Iraqi Journal of Statistical Sciences (24) 2013
PP[1-16]

A Nonlinear Conjugate Gradient Methods Based on a Modified Secant Condition

Dr. Basim A. Hassan*   Hameed M. Sadiq**

Abstract

In this paper, a new nonlinear conjugate gradient methods based on the modified secant condition is derived which are given by Li and Fukushima (Li and Fukushima, 2001). These methods showed global convergent under some assumptions. Numerical results indicate the efficiency of these methods to solve the given test problems.

طرائق تدرج مترافق غير خطية معتمدة على شرط قاطع معدل

الملخص

تم في هذا البحث اشتقاق طرائق جديدة غير خطية للتدرج المترافق معتمدة على شرط قاطع معدل الذي معطى من قبل لي وفيوكيوشيما (Li and Fukushima, 2001). وقد أظهرت هذه الطرائق تقاربا شاملا تحت بعض الفرضيات. وأشارت النتائج العددية إلى كفاءة هذه الطرائق في حل دوال الاختبار المعطاة.

* *Assistant Prof* / *Dept. of Mathematics* / *Computer Sciences and Mathematics College / Mosul* University/ *albasim977@yahoo.com*
** *M.Sc. Student* / *Dept. of Mathematics* / *Computer Sciences and Mathematics College / Mosul* University

Received Date 21/1/2013 —————— Accept Date 14/3/2013

1. Introduction

There are now many conjugate gradient schemes for solving unconstrained optimization problem of the form :

$$\min f(x)\ ,\ x \in R^n \qquad \text{..........(1)}$$

where $f : R^n \to R$ is a continuously differentiable function whose gradient is denoted by g.

Conjugate gradient methods are very efficient for solving (1) especially when the dimension n is large. A nonlinear conjugate gradient method generates a sequence $\{x_k\}$, starting from an initial point $x_\circ \in R^n$, using the recurrence

$$x_{k+1} = x_k + \alpha_k d_k\ ,\quad k = 0,1,\ldots \qquad \text{..........(2)}$$

where x_k is the current iterate, α_k is a positive scalar and called the steplength which is determined by some line search, and d_k is the search direction generated by the rule

$$d_{k+1} = \begin{cases} -g_{k+1} & \text{if } k = 0 \\ -g_{k+1} + \beta_k d_k & \text{if } k > 0 \end{cases} \qquad \text{..........(3)}$$

where g_{k+1} is gradient of f at x_{k+1}, and β_k is a parameter such that the method reduces to the linear conjugate gradient method in the case when f is strictly convex quadratic function and the line search is exact. Well-known conjugate gradient methods include the Fletcher-Reevess (FR) method (Fletcher and Reeves, 1964), the Polak-Ribiere-Polyak (PRP) method (Polak and Ribiere, 1969), the Hestenes-Stiefel (HS) method (Hestenes and Stiefel, 1952), the Conjugate descent (CD) method (Fletcher, 1989), the Dai-Yuan (DY) method (Dai and Yuan, 1999) and the Liu-Storey (LS) method (Liu and Storey, 1991), They are specified by

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}\ ,\ \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}\ ,\ \beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \qquad \text{..........(4)}$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{d_k^T g_k}\ ,\ \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}\ ,\ \beta_k^{LS} = -\frac{g_{k+1}^T y_k}{d_k^T g_k} \qquad \text{..........(5)}$$

respectively, where $y_k = g_{k+1} - g_k$ and $\|.\|$ stands for the Euclidean norm of vectors.

In the convergence analysis and implementation of conjugate gradient method, one often requires the exact and inexact line search such as the Wolfe conditions or the strong Wolfe conditions. The Wolfe line search is to find α_k such that

$$f(x_k+\alpha_k d_k)\le f(x_k)+\delta\alpha_k g_k^T d_k \qquad ....(6)$$

$$d_k^T g(x_k+\alpha_k d_k)\ge \sigma\, d_k^T g_k \qquad ....(7)$$

with $0<\delta<\sigma$. The strong Wolf line search is to find α_k such that

$$f(x_k+\alpha_k d_k)\le f(x)+\delta\alpha_k g_k^T d_k \qquad ....(8)$$

$$\left|d_k^T g(x_k+\alpha_k d_k)\right|\le -\sigma\, d_k^T g_k \qquad ....(9)$$

where $0<\delta<\sigma<1$ are constants, by Li and Weijun (Li and Weijun, 2008).

