Pure Mathematics 理论数学, 2011, 1, 1-7
http://dx.doi.org/10.12677/pm.2011.11001 Published Online April 2011 (http://www.hanspub.org/journal/pm/)

Modified DY and HS Conjugate Gradient Algorithms and Their Global Convergence

Xiangrong Li
College of Mathematics and Information Science, Guangxi University, Nanning
Email: xrli68@163.com
Received: Mar. 12th, 2011; revised: Mar. 20th, 2011; accepted: Mar. 21st, 2011.

Abstract: Yuan[15] proposed a modified PRP conjugate gradient method which can ensure that the scalar $\beta_k \ge 0$ holds and the search direction possesses the sufficient descent property without any line search. This technique has been extended to other conjugate gradient methods, but the convergence has been not given. In this paper, our purpose is to analyze the property of DY and HS: sufficient descent property and global convergence, moreover numerical results are shown.

Keywords: Conjugate Gradient Method; Sufficient Descent Property; Global Convergence

修改的 DY 和 HS 共轭梯度算法及其全局收敛性

李向荣
广西大学数学与信息科学学院，南宁
Email: xrli68@163.com
收稿日期：2011 年 3 月 12 日；修回日期：2011 年 3 月 20 日；录用日期：2011 年 3 月 21 日

摘　要：Yuan[15]提出了修改的 PRP 共轭梯度方法，该方法能保证参数 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。作者也将此技术推广到其它共轭梯度方法中，并给出了修改的公式，但是没有给出具体的收敛性证明。本文的主要工作就是分析修改的 DY 和 HS 共轭梯度方法的性质：充分下降性和全局收敛性，同时给出数值检验结果。

关键词：共轭梯度方法；分下降性；全局收敛性

1. 引言

求解下述问题

$$\min f(x)$$
$$x \in \Re^n \tag{1.1}$$

其中 $f(x)$ 连续可微。共轭梯度法拥有结构简单且效果明显的性质常被用于求解此问题，该方法采用下面的迭代公式

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

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 的选取不同而称为不同的共轭梯度法，下面我们给出几个著名的 β_k 的选取方法：

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

$$\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}, [4], \quad \beta_k^{LS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{-d_k^T g_k}, [5]$$

$\beta_k^{HS}=\dfrac{g_{k+1}^T(g_{k+1}-g_k)}{(g_{k+1}-g_k)^Td_k}$,[6], $\beta_k^{DY}=\dfrac{g_{k+1}^Tg_{k+1}}{(g_{k+1}-g_k)^Td_k}$,[7],

$g_k=\nabla f(x_k)$ 和 $g_{k+1}=\nabla f(x_{k+1})$，分别表示函数 $f(x)$ 在 x_k 和 x_{k+1} 的梯度值，$\|.\|$ 是欧氏向量范数。在上述方法中，**PRP**，**LS** 和 **HS** 的数值表现优越但收敛性不理想，**FR**，**CD** 和 **DY** 方法的收敛性好但数值表现不优越。其中，**PRP** 方法的数值表现最为理想，常常被人们用于实际的问题求解。许多学者都希望找到数值表现可与 **PRP** 相媲美同时性质又比其好的方法，已取得许多成果（见[8-18]等）。在文献[15]中，Yuang 给出了下述 **PRP** 修改公式：

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

其中 $\mu>\dfrac{1}{4}$ 是常数，$y_k=g_{k+1}-g_k$。该方法拥有充分下降性和全局收敛性，且数值表现优于 **PRP** 方法。在此公式的基础上，作者将此思想进行了推广，得到了下面的修改的 **DY** 和 **HS** 公式：

$$\beta_k^{MDY}=\beta_k^{DY}-\min\left\{\beta_k^{DY},\frac{\mu\|g_{k+1}\|^2}{\left(d_k^Ty_k\right)^2}g_{k+1}^Td_k\right\}\qquad(1.3)$$

和

$$\beta_k^{MHS}=\beta_k^{HS}-\min\left\{\beta_k^{HS},\frac{\mu\|y_k\|^2}{\left(d_k^Ty_k\right)^2}g_{k+1}^Td_k\right\}\qquad(1.4)$$

