Development a Special Conjugate Gradient Algorithm for Solving Unconstrained Minimization Problems

Basim A. Hassan

basimabas39@gmial.com

College of Computer Sciences and Mathematics
University of Mosul

Received on: 20/04/2011 **Accepted on: 02/11/2011**

ABSTRACT

This paper develops a special conjugate gradient algorithm for solving unconstrained minimized problems. This development can be regarded as some kind of convex combination of the MPR and MLS methods. Experimental results indicate that the new algorithm is more efficient than the Polak and Ribiere - algorithm.

Keywords: Conjugate gradient methods, A special of conjugate gradient, Experimental results

استحداث خوارزمية للتدرج المترافق الطيفي لحل المسائل التصغيرية غير المقيدة

باسم عباس حسن

كلية علوم الحاسبات والرياضيات

جامعة الموصل

تاريخ قبول البحث: ٢٠١١/١١/٢ **تاريخ استلام البحث: ٢٠١١/4/٢٠**

الملخص

في هذا البحث تم استحداث خوارزمية من خوارزميات التدرج المترافق الطيفي لحل المسائل التصغيرية غير المقيدة. هذا الاستحداث يمكن إن يعتبر نوعاً من المجاميع المحدبة لطريقتي MPR وMLS. النتائج العددية أثبتت كفاءة الخوارزمية الجديدة مقارنة بخوارزمية Polak وRibiere.

الكلمات المفتاحية : طرائق التدرج المترافق ، التدرج المترافق الطيفية، النتائج العددية.

ABSTRACT

This paper develops a special conjugate gradient algorithm for solving unconstrained minimized problems. This development can be regarded as some kind of convex combination of the MPR and MLS methods. Experimental results indicate that the new algorithm is more efficient than the Polak and Ribiere - algorithm .

1. Introduction

Let us consider the unconstrained optimization problem

$$\min\left\{ f(x) \;\middle|\; x \in R^n \right\} \qquad ....(1)$$

where $f : R^n \to R$ is a continuously differentiable function, bounded from below. For solving this problem, starting from an initial guess $x_0 \in R^n$, a nonlinear conjugate gradient method, generates a sequence $\{x_k\}$ as :

$$x_{k+1} = x_k + \alpha_k d_k \qquad ....(2)$$

where $\alpha_k > 0$ is obtained by line search, and the direction d_k is generated as $d_{k+1} = -g_{k+1} + \beta_k d_k, \; d_0 = -g_0$ .... (3)

where β_k is known as the conjugate gradient parameter, $v_k = x_{k+1} - x_k$ and $g_k = \nabla f(x_k)$. Consider $\| . \|$ the Euclidean norm and $y_k = g_{k+1} - g_k$ [1]. The step size α_k is chosen in such a way that $\alpha_k > 0$ and satisfies the strong Wolfe (SW) conditions

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta_1 \alpha_k d_k^T g_k \quad .... (4)$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \le -\delta_2 d_k^T g_k \quad .... (5)$$

with $0 < \delta_1 < \delta_2 < 1$, where $f_k = f(x_k)$, $g_k = g(x_k)$, g_k are the gradient of f evaluated at the current iterate x_k [7]. Where d_k is a descent direction. Different conjugate gradient algorithms correspond to different choices for the scalar parameter β_k. Some of these methods as Fletcher and Reeves (FR) [4], Dai and Yuan (DY) [2] and Conjugate Descent (CD) [3] :

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \; , \beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{y_k^T d_k} \; , \beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{\left| g_k^T d_k \right|} \, , \quad .... (6)$$

They have strong convergence properties, but they may have modest practical performance due to jamming. On the other hand, the methods of Polak and Ribiere (PR) [8], Hestenes and Stiefel (HS) [5], or Liu and Storey, (LS) [6] :

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k} \; , \beta_k^{HS} = \frac{g_{k+1}^T y_k}{y_k^T d_k} \; , \beta_k^{LS} = \frac{g_{k+1}^T y_k}{\left| g_k^T d_k \right|} \, , \quad .... (7)$$

in general, may not be convergent, but they often have better computational performances.

