Applied Mathematics, **2011, 2, 779-782**
doi:10.4236/am.2011.26104 Published Online June 2011 (http://www.SciRP.org/journal/am)

Modified LS Method for Unconstrained Optimization*

Jinkui Liu[1], Li Zheng[2]
[1]*College of Mathematics and Computer Science, Chongqing Three Gorges University, Chongqing, China*
[2]*Chongqing 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 β_k^{VLS} and its algorithm for solving unconstrained optimization problems are proposed. The given formula β_k^{VLS} satisfies $\beta_k^{VLS} \geq 0$ with d_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:

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

where $f: R^n \to R$ is smooth and its gradient g is available. LS conjugate gradient method for solving unconstrained optimization problem is iterative formulas of the form

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.1}$$

$$d_k = \begin{cases} -g_k, & \text{for } k = 1; \\ -g_k + \beta_k d_{k-1}, & \text{for } k \geq 2, \end{cases} \tag{1.2}$$

where x_k is the current iterate, α_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 β_k is a scalar and

$$\beta_k^{LS} = -\frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{d_{k-1}^{\mathrm{T}} g_{k-1}}, \text{ (Liu-Storey (LS) [1])},$$

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

$$\alpha_k = \max\left\{\rho^j \frac{\tau \left|g_k^{\mathrm{T}} d_k\right|}{\|d_k\|^2}, j = 0,1,2,\cdots\right\} \tag{1.3}$$

satisfying :

$$f(x_k + \alpha_k d_k) - f(x_k) \leq -\delta \alpha_k^2 \|d_k\|^2, \tag{1.4}$$

$$-c_2 \|g_{k+1}\|^2 \leq g_{k+1}^{\mathrm{T}} d_{k+1} \leq -c_1 \|g_{k+1}\|^2, \tag{1.5}$$

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:

$$\beta_k^{\mathrm{VLS}} = -\frac{g_k^{\mathrm{T}}(g_k - t_k g_{k-1})}{d_{k-1}^{\mathrm{T}} g_{k-1}}, \tag{1.6}$$

where $t_k = \dfrac{\|g_k\|}{\|g_{k-1}\|}$,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

*This work was supported by The Nature Science Foundation of Chongqing Education Committee (KJ091104, KJ101108).

the following assumption.

Assumption (H):

1) The level set $N = \{x | f(x) \le f(x_1)\}$ is bounded, where x_1 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

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \text{for all } x, y \in W. \tag{2.1}$$

Lemma 2.1 [5]. Suppose Assumption (H) holds. Consider any iteration in the form (1.1) and (1.2), where d_k satisfies $g_k^{\mathrm{T}} d_k < 0$ for $k \in N^+$ and α_k satisfies Grippo-Lucidi line search. Then

$$\sum_{k \ge 1} \cos^2 \theta_k \|g_k\|^2 < +\infty . \tag{2.2}$$

where $\cos\theta_k = -g_k^{\mathrm{T}} d_k / (\|g_k\| \cdot \|d_k\|)$ and θ_k is the angle between $-g_k$ and d_k.

The following Lemma shows that the Grippo-lucidi line search is suitable for the new formula.

Lemma 2.2. Suppose that Assumption (H) holds. Consider the method of form (1.1) and (1.2), where $\beta_k = \beta_k^{\mathrm{VLS}}$, and where α_k satisfies Grippo-Lucidi line search. Then $\forall k$, there exists a constant $c > 0$ such that $\alpha_k \ge c \dfrac{|g_k^{\mathrm{T}} d_k|}{\|d_k\|^2}$.

Proof. Since $d_1 = -g_1$, (1.5) holds for $k = 1$. Suppose that (1.5) holds for $k \ge 1$.

