International Journal of Pure and Applied Mathematics
Volume 51 No. 2 2009, 249-256

CONVERGENCE OF LIU-STOREY CONJUGATE METHOD WITH NONMONOTONE ARMIJO LINE SEARCH

Zhijun Luo[1 §], Zhibin Zhu[2]
[1]Department of Mathematics and Applied Mathematics
Technological Institute of Hunan
Loudi, 417000, P.R. CHINA
e-mail: ldlzj123@163.com
[2]School of Computing Science and Mathematics
Guilin University of Electronic Technology
Guilin, 541004, P.R. CHINA

Abstract: In this paper, we develop a new nonmonotone Armijo-type line search for LS (Liu-Storey) conjugate gradient method for minimizing functions having Lipschitz continuous partial derivatives. The nonmonotone line search can guarantee the global convergence of original LS method under some mild conditions.

AMS Subject Classification: 90C30, 65K05
Key Words: unconstrained optimization, Liu-Storey conjugate gradient, nonmonotone Armijo-type line search, global convergence

1. Introduction

In this paper, we consider the following an unconstrained minimization problem:

$$\min f(x), \qquad x \in R^n, \tag{1}$$

where R^n denotes an n-dimensional Euclidean space and $f : R^n \longrightarrow R$ is a smooth and nonlinear function.

It is well know, conjugate gradient method is a line search method that takes the form

Received: January 21, 2009

§Correspondence author

$$x^{k+1} = x^k + \alpha_k d^k, \qquad k = 0, 1, 2, \cdots, \tag{2}$$

where d^k is a descent direction of $f(x)$ at x^k is a step size. If x_k is the current iterate, we denote $f(x^k)$ *by* $f_k, \nabla f(x^k)$ *by* g_k, $\bigtriangledown f^2(x^k)$ by G_k and $f(x^*)$ by f^*, respectively. If G_k is available and inverse, then $d^k = -G_k^{-1} g_k$ leads to the Newton method and $d^k = -g_k$ results in the steepest descent method [6]. The search direction d^k is generally required to satisfy

$$g_k^T d^k < 0,$$

which guarantees that d^k is a descent direction of $f(x)$ at x^k [11]. In order to guarantee the global convergence, we sometimes require d^k to satisfy a sufficient descent condition

$$g_k^T d^k \le -c\|g_k\|^2,$$

where $c > 0$ is a constant. In line search methods, the well-known conjugate gradient method has the form (2) in which

$$d^k = \begin{cases} -g^k, & \text{if } k = 0, \\ -g^k + \beta_k d^{k-1}, & \text{if } k \ge 1. \end{cases} \tag{3}$$

where

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}, \quad \beta_k^{LS} = -\frac{g_k^T(g_k - g_{k-1})}{d^{(k-1)T} g_{k-1}},$$

or β_k is represented by other formulae [3]. The corresponding methods are called FR (Fletcher- Reeves [4]), PRP (Polak-Ribiére-Polyak [8]-[7]) and LS (Liu-Storey [5]) conjugate gradient method, respectively.

Although the above mentioned conjugate gradient algorithms are equivalent to each other for minimizing strong convex quadratic functions under exact line search, they have different performance when using them to minimize non-quadratic functions or using inexact line searches. For non-quadratic objective functions, the FR method has global convergence when exact line search or strong Wolfe line search [1]-[2] is used. The PRP method and LS methos have no global convergence under some traditional line searches. Recently, many authors proposed some new methods for the original PRP method [9]-[10]. In their methods, some convergence properties are given under some mild conditions.

In this paper, we devote to the global convergence of original LS method. A new nonmonotone Armijo-type line search is proposed for the original LS conjugate gradient method. The new line search is a novel scheme of the nonmonotone Armijo line search and allows one to find a larger accepted step size and possibly reduces the function evaluations at each iteration. It needs one to estimate the Lipschitz constant of the derivative of the objective functions in practical computation. Moreover, the nonmonotone line search can guarantee the global convergence of original LS method under some mild conditions.

The rest of this paper is organized as follows. In the next section we introduce a new nonmonotone Armijo-type line search and establish a convergent version of LS method. In Sections 3 the global convergence is analyzed.

2. Nonmonotone Armijo-Type Line Search

We first assume that

Assumtion H1. The objective function $f(x)$ is continuously differentiable on R^n and has a lower bound.

