Journal of Computational Science & Engineering 2 (2011) 65-68

Journal of Computational Science & Engineering

Available at http://www.asocse.org ISSN 1710-4068

Global Convergence of a Modified DY Conjugate Gradient Method With Wolfe Type Line Search

Yuan-yuan Chen *

College of Mathematics, Qingdao University, Qingdao, Shandong 266071, China

Article Information

Article history:
Received 12 March 2011
Revised 10 May 2011
Accepted 25 May 2011
Available online 10 June 2011

Keywords:
Unconstrained optimization
DY conjugate gradient method
Line search
Global convergence

Abstract

In this paper, a new nonlinear modified DY conjugate gradient method with wolfe type line search for solving unstrained optimization problems is proposed. The direction generated by the method is a descent direction. Under some conditions, we prove that the modified DY conjugate gradient method with Wolfe type line search is global convergence. Finally, the numerical results show that the new method is efficient for unconstrained optimizations.

1. Introduction

In this paper, we consider the following large scale unconstrained optimization problem

$$\min_{x\in R^n} f(x) \tag{1.1}$$

where $f: R^n \to R$ is continuously differentiable and its gradient $g(x)=\nabla f(x)$ is available. Iterative methods are widely used for solving (1.1) and the iterative formula is given by

$$x_{k+1} = x_k + \alpha_k d_k \tag{1.2}$$

where $x_k \in R^n$ is the k-th approximation to the solution of (1.1), α_k is a stepsize obtained by carrying out a line search and d_k is a search direction.

There are many kinds of iterative methods for solving (1.1), such as Newton method, steepest descent method and nonlinear conjugate gradient methods.Due to the simplicity of their iteration and low memory requirements, the nonlinear conjugate gradient methods are one of the most famous methods for solving (1.1), especially in case of the large scale optimization problems, which are often proposed in scientific and engineering computation. The search direction

$$d_k = \begin{cases} -g_k, & k=1, \\ -g_k + \beta_k d_{k-1}, & k \geq 2, \end{cases} \tag{1.3}$$

where β_k is a scalar and $g_k = g(x_k)$ .Different conjugate gradient methods correspond to different scalar β_k , i.e. [1-4]

* Corresponding author. E-mail: *qddsq1@163.com*

2011 © The American Computational Science Society. All rights reserved.

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \text{ (Fletcher-Reeves)}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_{k+1}-g_k)}{\|g_{k-1}\|^2}, \text{ (Polak-Ribiere-Polyak)}$$

$$\beta_k^{HS} = \frac{g_k^T(g_{k+1}-g_k)}{d_{k-1}^T(g_{k+1}-g_k)}, \text{ (Hestenes-Stiefel)}$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k-g_{k-1})^T d_{k-1}}, \text{ (Dai-Yuan).}$$

The DY conjugate gradient method is a relatively new nonlinear conjugate gradient method, but it has got more attention for the property of producing a descent direction at each iteration. In [4] the global convergence of DY method with wolfe line search was given, that is, α_k satisfying

$$f(x_k+\alpha_k d_k) \le f(x_k)+\delta\alpha_k g_k^T d_k,$$
$$|d_k^T g(x_k+\beta\alpha_k d_k)| \le \sigma |g_k^T d_k|,$$

where $0<\delta<\sigma<\frac{1}{2}$. Other nonlinear conjugate gradient methods and the global convergence results about Fletcher-Reeves(FR) method, Polak-Ribiere-Polyak(PRP) method and Hestenes-Stiefel(HS) method can see [5-9].

Quite recently, Zhang proposed two modified version of the DY nonlinear conjugate gradient method in [10]. The scalar β_k is given by the following way

$$\beta_k = \frac{\|g_k\|^2}{d_{k-1}^T z_{k-1}} \tag{1.4}$$

where

$$z_{k-1} = y_{k-1} + t_k s_{k-1}$$

$$t_k = \varepsilon_1 \|g_k\|^2 + \max\{0, -\frac{s_{k-1}^T y_{k-1}}{s_{k-1}^T s_{k-1}}\}$$

$y_{k-1} = g_k - g_{k-1}$, $s_{k-1} = x_k - x_{k-1}$ and ε_1 is a positive constant, $\|\cdot\|$ stands for the Euclidean norm.

In this paper, based the modified DY conjugate gradient method in [10], under some mild conditions, we give the global convergence of the modified DY method with Wolfe type line search.