但没给出它们的性质分析，本文就是对上述两种方法进行分析，得到它们的充分下降性和全局收敛性。下一节我们将给出算法步骤，在第三部分分析收敛性，数值检验结果将在最后一节给出。

2. 算法

算法 1（修改的 DY 和 HS 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：

$$f(x_k+\alpha_kd_k)\le f_k+\delta\alpha_kg_k^Td_k\qquad(2.1)$$

和

$$g(x_k+\alpha_kd_k)^Td_k\ge\sigma g_k^Td_k\qquad(2.2)$$

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

步骤 4：利用下面公式计算搜索方向

$$d_{k+1}=-g_{k+1}+\beta_k^{MDY}d_k\qquad(2.3)$$

或

$$d_{k+1}=-g_{k+1}+\beta_k^{MHS}d_k\qquad(2.4)$$

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 DY 和 HS 方向具有充分下降性。

引理 3.1. 对 $k\ge0$，修改的 DY 和 HS 搜索方向对下式

$$d_k^Tg_k\le-c\|g_k\|^2\qquad(3.1)$$

和

$$d_k^Ty_k\ge c(1-\sigma)\|g_k\|^2\qquad(3.2)$$

满足，$c>0$ 是常数。

证明：If $k=0$，则 $g_0^Td_0=-\|g_0\|^2$，则(3.1)成立。假设当 $k\ge1$，(3.1)对修改的 DY(2.3)和 HS(2.4)均满足，对 $k+1$，利用(2.3)和(2.4)，分别得到

$$\begin{aligned}g_{k+1}^Td_{k+1}&=-\|g_{k+1}\|^2+\beta_k^{MDY}d_k^Tg_{k+1}=-\|g_{k+1}\|^2\\&+\left[\frac{\|g_{k+1}\|^2}{d_k^Ty_k}-\min\left\{\frac{\|g_{k+1}\|^2}{d_k^Ty_k},\frac{\mu\|g_{k+1}\|^2}{\left(d_k^Ty_k\right)^2}g_{k+1}^Td_k\right\}\right]d_k^Tg_{k+1}\end{aligned}\qquad(3.3)$$

和

$$\begin{aligned}g_{k+1}^Td_{k+1}&=-\|g_{k+1}\|^2+\beta_k^{MHS}d_k^Tg_{k+1}=-\|g_{k+1}\|^2\\&+\left[\frac{g_{k+1}^Ty_k}{d_k^Ty_k}-\min\left\{\frac{g_{k+1}^Ty_k}{d_k^Ty_k},\frac{\mu\|y_k\|^2}{\left(d_k^Ty_k\right)^2}g_{k+1}^Td_k\right\}\right]d_k^Tg_{k+1}\end{aligned}\qquad(3.4)$$

首先分析(3.4)，利用 y_k 的定义式(3.1)和(2.2)，有下式成立

$$d_k^Ty_k=d_k^T(g_{k+1}-g_k)\ge-(1-\sigma)g_k^Td_k>0\qquad(3.5)$$

取 $u=\dfrac{\sqrt{d_k^Ty_k}}{\sqrt{2\mu}}g_{k+1},v=\dfrac{\sqrt{2\mu}g_{k+1}^Td_k}{\sqrt{d_k^Ty_k}}y_k$。下面分两种

情况讨论(3.4)：

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

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

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

最后一个不等式利用了 $u^T v \le \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)$。令 $c = 1 - 1/4\mu$ 且联立(3.5)，则(3.1)和(3.2)均成立。综上所述，关于修改的 HS 的方法，此引理成立。对于(3.3)式，与上述证明过程类似，只需要在情形 1)和 2)的分析中，取

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

即可。所以此命题成立，证毕。

为了证明算法 1 的收敛性，需要下面的假设条件：

(A): (1) 水平集 $\Omega = \{x \in \Re^n : f(x) \le f(x_0)\}$ 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

