Modified LS Method for Unconstrained Optimization*  

Jinkui Liu1, Li Zheng2  
1College of Mathematics and Computer Science, Chongqing Three Gorges University, Chongqing, China  
2Chongqing Energy College, Chongqing, China  
E-mail: liujinkui2006@126.com  
Received December 28, 2010; revised May 6, 2011; accepted May 9, 2011

Abstract  
In this paper, a new conjugate gradient formula and its algorithm for solving unconstrained optimization problems are proposed. The given formula satisfies \( k \) satisfying the descent condition. Under the Grippo-Lucidi line search, the global convergence property of the given method is discussed. The numerical results show that the new method is efficient for the given test problems.

Keywords: Unconstrained Optimization, Conjugate Gradient Method, Grippo-Lucidi Line Search, Global Convergence

1. Introduction  
The primary objective of this paper is to study the global convergence properties and practical computational performance of a new conjugate gradient method for nonlinear optimization without restarts, and with suitable conditions.

Consider the following unconstrained optimization problem:

\[
\min_{x \in \mathbb{R}^n} f(x),
\]

where \( f: \mathbb{R}^n \rightarrow \mathbb{R} \) is smooth and its gradient \( g \) is available. LS conjugate gradient method for solving unconstrained optimization problem is iterative formulas of the form

\[
x_{k+1} = x_k + \alpha_k d_k, \quad (1.1)
\]

\[
d_k = \begin{cases} -g_k, & \text{for } k = 1; \\ -g_k + \beta_k d_{k-1}, & \text{for } k \geq 2, \end{cases} \quad (1.2)
\]

where \( x_k \) is the current iterate, \( \alpha_k \) is a positive scalar and called the steplength which is determined by some line search, \( d_k \) is the search direction, \( g_k \) is the gradient of \( f \) at \( x_k \), and \( \beta_k \) is a scalar and

\[
\beta_k = -\frac{g^T_k (g_k - g_{k-1})}{d^T_k g_{k-1}}, \quad (\text{Liu-Storey (LS)}), \tag{1.6}
\]

[2] proved the global convergence of the LS method with Grippo-Lucidi line search. And the Grippo-Lucidi line search is to compute

\[
\alpha_k = \max \left\{ \rho^j \frac{\tau \|d_k\|^2}{\|g_k\|^2}, j = 0, 1, 2, \ldots \right\} \quad (1.3)
\]

satisfying:

\[
f(x_k + \alpha_k d_k) - f(x_k) \leq -\delta \alpha_k^2 \|d_k\|^2, \quad (1.4)
\]

\[-c_2 \|g_{k+1}\| \leq g^T_k d_{k+1} \leq -c_1 \|g_{k+1}\|^2, \quad (1.5)\]

where \( \delta > 0 \), \( \tau > 0 \), \( \rho \in (0, 1) \) and \( 0 < c_1 < 1 < c_2 \).

It is well known that some other people have studied many of the variants of the LS method, for example [3-4]. In this paper, a kind of the LS method is proposed:

\[
\beta_{k+1} = \frac{g^T_k (g_k - t_k g_{k+1})}{d^T_{k+1} g_{k+1}}, \quad (1.6)
\]

where \( t_k = \frac{\|g_k\|}{\|g_{k+1}\|} \) and \( \|\cdot\| \) is the Euclidean norm.

In the next section, we prove the global convergence of the new method for nonconvex functions with the Grippo-Lucidi line search. In Section 3, numerical experiments are given.

2. Global Convergence of the New Method

In order to prove the global convergence of the new method, we assume that the objective function satisfies
the following assumption.