In order to introduce our method, let us simply recall the conjugcy condition proposed by Dai and Liao (Dai and Liao, 2001). Linear conjugate gradient methods generate a search direction such that the conjugacy condition holds. Namely

$$d_i Q d_j = 0 \quad \forall\ \mathrm{i}\neq \mathrm{j} \qquad ....(10)$$

where Q is the symmetric and positive definite Hessian matrix of the quadratic objective function $f(x)$. For general nonlinear function, it follows from the mean-value theorem that there exists some $\tau\in(0,1)$ such that

$$d_{k+1}^T y_k = \alpha_k d_{k+1}^T \nabla f^2(x_k+\tau\alpha_k d_k)d_k \qquad ....(11)$$

Therefore it is reasonable to replace (11) by the following conjugacy condition:

$$d_{k+1}^T y_k = 0 \qquad ....(12)$$

Dai and Liao (Dai and Liao, 2001) used the secant condition of quasi-Newton methods that is

$$B_{k+1}s_k = y_k \qquad ....(13)$$

where B_{k+1} is some $n\times n$ symmetric and positive definite matrix. For quasi-Newton methods, the search direction d_{k+1} can be calculated in the form

$$d_{k+1} = -B_{k+1}^T g_{k+1} \qquad \text{..........(14)}$$

by the use (13) and (14) we get that

$$d_{k+1}^T y_k = d_{k+1}^T (B_{k+1} s_k) = d_{k+1}^T B_{k+1} s_k = -g_{k+1}^T s_k \qquad \text{..........(15)}$$

the above relation implies that (12) holds if the line search is exact since in this case
$g_{k+1}^T s_k = 0$, however, practical numerical algorithms normally adopt inexact line searches instead of exact ones. For this reason, Dai and Liao replaced the above conjugacy condition by

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k \qquad \text{..........(16)}$$

where t is a scalar.

To ensure that the search direction d_k satisfies this condition, substituting (3) into condition (16), we have

$$-g_{k+1}^T y_k + \beta_k d_k^T y_k = -t g_{k+1}^T s_k \ . \qquad \text{..........(17)}$$

Motivated by idea of Li and Fukushima (Li and Fukushima, 2001), based on the modified secant condition

$$\bar{G}_{k+1} s_k = (G_{k+1} + \eta I) s_k = y_k + \eta s_k \ , \eta > 0$$
$$B_{k+1} s_k = y_k + \eta s_k = z_k \qquad \text{..........(18)}$$

we can present the new conjugacy condition :

$$d_{k+1}^T z_k = -t g_{k+1}^T s_k \qquad \text{..........(19)}$$

Substituting (3) into condition (18), we have

$$-g_{k+1}^T z_k + \beta_k d_k^T z_k = -g_{k+1}^T s_k \qquad \text{..........(20)}$$

These line search strategies require the descent condition

$$g_{k+1}^T d_{k+1} < 0 \ , \ \text{for all k} \qquad \text{..........(21)}$$

however most of conjugate gradient methods don't always generate a descent condition, so condition (21) is usually assumed in the analyses and implementations, Zhi-Feng Dai (Zhi, 2011).

The structure of the paper is as follows. In section (2) we present the new formulas β_k^{New1} and β_k^{New2}. Section (3) show that the search direction generated by this proposed algorithms at each iteration satisfies the sufficient descent condition and descent algorithm. Section (4) establishes the global convergence property for the new CG-methods. Section (5) establishes some numerical results to show the effectiveness of the proposed CG-method and Section (6) gives a brief conclusions and discussions.

2. The new formulas and Algorithms :

In this section, we derive a new conjugacy condition based on modified secant condition. From modified update formule, we have the following secant condition

$$B_{k+1}s_k = z_k \qquad \text{..........(22)}$$

Since (22) is true for all positive and symmetric matrix, in practical true for diagonal matrix then we can assume

$$B_{k+1} = \gamma\, \mathrm{I} \qquad \text{..........(23)}$$

On the other hand, from the equation (18), we have

$$\begin{aligned} y_k^T B_{k+1} s_k &= y_k^T y_k + \eta\, y_k^T s_k \\ y_k^T (\gamma\, \mathrm{I})\, \mathrm{s_k} &= \mathrm{y_k^T} y_k + \eta\, \mathrm{y_k^T} s_k \qquad \text{..........(24)} \\ \gamma\ y_k^T\, \mathrm{s_k} &= \mathrm{y_k^T} y_k + \eta\, \mathrm{y_k^T} s_k \end{aligned}$$

then we get

$$\gamma = \frac{\|y_k\|^2 + \eta\, \mathrm{y_k^T} s_k}{y_k^T s_k} I_{n\times n} \quad \mathrm{s} \qquad \text{..........(25)}$$