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

我们采用反证法来得到全局收敛性，即假设存在常数 $\gamma > 0$ 满足

$$
\|g_k\| \ge \gamma, \forall k \ge 0, \tag{3.7}
$$

从上式导出矛盾，从而证明结论。

引理 3.2. 设假设(A)满足，序列 $\{g_k\}$ 和 $\{d_k\}$ 由算法 1 产生。如果不等式(3.7)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)，有

$$
-(1-\sigma) g_k^T d_k \le (g_{k+1} - g_k)^T d_k \le \alpha_k L \|d_k\|^2,
$$

结合(3.1)，得

$$
\alpha_k \ge \frac{1-\sigma}{L} \frac{|g_k^T d_k|}{\|d_k\|^2},
$$

将上式代入(2.1)，并利用假设(A)中函数 f 有下界，有下式

$$
\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \tag{3.8}
$$

成立。关系式(3.7)和引理 3.1 隐含着 $d_k \ne 0$，否则 $g_k = 0$，全局收敛性得证。因此 $u_k = \frac{d_k}{\|d_k\|}$ 是有意义的。令

$$
\delta_k = \beta_k^m \frac{\|d_k\|}{\|d_{k+1}\|}, r_{k+1} = -\frac{g_{k+1}}{d_{k+1}}.
$$

其中 $\beta_k^m = \beta_k^{MDY}$ 或 $\beta_k^m = \beta_k^{MHS}$。根据(2.3)(或(2.4))，当 $k \ge 1$，均有

$$
u_{k+1} = r_{k+1} + \delta_k u_k,
$$

再利用 $\|u_{k+1}\| = \|u_k\| = 1$，有

$$
\|r_{k+1}\| = \|u_{k+1} - \delta_k u_k\| = \|\delta_k u_{k+1} - u_k\| \tag{3.9}
$$

由 $\beta_k^m \ge 0$，得 $\delta_k \ge 0$。利用(3.9)和三角不等式，得到

$$
\begin{aligned}
\|u_{k+1} - u_k\| &\le \|(1+\delta_k) u_{k+1} - (1+\delta_k) u_k\| \\
&\le \|u_{k+1} - \delta_k u_k\| + \|\delta_k u_{k+1} - u_k\| \\
&= 2\|r_{k+1}\|
\end{aligned} \tag{3.10}
$$

利用(3.1)和(3.8)，有

$$
\sum_{k\ge1} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} = \sum_{k\ge1} \|r_{k+1}\|^2 \|g_{k+1}\|^2 < \infty
$$

考虑(3.7)，得

$$
\sum_{k\ge1} \|r_{k+1}\|^2 < \infty
$$

联立上述不等式和(3.10)，即获得此引理结论。证毕。

下面的性质(*)由 Gilbert 和 Nocedal[19]给出，具体内容为：

性质(*):如果

$$0 < \gamma_1 \le \|g_k\| \le \gamma_2 . \tag{3.11}$$

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 $|\beta_k| \le \frac{1}{2b}$，我们说该方法满足性质(*)。

引理 3.3. 设假设(A)满足，序列$\{g_k\}$和$\{d_k\}$由算法 1 产生，如果存在常数$M>0$满足$\|d_k\| \le M$，则β_k^m满足性质(*)，其中$\beta_k^m = \beta_k^{MDY}$ (或$\beta_k^m = \beta_k^{MHS}$)。

证明：对于修改的 HS 方法，即 $\beta_k^m =_k^{MHS}$，如果 $\frac{g_{k+1}^T y_k}{d_k^T y_k} \le \frac{\mu\|y_k\|^2}{(d_k^T y_k)^2} g_{k+1}^T d_k$ 成立，结论显然得证。否则，利用假设(A)(1),存在常数 $M_1>0$ 使得下式成立

