DOI: 10.21276/sjet.2016.4.12.5

Scholars Journal of Engineering and Technology (SJET)
Sch. J. Eng. Tech., 2016; 4(12):588-593
©Scholars Academic and Scientific Publisher
(An International Publisher for Academic and Scientific Resources)
www.saspublisher.com

ISSN 2321-435X (Online)
ISSN 2347-9523 (Print)

Original Research Article

A PRP-FR Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization

Lirong Wang[1], Xiaoli Tian[2]
[1]School of Information science, 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**
Lirong Wang
Email: ldlzj11@163.com

Abstract: In this paper, a modified hybrid conjugate gradient algorithm is proposed for solving unconstrained optimization problems, which avoid the drawbacks of PRP and FR. The global convergence of this method is established under strong Wolfe line search conditions. The numerical results show that the proposed method is effective.
Keywords: Unconstrained optimization; Conjugate gradient method; Global convergence.

INTRODUCTION

It is well known that the conjugate gradient method is an effective method to solve largescale minimization problems. In this paper, we consider the following unconstrained optimization problem

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

where R^n denotes an n-dimensional Euclidean space and $f(x):R^n \to R$ is a continuously differentiable function. In usual, the iterative formula of the conjugate gradient method is given as follows:

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

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

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

where $g_k = \nabla f(x_k)$, and β_k is a scalar. 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. The step length α_k usually 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

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

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

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:

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

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

$$\beta_k^{DY} = \frac{\| g_k \|^2}{d_{k-1}^T y_{k-1}}, \quad DY\ (Dai\text{-}Yuan)\ [3],$$

$$\beta_k^{CD} = -\frac{\| g_k \|^2}{d_{k-1}^T g_{k-1}}, \quad CD\ (conjugate\ descent)\ [4],$$

$$\beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \quad LS\ (Liu\text{-}Storey)\ [5],$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \quad HS\ (Hestenes\text{-}Stiefel)\ [6].$$

(5)

where $y_{k-1} = g_k - g_{k-1}$, the symbol $\|\cdot\|$ be the Euclidean norm. As is well known, the CG methods β_k^{FR}, β_k^{CD} and β_k^{DY} possess strong global convergence properties, but have less computational performance. On the other hand, the β_k^{PRP}, β_k^{LS} and β_k^{HS} methods have been shown that although they may not always converge, they often offer better computational performance. In most cases, hybrid conjugate gradient methods are more efficient than basic conjugate gradient methods. Recently, Dai & 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. They established the global convergence of these hybrid computational schemes under the weak Wolfe conditions. N. Andrei [9] Combined between PRP and DY conjugate gradient methods, proposed the following hybrid method:

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

where the parameter in the convex combination is computed in such a way that the conjugacy condition is satisfied, independently of the line search. some kinds of new hybrid conjugate gradient methods are given in [10, 11].

In this paper, we propose another hybrid conjugate gradient as a convex combination of PRP and FR conjugate gradient algorithms. By this method, we hope to obtain a more efficient conjugate gradient algorithm. 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.

DESCRIPTION OF ALGORITHM

Based on the ideas of N. Andrei [9], we propose another hybrid of β^{PRP} and β^{FR} as following:

$$b = u b^{FR} + (1 -\ u) b^{PRP}, 0 \pounds\ u \pounds\ 1, \tag{6}$$

where u is a scalar parameter. Obviously, if $u=1$, then $\beta = \beta^{PRP}$, and if $u = 0$,then $b = b^{FR}$. On the other hand, if $0 < u < 1$, then, b is a convex combination of b^{PRP} and b^{FR}. Hence, from $y_{k-1}^T d_k = 0$ (the conjugacy condition), after some algebra, we get

$$u_k = \frac{(y_{k-1}^T g_k)(\|g_{k-1}\|^2 - y_{k-1}^T d_{k-1})}{(g_k^T g_{k-1})(y_{k-1}^T d_{k-1})}. \tag{7}$$

Further, when one of the following two groups of inequality holds,

$$\begin{cases} g_k^T g_{k-1} \geq 0 \\ y_{k-1}^T g_k \geq 0 \end{cases} \text{or} \begin{cases} g_k^T g_{k-1} < 0 \\ g_k^T d_{k-1} \leq 0 \end{cases}, \tag{8}$$

it holds that $u \in [0,1)$. Hence, we constructed the following parameters

$$\beta_k = \begin{cases} u_k \dfrac{\|g_k\|^2}{\|g_{k-1}\|^2} + (1-u_k)\dfrac{g_k^T y_{k-1}}{\|g_{k-1}\|^2} = u_k \beta^{FR} + (1-u_k)\beta^{PRP}, \text{if } g_k \text{ satisfies (8)}, \\ 0, \quad \text{otherwise.} \end{cases} \tag{9}$$

Now we state our algorithm as follows.
Algorithm PRPFR:

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

$$\begin{cases} d_0 = -g_0 \\ d_k = -g_k + \beta_k d_{k-1}, k \geq 1, \ \ \textit{where } \beta_k \textit{ is followed by } (9). \end{cases}$$

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 $L_0 = \{x \in R^n \mid f(x) \leq f(x_0)\}$, 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

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

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 PRPFR, then $g_k^T d_k < 0$.

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

Suppose the assertion has been proved for order $n = k-1$, i.e. $g_{k-1}^T d_{k-1} < 0$.

We shall show that it is then valid for order $n = k$,

$g_k^T d_k = g_k^T(-g_k + \beta_k d_{k-1})$.