In [9] modified methods Polak and Ribiere (MPR), Liu and Storey (MLS) are given by the rule

$$d_{k+1} = -\varphi_k g_{k+1} + \beta_k d_k, \; d_0 = -g_0 \quad .... (8)$$

where the values $\beta_k^{PR}, \beta_k^{LS}$ are determined by (7) and

$$\varphi_k^{PR} = \frac{y_k^T d_k}{g_k^T g_k} \; , \varphi_k^{LS} = \frac{y_k^T d_k}{\left| g_k^T d_k \right|} \, , \quad .... (9)$$

If the minimized function is quadratic, then its gradients are mutually orthogonal, and so

$$y_k^T g_{k+1} = g_{k+1}^T g_{k+1} - g_k^T g_{k+1} = g_{k+1}^T g_{k+1} \; . \quad .... (10)$$

Another conjugate gradient method can combine the Polak-Ribiere and Liu-Story which is defined by

$$\beta_{k+1}^{SG} = \frac{g_{k+1}^T y_k}{(1-u) g_k^T g_k - u d_k^T g_k}, \quad .... (11)$$

Where $u \in [0 \; , \; 1]$ is a constant. Obviously, $\beta_k^{SG} = \beta_k^{PRP}$ for $u = 0$, and $\beta_k^{SG} = \beta_k^{LS}$ for $u = 1$, (for more details see [9]).

The structure of the paper is as follows. In section 2, we present the new special conjugate gradient . Section 3 new Algorithm and Convergence. Section 4 numerical results are presented and Section 5 we also give brief conclusions and discussions.

2. A Special of Conjugate Gradient Method

The modified Polak-Ribiere and modified Liu-Story conjugate gradient methods are special cases of the new class of conjugate gradient methods which is defined by

$$\varphi_k^{NEW} = \frac{y_k^T d_k}{(1-u) g_k^T g_k + u \left| d_k^T g_k \right|}, \qquad \text{..........(12)}$$

Where $u \in [0\ ,\ 1]$ is a constant. Obviously, $\varphi_k^{NEW} = \varphi_k^{PRP}$ for $u = 0$, and $\varphi_k^{NEW} = \varphi_k^{LS}$ for $u = 1$. The search direction generated by the method at each iteration satisfies the sufficient descent condition. Special attention must be paid to how to keep the descent property of conjugate gradient methods. Let us consider the method (8) with the step length α_k satisfying the strong Wolfe conditions $(4)-(5)$. Assume that the search direction d_k is downhill, namely,

$$g_k^T d_k < 0\ . \qquad \text{..........(13)}$$

It follows from (8) that

$$g_{k+1}^T d_{k+1} = -\varphi_k \|g_{k+1}\|^2 + \beta_k g_{k+1}^T d_k\ . \qquad \text{..........(14)}$$

Then, the descent property of d_k requires

$$\varphi_k \|g_{k+1}\|^2 > \beta_k g_{k+1}^T d_k\ . \qquad \text{..........(15)}$$

Our motivation to get a good algorithm for solving (1) is to choose the parameter u in (12) in such a way so that for every $k \geq 1$ the direction d_{k+1} given by (8) is the Newton direction. This is motivated by the fact that when the initial point x_0 is near the solution of (1) and the Hessian is a nonsingular matrix then the Newton direction is the best line search direction. Therefore, from the equation

$$-G^{-1} g_{k+1} = -\varphi_k g_{k+1} + \beta_k^{PR} d_k\ . \qquad \text{..........(16)}$$

Multiplying (16) by y_k^T, we have

$$-G^{-1} y_k^T g_{k+1} = -\varphi_k y_k^T g_{k+1} + \frac{g_{k+1}^T y_k}{\|g_k\|^2} y_k^T d_k \qquad \text{..........(17)}$$