$$\|s_k\| \le M_1 \tag{3.12}$$

利用 β_k^m，$\|d_k\| \le M, (3.1), (3.2), (3.6), (3.11)$和$(3.12)$，可得

$$\begin{aligned}
|\beta_k^m| &\le \left|\frac{g_{k+1}^T y_k}{d_k^T y_k}\right| + \frac{\mu\|y_k\|^2}{\left(d_k^T y_k\right)^2}\left|g_{k+1}^T d_k\right| \\
&\le \frac{\|g_{k+1}\|\|y_k\|}{c(1-\sigma)\|g_k\|^2} + \frac{\mu\|g_{k+1}\|\|d_k\|\|y_k\|^2}{c^2(1-\sigma)^2\|g_k\|^4} \\
&\le \frac{L\gamma_2\|s_k\|}{c(1-\sigma)\gamma_1^2} + \frac{\mu\gamma_2 ML^2 M_1\|s_k\|}{c^2(1-\sigma)^2\gamma_1^4} \\
&= \left(\frac{cL\gamma_2\gamma_1^2(1-\sigma)+\mu\gamma_2 ML^2 M_1}{c^2(1-\sigma)^2\gamma_1^4}\right)\|s_k\|
\end{aligned} \tag{3.13}$$

取 $b = \max\left\{2, \left(\frac{cL\gamma_2\gamma_1^2(1-\sigma)+\mu\gamma_2 ML^2 M_1}{c^2(1-\sigma)^2\gamma_1^4}\right)M_1\right\}$

和 $\lambda = \frac{c^2(1-\sigma)^2\gamma_1^4}{b\left(2cL\gamma_2\gamma_1^2(1-\sigma)+\mu\gamma_2 ML^2 M_1\right)}$

联立(3.13)和上述b和λ的取法以及$b>1$，得到

$$|\beta_k^m| \le b$$

和

$$\begin{aligned}
|\beta_k^m| &\le \left(\frac{cL\gamma_2\gamma_1^2(1-\sigma)+\mu\gamma_2 ML^2 M_1}{c^2(1-\sigma)^2\gamma_1^4}\right)\|s_k\| \\
&\le \left(\frac{cL\gamma_2\gamma_1^2(1-\sigma)+\mu\gamma_2 ML^2 M_1}{c^2(1-\sigma)^2\gamma_1^4}\right)\lambda = \frac{1}{2b}
\end{aligned}$$

对于修改的 DY 方法，即 $\beta_k^m = \beta_k^{MDY}$，我们只需得到 β_k^{MDY} 也具有式(3.13)的形式即可。

同样利用$\|d_k\| \le M, (3.1), (3.2), (3.6), (3.11)$和$(3.12)$得

$$\begin{aligned}
|\beta_k^m| = \beta_k^{MDY} &\le \left|\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right| + \frac{\mu\|g_{k+1}\|^2}{\left(d_k^T y_k\right)^2}\left|g_{k+1}^T d_k\right| \\
&\le \frac{\|g_{k+1}\|^2}{c(1-\sigma)\|g_k\|^2} + \frac{\mu\|g_{k+1}\|\|d_k\|\|g_{k+1}\|^2}{c^2(1-\sigma)^2\|g_k\|^4} \\
&\le \frac{\gamma_2^2\|s_k\|}{c(1-\sigma)\gamma_1^2} + \frac{\mu\gamma_2^3 MM_1\|s_k\|}{c^2(1-\sigma)^2\gamma_1^4} \\
&= \left(\frac{c\gamma_2^2\gamma_1^2(1-\sigma)+\mu\gamma_2^3 MM_1}{c^2(1-\sigma)^2\gamma_1^4}\right)\|s_k\|
\end{aligned} \tag{3.14}$$

余下的证明与修改的 HS 方法相同。综上所述，修改的 HS 和 DY 方法都具有性质(*)。利用假设(A)，引理 3.1～3.3，类似于文献 9 中的定理 3.2 的证明，我们不难得到算法 1 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$\{d_k, x_k, g_k\}$由算法 1 产生，则$\lim_{k\to\infty}\inf\|g_k = 0\|$成立。

本节主要对修改的 DY 和 HS 算法在理论上进行了分析，证明了算法的充分下降性和全局收敛性。下一节将从数值检验上说明修改的算法是有效的。

4. 数值结果

此部分给出数值结果，检验函数可从下面网页中找到：www.ici.ro/camo/neculai/SCALCG/testuo.pdf.参数的选取和终止条件均与文献[15]中的取值相同，终止条件是：$\|g(x_k)\| \le \varepsilon$ 或 $\|g(x_k)\| \le \varepsilon(1+|f(x_k)|)$，$\varepsilon = 1.0D-5$。参数取值：$\delta = 1.0D-4$，$\sigma = 1.0D-1$和$\mu = 0.5$。数值结果 Table 2 中的参数含义为：

P：问题名；fail：线搜索失败；Dim：问题维数；NI：迭代次数；NFN：函数和梯度次数和。

下面的结果也可从 http://blog.sina.com.cn/gonglinyuan 找到。

下面表格列举的是检验函数名称。

Table 1.

