Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 3 Issue 12, December - 2016

A Modified Conjugate Gradient Method with Armijo-Type Linear Search

Qi Liu, Zhijun Luo(Corresponding author) and **Guohua Chen**,
School of Mathematics and Finance,
Hunan University of Humanities, Science and Technology,
Loudi, 417000, P.R.China

***Abstract*—In this paper, we have presented a new conjugate gradient algorithm based on the Armijo-type line search for solving unconstrained optimization problems. Under some suitable conditions, we proved the global convergence of the algorithm. The numerical results show that the proposed methods are effective.**

Keywords—Unconstrained optimization; Conjugate gradient method; Armijo-type line search ; Global convergence

I. INTRODUCTION

Various numerical methods have been introduced to solve unconstrained optimization problems as they appear in engineering, mathematics and computer science. The conjugate gradient method is an efficient algorithm for the numerical solution of unconstrained optimization problems. In this paper, we consider the following unconstrained optimization problems

$$\min_{x\in R^n} f(x) \tag{1}$$

where $f(x): R^n \to R$ is a smooth, nonlinear function whose gradient will be denoted by $g(x)$. the conjugate gradient method for solving problem (1) always generates a sequence $\{x_k\}$ as

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

where $\alpha_k > 0$ is obtained by line search and the directions d_k are generated as

$$d_k = \begin{cases} -g_k, & if \quad k = 0 \\ -g_k + \beta_k d_{k-1}, & if \quad k \ge 1, \end{cases} \tag{3}$$

where $g_k = \nabla f(x_k)$, and β_k is a scalar, different conjugate gradient algorithms correspond to different choices for the scalar parameter β_k. There are some well-known formulas for β_k which are given below(see[1-6]):

$$\beta_k^{FR} = \frac{\| g_k \|^2}{\| g_{k-1} \|^2}, \quad \beta_k^{DY} = \frac{\| g_k \|^2}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{CD} = -\frac{\| g_k \|^2}{d_{k-1}^T g_{k-1}},$$

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\| g_{k-1} \|^2}, \quad \beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad \beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}},$$

or by other formulae(where $y_{k-1} = g_k - g_{k-1}$, the symbol $\|\cdot\|$ be the Euclidean norm.). The stepsize α_k is determined by exact or inexact line search. The Armijo-type line search: Let $s > 0$ be a constant, $\mu \in (0,1), \rho \in (0,1)$, choose α_k to be the largest α in $\{s, s\rho, s\rho^2, \cdots,\}$ such that

$$f_k - f(x_k + \alpha d_k) \ge -\alpha \mu g_k^T d_k .$$

In this paper, we devote to the global convergence of a modified Conjugate Gradient Method for unconstrained optimization problems. This method uses a new Armijo-type line search which allows one to find a larger accepted step size and possibly reduces the function evaluations at each iteration. It can guarantee the global convergence of modified method under some mild conditions.

II. DESCRIPTION OF ALGORITHM.

We first assume that

H2.1

i) The objective function $f(x)$ is continuously differentiable and has a lower bound on the level set $L_0 = \{x \in R^n \mid f(x) \le f(x_0)\}$, where x_0 is the starting point.

ii) The gradient $g(x)$ of $f(x)$ is Lipschitz continuous in some neighborhood U of L0, namely, there exists a constant L > 0 such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in U.$$

Throughout this paper we suppose that the Lipschtiz constant L of g(x) is a known priori or easy to estimate in practical computation. There are some estimations L_k for the Lipschitz constant L [6-7]. Motivated by [6], we propose the following New Armijo-type line search:

Given $\mu \in \left(0, \frac{1}{2}\right), \rho \in (0,1), c \in (0,1), \varepsilon \in (0,1)$, set $l_k = \frac{1-c}{L_k} \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|}$ and α_k is the first α in $\{l_k, l_k\rho, l_k\rho^2, \cdots,\}$ such that :

$$f_k - f(x_k + \alpha d_k) \geq -\alpha\mu g_k^T d_k, \qquad (4)$$

where $L_k = \max\left(L_{k-1}, \frac{\|y_{k-1}\|}{\|s_{k-1}\|}\right)$.