Since $G y_k = v_k$ then we have

$$-v_k^T g_{k+1} = -\varphi_k y_k^T g_{k+1} + \frac{g_{k+1}^T y_k}{\|g_k\|^2} y_k^T d_k \qquad \text{..........(18)}$$

from (18) we get :

$$\varphi_k = \frac{v_k^T g_{k+1}}{y_k^T g_{k+1}} + \frac{y_k^T d_k}{\|g_k\|^2}$$

$$\frac{y_k^T d_k}{(1-u) g_k^T g_k + u \left| d_k^T g_k \right|} = \frac{v_k^T g_{k+1}}{y_k^T g_{k+1}} + \frac{y_k^T d_k}{\|g_k\|^2} \qquad \text{..........(19)}$$

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

$$y_k^T d_k (y_k^T g_{k+1})(\|g_k\|^2) = (\|g_k\|^2 v_k^T g_{k+1} + y_k^T g_{k+1} y_k^T d_k) \quad ((1-u) g_k^T g_k + u\ |d_k^T g_k|)$$

$$y_k^T d_k (y_k^T g_{k+1})(\|g_k\|^2) = \|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} + \|g_k\|^2 y_k^T g_{k+1} y_k^T d_k - u\|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} \quad \text{......(20)}$$

$$- u\|g_k\|^2 y_k^T g_{k+1} y_k^T d_k + u\ |d_k^T g_k| \|g_k\|^2 v_k^T g_{k+1} + u\ |d_k^T g_k| y_k^T g_{k+1} y_k^T d_k$$

$$y_k^T d_k (y_k^T g_{k+1})(\|g_k\|^2) - \|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} - \|g_k\|^2 y_k^T g_{k+1} y_k^T d_{kk+1} =$$
$$- u\|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} - u\|g_k\|^2 y_k^T g_{k+1} y_k^T d_k + u\ |d_k^T g_k| \|g_k\|^2 v_k^T g_{k+1} + u\ |d_k^T g_k| y_k^T g_{k+1} y_k^T d \quad \text{......(21)}$$

$$y_k^T d_k (y_k^T g_{k+1})(\|g_k\|^2) - \|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} - \|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} =$$
$$u\,(- \|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} - \|g_k\|^2 y_k^T g_{k+1} y_k^T d_k + |d_k^T g_k| \|g_k\|^2 v_k^T g_{k+1} + |d_k^T g_k| y_k^T g_{k+1} y_k^T d\ ) \quad \text{......(22)}$$

and from (22) we get :

$$u = \frac{-\|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1}}{-\|g_k\|^2 \|g_k\|^2 v_k^T g_{k+1} + |d_k^T g_k| \|g_k\|^2 v_k^T g_{k+1} - \|g_k\|^2 (y_k^T g_{k+1}) y_k^T d_k + |d_k^T g_k| y_k^T g_{k+1} y_k^T d_k} \quad \text{....(23)}$$

3. New Algorithm and Convergence

First, we will give the following assumptions on objective function $f(x)$, which have been used often in the literature to analyze the global convergence of conjugate gradient methods with inexact line searches.

Assumptions

i- The level set $L = \{x \in R^n \,|\, f(x) \le f(x_0)\}$ is bounded.

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

$$\|g(x_{k+1}) - g(x_k)\| \le \mu \|x_{k+1} - x_k\|, \ \forall x_{k+1}, x_k \in U . \quad \text{..........(24)}$$

3.1. The Algorithm has the Following Steps :

Step 0 : Given parameters $\varepsilon = 1*10^{-5}$, $\delta_1 \in (0,1)$, $\delta_2 \in (0,1/2)$ choose initial point $x_0 \in R^n$.

Step 1 : Computing g_k ; if $\|g_k\| \le \varepsilon$ then stop ; else continue .

Step 2 : Set $\beta_k = \beta_k^{PR}$, $\varphi_k^{NEW} = \dfrac{y_k^T d_k}{(1-u) g_k^T g_k + u\ |d_k^T g_k|}$, , u is defined by (23) .