Denote

$$c_3 = \frac{\min(1 - c_1, c_2 - 1)}{2L} > 0 . \tag{2.3}$$

By (1.2), Lipschitz condition (2.1) and (1.5), for any $\alpha_k \in \left(0, c_3 \dfrac{|g_k^{\mathrm{T}} d_k|}{\|d_k\|^2}\right)$, we have

$$\begin{aligned}
\left| g_{k+1}^{\mathrm{T}} d_{k+1} + \|g_{k+1}\|^2 \right| &\le \left|\beta_{k+1}^{VLS}\right| \cdot \left|g_{k+1}^{\mathrm{T}} d_k\right| = \frac{\left|g_{k+1}^{\mathrm{T}}(g_{k+1} - t_{k+1} g_k)\right|}{\left|-d_k^{\mathrm{T}} g_k\right|} \cdot \left|g_{k+1}^{\mathrm{T}} d_k\right| \le \frac{\|g_{k+1}\|^2 \cdot (\|g_{k+1} - g_k + g_k - t_{k+1} g_k\|) \cdot \|d_k\|}{\left|-d_k^{\mathrm{T}} g_k\right|} \\
&\le \frac{\|g_{k+1}\|^2 \cdot (\|g_{k+1} - g_k\| + \|g_k - t_{k+1} g_k\|) \cdot \|d_k\|}{\left|-d_k^{\mathrm{T}} g_k\right|} \le \frac{\|g_{k+1}\|^2 \cdot (\|g_{k+1} - g_k\| + \left|\|g_{k+1}\| - \|g_k\|\right|) \cdot \|d_k\|}{\left|-d_k^{\mathrm{T}} g_k\right|} \\
&\le \frac{\|g_{k+1}\|^2 \cdot 2\|g_{k+1} - g_k\| \cdot \|d_k\|}{\left|-d_k^{\mathrm{T}} g_k\right|} \le \frac{\|g_{k+1}\|^2 \cdot 2L\alpha_k \cdot \|d_k\|^2}{\left|-d_k^{\mathrm{T}} g_k\right|} \le \min(1 - c_1, c_2 - 1) \cdot \|g_{k+1}\|^2 .
\end{aligned}$$

So (1.5) holds, for any $\alpha_k \in \left(0, c_3 \dfrac{|g_k^{\mathrm{T}} d_k|}{\|d_k\|^2}\right)$.

On the other hand, by the mean value theorem and Lipschitz condition (2.1), we have

$$\begin{aligned}
& f(x_k + \alpha_k d_k) - f(x_k) \\
&= \int_0^1 g(x_k + t\alpha_k d_k)^{\mathrm{T}} (\alpha_k d_k) \mathrm{d}t \\
&= \alpha_k g_k^{\mathrm{T}} d_k + \int_0^1 \left[g(x_k + \alpha_k d_k) - g_k \right]^{\mathrm{T}} (\alpha_k d_k) \mathrm{d}t \\
&\le \alpha_k g_k^{\mathrm{T}} d_k + \frac{1}{2} L \alpha_k^2 \|d_k\|^2 .
\end{aligned}$$

We can test (1.4) holds, for $\alpha_k \in \left(0, \dfrac{2}{L + 2\delta} \dfrac{|g_k^{\mathrm{T}} d_k|}{\|d_k\|^2}\right)$.

The existence of α_k satisfying (1.4) and (1.5) has bean proved. Furthermore, the conclusion holds for $c = \min\left(\tau, c_3, \dfrac{2}{L + 2\delta}\right)$.

Theorem 2.1. Suppose that Assumption (H) holds. Consider the method of form (1.1) and (1.2), where $\beta_k = \beta_k^{\mathrm{VLS}}$, and where α_k satisfies Grippo-Lucidi line search. Then

$$\liminf_{k \to +\infty} \|g_k\| = 0 .$$

Proof. By Lipschitz condition (1.2), (1.3), (1.5) and (2.1), we can obtain

$$\begin{aligned}
\|d_k\| &\le \|g_k\| + \left|\beta_k^{\mathrm{VLS}}\right| \cdot \|d_{k-1}\| \\
&\le \|g_k\| \cdot \left(1 + \frac{\|g_k - t_k g_{k-1}\|}{\left|-d_{k-1}^T g_{k-1}\right|} \cdot \|d_{k-1}\|\right) \\
&\le \|g_k\| \cdot \left(1 + \frac{(\|g_k - g_{k-1}\| + \left|\|g_k\| - \|g_{k-1}\|\right|) \cdot \|d_{k-1}\|}{\left|-d_{k-1}^{\mathrm{T}} g_{k-1}\right|}\right) \\
&\le \|g_k\| \cdot \left(1 + \frac{2L\alpha_{k-1}}{\left|-d_{k-1}^{\mathrm{T}} g_{k-1}\right|} \|d_{k-1}\|^2\right) \\
&= (1 + 2L\tau) \cdot \|g_k\|.
\end{aligned} \tag{2.4}$$

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.