1	Extended Freudenstein and Roth	28	Extended Maratos Function
2	Extended Trigonometric Function	29	Extended Cliff
3	Extended White and Holst function	30	Quadratic Diagonal Perturbed Function
4	Diagonal3 Function	31	Extended Wood Function
5	Hager Function	32	Extended Hiebert Function
6	Extended Three Exponential Terms	33	Quadratic Function QF1
7	Generalized PSCI Function	34	Extended Quadratic Penalty QP2 Function
8	Extended Powell	35	A Quadratic Function QF2
9	Extended Quadratic Penalty QPI Function	36	Extended EPI Function
10	DIXMAANB (CUTE)	37	BDQRTIC(CUTE)
11	DIXMAANC (CUTE)	38	TRIDIA(CUTE)
12	Extended White and Holst function	39	ARWHEAD(CUTE)
13	Extended Beale Function U63 (MatrixRom)	40	NONDIA (Shanno-78) (CUTE)
14	Extended Penalty Function	41	DQDRTIC
15	Perturbed Quadratic function	42	EG2(CUTE)
16	Raydan 1 Function	43	DIXMAANA (CUTE)
17	Raydan 2 Function	44	DIXMAANE (CUTE)
18	Diagonal 1 Function	45	Partial Perturbed Quadratic
19	Diagonal 2 Function	46	Broyden Tridiagonal
20	Generalized Tridiagonal-1 Function	47	Almost Perturbed Quadratic
21	Extended Tridiagonal-1 Function	48	Tridiagonal Perturbed Quadratic
22	Generalized Tridiagonal-2	49	EDENSCH Function (CUTE)
23	Diagonal4 Function	50	LIARWHD (CUTE)
24	Diagonal5 Function (MatrixRom)	51	DIXMAANG (CUTE)
25	Extended Himmelblan Function	52	DIXMAANI (CUTE)
26	Extended PSCI Function	53	DIXMAANK (CUTE)
27	Extended Block Diagonal BDI Function	54	DIXMAANL (CUTE)
55	ENGVALI (CUTE)	56	FLETCHCR (CUTE)
57	COSINE (CUTE)	58	DENSCHNB (CUTE)
59	SINQUAD(CUTE)	60	Scaled Quadratic SQI
61	Scaled Quadratic SQ2		

Table 2.

		DY	MDY	HS	MHS			DY	MDY	HS	MHS
P	Dim	NI/NHN	NI/NHN	NI/NFN	NI/NFN	P	Dim	NI/NFN	NI/NFN	NI/NFN	NI/NFN
	500	8/27	9/30	7/18	9/26	29	1000	5/15	2/9	fail	2/9
1	5000	9/23	15/36	5/13	6/19		10000	4/11	2/9	5/13	3/10
	100000	6/21	11/31	5/14	5/18	30	1000	86/201	95/232	83/186	96/23 0
	500	85/190	188/393	21/64	47/116		10000	376/922	471/1164	713/1639	317/764
2	5000	120/264	106/237	60/164	31/87	31	1000	1030/2072	278/570	292/612	291/602
	100000	128/306	64/161	56/155	42/117		10000	fail	1176/2362	65/161	275/570
	500	176/380	276/584	fail	22/61	32	1000	55/134	27/84	fail	33/118
3	5000	17/49	165/343	fail	16/52		10000	10/33	9/34	8/34	9/33
	10000	380/789	322/677	fail	21/64	33	1000	166/335	166/335	166/335	166/335
P	Dim	NI/NHN	NI/NHN	NI/NFN	NI/NFN	P	Dim	NI/NFN	NI/NFN	NI/NFN	NI/NFN
	500	450/904	346/696	60/134	58/119		10000	542/1088	542/1088	527/1058	527/1058
4	5000	35/76	35/76	14/34	14/34	34	1000	2013/4068	1909/3849	fail	20/65
	10000	92/190	25/56	30/64	9/24		10000	fail	fail	fail	28/92
	500	38/82	24/54	17/42	17/40	35	1000	1029/2064	1038/2082	209/421	221/453
5	5000	15/36	11/29	9/32	8/23		10000	1242/2488	811/1628	685/1375	691/1387
	10000	23/55	6/21	7/28	5/19	36	1000	1/3	1/3	1/3	1/3

