New Scaled Proposed formulas For Conjugate Gradient Methods in Unconstrained Optimization

Abbas Y. Al-Bayati
University of Telafer, Iraq
profabbasalbayati@yahoo.com

Marwan S. Jameel
University of Mosul/Iraq

Received on: 10/6/2013 **Accepted on: 16/9/2013**

ABSTRACT

In this paper, three efficient Scaled Nonlinear Conjugate Gradient (CG) methods for solving unconstrained optimization problems are proposed. These algorithms are implemented with inexact line searches (ILS). Powell restarting criterion is applied to all these algorithms and gives dramatic saving in the computational efficiency. The global convergence results of these algorithms are established under the Strong Wolfe line search condition. Numerical results show that our proposed CG-algorithms are efficient and stationary by comparing with standard Fletcher-Reeves (FR); Polak-Ribiere (PR) CG-algorithms, using 35-nonlinear test functions.

Keywords: Scaled Proposed formulas, Conjugate Gradient,Unconstrained Optimization

خوارزميات مقيسة جديدة للتدرج المترافق في مجال الأمثلية غير المقيدة

مروان صالح جميل
جامعة الموصل/كلية علوم الحاسوب والرياضيات

عباس يونس البياتي
جامعة تلعفر

تاريخ قبول البحث : 2013/9/16 **تاريخ استلام البحث : 2013/6/10**

الملخص

تم في هذا البحث اقتراح ثلاث خوارزميات مقيسة جديدة في مجال الأمثلية غير المقيدة، وقد تم استخدام خط بحث غير تام ومقياس Powell للاسترجاع على جميع الصيغ المستعملة إذ يَعطي توفير جيد في الكفاءةِ الحسابية، خاصية التقارب الشاملة والانحدار الكافي درست بوجود شرطي وولف القوية، كما إن النتائج التي تم التوصل أليها عمليا أثبتت إن الخوارزميات الجديدة هي أكثر كفاءة من الخوارزميات المقارنة (FR & PR) باستخدام 35 دالة غير خطية.

الكلمات المفتاحية: خوارزميات مقيسة ، التدرج المرافق، الامثلية الغير مقيدة.

1. Introduction

We consider the following unconstrained optimization problem:

$$\min f(x), \quad x \in R^n \qquad \ldots(1)$$

where $f : R^n \to R$ is continuously differentiable and the gradient of f at x is denoted by $g(x) = \nabla f(x)$ is available. There are several kinds of numerical methods for solving equation (1), which include the Steepest Descent (SD) method; Newton method; CG and Quasi-Newton (QN) methods. Due to its simplicity and its very low memory requirement, CG-method plays a very important role, especially when the scale is large; the CG-method is very efficient. Let $x_0 \in R^n$ be the initial guess of the solution of problem (1). A nonlinear CG-method is usually designed by the iterative form:[1]

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

where x_k is the current iterate point, $\alpha_k > 0$ is a step length which is determined by some line search, and d_k is the search direction defined by:

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

where g_{k+1} denotes $g(x_{k+1})$ and β_k is a parameter ($0 < \beta_k < 1$). There are some well-known formulas for β_k which are given as follows[2]:

$\beta_k^{FR} = \dfrac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$	(Fletcher-Reeves (FR), 1964)
$\beta_k^{HS} = \dfrac{g_{k+1}^T y_k}{d_k^T y_k}$	(Hestenes -Stiefel (HS), 1952)
$\beta_k^{PR} = \dfrac{g_{k+1}^T y_k}{g_k^T g_k}$	(Polak- Ribiere (PR), 1969)
$\beta_k^{CD} = \dfrac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}$	(Fletcher (CD), 1987)
$\beta_k^{DY} = \dfrac{g_{k+1}^T g_{k+1}}{d_k^T y_k}$	(Dai-Yuan (DY), 1999)
$\beta_k^{BA1} = \dfrac{y_k^T y_k}{-d_k^T g_k}$	(Al−Bayati & Al−Assady (BA1), 1986)
$\beta_k^{BA2} = \dfrac{y_k^T y_k}{g_k^T g_k}$	(Al−Bayati & Al−Assady (BA2), 1986)
$\beta_k^{BA3} = \dfrac{y_k^T y_k}{d_k^T y_k}$	(Al−Bayati & Al−Assady (BA3), 1986)