The rest of the paper is organized as follows. In Section 2, we present the Wolfe type line search and the modified DY conjugate gradient method. The global convergence of the modified DY conjugate gradient method will be given in Section 3. Numerical results are given in the last section.

2. Modified DY conjugate gradient method

In this section, we will give the following assumptions on objective function $f(x)$ in (1.1), which have been often used in the literature to analyze the global convergence of nonlinear conjugate gradient methods with inexact line searches for unconstrained optimization problems.

Assumption 2.1.

(1) The level set $L_0 = \{x \in R^n \mid f(x) \le f(x_0)\}$ is bounded.

(2) In some neighborhood U of L_0, $f(x)$ is continuously differentiable and its gradient is Lipschitz continuous, namely, there exists a constant $L>0$ such that

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

Now we consider the modified DY conjugate gradient method for (1.1) with Wolfe type line search. Firstly, we give the following Wolfe type line search (we have proposed in [11]).

The line search is to choose $\alpha_k > 0$ such that

$$f(x_k) - f(x_k+\alpha_k d_k) \ge \rho\alpha_k^2 \|d_k\|^2, \tag{2.1}$$

$$g(x_k+\alpha_k d_k)^T d_k \ge -2\sigma\alpha_k \|d_k\|^2, \tag{2.2}$$

where $0<\rho<\sigma<1$.

Now we present the new modified DY conjugate gradient method as follows.

Algorithm 2.1.

Step 0. Given $x_0 \in R^n$, set $d_0 = -g_0$, $k := 0$. If $g_0 = 0$, then stop.

Step1. Find $\alpha_k > 0$ satisfying the Wolfe type

2011 © The American Computational Science Society. All rights reserved.

line search (2.1), (2.2), by (1.2), x_{k+1} is given. If $g_{k+1}=0$, then stop.

Step 2. Compute d_k by (1.3) and (1.4).

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

In order to establish the global convergence of the Algorithm 2.1, we give the following results.

Lemma 2.1. Suppose that Assumption 2.1 holds, then the Wolfe type line search (2.1), (2.2) is feasible.

The proof is essentially the same as Lemma 1 of [11], hence we omit it.

Lemma2.2. Suppose direction d_k is given by (1.3) and (1.4), then we have

$$d_k^T g_k < 0, \tag{2.3}$$

holds for any $k \geq 0$, and $\beta_k \leq \dfrac{g_k^T d_k}{g_{k-1}^T d_{k-1}}$.

Proof. We show by induction that (2.3) holds for Theorem 2.1 of [10], hence we omit it.

Lemma 2.3. Suppose that Assumption 2.1 holds, we have

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \tag{2.4}$$

Proof. From (2.1), (2.2), (2.3) and Assumption 2.1, we obtain

$$(2\sigma + L)\alpha_k \|d_k\|^2 \geq -g_k^T d_k .$$

Then, we know

$$\alpha_k \|d_k\| \geq \frac{1}{2\sigma + L}\left(\frac{-g_k^T d_k}{\|d_k\|}\right).$$

Squaring both sides of above formula, we get

$$\alpha_k^2 \|d_k\|^2 \geq \left(\frac{1}{2\sigma + L}\right)^2 \left(\frac{-g_k^T d_k}{\|d_k\|}\right)^2.$$

From (2.1), we have

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \leq (2\sigma + L)^2 \sum_{k=1}^{\infty} \alpha_k^2 \|d_k\|^2$$

$$\leq \frac{(2\sigma + L)^2}{\rho} \sum_{k=1}^{\infty} \{ f(x_k) - f(x_{k+1}) \}$$

$< +\infty$.

So we get (2.4), this completes the proof.

3. Global convergence result for the Algorithm 2.1

From the above analysis of Section 2, we give the following result of global convergence for the modified DY conjugate gradient method. The general idea for proving the global convergence of the modified DY conjugate gradient method is to assume first by contradiction, then derive contradictory relations.

Theorem 3.1. Consider the modified DY conjugate gradient Algorithm 2.1, suppose that Assumptions 2.1 holds. Then we have

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

Proof. Suppose by contradiction that there exists a positive constant $\varepsilon > 0$ such that

$$\|g_k\| \geq \varepsilon, \tag{3.1}$$

holds for all $k \geq 0$.