Now we state our algorithm as follows.

Algorithm A:

Step 0 Initialization:

Given a starting point $x_0 \in R^n$, choose parameters $0 < \varepsilon << 1, \ \mu \in \left(0, \frac{1}{2}\right), \rho \in (0,1)$, $c \in (0,1), \varepsilon \in (0,1)$. Set $k := 0$;

Step 1 If $\| g_k \| < \varepsilon$, STOP, else go to Step 2;

Step 2 Compute the search direction d_k by (3), where

$$\beta_k = \frac{g_k^T (g_k - g_{k-1})}{|y_{k-2}^T d_{k-2}| \|d_{k-1}\| + \varepsilon \|d_{k-1}\|^2}.$$

Step 3 α_k is defined by the new Armijo-type line search (4).

Step 4 Let $x_{k+1} = x_k + \alpha_k d_k$, k := k + 1, and go to Step 2.

Lemma 2.1 Assume thatH2.1 hold, the infinite sequence $\{x_k\}$ is generated by Algorithm A. Then,

1) there exist $m_0 > 0$ *and* $M_0 > 0$ such that $m_0 \leq L_k \leq M_0$.
2) for $k \geq 1$, if

$$\alpha_k \leq \frac{1-c}{L_k} \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|},$$

then,

$$g_{k+1}^T d_{k+1} \leq -c\|g_{k+1}\|^2.$$

3) $\|d_k\| \leq \left(1 + \frac{(1-c) L_k}{m_0}\right) \|g_k\|, \forall k.$

III. Global convergence of Algorithm

In this section, we analyze the global convergence of the Algorithm.

Lemma 3.1. Assume thatH2.1 hold, then the new Armijo−type linear search is well-defined.

Proof. 1) If $\alpha_k = \frac{1-c}{L} \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|}$ then we have

$$\alpha_k = \frac{1-c}{L} \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|} \geq \frac{1-c}{L} \varepsilon > 0.$$

2) If $\alpha_k < \frac{1-c}{L} \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|}$, then, set $\alpha = \rho^{-1}\alpha_k$, the following inequality does not hold,

$$f_k - f(x_k + \alpha d_k) \geq -\alpha\mu g_k^T d_k,$$

i.e.

$$f_k - f(x_k + \alpha d_k) < -\alpha\mu g_k^T d_k \circ$$

Considering mean value theorem on the left-hand side of the above inequality, there exists $t_k \in (0,1)$, such that

$$-\alpha g(x_k + t_k \alpha d_k)^T d_k < -\alpha\mu g_k^T d_k,$$

i.e.

$$g(x_k + t_k \alpha d_k)^T d_k > \mu g_k^T d_k.$$

By H2.1, and Lemma 2.1, we have

$$L_k \alpha \|d_k\|^2 \geq \|g(x_k + t_k \alpha d_k) - g_k\| \|d_k\|$$
$$\geq g(x_k + t_k \alpha d_k)^T d_k \geq -(1-\mu) g_k^T d_k$$
$$\geq c(1-\mu)\|g_k\|^2.$$

From Lemma 2.1, we can get that

$$\alpha_k \geq \frac{c\rho(1-\mu)}{L_k} \frac{\|g_k\|^2}{\|d_k\|^2} \geq \left(1 + \frac{(1-c) L_k}{m_0}\right)^{-2} > 0.$$

Theorem 3.1 Under the assumption H2.1, the sequence $\{x_k\}$ generated by Algorithm A is global convergent. That is

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

Proof: Set $\eta_0 = \inf_{\forall k} \{\alpha_k\}$, if $\eta_0 > 0$, then we have

$$f_k - f(x_k + \alpha d_k) \geq -\alpha\mu g_k^T d_k \geq \mu\eta_0 c\|g_k\|^2.$$