Therefore, from the above equation, we have

$$B_{k+1}^{-1} \approx \frac{1}{\gamma} = \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, \mathrm{y_k^T} s_k} I_{n\times n} \qquad \text{..........(26)}$$

Then the Newton direction $d_{k+1} = -\gamma^{-1} g_{k+1}$ can be written as :

$$d_{k+1}^N = -\left(\frac{y_k^T s_k}{\|y_k\|^2 + \eta\, \mathrm{y_k^T} s_k} \right) g_{k+1} \qquad \text{..........(27)}$$

use the conjugacy condition (10) because Newton direction are conjugate gradient with exact line searches :

$$-\frac{y_k^T s_k}{\|y_k\|^2 + \eta\, \mathrm{y_k^T} s_k} g_{k+1}^T y_k = -g_{k+1}^T y_k + \beta_k d_k^T y_k \qquad \text{..........(28)}$$

Then we have

$$\beta_k = \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, \mathrm{y_k^T} s_k} \right) \frac{g_{k+1}^T y_k}{d_k^T y_k} \qquad \text{..........(29)}$$

$$d_{k+1} = -g_{k+1} + \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k \qquad \text{..........(30)}$$

Since, $s_k = \alpha_k d_k$ then:

$$d_{k+1} = -g_{k+1} + \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{y_k^T s_k} s_k \qquad \text{..........(31)}$$

where new formulae denote by β_k^{New1} is defined by :

$$\beta_k^{New1} = \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} \qquad \text{..........(32)}$$

We can therefore modify the Eq. (32) and Eq. (31) by using the idea of Dai and Laio (Dai and Liao, 2001) and combining the quasi-Newton condition with conjugace condition:

$$-\frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k} g_{k+1}^T y_k + g_{k+1}^T s_k = 0 \qquad \text{..........(33)}$$

and

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \beta_k d_k^T y_k = 0 \qquad \text{..........(34)}$$

From (33) and (34) we get :

$$\beta_k = \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} + \frac{g_{k+1}^T s_k}{d_k^T y_k} \qquad \text{..........(35)}$$

Since, $s_k = \alpha_k d_k$ then :

$$d_{k+1} = -g_{k+1} + \left\{\left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{y_k^T s_k} + \frac{g_{k+1}^T y_k}{y_k^T s_k}\right\} s_k \qquad \text{..........(36)}$$

where new formulae denote by β_k^{New2} is defined by :

$$\beta_k^{New2} = \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} + \frac{g_{k+1}^T s_k}{d_k^T y_k} \qquad \text{..........(37)}$$

It seems from (37) if exact line search used $s_k^T g_{k+1} = 0$ then (37) reduces (32).

Now we can obtain the new conjugate gradient algorithms, as follows:

New Algorithm :

Step 1. Initialization. Select $x_1 \in R^n$ and the parameters $0<\delta_1<\delta_2<1$. Compute $f(x_1)$ and g_1. Consider $d_1=-g_1$ and set the initial guess $\alpha_1 = 1/\|g_1\|$.

Step 2. Test for continuation of iterations. If $\|g_{k+1}\| \le 10^{-6}$, then stop.

Step 3. Line search. Compute $\alpha_{k+1}>0$ satisfying the Wolfe line search condition (8) and (9) and update the variables $x_{k+1} = x_k + \alpha_k d_k$.

Step 4. β_k conjugate gradient parameter which defined in (32) and (37) .

Step 5. Direction computation. Compute $d_{k+1} = -g_{k+1} + \beta_k d_k$. If the restart criterion of Powell $\left|g_{k+1}^T g_k\right| \ge 0.2\,\|g_{k+1}\|^2$, is satisfied, then set $d_{k+1} = -g_{k+1}$. Otherwise define $d_{k+1} = d$. Compute the initial guess $\alpha_k = \alpha_{k-1}\|d_{k-1}\|/\|d_k\|$, set $k=k+1$ and continue with step2 .