Assumtion H2. The gradient $g(x)$ of $f(x)$ is Lipschitz continuous on an open convex set B that contains the level set $L_0 = \{x \in R^n \mid f(x) \leq f(x^0)\}$ with x^0 being given, i.e., there exists $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall x, y \in B.$$

Throughout this paper we suppose that the Lipschtiz constant L of $g(x)$ is known a priori or easy to estimate in practical computation. There are some estimations L_k for the Lipschitz constant L.

Given $L_0 > 0$, in the k-th iteration we can take the Lipschitz constant as

$$L_k = \max(L_{k-1}, \frac{\|y_{k-1}\|}{\|\delta_{k-1}\|}), \qquad k = 1, 2, \cdots, \tag{4}$$

$$L_k = \max(L_{k-1}, \frac{\delta_{k-1}^T y_{k-1}}{\|\delta_{k-1}\|^2}), \qquad k = 1, 2, \cdots, \tag{5}$$

with $\delta_{k-1} = x^k - x^{k-1}$ and $y_k = g_k - g_{k-1}$.

Nonmonotone Armijo-type line search rule is stated as follows.

Nonmonotone Armijo-Type Line Search. Given $\mu \in (0, \frac{1}{2})$, $\rho \in (0, 1)$, and $c \in (\frac{1}{2}, 1)$. Let M be a nonnegative integer. For each k, let $m(k)$ satisfy

$$m(0) = 0, \; 0 \leq m(k) \leq \min(m(k-1), M), \tag{6}$$

Set $s_k = -\frac{1-c}{L_k} \cdot \frac{g_k^T d^k}{\|d^k\|^2}$ and α_k is the largest $\alpha \in \{s_k, s_k\rho, s_k\rho^2, \cdots\}$ such that

$$\max_{0 \leq j \leq m(k)} [f_{k-j}] - f(x^k + \alpha d^k) \geq -\alpha\mu g_k^T d^k, \tag{7}$$

and

$$g(x^k + \alpha d^k)^T d(x^k + \alpha d^k) \leq -c\|g(x^k + \alpha d^k)\|^2, \tag{8}$$

where

$$d(x^k + \alpha d^k) = -g(x^k + \alpha d^k) + \frac{g(x^k + \alpha d^k)^T (g(x^k + \alpha d^k) - g_k)}{-g_k^T d^k} d^k,$$

and L_k is estimated by (4) or (5).

LS method with nonmontone Armijo-type line search.

Algorithm.

Step 0. Choose $x^0 \in R^n$ and set $d^0 = -g_0$, $L_0 > 0$, $k := 0$.

Step 1. If $\|g_k\| = 0$ then stop else go to Step 2.

Step 2. Set $x^{k+1} = x^k + \alpha_k d^k$ where d^k is defined by (3), $\beta_k = \beta_k^{LS}$ and $\alpha_k$is decided by (7) and (8).

Step 3. Set $k := k + 1$ and go to Step 1.

By (4) and (5) and using Cauchy-Schwartz Inequality, we can prove the following result easily.

Lemma 2.1. *Assume that H1 and H2 hold. LS conjugate gradient method with the nonmonotone Armijo-type line search generates an infinite sequence* $\{x^k\}$. *Then*

$$L_0 \leq L_k \leq \max(L_0, L).$$

3. Global Convergence of Algorithm

In this section, we analyze the global convergence of the Algorithm. Firstly, we show the Algorithm is well defined.

Lemma 3.1. *If H1 and H2 hold, then, the nonmonotone Armijo-type line search is well defined.*

Proof. On the one hand, since

$$\lim_{\alpha \to 0} \frac{\max\limits_{0 \leq j \leq m(k)} [f_{k-j}] - f(x^k + \alpha d^k)}{\alpha} \geq \lim_{\alpha \to 0} \frac{f_k - f(x^k + \alpha d^k)}{\alpha}$$
$$= -g_k^T d^k > -\mu g_k^T d^k,$$

there is an $\alpha_k^{'}$ such that

$$\frac{\max\limits_{0 \leq j \leq m(k)} [f_{k-j}] - f(x^k + \alpha d^k)}{\alpha} \geq -\mu g_k^T d^k, \quad \forall \alpha \in [0, \alpha_k^{'}].$$