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

$$\infty > \sum_{k\ge1}\cos^2\theta_k\|g_k\|^2 = \sum_{k\ge1}\frac{\left(g_k^{\mathrm{T}}d_k\right)^2}{\|d_k\|^2} \ge c_1^2\left(1+2L\tau\right)^{-2}\sum_{k\ge1}\|g_k\|^2.$$

This result implies $\liminf_{k\to+\infty}\|g_k\|=0$.

3. Numerical Reusults

In this section, we give the new algorithm.

Algorithm 3.1:

Step 1: Data: $x_1\in R^n$, $\varepsilon\ge0$. Set $d_1=-g_1$, if $\|g_1\|\le\varepsilon$, then stop.

Step 2: Compute α_k by the Grippo-Lucidi line searches.

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

Step 4: Compute β_{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 α_k is computed by

$$f\left(x_k+\alpha_k d_k\right)\le f\left(x_k\right)+\delta\alpha_k g_k^{\mathrm{T}}d_k, \tag{3.1}$$

$$\left|g\left(x_k+\alpha_k d_k\right)^{\mathrm{T}}d_k\right|\le-\sigma g_k^{\mathrm{T}}d_k. \tag{3.2}$$

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 $\|g_k\|\le10^{-6}$, 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: $\beta_k=\beta_k^{\mathrm{VLS}}$, α_k by the Grippo-Lucidi line searches

LS method: $\beta_k=\beta_k^{\mathrm{LS}}$, α_k by the strong Wolfe line searches.

DY method: α_k by the strong Wolfe line searches, β_k is computed by $\beta_k^{\mathrm{LS}}=\dfrac{\|g_k\|^2}{d_{k-1}^{\mathrm{T}}\left(g_k-g_{k-1}\right)}$.

In the following, we give the tested functions:

1) Beale Test Function:

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

the initial point $\left(1,1\right)^{\mathrm{T}}$.

2) Box Three-Dimensional Test Function:

$$f\left(x\right)=\sum_{i=1}^{3}\left[\mathrm{e}^{-0.1ix_1}-\mathrm{e}^{-0.1ix_2}-x_3\left(\mathrm{e}^{-0.1i}-\mathrm{e}^{-i}\right)\right]^2,$$

the initial point $\left(0,10,20\right)^{\mathrm{T}}$.

3) Penalty Test Function I:

$$f\left(x\right)=10^{-5}\sum_{i=1}^{n}\left(x_i-1\right)^2+\left(\sum_{i=1}^{n}x_i^2-0.25\right)^2,$$

the initial point $\left(1,2,\cdots,m\right)^{\mathrm{T}}$.

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

[1] Y. Liu and C. Storey, "Efficient Generalized Conjugate Gradient Algorithms. Part 1: Theory," *Journal of Optimization Theory and Applications*, Vol. 69, No. 1, 1992, pp. 129-137. doi:10.1007/BF00940464

[2] Z. F. Li, J. Chen and N. Y. Deng, "A New Conjugate Gradient Method and Its Global Convergence Properties," *Mathematical Programming*, Vol. 78, 1997, pp. 375-391. doi:10.1007/BF02614362

[3] G. H. Yu, Y. L. Zhao and Z. X. Wei, "A Descent Nonlinear Conjugate Gradient Method for Large-Scale Unconstrained Optimization," *Applied Mathematics and Computation*, Vol. 187, No. 2, 2007, pp. 636-643. doi:10.1016/j.amc.2006.08.087

[4] J. K. Liu, X. L. Du and K. R. Wang, "Convergence of Descent Methods with Variable Parameters," *Acta Mathematicae Applicatae Sinica*, Vol. 33, No. 2, 2010, pp. 222-232.

[5] Z. F. Li, J. Chen and N. Y. Deng, "Convergence Properties of Conjugate Gradient Methods with Goldstein Line Searches," *Journal of China Agricultural University*, Vol. 1, No. 4, 1996, pp. 15-18.