3. The Descent Property of the New Methods :

Below we have to show the sufficient descent property for our proposed new conjugate gradient methods, denoted by β_k^{New1} and β_k^{New2} . For the sufficient descent property to hold, then

$$g_k^T d_k \le -c\|g_k\|^2 \quad \text{for} \quad k \ge 0 \text{ and } c>0 \qquad \text{..........(38)}$$

Assumption(1):

Assume f is bound below in the level set $S=\{x \in R^n : f(x) \le f(x_\circ)\}$; In some neighborhood N of S, f is continuously differentiable and its gradient is
Lipshitz continuos, there exist $L>0$ such that:

$$\|g(x)-g(y)\| \le L\|x-y\| \quad \forall \mathrm{x}, \mathrm{y} \in \mathrm{N} \;. \qquad \text{..........(39)}$$

From (39), we get

$$y_k^T y_k \le L y_k^T s_k \,. \qquad \text{..........(40)}$$

Theorem (3.1) :

If $m > 1$ then the search direction (3) and β_k^{New1} given in equation (32) then, condition (38) will holds for all $k \geq 1$.

Proof :

Since $d_0 = -g_0$, we have $g_0^T d_0 = -\|g_0\|^2$, which satisfy (34). Multiplying (27) by g_{k+1}, we have

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left(1 - \frac{y_k^T s_k}{\|y_k\|^2 + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} \qquad ....(41)$$

$$\leq -\|g_{k+1}\|^2 + \left(1 - \frac{y_k^T s_k}{L s_k^T y_k + \eta\, y_k^T s_k}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} \qquad ....(42)$$

$$\leq -\|g_{k+1}\|^2 + \left(1 - \frac{1}{L + \eta}\right) \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} \qquad ....(43)$$

Let $m = \dfrac{1}{L+\eta}$, where m is constant, then :

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + (1-m) \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} \qquad ....(44)$$

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + (1-m) \frac{g_{k+1}^T y_k}{(d_k^T y_k)^2} (d_k^T y_k) d_k^T g_{k+1} \qquad ....(45)$$

Applying the inequality $u^T v \leq \frac{1}{2}(\|u\|^2 + \|v\|^2)$ to the second term of the right hand side of the above equality, with $u = (y_k^T d_k) g_{k+1}$ and $v = (g_{k+1}^T d_k) y_k$ we get :

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\| + \frac{(1-m)}{(d_k^T y_k)^2} \left(\frac{1}{2}\left[\|g_{k+1}\|^2 (y_k^T d_k)^2 + (g_{k+1}^T d_k)^2 \|y_k\|^2\right]\right) \qquad ....(46)$$

$$d_{k+1}^T g_{k+1} \leq \left[\frac{(1-m)}{2} - 1\right] \|g_{k+1}\|^2 + \frac{(1-m)}{2(d_k^T y_k)^2} (g_{k+1}^T d_k)^2 \|y_k\|^2$$

$$\leq \left[\frac{1}{2} - \frac{m}{2} - 1\right] \|g_{k+1}\|^2 + \frac{(1-m)}{2(d_k^T y_k)^2} (g_{k+1}^T d_k)^2 \|y_k\|^2 \qquad ....(47)$$

From (47) we get :

$$d_{k+1}^T g_{k+1} \leq \left[-\frac{1}{2} - \frac{m}{2}\right] \|g_{k+1}\|^2 + \frac{(1-m)}{2(d_k^T y_k)^2} (g_{k+1}^T d_k)^2 \|y_k\|^2 \qquad ....(48)$$

$$d_{k+1}^T g_{k+1} \le -\left[\frac{1}{2}+\frac{m}{2}\right]\|g_{k+1}\|^2 + \frac{(1-m)}{2(d_k^T y_k)^2}(g_{k+1}^T d_k)^2\|y_k\|^2 \qquad \text{..........(49)}$$

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2\left[\frac{1}{2}+\frac{m}{2}\right] + \frac{(g_{k+1}^T d_k)^2}{(d_k^T y_k)^2}\left(\frac{1}{2}(1-m)\|y_k\|^2\right) \qquad \text{..........(50)}$$

Therefore, when $\frac{1}{2}+\frac{m}{2} > 0$ and $1-m<0$, we get

$$\begin{aligned} d_{k+1}^T g_{k+1} &\le -\left(\frac{1}{2}+\frac{m}{2}\right)\|g_{k+1}\|^2 \\ &\le -c\|g_{k+1}\|^2 \end{aligned} \qquad \text{..........(51)}$$

where $c=\frac{1}{2}+\frac{m}{2}$

Theorem (3.2) :