Step 3 : Set $x_{k+1} = x_k + \alpha_k d_k$, (Use strong Wolfe line search technique to compute the parameter α_k)

Step 4 : Compute $d_{k+1} = -\varphi_k^{NEW} g_{k+1} + \beta_k d_k$,

Step 5 : Go to step (1) with new values of x_{k+1} and g_{k+1}.

Theorem (3.1)

Suppose that α_k in (2) satisfies the strong Wolfe conditions (4) − (5) , then the direction d_{k+1} given by (8) is a sufficient descent direction provided that $g_{k+1}^T y_k > 0$.

Proof.

Since $d_0 = -g_0$, we have $g_0^T d_0 \le -\|g_0\|^2 < 0$. Assume by induction that

$$g_k^T d_k \le -c\|g_k\|^2 < 0 \text{ where } 0 < c < 1 \quad \text{..........}(25)$$

which is a sufficient descent direction. To complete the proof, we have to show that the theorem is true for all $k+1$. first we have to prove that $\varphi_k \ge 0$

$$\varphi_k = \frac{y_k^T d_k}{(1-u)g_k^T g_k + u\,|d_k^T g_k|} \quad \text{..........}(26)$$

From (25), we have

$$\varphi_k \ge \frac{y_k^T d_k}{(1-u)g_k^T g_k + u\,c\|g_k\|^2} \quad \Rightarrow \quad \varphi_k \ge \frac{y_k^T d_k}{(1-u+uc)\|g_k\|^2} \quad \text{..........}(27)$$

multiplying (27) by α_k we get

$$\varphi_k \ge \frac{\alpha_k y_k^T d_k}{\alpha_k (1-u+uc)\|g_k\|^2} = \frac{y_k^T v_k}{\alpha_k (1-u+uc)\|g_k\|^2} = \frac{y_k^T v_k}{\alpha_k \bar{c}\|g_k\|^2} \quad \text{..........}(28)$$

Since $\alpha_k > 0$ and $c \le \bar{c} \le 1$ are positive and $y_k^T v_k$ is always positive, now (28), this yields :

$$\varphi_k \ge 0 \quad \text{..........}(29)$$

Multiplying (8) by g_{k+1}^T, we have

$$g_{k+1}^T d_{k+1} = -\varphi_k \|g_{k+1}\|^2 + \beta_k g_{k+1}^T d_k$$

$$= -\frac{y_k^T d_k}{(1-u)g_k^T g_k + u\,|d_k^T g_k|}\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{\|g_k\|^2} g_{k+1}^T d_k \quad \text{..........}(30)$$

By using (5) and (25), we have

$$g_{k+1}^T d_{k+1} \le -\frac{y_k^T d_k}{(1-u)g_k^T g_k + u\,c\|g_k\|^2}\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{\|g_k\|^2} g_{k+1}^T d_k$$

$$= -\frac{y_k^T d_k}{(1-u+u\,c)\,\|g_k\|^2}\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{\|g_k\|^2}(-\delta_2 g_k^T d_k)$$

$$= -\frac{y_k^T d_k}{\bar{c}\,\|g_k\|^2}\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k}{\|g_k\|^2}(c\delta_2\|g_k\|^2)$$

$$= -\frac{y_k^T d_k}{\bar{c}\,\|g_k\|^2}\|g_{k+1}\|^2 + \delta_3 g_{k+1}^T y_k$$

$$\le -\frac{\|y_k\|\|d_k\|}{\bar{c}\,\|g_k\|^2}\|g_{k+1}\|^2 + \delta_3\|g_{k+1}\|\|y_k\| \quad \text{.........}(31)$$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -\left[\frac{\|d_k\|}{\bar{c}\,\|g_k\|^2} - \frac{\delta_3}{\|g_{k+1}\|}\right]\|y_k\|$$
$$\le -\left[\frac{\|d_k\|}{\bar{c}\,\|g_k\|^2} - \frac{\delta_3}{\|g_{k+1}\|}\right]\|y_k\| \qquad \text{..........(32)}$$

where $\delta_3 = c\delta_2$. Since c and δ_2 are small positive values, then δ_3 is very small and $\bar{c} = 1 - u + uc$ and $u \in [0,1]$, we get $c \le \bar{c} \le 1$ which is also very small because $\frac{1}{\bar{c}}$ is very large. Now the first part of (32) is larger than the second part, hence we have