	500	6/18	6/18	5/20	5/18		10000	1/3	1/4	1/3	1/4
6	5000	7/20	4/14	6/18	4/14	37	1000	144/303	94/229	243/503	43/121
	10000	4/16	4/16	5/17	3/14		10000	130/294	85/225	392/817	187/485
	500	13/34	19/49	26/79	15/41	38	1000	326/655	327/657	326/655	326/655
7	5000	13/38	13/37	12/56	12/35		10000	1071/2146	1072/2148	1072/2148	1073/2150
	10000	11/35	9/34	7/39	8/30	39	1000	91/209	34/84	12/31	12/42
	500	503/1011	503/1009	47/105	216/462		10000	fail	29/90	11/30	fail
8	5000	1040/2084	1121/2246	222/495	58/125	40	1000	499/1010	18/60	19/47	16/43
	10000	882/1768	222/472	90/195	148/323		10000	48/149	fail	fail	23/73
	500	11/33	2/18	18/120	2/18	41	1000	5/14	5/14	5/14	5/14
9	5000	4/20	fail	fail	3/13		10000	6/17	6/17	6/17	6/17
	10000	4/16	fail	fail	4/26	42	1000	179/365	85/180	fail	27/77
	500	6/22	6/22	7/24	6/23		10000	26/65	30/76	fail	15/51
10	5000	8/21	6/20	9/41	5/20	43	1000	9/22	8/23	9/28	7/22
	10000	14/34	5/21	fail	6/22		10000	11/28	6/21	8/26	6/20
	500	7/21	7/22	11/31	7/21	44	1000	101/207	99/203	70/146	118/241
11	5000	7/22	6/21	9/33	6/23		10000	279/563	272/549	283/572	277/559
	10000	14/35	7/25	fail	8/25	45	1000	133/269	132/267	132/267	132/267
12	1000	69/160	108/252	27/76	33/97	46	10000	fail	fail	fail	fail
	10000	101/226	76/185	25/73	33/100		1000	25/57	92/193	29/69	30/70
13	1000	1015/2051	64/166	fail	36/120	47	10000	62/135	55/124	30/74	27/70
	10000	159/327	25/70	41/134	14/41		1000	170/343	170/343	170/343	170/343
14	1000	7/29	7/30	12/62	7/30	48	10000	550/1104	550/1104	541/1086	541/1086
	10000	8/29	3/24	11/45	5/26		1000	160/323	160/323	16//325	161/325
15	1000	170/343	170/343	170/343	170/343	49	10000	511/1026	511/1026	512/1028	512/1028
	10000	549/1102	549/1102	540/1084	540/1084		1000	9/25	8/26	9/31	7/24
16	1000	33/71	29/64	25/56	23/52	50	10000	7/25	5/23	8/40	5/23
	10000	9/24	8/23	11/29	6/19		1000	49/126	41/108	18/41	15/42
17	1000	3/11	1/8	3/11	1/8	51	10000	201/444	35/101	17/44	18/52
	10000	1/8	1/9	1/8	1/9		1000	fail	134/278	104/218	104/217
18	1000	72/153	47/105	26/65	23/58	52	10000	287/579	287/579	273/556	252/509
	10000	69/151	20/52	35/93	9/31		1000	99/203	97/199	65/136	118/242
19	1000	191/389	240/487	95/198	75/159	53	10000	275/555	269/543	277/562	268/541
	10000	3263/6568	670/1368	218/456	208/442		1000	182/380	202/410	fail	110/226
20	1000	18/43	18/44	16/39	16/39	54	10000	417/840	612/1233	fail	262/531
	10000	15/39	15/39	13/37	13/35		1000	1002/2009	1004/2013	1003/2014	1002/2010
21	1000	28/65	19/47	7/25	7/25	55	10000	711/1426	799/1603	1143/2298	586/1178
	10000	57/119	19/47	15/38	5/15		1000	11/31	10/29	10/31	10/29
22	1000	306/635	223/451	34/74	41/87	56	10000	10/29	6/28	11/59	6/28
	10000	151/313	49/112	39/88	63/137		1000	45/97	40/87	33/77	39/85
23	1000	2/7	2/7	2/7	2/7	57	10000	3157/6323	2659/5328	2494/5006	3378/6766
	10000	2/8	2/8	2/8	2/8		1000	173/356	4/13	fail	4/13
24	1000	2/7	1/8	Fail	1/8	58	10000	3/11	3/14	17/51	3/12
	10000	2/9	1/7	2/9	1/7		1000	13/30	7/21	fail	10/24
25	1000	20/46	20/46	8/22	7/20	59	10000	69/145	62/129	10/26	10/28
	10000	20/47	20/47	9/25	7/21		1000	563/1346	162/439	fail	95/267
26	1000	101/211	6/21	8/32	7/23	60	10000	1227/2505	100/266	fail	1093/2877
	10000	374/757	179/371	fail	5/24		1000	166/335	166/335	166/335	166/335
27	1000	38/87	13/59	fail	18/63	61	10000	541/1086	541/1086	166/335	527/1058
	10000	fail	fail	12/31	22/73		1000	39/81	39/81	39/81	39/81
28	1000	213/445	243/507	fail	27/84		10000	129/262	129/262	129/262	129/262
	10000	17/45	65/155	fail	28/92						