For the line search directions defined by

$$d_{k+1}^T = -g_{k+1} + \left(1-\frac{y_k^T s_k}{\|y_k\|^2+\eta\, y_k^T s_k}\right)\frac{g_{k+1}^T y_k}{d_k^T y_k} d_k + \frac{g_{k+1}^T s_k}{d_k^T y_k} d_k \qquad \text{..........(52)}$$

If $\frac{y_k^T s_k}{\|y_k\|^2+\eta\, y_k^T s_k} = t \ge 1+\frac{2 s_k^T y_k}{\|y_k\|^2}$ then

$$d_{k+1}^T g_{k+1} \le -c\|g_{k+1}\|^2 \ . \qquad \text{..........(53)}$$

Proof :

The inequality (53) holds for $k=0$, clearly. Now, we let $k \ge 1$. From the following inequality

$$u^T v \le \frac{1}{2}\left(\|u^2\|+\|v\|^2\right), \quad \text{and} \quad u, v \in R^n \qquad \text{..........(54)}$$

It can be derived that

$$\left((g_{k+1}^T s_k) y_k\right)^T \left((y_k^T s_k) g_{k+1}\right) \le \frac{1}{2}\left((g_{k+1}^T s_k)^2\|y_k\|^2 + (y_k^T s_k)^2\|g_{k+1}\|^2\right) \qquad \text{..........(55)}$$

So, it follows from (34) , (36) and (55) that

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left(1-\frac{y_k^T s_k}{\|y_k\|^2+\eta\, y_k^T s_k}\right)\frac{g_{k+1}^T y_k}{s_k^T y_k} s_k^T g_{k+1} + \frac{g_{k+1}^T s_k}{s_k^T y_k} s_k^T g_{k+1} \qquad \text{..........(56)}$$

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + (1-t)\frac{1}{2(s_k^T y_k)^2}\left((g_{k+1}^T s_k)^2 \|y^2\| + (y_k^T s_k)^2 \|g_{k+1}\|^2\right) + \frac{g_{k+1}^T s_k}{s_k^T y_k} s_k^T g_{k+1}$$

$$\le -\|g_{k+1}\|^2 + (1-t)\frac{1}{2(s_k^T y_k)^2}\left((g_{k+1}^T s_k)^2 \|y_k\|^2 + (y_k^T s_k)^2 \|g_{k+1}\|^2\right) + \frac{(g_{k+1}^T s_k)^2}{(s_k^T y_k)^2} s_k^T y_k \quad \text{..........(57)}$$

$$\le -\|g_{k+1}\|^2 \left(\frac{1}{2} + \frac{t}{2}\right) + \frac{(g_{k+1}^T s_k)^2}{(s_k^T y_k)^2}\left(s_k^T y_k - \frac{t\|y\|^2}{2} + \frac{\|y\|^2}{2}\right) \quad \text{..........(58)}$$

Therefore, when $c = \frac{1}{2} + \frac{t}{2} > 0$ and $s_k^T y_k - \frac{t\|y\|^2}{2} + \frac{\|y\|^2}{2} \le 0$, we get

$$d_{k+1}^T g_{k+1} \le -\left(\frac{1}{2} + \frac{t}{2}\right)\|g_{k+1}\|^2$$

$$\le -c\|g_{k+1}\|^2. \quad \text{..........(59)}$$

4. Global convergence :

Next we will show that CG method with β_k^{New1} and β_k^{New2} converges globally. We come to study the convergence for uniformly convex function, then there exists a constant $\mu > 0$ such that

$$(\nabla f(x) - \nabla f(y))^T (x - y) \ge \mu \|x - y\|^2 \text{ , for any } x, y \in S \quad \text{..........(60)}$$

or equivalently

$$y_k^T s_k \ge \mu \|s_k\|^2 \quad \text{and} \quad \mu \|s_k\|^2 \le y_k^T s_k \le L\|s_k\|^2 \quad \text{..........(61)}$$

On the other hand, under Assumption(1), It is clear that there exist positive constants B such

$$\|x\| \le B \ , \forall x \in S \quad \text{..........(62)}$$

Proposition:

Under Assumption1 and equation (62) on f, there exists a constant $\bar{\gamma} > 0$ such that

$$\|\nabla f(x)\| \le \bar{\gamma} \ , \ \forall x \in S \quad \text{..........(63)}$$

Lemma(1):

Suppose that Assumption(1) and equation (62) hold. Consider any conjugate gradient method in from (2) and (3), where d_k is a descent direction and α_k is obtained by the strong Wolfe line search. If

$$\sum_{k>1} \frac{1}{\|d_{k+1}\|^2} = \infty \qquad \text{..........(64)}$$

then we have

$$\liminf_{k\to\infty} \|g_k\| = 0 . \qquad \text{..........(65)}$$

More details can be found in (Dai and Liao, 2001) and (Tomizuka and Yabe, 2004).