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -c \ . \qquad \text{..........(33)}$$

Theorem (3.2)

Suppose that assumption holds. Let $\{g_k\}$ and $\{d_k\}$ be generated by Algorithm, then we have

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \ . \qquad \text{..........(34)}$$

Proof :

From Theorem (3.1), we have $g_k^T d_k < 0$ for all $k \ge 1$. We also have from (5) and assumption (ii) that

$$-(1-\delta_2) d_k^T g_k \le (g_{k+1} - g_k)^T d_k \le \mu \alpha_k \|d_k\|^2 \ . \qquad \text{..........(35)}$$

Thus,

$$\alpha_k \ge -\frac{1-\delta_2}{\mu} \frac{g_k^T d_k}{\|d_k\|^2} \ , \qquad \text{..........(36)}$$

Which combines (4), we get

$$f(x_k) - f(x_{k+1}) \ge -\delta_1 \alpha_k g_k^T d_k \ge \delta_1 \frac{1-\delta_2}{\mu} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \ . \qquad \text{..........(37)}$$

Further, from assumption (i) we have $\{f(x_k)\}$ which is a decreasing sequence and has a bound below in L, and shows $\lim_{k\to\infty} f(x_{k+1}) < \infty$, this together with (37) shows

$$\infty > f(x_1) - \lim_{k\to\infty} f(x_{k+1}) = \sum_{k\ge 1}\left[f(x_k) - f(x_{k+1})\right] \ge \delta_1 \frac{1-\delta_2}{\mu} \sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2}. \ , \qquad \text{..........(38)}$$

We can conclude that (34) holds.**[10]**

Property (*)

Consider a method of the from (2) and (8), and suppose that

$$0 < \gamma \le \|g_{k+1}\| \le \bar{\gamma} \qquad \text{..........(39)}$$

for all $k \ge 1$. Under this assumption, we say that the method has the property (*) if there exists constants $b > 0$ and $\lambda > 0$ such that for all k **:**

$|\beta_k| \le b$, ...(40)

and

$\|v_k\| \le \lambda \quad \Rightarrow \quad |\beta_k| \le \dfrac{1}{2b}$ ....(41)

It is easy to see that under assumptions (i) and (ii) the PR method has the property (*).

For the PR method, using the constants γ and $\bar{\gamma}$ in (39), we can choose $b = 2\bar{\gamma}/\gamma^2$ and $\lambda = \gamma^2 / 2(L\bar{\gamma} b)$. Then, we have from β_k^{PR} and (39),

$$|\beta_k| \le \frac{(\|g_{k+1}\| + \|g_k\|)\|g_{k+1}\|}{\|g_k\|^2} \le \frac{2\bar{\gamma}}{\gamma^2} = b, \qquad \text{..........(42)}$$

and when $\|v_k\| \le \lambda$, we have from (24),

$$|\beta_k| \le \frac{(\|y_k\| + \|g_k\|)\|g_{k+1}\|}{\|g_k\|^2} \le \frac{\mu\lambda\bar{\gamma}}{\gamma^2} = \frac{1}{b}. \qquad \text{..........(43)}$$

Lemma (3.1)

Suppose that assumptions hold. Let $\{d_{k+1}\}$ be generated by the algorithm(3.1). If there exists a constant $\gamma > 0$, such that $\|g_{k+1}\| \ge \gamma$ for all $k+1$, we have

$$\sum_{k \ge 1} \|\mu_{k+1} - \mu_k\|^2 < \infty. \quad \text{where} \quad \mu_{k+1} = d_{k+1}/\|d_{k+1}\|. \qquad \text{..........(45)}$$

Proof :

