4.360934e + 0/0	3/9/2.647550e − 6/0	5/31/1.198326e + 0/0	3/9/5.618497e − 6/0
Dim	100	3/9/2.795382e − 5/0.0312	2/6/0.000000e + 0/0	3/9/3.217599e − 5/0.0312	2/6/0.000000e + 0/0
	300	2/6/0.000000e + 0/0.0468	2/6/0.000000e + 0/0.0468	2/6/0.000000e + 0/0.0624	2/6/0.000000e + 0/0.0468
总的 CUP 时间		5.6968		5.933	

为更直观的比较修改的 LS 算法与原 LS 算法的效率，我们采用下述技术即两算法的所有 NFG 的乘积的比值在开二十四分之一(结果的总个数)次方，以通常 LS 算法作为底，定义为 1，此种技术详细可见文献[16]，则其比值如表 2。

Table 2. The efficiency of NFG comparison
表 2. 关于 NFG 的效率比较

修改的 LS 算法	通常的 LS 算法
0.9122	1

从表 1 的数值结果可以看出，修改的 LS 算法和通常的 LS 算法均能对上述 4 个问题有效地求解，迭代次数不是很多且非常接近最优函数值，新方法所需的 CPU 时间更少一些。从表 2 中 NFG 的效率来看，修改的 LS 算法相对于通常的 LS 算法，效率能够提高 8.78%。可以说，新方法更具竞争力。

致谢：本文受广西高等学校特色专业及课程一体化建设项目(管理科学专业)，广西自然科学基金项目(编号：2012GXNSFAA053002)和国家自然科学基金项目(编号：11261006 和 11161003)资助。同时感谢编者对论文提出的宝贵意见，使此论文更加完整。

参考文献 (References)

[1] E. Polak, G. Ribiere. Note sur la convergence de méthodes de directions conjugées. Revue Française d'Informatique et Recherche Opératinelle, Série Rouge Tome 3, 1969, 16(1): 35-43.
[2] B. T. Polyak. The conjugate gradient method in extreme problems. USSR Computational Mathematics and Mathematical Physics, 1969, 9: 94-112.
[3] R. Fletcher, C. M. Reeves. Function minimization by conjugate gradients. Computer Journal, 1964, 7(2): 149-154.
[4] R. Fletcher. Practical method of optimization, Volume I: Unconstrained optimization (2nd Edition). Hoboken: John Wiley & Sons, 2007.
[5] Y. Liu, C. Storey. Efficient generalized conjugate gradient algorithms, part 1: Theory. Journal of Optimization Theory and Application, 1992, 69(1): 17-41.
[6] M. R. Hestenes, E. Stiefel. Method of conjugate gradient for solving linear equations. Journal of Research of the National Bureau Standards, 1952, 49(6): 409-436.
[7] Y. Dai, Y. Yuan. A nonlinear conjugate gradient with a strong global convergence property. SIAM Journal on Optimization, 2000, 10(1): 177-182.
[8] Y. Dai, L. Z. Liao. New conjugacy conditions and related nonlinear conjugate methods. Applied Mathematics and Optimization, 2001, 43(1): 87-101.
[9] W.W. Hager, H. Zhang. A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM Journal on Optimization, 2005, 16(1): 170-192.
[10] W.W. Hager, H. Zhang. Algorithm 851: CGDESENT, a conjugate gradient method with guaranteed descent. ACM Transactions on Mathematical Software, 2006, 32: 113-137.
[11] G. Li, C. Tang and Z. Wei. New conjugacy condition and related new conjugate gradient methods for unconstrained optimization problems. Journal of Computational and Applied Mathematics, 2007, 202(2): 532-539.
[12] Z. Wei, G. Li and L. Qi. New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems. Applied Mathematics and Computation, 2006, 179(2): 407-430.
[13] Z. Wei, G. Li and L. Qi. Global convergence of the PRP conjugate gradient methods with inexact line search for nonconvex unconstrained optimization problems. Mathematics of Computation, 2008, 77: 2173-2193.
[14] Z. Wei , S. Yao and L. Lin. The convergence properties of some new conjugate gradient methods. Applied Mathematics and Computation, 2006, 183(2): 1341-1350.
[15] G. L. Yuan. Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. Optimization Letters, 2009, 3(1): 11-21
[16] G. L. Yuan, X. W. Lu. A modified PRP conjugate gradient method. Annals of Operations Research, 2009, 166(1): 73-90.
[17] G. L. Yuan, X. W. Lu and Z. X. Wei. A conjugate gradient method with descent direction for unconstrained optimization. Journal of Computational and Applied Mathematics, 2009, 233(2): 519-530.
[18] J.C. Gilbert, J. Nocedal. Global Convergence properties of conjugate gradient methods for optimization. SIAM Journal on Optimization, 1992, 2(1): 21-42.
[19] http://www.cs.cmu.edu/afs/cs/project/jair/pub/volume24/ortizboyer05a-html/node6.html