By H2.1 we have $\sum_{k=0}^{+\infty} \|g_k\|^2 < +\infty$, thus we obtain

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

If the conclusion doesn't hold, assume that $\eta_0 = 0$. Then, there exists an infinite subset $K \subseteq \{0,1,2,...,\}$ such that

$$\lim_{k \in K, k \to \infty} \alpha_k = 0, \qquad (5)$$

By Lemma 2.1 and Lemma 3.1 we obtain

$$l_k = \frac{(1-c)}{L} \cdot \frac{|y_{k-1}^T d_{k-1}| + \varepsilon \|d_k\|}{\|d_k\|} > 0.$$

Therefore, there is a k' such that $\rho^{-1}\alpha_k \le l_k, \forall k \ge k', k \in K$. Let $\alpha = \rho^{-1}\alpha_k$, by Lemma 3.1 we obtain

$$\alpha_k \ge \frac{c\rho(1-\mu)}{L}\frac{\|g_k\|^2}{\|d_k\|^2} \ge \frac{c\rho(1-\mu)}{L}\left(1+\frac{(1-c)L}{m_0}\right)^{-2} > 0$$

$k \ge k', k \in K$.

Which contradicts (5). This shows that $\eta_0 > 0$. The whole proof is completed.

IV. Numerical Experiments

In the following numerical experiment, the code of the proposed algorithm is written by using MATLAB 7.0. We stop the iteration if the inequality $\|g_k\| \le 10^{-6}$ is satisfied. The problems that we tested are from [8] and [9]. Table 1 show the computation results, where the columns have the following meanings:

Dim —the dimension of the problems;
NI—the number of iterations;
NF—the number of function evaluations;
NG—the dimension of gradient evaluations;

Table 1 The numerical results of Algorithm A

Problem	Dim	Algorithm A NI/NF/NG
Penalty1	4	5/128/89
Trigonometric	100	18/72/67
Brown	2	13/148/51
Axis hyper	500	183/6168/705
DeJong	2	3/18/13
	500	4/24/15

Acknowledgment

This work was supported in part by Scientific Research Fund of Hunan Provincial Education Department (NO. 12A077 and 14C0609), and the Educational Reform Research Fund of Hunan University of Humanities, Science, and Technology(no. RKJGY1526).

References

[1] R. Fletcher, C. Reeves. “Function minimization by conjugate gradients,” The computer journal. vol. 7, pp. 149-154, 1964.

[2] E. Polak and G. Ribi´ere, “Note sur la convergence de m´ethodesde directions conjugu´ees,” Revue Franc¸aise de RechercheOp´erationnelle, no. 16, pp. 35–43, 1969.

[3] Y. H. Dai and Y. Yuan, “A nonlinear conjugate gradient method with a strong global convergence property,” SIAM Journal on Optimization, vol. 10, no. 1, pp. 177–182, 1999.

[4] R. Fletcher, Unconstained Optimization: Practical Methods of Optimization, vol. 1, John Wiley & Sons, New York, NY, USA, 1987.

[5] Y. Liu and C. Storey, “Efficient generalized conjugate gradient algorithms, Part 1,” Journal of Optimization Theory and Applications, vol. 69, no. 1, pp. 129–137, 1991.

[6]M. R. Hestenes and E. Stiefel, “Methods of conjugate gradients for solving linear systems,” Journal of Research of the National Bureau of Standards, vol. 49, pp. 409–436, 1952.

[6] Shi Z J, Shen J. Convergence of Liu–Storey conjugate gradient method [J]. European Journal of Operational Research, 2007, 182(2):552-560.

[7] Z.J. Shi, J. Shen, Step-size estimation for unconstrained optimization methods, Computational and Applied Mathematics24 (3) (2005) 399–416.

[8]M. Molga, C. Smutnicki.“Test functions for optimization needs,” http://www.zsd.ict.pwr.wroc.pl/files/ docs/functions.pdf, 2005

[9]W. Hock and K. Schittkowski, “Test examples for nonlinear programming codes,” Journal of Optimization Theory and Applications, vol. 30, no. 1, pp. 127–129, 1981.