Assumption (H):
1) The level set \( N = \{ x | f(x) \leq f(x_0) \} \) is bounded, where \( x_0 \) is the starting point.
2) In some neighborhood \( W \) of \( N \), the objective function is continuously differentiable, and its gradient is Lipschitz continuous, \( i.e. \), there exists a constant \( L > 0 \) such that
\[
\| g(x) - g(y) \| \leq L \| x - y \|, \quad \text{for all } x, y \in W. \tag{2.1}
\]

Lemma 2.1 \([5]\). Suppose Assumption (H) holds. Consider any iteration in the form (1.1) and (1.2), where \( d_k \) satisfies \( g_k^T d_k < 0 \) for \( k \in \mathbb{N}^+ \) and \( \alpha_k \) satisfies Grippo-Lucidi line search. Then
\[
\sum_{k=1}^{\infty} \cos^2 \theta_k \| g_{k+1} \|^2 < +\infty. \tag{2.2}
\]
where \( \cos \theta_k = -g_k^T d_k / (\| g_k \| \cdot \| d_k \|) \) and \( \theta_k \) is the angle between \( -g_k \) and \( d_k \).

\[
\begin{align*}
 g_{k+1}^T d_{k+1} + \| g_{k+1} \|^2 &\leq \beta_{k+1}^2 \| g_{k+1} \|^2 \| g_{k+1} \| \| d_k \| \\
 &\leq \| g_{k+1} \|^2 \cdot (\| g_{k+1} - g_k + g_k - g_k \|) \| d_k \| \\
 &\leq \| g_{k+1} \|^2 \cdot (\| g_{k+1} - g_k \| + \| g_k \| \cdot \| d_k \|) \| d_k \|
\end{align*}
\]

So (1.5) holds, for any \( \alpha_k \in \left( 0, c_1 \right] \).\( ^T \)

On the other hand, by the mean value theorem and Lipschitz condition (2.1), we have
\[
\begin{align*}
f(x_k + \alpha_k d_k) - f(x_k) &= \int_0^1 g(x_k + \tau \alpha_k d_k) \tau d\tau \\
&= \int_0^{\alpha_k} g(x_k + \tau \alpha_k d_k) \tau d\tau \\
&= \alpha_k g_k^T d_k + \int_0^{\alpha_k} \left( g(x_k + \tau \alpha_k d_k) - g_k \right) \tau d\tau \\
&\leq \alpha_k g_k^T d_k + \frac{1}{2} L \alpha_k^2 \| d_k \|^2.
\end{align*}
\]

We can test (1.4) holds, for \( \alpha_k \in \left( 0, 2 / (L + 28) \right] \).

Theorem 2.1. Suppose that Assumption (H) holds. Consider the method of form (1.1) and (1.2), where

\[
\begin{align*}
\beta_k &= \beta_k^{VLS}, \quad \text{and where } \alpha_k \text{ satisfies Grippo-Lucidi line search. Then}
\end{align*}
\]

\[
\begin{align*}
\liminf_{k \to \infty} \| g_k \| = 0.
\end{align*}
\]

Proof. By Lipschitz condition (1.2), (1.3), (1.5) and (2.1), we can obtain
\[
\begin{align*}
\| d_k \| &\leq \| g_k \| + \| g_{k+1} - g_k \| \\
&\leq \| g_k \| \left( 1 + \frac{\| g_{k+1} - g_k \|}{\| d_{k+1} \|} \right) \\
&\leq \| g_k \| \left( 1 + \frac{\| g_{k+1} - g_k \| + \| g_k \|}{\| d_{k+1} \|} \right) \\
&\leq \| g_k \| \left( 1 + \frac{2L \alpha_k}{\| d_{k+1} \|} \right)
\end{align*}
\]

By the Assumption (H), we know that Lemma 3.1 holds. From (1.5), (2.2) and (2.4), we have
Table 1. The performance of DY method, LS method and VLS method.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Dim</th>
<th>DY</th>
<th>LS</th>
<th>VLS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beale</td>
<td>2</td>
<td>75/186/164</td>
<td>18/65/55</td>
<td>25/72/64</td>
</tr>
<tr>
<td>Box Three-Dimensional</td>
<td>3</td>
<td>1/1/1</td>
<td>1/1/1</td>
<td>1/1/1</td>
</tr>
<tr>
<td>Penalty1</td>
<td>50</td>
<td>1727/2117/2043</td>
<td>85/426/315</td>
<td>65/112/98</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>31/157/121</td>
<td>18/120/83</td>
<td>22/146/119</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>26/160/121</td>
<td>28/157/114</td>
<td>20/124/93</td>
</tr>
</tbody>
</table>

This result implies \( \lim \inf_{k \to +\infty} \| g_k \| = 0 \).