First, note that $d_{k+1} \ne 0$. Therefore, μ_{k+1} is well defined. Now, let us define

$$r_{k+1} = \frac{-\varphi_k^{NEW} g_{k+1}}{\|d_{k+1}\|} \quad \text{and} \quad \delta_{k+1} = \frac{\beta_k \|d_k\|}{\|d_{k+1}\|} \qquad \text{..........(46)}$$

Form (8), we have for $k \ge 1$:

$$\mu_{k+1} = r_{k+1} + \delta_{k+1}\mu_k . \qquad \text{..........(47)}$$

Using the identity $\|\mu_{k+1}\| = \|\mu_k\|$ and (47), we have

$$\|r_{k+1}\| = \|\mu_{k+1} - \delta_{k+1}\mu_k\| = \|\delta_{k+1}\mu_{k+1} - \mu_k\| \qquad \text{..........(48)}$$

(the last equality can be verified by squaring both sides). Using the condition $\delta_k \ge 0$, the triangle inequality, and (48), we obtain

$$\begin{aligned} \|\mu_{k+1} - \mu_k\| &\le \|(1+\delta_{k+1})\mu_{k+1} - (1+\delta_{k+1})\mu_k\| \\ &\le \|\mu_{k+1} - \delta_{k+1}\mu_k\| + \|\delta_{k+1}\mu_{k+1} - \mu_k\| \qquad \text{..........(49)} \\ &= 2\|r_{k+1}\|. \end{aligned}$$

Now, by $\cos\theta_{k+1} = -\langle g_{k+1}, d_{k+1}\rangle / \|g_{k+1}\|\|d_{k+1}\|$ and (33), we have

$$\cos\theta_{k+1} \ge c\|g_{k+1}\| / \|d_{k+1}\|. \qquad \text{..........(50)}$$

This relation, Zoutendijk's condition and (46) imply

$$\sum_{k \ge 1} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} = \sum_{k \ge 1} \|r_{k+1}\|^2 \|g_{k+1}\|^2 < \infty. \qquad \text{..........(51)}$$

Using $\|g_{k+1}\| \geq \gamma$, we obtain

$$\sum_{k\geq 1} \|r_{k+1}\|^2 < \infty. \qquad ....(52)$$

Which together with (49) completes the proof.

Lemma (3.2)

Suppose those assumptions and (33) hold. Let $\{v_{k+1}\}$ and $\{d_{k+1}\}$ be generated by the algorithm(3.1). We have β_k^{PR} has property (*), if there exists a constant $\gamma > 0$, such that $\|g_{k+1}\| \geq \gamma$ for all $k+1$, then, for any $\gamma > 0$, there exist $\Delta \in Z^+$ and $k_1 \in Z^+$, for all $k \geq k_1$, such that

$$\left|\kappa_{K,\Delta}^{\lambda}\right| \geq \frac{\Delta}{2}, \qquad ....(46)$$

where $\kappa_{K,\Delta}^{\lambda} = \{i \in Z^+ : k+1 \leq i \leq k+\Delta, \|v_k\| > \lambda\}, \left|\kappa_{K,\Delta}^{\lambda}\right|$ denotes the number of the $\kappa_{K,\Delta}^{\lambda}$.
If (39) hold and the methods have Property (*),then, the small step length should not be too many. The above lemma shows this property.

Lemma (3.3)

Suppose that assumptions and (39) hold. Let $\{x_{k+1}\}$ be generated by the (1) and (2), α_{k-1} satisfies SWP, and $\beta_k^{PR} > 0$ has property (*). Then

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

The proofs of Lemma (3.2) and Lemma (3.3) had been given in [3,10].

4. Numerical Results

In this section, we reported some numerical results obtained with the implementation of the new algorithm on a set of unconstrained optimization test problems. We have selected (8) large scale unconstrained optimization problems in extended or generalized form, for each test function, we have considered numerical experiment with the number of variable n=100-1000. Using the standard Wolfe line search conditions (4) and (5) with $\delta_1 = 0.0001$ and $\delta_2 = 0.1$ In the all these cases, the stopping criteria is the $\|g_k\| \leq 10^{-5}$. The programs were written in Fortran 90. The test functions were commonly used for unconstrained test problems with standard starting points and a summary of the results of these test functions was given in Tables (3.1) and (3.2) . We tabulate for comparison of these algorithms, the number of function evaluations (NOF) and the number of iterations (NOI) .