Theorem (4.1):

Suppose that Assumption (1) and equation (62) and the descent condition hold. Consider a conjugate gradient method in the form $(2)-(3)$ with β_k^{New1} as in (32), where α_k is computed from Wolf line search condition (8) and (9). If the objective function is uniformly on S, then $\liminf_{n\to\infty} \|g_k\| = 0$.

Proof :

Firstly, we need simplify our new β_k^{New1}, So that our convergence proof will be much easier. Let $w = \dfrac{y_k^T s_k}{\|y_k\|^2 + \eta\, \mathrm{y}_\mathrm{k}^\mathrm{T} s_k}$, where w is constant, we obtain :

$$\beta_\mathrm{k}^{\mathrm{New1}} = (1-w)\frac{g_{k+1}^T y_k}{d_k^T y_k} \qquad \text{..........(66)}$$

Now, we get

$$\|\mathrm{d}_{\mathrm{k+1}}\| = \left\| - g_{k+1} + (1-w)\frac{g_{k+1}^T y_k}{d_k^T y_k} d_k \right\| \qquad \text{..........(67)}$$

$$\begin{aligned}
\|\mathrm{d}_{\mathrm{k+1}}\| &= \left\| - g_{k+1} + \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k - w\frac{g_{k+1}^T y_k}{d_k^T y_k} d_k \right\| \\
&= \left\| - g_{k+1} + \frac{g_{k+1}^T y_k}{y_k^T s_k} s_k - w\frac{g_{k+1}^T y_k}{y_k^T s_k} s_k \right\| \\
&\le \|g_{k+1}\| + \frac{\|g_{k+1}\|\|L\|\|s\|^2}{\mu\|s\|^2} - w\frac{\|g_{k+1}\|\|L\|\|s\|^2}{\mu\|s\|^2} \\
&\le \|g_{k+1}\|\left(1 + \frac{L}{\mu} - w\frac{L}{\mu}\right) \qquad \text{..........(68)}
\end{aligned}$$

$$\left\|d_{k+1}\right\| \le \left(\frac{\mu+(1-w)L}{\mu}\right)\bar{\gamma} \qquad \text{..........(69)}$$

This relation shows that

$$\sum_{k\ge 1}\frac{1}{\left\|d_k\right\|^2} \ge \left(\frac{\mu}{\mu+(1-2c)L}\right)^2 \frac{1}{\bar{\gamma}^2}\sum_{k\ge 1} 1 = \infty \qquad \text{..........(70)}$$

Therefore, from Lemma 1 we have $\liminf_{k\to\infty}\left\|g_k\right\| = 0$, which for uniformly convex function equivalent to $\lim_{k\to\infty}\left\|g_k\right\| = 0$.

Theorem (4.2):

Suppose that the assumption 1 holds and consider conjugate gradient algorithm in the form $(2)-(3)$ with β_k^{New2} as (37), where d_k is descent direction and α_k is obtained by Wolf line search. If

$$\sum_{k>1}\frac{1}{\left\|d_{k+1}\right\|^2} \le \infty \qquad \text{..........(71)}$$

Then

$$\liminf_{k\to\infty}\left\|g_k\right\| = 0. \qquad \text{..........(72)}$$

Proof :

From (40), we obtain :

$$\left|\beta_k^{New2}\right| = \left|\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{y_k^T s_k}{\left\|y_k\right\|^2 + \eta\, y_k^T s_k}\frac{g_{k+1}^T y_k}{d_k^T y_k} + \frac{g_{k+1}^T s_k}{d_k^T y_k}\right| \qquad \text{..........(73)}$$

$$\left|\beta_k^{New2}\right| \le \left\|g_{k+1}\right\|\left[\frac{L\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|} - w\frac{L\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|} + \frac{\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|}\right] \qquad \text{..........(74)}$$