3. Numerical Results

In this section, we give the new algorithm.

Algorithm 3.1:

Step 1: Data: \( x_0 \in \mathbb{R}^n \), \( \varepsilon \geq 0 \). Set \( d_1 = -g_1 \), if \( \| g_1 \| \leq \varepsilon \), then stop.

Step 2: Compute \( \alpha_k \) by the Grippo-Lucidi line searches.

Step 3: Let \( x_{k+1} = x_k + \alpha_k d_k, g_{k+1} = g(x_{k+1}) \), if \( \| g_{k+1} \| \leq \varepsilon \), then stop.

Step 4: Compute \( \beta_{k+1} \) by (1.6), and generate \( d_{k+1} \) by (1.2).

Step 5: Set \( k = k + 1 \), go to step 2.

We test the Algorithm 3.1 on the following problems,

and compare its performance to that of the DY method and LS method with the strong Wolfe line searches where \( \alpha_k \) is computed by

\[
\begin{align*}
&f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (3.1) \\
&\| g(x_k + \alpha_k d_k) \| \leq - \sigma g_k^T d_k. \quad (3.2)
\end{align*}
\]

In algorithm, the parameters: \( \tau = 1.5 \), \( \rho = 0.5 \), \( c_1 = 0.25 \), \( c_2 = 1.5 \), \( \delta = 0.01 \), \( \sigma = 0.1 \). The termination condition is \( \| g_k \| \leq 10^{-6} \), or It-max > 9999. It-max denotes the Maximum number of iterations.

The numerical results of our tests are reported in Table 1. The column “Problem” represents the problem’s name; “Dim” denotes the dimension of the tested problems. The detailed numerical results are listed in the form NI/NF/NG, where NI, NF, NG denote the number of iterations, function evaluations, and gradient evaluations, respectively.

VLS method: \( \beta_k = \beta_k^{\text{VLS}}, \alpha_k \) by the Grippo-Lucidi line searches

LS method: \( \beta_k = \beta_k^{\text{LS}}, \alpha_k \) by the strong Wolfe line searches.

DY method: \( \alpha_k \) by the strong Wolfe line searches,

\[
\beta_k \text{ is computed by } \beta_k^{\text{LS}} = \frac{\| g_k \|}{d_k^T (g_k - g_{k-1})}.
\]

In the following, we give the tested functions:

1) Beale Test Function:

\[
f(x) = \left[ 1.5 - x_1 (1 - x_2)^2 \right] + \left[ 2.25 - x_1 (1 - x_2^2)^2 \right] + \left[ 2.625 - x_1 (1 - x_2^3)^2 \right],
\]

the initial point \( (1,1)^T \).

2) Box Three-Dimensional Test Function:

\[
f(x) = \sum_{i=1}^{3} \left[ e^{-0.1(x_i - e^{-0.1x_i})} - x_i (e^{-0.1x_i} - e^{-0.1}) \right]^2,
\]

the initial point \( (0,10,20)^T \).

3) Penalty Test Function I:

\[
f(x) = 10^3 \sum_{i=1}^{m} \left( x_i - 1 \right)^2 + \left( \sum_{i=1}^{m} x_i^2 - 0.25 \right)^2,
\]

the initial point \( (1,2,\ldots,m)^T \).

From the numerical results, we know that the new method is efficient for the given problems under the Grippo-Lucidi line searches.

4. Acknowledgements

We are grateful to anonymous referees and editors for their useful suggestions and comments on this paper.

5. References