Thus, letting $\alpha_k^{''} = \min(s_k, \alpha_k^{'})$ yields

$$\frac{\max\limits_{0\le j\le m(k)}[f_{k-j}] - f(x^k + \alpha d^k)}{\alpha} \ge -\mu g_k^T d^k, \quad \forall \alpha \in [0, \alpha_k^{''}].$$

On the other hand, we can obtain

$$g(x^k + \alpha d^k)^T d(x^k + \alpha d^k) \le -c\|g(x^k + \alpha d^k)\|^2,$$

when $\alpha < -\frac{1-c}{L} \cdot \frac{g_k^T d^k}{\|d_k\|^2}$. In fact, the above inequality holds if and only if

$$\frac{g(x^k + \alpha d^k)^T (g(x^k + \alpha d^k) - g_k)}{-g_k^T d^k} g(x^k + \alpha d^k)^T d^k \le (1-c)\|g(x^k + \alpha d^k)\|^2.$$

Using Cauchy-Schwartz inequality and H2, we have

$$\begin{aligned}
&\frac{g(x^k + \alpha d^k)^T (g(x^k + \alpha d^k) - g_k)}{-g_k^T d^k} g(x^k + \alpha d^k)^T d^k \\
&\quad \le \|g(x^k + \alpha d^k)\|^2 \|g(x^k + \alpha d^k) - g_k)\| \|d^k\| / (-g_k^T d^k) \\
&\qquad \le \|g(x^k + \alpha d^k)\|^2 \alpha L \|d^k\|^2 \le (1-c)\|g(x^k + \alpha d^k)\|^2.
\end{aligned}$$

Let

$$\bar{\alpha}_k = \min\left(\alpha^{''}, -\frac{1-c}{L} \cdot \frac{g_k^T d^k}{\|d_k\|^2}\right),$$

we can prove that the nonmonotone Armjo-type line search is well defined when $\alpha \in [0, \bar{\alpha}_k]$. □

Lemma 3.2. *If H1 and H2 hold, LS method with the nonmonotone Armijo-type line search generates an infinite sequence $\{x^k\}$. Then,*

$$\|d^k\| \le \left(1 + \frac{L(1-c)}{L_0}\right)\|g_k\|, \quad \forall k. \tag{9}$$

Proof. For $k = 0$, it is easy to see that

$$\|d^0\| = \|g_0\| \le \left(1 + \frac{L(1-c)}{L_0}\right)\|g_0\|.$$

For $k \ge 1$, by Lemma 2.1, we obtain

$$\alpha_k \le -\frac{1-c}{L_k}\frac{g_k^T d^k}{\|d^k\|^2} \le -\frac{1-c}{L_0}\frac{g_k^T d^k}{\|d^k\|^2}.$$

By Cauchy-Schwartz Inequality and the above inequality, noting the LS formula and H2 , we have

$$\|d^{k+1}\| = \| - g_{k+1} + \beta_{k+1}^{LS} d^k\|$$

$$\leq \|g_{k+1}\| + \frac{\|g_{k+1}(g_{k+1}-g_k)\|}{-g_k^T d^k}\|d^k\| \leq \|g_{k+1}\|\big(1-\alpha_k L\|d^k\|^2/(g_k^T d^k)\big)$$
$$\leq \left(1+\frac{L(1-c)}{L_0}\right)\|g_{k+1}\|.$$

This proof is completed. □

From Lemma 3.1, Lemma 3.2 and imitating the analysis of Lemma 3.2 in [10], we can obtain the following result.

Lemma 3.3. *Assume that H1 and H2 hold, LS method with the nonmonotone Armijo-type line search generates an infinite sequence* $\{x^k\}$. *Then, exists* $\eta > 0$ *such that*

$$\max_{0\leq j\leq m(k)}[f_{k-j}] - f_{k+1} \geq \eta\|g_k\|^2\,. \tag{10}$$

Lemma 3.4. *If the conditions of Lemma 3.3 hold, then,*

$$\max_{1\leq j\leq M}[f(x^{Ml+j})] \leq \max_{1\leq i\leq M} f_{M(l-1)+i} - \eta \min_{1\leq j\leq M}\|g_{Ml+j-1}\|^2, \tag{11}$$

and

$$\sum_{l=1}^{\infty}\min_{1\leq j\leq M}\|g_{Ml+j-1}\|^2 < +\infty. \tag{12}$$

Theorem 3.1. *Assume that H1 and H2 hold. LS method with the nonmonotone Armijo-type line search generates an infinite sequence* $\{x^k\}$. *Then*