From (1.4) and Lemma 2.2, we have

$$\|d_k\|^2 = (\beta_k)^2 \|d_{k-1}\|^2 - 2 g_k^T d_k - \|g_k\|^2 \leq$$

$$\left(\frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}\right)^2 \|d_{k-1}\|^2 - 2 g_k^T d_k - \|g_k\|^2 .$$

Dividing both sides of the above inequality by $(g_k^T d_k)^2$ and by (3.1), we have

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{2}{d_k^T g_k} - \frac{\|g_k\|^2}{(g_k^T d_k)^2}$$

$$= \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{1}{\|g_k\|^2}\left(1 - \frac{\|g_k\|^2}{d_k^T g_k}\right)^2 + \frac{1}{\|g_k\|^2}$$

$$\leq \sum_{i-0}^{k-1} \frac{1}{\|g_i\|^2}$$

$$\leq \frac{k}{\varepsilon^2}.$$

So we obtain

2011 © The American Computational Science Society. All rights reserved.

$$\sum_{k=1}^{\infty}\frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \varepsilon^2 \sum_{k\geq 1}\frac{1}{k}$$

$= +\infty$.

Which contradicts (2.4). Therefore, we have $\liminf_{k\to\infty}\|g_k\| = 0$.

Table 1 Test results for the Algorithm 2.1 and DY method

Name	Dim	NI	NF	NG(Algorithm 2.1)
BV	3	1000	2036	1019
BV	10	1000	3388	1956
IE	3	1000	1504	1002
Name	Dim	NI	NF	NG(DY method)
BV	3	1000	3945	1890
BV	10	1000	2896	1897
IE	3	1000	6096	1222

4. Numerical experiments

In this section, we give some numerical experiments for the above new nonlinear conjugate gradient method. The problem that we tested are from [12]. Our line search subroutine computes α_k such that the Wolfe type line search hold with $\rho = 0.001$ and $\sigma = 0.01$. We also use the condition $\|g_{k+1}\| \leq 10^{-6}$ as the stopping criterion. We use MATLAB7.0 to test the chosen problems. To show the efficient of our method, we also give the numerical experiments for DY method under the Wolf type line search. The numerical results of Algorithm 2.1 and DY method are listed in Table 1, where the items in each column have the following meanings:

Name: the name of the test problem;
Dim: the dimension of the problem;
NI: the number of iterations;
NF: the number of function evaluations;
NG: the number of gradient evaluations.

Acknowledgement

The authors are grateful for the financial support provided national science foundation of China(10971118,70971070), Shandong province higher educational science and technology program (J10LA05), Science foundation of Shandong (ZR2010GM006), Innovation program of Shanghai municipal education commis sion(10YZ99).

References

[1] R.Fletcher, C.Reeves, Function Minimization by Conjugate Gradients. J. Comput. 7(1964) 149-154.
[2] T.Polyak,The conjugate gradient method in extreme problems. USSR Comp.Math. Math. Phys. 9(1969) 94-112.
[3] M.R.Hestenes,E.Stiefel, Mothod of conjugate gradient for solvoing linear equations, J.Res.Nat.Bur.Stand. 49(1952) 409-436.
[4] Y.H.Dai,Y.X.Yuan, A nonlinear conjugate gradient method with nice global convergence properties, SIAM Journal on Optimization 10(1999) 177-182.
[5] J.C. Gilbert, J. Nocedal, Global Convergence properties of conjugate gradient methods for optimization, SIAM Journal on Optimization 2(1992) 21-24.
[6] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, Part 1: Theory, Journal of Optimization Theory and Applications 69(1991) 129-137.
[7] C. Pan, L. Chen, A class of efficient new descent methods, ACTA Math. Appl. 30(2007) 88-98.(in Chinese)
[8] J. Sun, J. Zhang, Convergence of conjugate gradient methods without line search, Annals Operations Research 103(2001) 161-173.
[9] M.Raydan, TheBarzilain and Borwein gradient method for the large unconstrained minimization problem, SIAM Journal on Optimization 7(1997) 26-33.
[10] L. Zhang, Two modified Dai-yuan nonlinear conjugate gradient methods, Numerical Algorithms 50(2009) 1-16.
[11] C. Wang, Y. Chen, S. Du, Further insight into the Shamanskii modification of Newton method, Applied Mathematics and Computation 180(2006) 46-52.
[12] J.J. More, B.S. Garbow, K.E. Hillstrom, Testing unconstrained optimization software, ACM Trans. Math. Software 7(1981) 17-41.

2011 © The American Computational Science Society. All rights reserved.