DOI: 10.21276/sjpms.2016.3.4.9

Scholars Journal of Physics, Mathematics and Statistics **ISSN 2393-8056 (Print)**
Sch. J. Phys. Math. Stat. 2016; 3(4):170-173 **ISSN 2393-8064 (Online)**

A Modified LS-CD Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization

Aiai Gong[1], Xiaoli Tian[2], Qi Liu[1], Zhijun Luo*[1]
[1]School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi, 417000, P.R. China
[2]The third middle school of Xinhua County, Xinhua, Hunan, 417600, P.R. China

***Corresponding Author:**
Zhijun Luo
Email: ldlzj11@163.com

Abstract: In this paper, a new hybrid conjugate gradient method for solving unconstrained optimization problems is presented. Under strong Wolfe line search conditions, the global convergence of this method is established. The numerical results show that the proposed method is effective.
Keywords: Unconstrained optimization; Conjugate gradient method; Global convergence

INTRODUCTION

Conjugate gradient method is an efficient algorithm for the numerical solution of unconstrained optimization. We consider the following unconstrained optimization problem

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

where $f(x): R^n \to R$ is a smooth nonlinear function, whose gradient will be denoted by $g(x) = \nabla f(x)$. The iterative formula is

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

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. In general, the step length α_k is chosen by the Wolfe line search or Armijo-type linear search. Here, we use the strong Wolf line search condition, i.e., the step size α_k satisfies

$$\begin{cases} f(x_k + \alpha d_k) - f(x_k) \le \delta \alpha_k g_k^T d_k, \\ \left|g(x_k + \alpha d_k)\right|^T d_k \le -\sigma g_k^T d_k, \end{cases} \tag{4}$$

where $0 < \delta < \frac{1}{2}, and \quad \delta < \sigma < 1.$

As you know, different choices of β_k result in different nonlinear conjugate gradient methods. Some famous formulae for β_k are defined as follows:

$$\begin{aligned}
&\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad FR\ (Fletcher\text{-}Reeves)\ [1],\\
&\beta_k^{PRP} = \frac{g_k^{T} y_{k-1}}{\|g_{k-1}\|^2}, \quad PRP\ (Polak\text{-}Ribiere\text{-}Polyak)\ [2],\\
&\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^{T} y_{k-1}}, \quad DY\ (Dai\text{-}Yuan)\ [3],\\
&\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^{T} g_{k-1}}, \quad CD\ (conjugate\ descent)\ [4],\\
&\beta_k^{LS} = -\frac{g_k^{T} y_{k-1}}{d_{k-1}^{T} g_{k-1}}, \quad LS\ (Liu\text{-}Storey)[5],\\
&\beta_k^{HS} = \frac{g_k^{T} y_{k-1}}{d_{k-1}^{T} y_{k-1}}, \quad HS\ (Hestenes\text{-}Stiefel)[6].
\end{aligned} \tag{5}$$

where $y_{k-1} = g_k - g_{k-1}$, the symbol $\|\cdot\|$ be the Euclidean norm. Although all these methods are equivalent in the linear case, namely, when $f(x)$ is a strictly convex quadratic function and α_k are determined by exact line search, their behaviors for general objective functions may be far different [7].

In recent years, hybrid conjugate gradient methods are regarded as the best performing conjugate gradient methods in practice because of dynamically adjustment of β_k as the iterations evolve. Dai and Yuan [8] combined the DY algorithm with the HS algorithm, proposing the following two hybrid methods

$$b^{hDY} = \max\{-\,c\,b^{DY},\ \min\{b^{DY}, b^{HS}\}\},$$
$$b^{hDYz} = \max\{0,\ \min\{b^{DY}, b^{HS}\}\},$$

where c is a scalar. For the weak Wolfe conditions, they established the global convergence of these hybrid computational schemes. Combining between PRP and DY conjugate gradient methods, N. Andrei [9] proposed the following hybrid method:

$$b = (1-q)b^{PRP} + qb^{DY},$$

where the parameter in the convex combination is computed in such a way that the conjugacy condition is satisfied, independently of the line search.

Because LS has good computational properties, on one side, and CD has strong convergence properties, on the other side. In this paper, we propose another hybrid conjugate gradient as a convex combination of LS and CD conjugate gradient algorithms. By this method, we hope to obtain a more efficient conjugate gradient algorithm. The iterates $x_0, x_1, x_2, \cdots,$ of our algorithms are computed by means of the recurrence (2) where the stepsize $\alpha_k > 0$ is obtained by Wolfe conditions, and the directions are generated as

$$\begin{cases} d_0 = -g_0 \\ d_k = -g_k + \beta_k^{LSCD} d_{k-1}, k \ge 1, \end{cases} \tag{6}$$

Where $\beta_k = \max\{0, \min\{\beta_k^{LS}, \beta_k^{CD}\}\}$.