From (63) and (74) we have

$$\begin{aligned}\left\|d_{k+1}\right\| &\le \left\|g_{k+1}\right\| + \left|\beta_k^{New2}\right|\left\|d_k\right\| \\ &\le \left\|g_{k+1}\right\| + \left\|g_{k+1}\right\|\left[\frac{L\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|} - w\frac{L\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|} + \frac{\alpha_k\left\|d_k\right\|}{\left|d_k^T y_k\right|}\right]\left\|d_k\right\| \\ &= \left\|g_{k+1}\right\|\left[1 + \frac{L\alpha_k\left\|d\right\|^2}{\left|d_k^T y_k\right|} - w\frac{L\alpha_k\left\|d_k\right\|^2}{\left|d_k^T y_k\right|} + \frac{\alpha_k\left\|d_k\right\|^2}{\left|d_k^T y_k\right|}\right] \qquad \text{.......(75)}\end{aligned}$$

From the strong Wolfe conditions (8), (9) and sufficient descent condition, we have

$$d_k^T y_k \geq (\sigma - 1) g_k^T d_k = (1-\sigma)\|g_k\|^2 > 0 \qquad \text{..........(76)}$$

$$\|d_{k+1}\| \leq \|g_{k+1}\| \left[1 + \frac{L\alpha_k \|d_k\|^2}{(1-\sigma)\|g_k\|^2} - w \frac{L\alpha_k \|d_k\|^2}{(1-\sigma)\|g_k\|^2} + \frac{\alpha_k \|d_k\|^2}{(1-\sigma)\|g_k\|^2} \right] \qquad \text{..........(77)}$$

$$\leq \|g_{k+1}\| \left[1 + \frac{L\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2} - w \frac{L\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2} + \frac{\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2} \right]$$

$$= \|g_{k+1}\| M \leq \bar{\gamma} M . \qquad \text{..........(78)}$$

where $M = 1 + \frac{L\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2} - w \frac{L\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2} + \frac{\alpha_k B^2}{(1-\sigma)\bar{\gamma}^2}$. This relation implies

$$\sum_{k \geq 1} \frac{1}{\|d_{k+1}\|^2} \geq \frac{1}{M^2 \bar{\gamma}^2} \sum_{k \geq 1} 1 = \infty \qquad \text{..........(79)}$$

Therefore, we have $\lim_{k \to \infty} \inf \|g_k\| = 0$.

5. Numerical Results:

In this section, we compare the performance of new formals β_k^{New1} and β_k^{New2} CG-method to other classical CG-method (Polak-Ribe`re (PR) algorithm). we have selected (15) large scale unconstrained optimization problem, for each test problems taken from (Andrie, 2008). for each test function we have considered numerical experiments with the number of variables $n = 100, 1000$.

These two new versions are compared with well-known CG-algorithm, the polak-Ribe`re (PR) algorithm. All these algorithms are implemented with standard Wolf line search conditions (8) and (9) with $\delta = 0.001$ and $\sigma = 0.9$.

All codes are written in double precision FORTRAN Language with F77 default compiler settings. We record the number of iterations calls (NOI), and the number of restart calls (IRS) for the purpose our comparisons.

Table (5.1) Comparison of the algorithms for $n = 100$

Test problems	β_k^{PR}		β_k^{New1}		β_k^{New2}	
	NOI	**IRS**	**NOI**	**IRS**	**NOI**	**IRS**

Extended Rosenbrock	**50**	**19**	**34**	**18**	**34**	**18**
Extended While & Holst	**38**	**16**	**34**	**18**	**34**	**18**
Extended PSC 1	**15**	**10**	**10**	**7**	**10**	**7**
Extended Maratos	**94**	**34**	**67**	**32**	**67**	**31**
Quadratic QF2	**111**	**33**	**109**	**36**	**110**	**35**
Arwhead	**10**	**5**	**9**	**5**	**13**	**7**
Nondia	**13**	**7**	**16**	**8**	**15**	**8**
Partial Perturbed Quad.	**85**	**28**	**79**	**22**	**77**	**22**
Liarwhd	**25**	**13**	**23**	**13**	**23**	**13**
Denschnc	**17**	**8**	**17**	**10**	**17**	**10**
Denschnf	**22**	**19**	**23**	**20**	**22**	**18**
Extended Block Diagonal	**122**	**62**	**16**	**10**	**13**	**8**
Generalized Quad. GQ1	**11**	**6**	**11**	**7**	**11**	**7**
Sincos	**15**	**10**	**10**	**7**	**10**	**7**
Generalized Quad. GQ2	**38**	**12**	**36**	**13**	**39**	**17**
Total	**666**	**282**	**494**	**226**	**495**	**226**