Table (4.1)

No.	n	NOF (NOI) u=0.9	NOF (NOI) u=0.5	NOF (NOI) u=0.1
1	100	219 (106)	216 (105)	217 (105)
	1000	227 (110)	226 (110)	227 (110)
2	100	256 (97)	298 (136)	246 (112)
	1000	446 (176)	226 (110)	227 (110)
3	100	23 (9)	23 (9)	22 (9)
	1000	23 (9)	23 (9)	22 (9)
4	100	103 (19)	144 (20)	154 (20)
	1000	133 (21)	144 (20)	169 (21)
5	100	59 (19)	55 (19)	56 (19)
	1000	59 (19)	55 (19)	56 (19)
6	100	44 (17)	52 (19)	49 (19)
	1000	46 (18)	52 (19)	49 (19)
7	100	57 (19)	60 (22)	58 (21)
	1000	59 (20)	60 (22)	61 (23)
8	100	99 (49)	99 (49)	99 (49)
	1000	141 (70)	141 (70)	141 (70)
	Total	**1994 (688)**	**1874 (758)**	**1853 (735)**

Table (4.2)

No.	n	New with u is defined by (23)	New with u=1.0	PR-CG
		NOF (NOI)	NOF (NOI)	NOF (NOI)
1	100	217 (105)	219 (106)	217 (105)
	1000	225 (109)	225 (109)	227 (110)
2	100	343 (160)	304 (125)	2783 (1389)

	1000	591 (284)	492 (198)	4035 (2015)
3	**100**	23 (9)	23 (9)	23 (9)
	1000	- - (- -)	23 (9)	23 (9)
4	**100**	136 (19)	107 (20)	115 (19)
	1000	151 (20)	200 (27)	173 (23)
5	**100**	52 (17)	58 (20)	61 (22)
	1000	- - (- -)	58 (20)	61 (22)
6	**100**	45 (14)	37 (15)	40 (15)
	1000	45 (14)	37 (15)	40 (15)
7	**100**	60 (22)	57 (20)	62 (22)
	1000	59 (22)	57 (20)	- - (- -)
8	**100**	99 (49)	99 (49)	99 (49)
	1000	141 (70)	141 (70)	145 (72)
	Total	**2197** **(914)**	**2137** **(850)**	**8104** **(3877)**

5. Conclusions and Discussion

In this paper, we have proposed a new CG-type method for solving unconstrained minimization problems. The computational experiments show that the new approaches given in this paper are successful.

Table (4.1) from the preliminary numerical results, we have for problems 4 and 7 the new method is efficient when u is little, and for problem 2,3,5,6 the new method is quite efficient when u is big the results are sensitive to the parameter u, which shows that the new methods are robust.

Table (4.2) gives a comparison between the new-algorithm and the Polak and Ribiere (PR)-algorithm for convex optimization , this table indicates that the new algorithm saves $(75-77)\%$ NOI and $(74-79)\%$ NOF, overall against the standard Polak and Ribiere (PR)-algorithm, especially for our selected test problems.

Relative Efficiency of the Different Methods Discussed in the Paper.

Tools	NOI	NOF
PR-CG	100 %	100 %
New with u defined in (23)	23.19 %	26.65 %
New with u=1.0	25.66 %	21.35 %

Appendix

1. *Generalized wood function*:

$$f(x)=\sum_{i=1}^{n/4} 4(x_{4i-2}-x_{4i-3}^2)^2+(1-x_{4i-3})^2+90(x_{4i}-x_{4i-1}^2)^2+(1-x_{4i-1})^2+ 10.1((x_{4i-2}-1)^2+(x_{4i}-1)^2+19.8((x_{4i-2}-1)+(x_{4i}-1))$$