If $\beta_k = 0$, then $g_k^T d_k < 0$. *If* g_k satisfies (8), then

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

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

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

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

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

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

By (4) shows, if the first group inequality holds of (8) , then there

$$g_k^T d_k \leq \frac{(g_{k-1}^T d_{k-1}) \| g_k \|^2 - (g_{k-1}^T g_k)(\sigma_1 g_{k-1}^T d_{k-1})}{y_{k-1}^T d_{k-1}}$$

$$= \frac{(\| g_k \|^2 - \sigma_1 g_{k-1}^T g_k)(g_{k-1}^T d_{k-1})}{y_{k-1}^T d_{k-1}}$$

$$< \frac{(\| g_k \|^2 - g_{k-1}^T g_k)(g_{k-1}^T d_{k-1})}{y_{k-1}^T d_{k-1}} \leq 0,$$

if the second group inequality holds of (8) , then there is

$$g_k^T d_k \leq \frac{\frac{1}{\sigma_1}(g_k^T d_{k-1}) \| g_k \|^2 - (g_k^T g_{k-1})(g_k^T d_{k-1})}{y_{k-1}^T d_{k-1}}$$

$$= \frac{(\frac{1}{\sigma_1} \| g_k \|^2 - g_k^T g_{k-1})(g_k^T d_{k-1})}{y_{k-1}^T d_{k-1}} < 0.$$

Thus, to sum up, $g_k^T d_{k-1} < 0$ holds for all $k \geq 1$. i.e. the theorem is proved.

In view of Theorem 3.1 and[12], we may obtain the following results.

Theorem 3.2 Suppose that Assumption H3.1 holds and the sequence $\{x_k\}$ is generated by Algorithm PRPFR. Then $\liminf_{k \to \infty} \| g_k \| = 0$.

NUMERICAL EXPERIMENTS

In this section, we give some numerical results of Algorithm A to show that the method is efficient for unconstrained optimization problems. The problems that we tested are from [13] and [14]. 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_*
Rosen	(1.00091143028257, 1.00182635749054)	8.314160330210927e-007
Freud	(11.41271934114850, -0.89680858859295)	48.98425368072392
Beale	(3.00323465930368, 0.50080623002464)	1.669350396112912e-006
Trigonometric	(0.24215550125275, 0.61293925994241)	3.566294149196800e-007
Brown	(0.99832798989672, 1.00265604787958)	1.432915886552999e-006

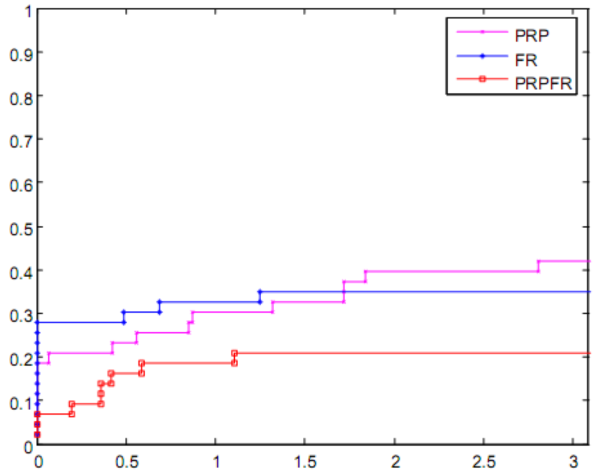

Fig-1: Performance profiles the number of function evaluations

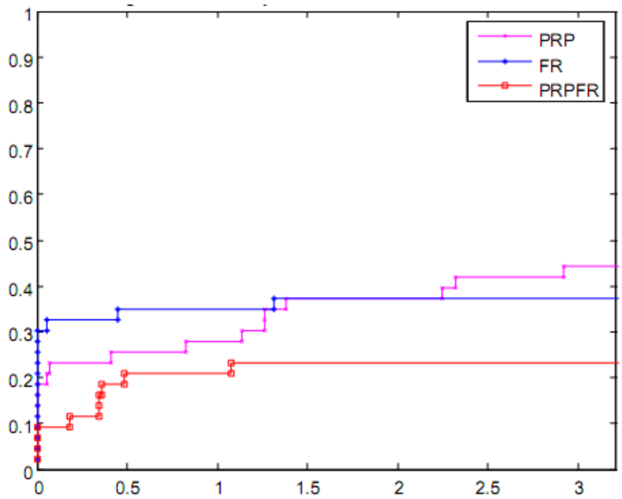

Fig-2: Performance profiles the number of iterations

Fig1-2 show the performance of the three methods relative to the function evaluations and iterations. All the methods successfully solved all the problems. From the figure, we see that the new method is very much competitive with the other methods.

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. Fletcher R, Reeves CM. Function minimization by conjugate gradients. The computer journal. 1964 Jan 1;7(2):149-54.
2. Polak E, Ribi´ere G. “Note sur la convergence de m´ethodesde directions conjugu´ees,” Revue Franc¸aise de RechercheOp´erationnelle. 1969;16:35–43.
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. Unconstained Optimization: Practical Methods of Optimization, vol. 1, John Wiley & Sons, New York, NY, USA, 1987.
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;409–436, 1952.
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. Zhou A, Zhu Z, Fan H, Qing Q. Three new hybrid conjugate gradient methods for optimization. Applied Mathematics. 2011 Mar 24;2(03):303.
11. 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.
12. Sun W, Yuan YX. Optimization theory and methods: nonlinear programming. Springer Science & Business Media; 2006 Aug 6.
13. Molga M, Smutnicki C. “Test functions for optimization needs,” http://www.zsd.ict.pwr.wroc.pl/files/ docs/functions.pdf, 2005.
14. Hock W, Schittkowski K. Test examples for nonlinear programming codes. Journal of Optimization Theory and Applications. 1980 Jan 1;30(1):127-9.