从上述数值结果可以看出，修改的方法确实优于原方法。

5. 致谢

本文受广西高校优秀人才资助计划项目，中国国家自然科学基金项目(10761001)和广西教育厅项目(201012MS013)资助。

参考文献 (References)

[1] E. Polak, G. Ribiere. Note sur la xonvergence de directions conjugees. Rev. Francaise informat Recherche Operatinelle 3e Annee, 1969, 16: 35-43.
[2] B. T. Polyak. The conjugate gradient method in extreme problems. USSR Comp Math Math Phys., 1969, 9(4): 94-112.
[3] R. Fletcher, C. Reeves. Function minimization by conjugate gradients. Computer Journal, 1964, 7(2): 149-154.
[4] R. Fletcher. Practical Method of Optimization. New York: John Wiley & Sons, 1998.
[5] Y. Liu, C. Storey. Efficient generalized conjugate gradient algorithms, part 1: theory. Journal of optimization theory and Application, 1992, 69(1):129-137.
[6] M. R. Hestenes, E. Stiefel. Method of conjugate gradient for solving linear equations. Res. Nat. Bur. Stand., 1952, 49(6): 409-436.
[7] Y. H. Dai, Y. Yuan. A nonlinear conjugate gradient method with a strong global convergence property. SIAM Journal on Optimization, 2000, 10(177): 177-182.
[8] Y. Dai, L. Z. Liao. New conjugacy conditions and related nonlinear conjugate methods. Appl. Math. Optim., 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(1): 113-137.
[11] G. Li, C. Tang, Z. Wei. New conjugacy condition and related new conjugate gradient methods for unconstrained optimization problems. Journal of Computational and Applied Mathemathics, 2007, 202(2): 532-539.
[12] Z. Wei, G. Li, 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, L. Qi. Global convergence of the PRP conjugate gradient methods with inexact line search for nonconvex unconstrained optimization problems. Mathematics of Computation, 2008, 77(264): 2173-2193.
[14] Z. Wei, S. Yao, 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, 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] L. Zhang, W. Zhou, D. Li. A descent modified Polak-Ribiere-Polyak conjugate method and its global convergence. IMA Journal on Numerical Analysis, 2006, 26(4): 629-649.
[19] J. C. Gilbert, J. Nocedal. Global Convergence properties of conjugate gradient methods for optimization. SIAM Journal on Optimization, 1992, 2(1): 21-42.