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

Proof. By Lemma 3.4 and letting,

$$\|g_{\omega(l)}\| = \min_{1\leq j\leq M}\|g_{Ml+j-1}\|,$$

we have

$$\sum_{l=1}^{\infty}\|g_{\omega(l)}\|^2 < +\infty,$$

it implies that

$$\lim_{l\to\infty}\|g_{\omega(l)}\| = 0. \tag{13}$$

The nonmonotone Armijo-type line search implies that

$$g_k^T d^k \leq -c\|g_k\|^2,$$

which results in

$$\|d^k\| \geq c\|g_k\| \tag{14}$$

by using Cauchy-Schwartz inequality. By Lemmas 2.1 and 3.2, (14), and the nonmonotone Armijo-type line search, we have

$$\|g_{k+1}\| = \|g_{k+1} - g_k + g_k\| \leq \|g_k\| + \|g_{k+1} - g_k\| \leq \|g_k\| + \alpha_k L \|d^k\|$$
$$\leq \|g_k\| + \alpha_k L\left(1 + \frac{L(1-c)}{L_0}\right)\|g_k\|$$
$$\leq \|g_k\|\left[1+\frac{1-c}{L_k}L\left(1+\frac{L(1-c)}{L_0}\right)\frac{-g_k^T d^k}{\|d^k\|^2}\right] \leq \|g_k\|\left[1+\frac{1-c}{cL_k}L\left(1+\frac{L(1-c)}{L_0}\right)\right]$$
$$\leq \|g_k\|\left[1 + \frac{1-c}{cL_0}L\left(1 + \frac{L(1-c)}{L_0}\right)\right] = \bar{M}\|g_k\|\,,$$

where

$$\bar{M} = 1 + \frac{1-c}{cL_0}L\left(1 + \frac{L(1-c)}{L_0}\right).$$

Which implies that

$$\|g_{M(l+1)+j}\| \leq \bar{M}\|g_{M(l+1)+j-1}\| \leq \cdots \leq \bar{M}^{2M}\|g_{\omega(l)}\|,$$

for $j = 1, 2, \cdots, M$. By (13) we obtain

$$\lim_{k\to\infty} \|g_k\| = 0. \qquad \square$$

Acknowledgements

This work was supported in part by the NNSF (No.10501009 and 0728206) of P.R. China and Guangxi Science Foundation.

References

[1] M. Al-Baali, New property and global convergence of the Fletcher-Reeves method with inexact line searches, *IMA Journal of Numerical Analysis*, **5** (1985), 122-124.

[2] Y.H. Dai, Y. Yuan, Convergence properties of the Fletcher-Reeves method, *IMA Journal of Numerical Analysis*, **16** (1996), 155-164.

[3] Y.H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim.*, **10** (1999), 177-182.

[4] R. Fletcher, C. Reeves, Function minimization by conjugate gradients, *Comput. J.*, **7** (1964), 149-154.

[5] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms. Part 1: Theory, *Journal of Optimization Theory and Applications*, **69** (1991), 129-137.

[6] J. Nocedal, J.S. Wright, *Numerical Optimization*, Springer-Verlag, New York (1999).

[7] B.T. Polyak, The conjugate gradient method in extreme problems, *USSR Comput. Math. Math. Phys.*, **9** (1969), 94-112.

[8] E. Polak, G. Ribiére, Note sur la convergence de directions conjuguées, *Rev. Francaise Infomat Recherche Operatonelle*, 3-e Année, **16** (1969), 35-43.

[9] Z.J. Shi, J. Shen, Convergence of PRP method with new nonmonotone line search, *Appl. Math. Comput.*, **181** (2006), 423-431.

[10] Z.J. Shi, J. Shen, Convergence of Polak-Ribiére-Polyak conjugate gradient method, *Nonliner Analysis*, **66** (2007), 1428-1441.

[11] Y. Yuan, *Numerical Methods for Nonlinear Programming*, Shanghai Scientific and Technical Publishers (1993).