Starting point: $(-3,-1,-3,-1,\ldots\ldots)^T$

2. *Generalized powell function*:

$$f(x)=\sum_{i=1}^{n/4}(x_{4i-3}-10x_{4i-2})^2+5(x_{4i-1}-x_{4i})^2+(x_{4i-1}-2x_{4i})^2+10(x_{4i-9}-x_{4i})^4+(x_{4i-2}-2x_{4i-1}-x_{4i})^2)$$

Starting point: $(3,1,0,1,\ldots\ldots)^T$

3. *Beale function*:

$$f(x)=(1.5-x_1(1-x_2))^2+(2.25-x_1(1-x_2^2)^2+(2.652-x_1(1-x_2^3)^2$$

Starting point: $(0,0,\ldots\ldots)^T$

4. Cantrell *function*:

$$f(x)=\sum_{i=1}^{n/4}[\exp(x_{4i-3})-x_{4i-2}]^4+100(x_{4i-2}-x_{4i-1})^6+[\tan^{-1}(x_{4i-1}-x_{4i})]^4+x_{4i-3}^8$$

Starting point: $(1,2,2,2,\ldots\ldots)^T$

5. *Rosenbrock function*:

$$f(x)=\sum_{i=1}^{n/2}(100(x_{2i}-x_{2i-1}^2)^2+(1-x_{2i-1})^2)$$

Starting point: $(-1.2,1,-1.2,1,\ldots\ldots)^T$

6. *Cubic function*:

$$f(x)=\sum_{i=1}^{n/2}(100(x_{2i}-x_{2i-1}^3)^2+(1-x_{2i-1})^2)$$

Starting point: $(-1.2,1,-1.2,1,\ldots\ldots)^T$

7. *Non-diagonal function*:

$$f(x)=\sum_{i=1}^{n/2}(100(x_i-x_i^3)^2+(1-x_i)^2)$$

Starting point: $(-1,\ldots\ldots)^T$

8. *Welfe function*:

$$f(x)=(-x_1(3-x_1/2)+2x_2-1)^2+\sum_{i=1}^{n-1}(x_{i-1}-x_i(3-x_i(3-x_i/2)+2x_{i+1}-1)^2+(x_{n+1}-x_n(3x_n/2-1)^2$$

Starting point: $(-1,\ldots\ldots)^T$

REFERENCES

[1] Andrie, N. (2008) ' An Unconstrained Optimization Test functions collection ' Advanced Modeling and optimization. 10, pp.147-161.

[2] Dai, Y. and Yuan Y. (1999) ' A nonlinear conjugate gradients method with a strong global convergence property' SIAM J. Optimization 10, pp.177-182

[3] Fletcher, R. (1987) ' Practical Methods of Optimization (second edition). John Wiley and Sons, New York.

[4] Fletcher, R. and Reeves C. (1964) ' Function minimization by conjugate gradients ' Computer J. 7, pp.149-154.

[5] Hestenes, M. R. and Stiefel E. L. (1952) ' Method of conjugate gradients for solving linear systems' J. Research Nat. Standards 49, pp.409-436.

[6] Liu Y. and Storey C. (1991) ' Efficient generalized conjugate gradients algorithms ' Part 1 : Theory. J. Optimization Theory and Applications 69, pp.129-137.

[7] Matonoha C., Luksan L. and Vlcek J. (2008) ' Computational experience with conjugate gradient methods for unconstrained optimization ' Technical report no.1038, pp.1-17.

[8] Polak, E. and Ribiere, G. (1969) ' Note for Convergence Direction Conjugate, Revue Francaise Informant, Reserche. Opertionelle, 16, 35-43.

[9] Sum M. and Liu J. (2010) ' A new family of conjugate gradient method with a rmijo line search ' Int. j. of Math. Analysis, 4, 987-993.

[10] Wei C. and Kai-rong W. (2010) ' Global convergence of a new conjugate gradient for modified Liu-Story formal' J.of Eest China Normal University (Natural Science), 1, pp.44-51.