Table (5.2) Comparison of the algorithms for $n=1000$

Test problems	β_k^{PR}		β_k^{New1}		β_k^{New2}	
	NOI	**IRS**	**NOI**	**IRS**	**NOI**	**IRS**
Extended Rosenbrock	**93**	**65**	**35**	**19**	**35**	**19**
Extended While & Holst	**348**	**317**	**36**	**19**	**32**	**18**
Extended PSC 1	**8**	**6**	**17**	**15**	**8**	**6**
Extended Maratos	**98**	**36**	**65**	**30**	**64**	**29**
Quadratic QF2	**394**	**174**	**325**	**92**	**416**	**127**
Arwhead	**39**	**23**	**17**	**12**	**21**	**14**
Nondia	**11**	**7**	**14**	**8**	**15**	**8**
Partial Perturbed Quad.	**506**	**264**	**278**	**55**	**224**	**52**
Liarwhd	**48**	**33**	**18**	**10**	**18**	**10**
Denschnc	**128**	**66**	**15**	**10**	**17**	**11**
Denschnf	**23**	**20**	**20**	**16**	**24**	**20**
Extended Block Diagonal	**130**	**66**	**17**	**10**	**13**	**8**
Generalized Quad. GQ1	**9**	**5**	**9**	**6**	**8**	**5**
Sincos	**8**	**6**	**17**	**15**	**8**	**6**
Generalized Quad. GQ2	**112**	**55**	**35**	**14**	**36**	**15**

Total	**1955**	**1143**	**918**	**331**	**939**	**348**

6. Conclusions and Discussions:

In this paper, we have proposed new a nonlinear CG-algorithms based on the modified secant condition defined by (38) and (43) respectively under some assumptions the two new algorithms have been shown to be globlly convergent for uniformly convex and satisfies the sufficient descent property. The computational experiments show that the new two kinds given in this paper are successful .

Table (5.1) gives a comparison between the new-algorithm and the Polak-Ribiere (PR) algorithm for convex optimization, this table indicates, see **Table (6.1)**, that the new algorithm saves $(26-34)\%$ NOI and $(20-72)\%$ IRS overall against the standard Polak-Ribiere (PR) algorithm, especially for our selected group of test problems.

Table(6.1): Relative efficiency of the new Algorithm $(n=100)$

Tools	NOI	IRS
PR Algorithm	100 %	100 %
New Algorithm with β_k^{New1}	74.174%	80.14%
New Algorithm with β_k^{New2}	74.324%	80.14%

Table(6.2): Relative efficiency of the new Algorithm $(n=1000)$

Tools	NOI	IRS
PR Algorithm	100 %	100 %
New Algorithm with β_k^{New1}	46.956%	28.95%
New Algorithm with β_k^{New2}	48.03%	30..44%

References

[1] Andrei, N. (2008), An Unconstrained Optimization test function collection. Adv. Model. Optimization. 10. pp.147-161.

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

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

[4] Hestenes, M. R. and Stiefel, E. L. (1952), 'Method of conjugate gradients for Solving linear systems ' Journal National Standards 49, pp. 409-436.

[5] Fletcher, R. (1989), 'Practical Method of Optimization', (2nd Edition), (John Wiley and Sons, New York).

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

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

[8] Dai, Y. and Liao, Z. (2001),' New conjugate conditions and related nonlinear conjugate gradient methods ' Appl. Math. Optim. 43, pp. 87-101.

[9] Li, H. and Fukushima, M. (2001), 'A modified BFGS method and its global convergence in nonconvex minimization' Journal comput Appl. Math.129, pp. 15-35.

[10] Li, Z. and Weijun, Z. (2008), 'Two descent hybrid conjugate gradient methods for optimization', Journal of computational and Appl. Math.216, pp251-264.

[11] Zhi, F. (2011), 'Tow modified HS type conjugate gradient methods for unconstrained optimization problems', Journal nonlinear analysis 74, pp. 927-936.

[12] Tomizuka, H. and Yabe, H. (2004), 'A Hybrid Conjugate Gradient method for unconstrained Optimization'.