The rest of this paper is organized as follows. The algorithm is presented in Section 2. In Sections 3 the global convergence is analyzed. We give the numerical experiments in Section 4

1. Description of algorithm

Now we state our algorithm as follows.

Algorithm A:

Step 0: Initialization: Given a starting point $x_0 \in R^n$, choose parameters

$$0<\varepsilon<<1, 0<\delta<\frac{1}{2}, \delta<\sigma<1, d_0=-g_0, k:=0$$

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

Step 2 : Let $x_{k+1}=x_k+\alpha_k d_k$, where d_k is followed by (6), and α_k is defined by the strong Wolf line search (4).

Step 3 : Let $k:=k+1.$, and go to Step 2.

Global convergence of Algorithm

At first, the following basic assumptions on the objective function are assumed, which have been widely used in the literature to analyze the global convergence of the conjugate gradient methods.

H3.1

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 L_0, namely, there exists a constant L > 0 such that

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

Lemma 3.1[7] Suppose that Assumption H3.1 holds. If the conjugate method satisfies $g_k^T d_k < 0$, then we have that

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\| d_k \|^2} < +\infty.$$

Theorem 3.1 Suppose that Assumption H3.1 holds and the sequence $\{x_k\}$ is generated by Algorithm A, then $g_k^T d_k < 0$.

Proof: For $n=0$, $g_0^T d_0 = -\| g_0 \|^2 < 0$.

When $k \ge 1$ multiplying g_k^T by $d_k = -g_k + \beta_k^{LS-CD} d_{k-1}$, we obtain that

$$g_k^T d_k = -\| g_k \|^2 + \beta_k^{LSCD} g_k^T d_{k-1},$$

it follows from $\beta_k^{LSCD} \ge 0$ and $g_k^T d_{k-1} \le 0$ that

$$g_k^T d_k \le -\| g_k \|^2 < 0.$$

Therefore, the result is true.

In view of Theorem 3.1 and[10] [11], we may obtain the following results.

Theorem 3.2 Suppose that Assumption H3.1 holds and the sequence $\{x_k\}$ is generated by Algorithm A. Then

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

Numerical experiments

In this section, we give the numerical results of Algorithm A to show that the method is efficient for unconstrained optimization problems. The problems that we tested are from [12] and [13]. **Table 1** show the computation results, where the columns have the following meanings:

x_k —the final point ;

f_* —the final value of the objective function ;

Table-1: Comparative numerical results of Algorithm A

Problem	x_k	f_*
Beale	(2.99998081854872, 0.50000717055051)	3.339569051726135e-009
Trigonometric	(0.24309754927761, 0.61287060083150)	4.316579730575097e-008
Brown	(0.99968107296011, 1.00056330601065)	6.519072115070798e-008

REFERENCES

1. Fletcher R, Reeves CM. Function minimization by conjugate gradients. The computer journal. 1964 Jan 1;7(2):149-54.
2. Polak E, Ribiere G. Note sur la convergence de méthodes de directions conjuguées. Revue française d'informatique et de recherche opérationnelle, série rouge. 1969;3(1):35-43.
3. Dai YH, Yuan Y. A nonlinear conjugate gradient method with a strong global convergence property. SIAM Journal on Optimization. 1999;10(1):177-82.
4. Fletcher R. Practical methods of optimization. John Wiley & Sons; 2013 Jun 6.
5. Liu Y, Storey C. Efficient generalized conjugate gradient algorithms, part 1: theory. Journal of Optimization Theory and Applications. 1991 Apr 1;69(1):129-37.
6. Hestenes MR, Stiefel E. Methods of conjugate gradients for solving linear systems. NBS; 1952 Dec 6.
7. Dai YH. Nonlinear conjugate gradient methods. Wiley Encyclopedia of Operations Research and Management Science. 2011.
8. Dai YH, Yuan Y. An efficient hybrid conjugate gradient method for unconstrained optimization. Annals of Operations Research. 2001 Mar 1;103(1-4):33-47.
9. Andrei N. A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes-Stiefel and Dai-Yuan. Studies in Informatics and Control. 2008 Mar 1;17(1):57.
10. Jian J, Han L, Jiang X. A hybrid conjugate gradient method with descent property for unconstrained optimization. Applied Mathematical Modelling. 2015 Feb 28;39(3):1281-90.
11. Sun W, Yuan YX. Optimization theory and methods: nonlinear programming. Springer Science & Business Media; 2006 Aug 6.
12. Molga M, Smutnicki C. Test functions for optimization needs. Test functions for optimization needs. 2005 May.
13. Hock W, Schittkowski K. Test examples for nonlinear programming codes. Journal of Optimization Theory and Applications. 1980 Jan 1;30(1):127-9.