where $y_k = g_{k+1} - g_k$ and $\|.\|$ stands for the Euclidean norm of vectors. Al-Bayati and Al-Assady [3] investigated three classical CG-methods such that in numerator $y_k^T y_k$ and three different well-known choices for denominator as follows: $(-d_k^T g_k,\ g_k^T g_k,\ d_k^T y_k)$ respectively. In this paper, we have proposed three scaled CG-methods which are based on Al-Bayati and Al-Assady 's CG-methods. Generally, in the convergence analysis of CG-methods, one hopes the ILS, such as the Strong Wolfe Conditions (SWC), which is showed as follows[16]:

- The Strong Wolfe line search is to findsuch that: α_k

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

$$\left| d_k^T g(x_k + \alpha_k d_k) \right| \le -\sigma d_k^T g_k \qquad \ldots(4)$$

$$0 \le \delta \le \frac{1}{2}, \quad \text{and} \quad \delta \le \sigma \le 1$$

This paper organized as follows: In the next section, New formulas for β_k with outline of our three new scaled CG-algorithms are presented. In Section 3, we have analyzed the global convergence properties for uniformly convex and general functions for the proposed new CG-methods. In Section 4, we have reported some numerical comparisons against FRCG and PRCG-methods by using 35-test problems in the CUTE [7] and general conclusions are given in Section 5.

2. New formulas for β_k

In this section, we have constructed three New Scaled CG-Methods with the search direction d_{k+1} as in (3) but β_k is derived by ideology treatment of Classical Al-Bayati and Al-Assady (BA(1,2,3)) 's CG-methods respectively as showed in introduction and we have explained the derivation only for New1 and others are completed in same way as follows, we start from β_k formula at which

$$\beta_k^{BA1} = \frac{y_k^T y_k}{-d_k^T g_k}$$

We notice weaken this method in numerator then action some algebraic operation and positing $y_k = g_{k+1} - g_k$,we get

$$\beta_k^{BA1} = \frac{y_k^T (g_{k+1} - g_k)}{-d_k^T g_k}$$

$$\beta_k^{BA1} = \frac{y_k^T g_{k+1} - y_k^T g_k}{-d_k^T g_k}$$

Again we set $y_k = g_{k+1} - g_k$, getting

$$\beta_k^{BA1} = \frac{g_{k+1}^T g_{k+1} + g_k^T g_k - 2g_{k+1}^T g_k}{-d_k^T g_k}$$

Now, we suggest distribution parameters (u,v,w) on terms existing in numerator and denominator. That is to obtaining on balance in terms after that in form β_k .

$$\beta_k^{new1} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}^T g_k}{-wd_k{}^T g_k} \ ; \ 0 < u, v, w \le 1 \qquad \ldots(5)$$

In same manner we can constricted New2 and New3 for the parameter β_k thus:

$$\beta_k^{new2} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}^T g_k}{w\|g_k\|^2} \ ; \ 0 < u, v, w \le 1 \qquad \ldots(6)$$

$$\beta_k^{new3} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}^T g_k}{wd_k{}^T y_k} \ ; \ 0 < u, v, w \le 1 \qquad \ldots(7)$$

2.1. Outline of the Three New Scaled CG-Algorithms

Step1: (Initializing). Given an initial point $x_0 \in R^n$ and positive parameters, $0 < u, v, w \le 1$, $\psi = 0.2$, $0 \le \delta \le 0.5$ and $\delta \le \sigma \le 1$. Set the initial search direction $d_0 = -g_0$ and Let k = 0 .

Step2: (Termination Criterion). If $\|g_k\| \le \varepsilon$**,** then stop.

Step3: **(Line search).** Determine step length $\alpha_k > 0$ satisfying the Strong Wolfe Condition (4) with **Acceleration scheme[5]:** compute

$z = x_k + \alpha_k d_k$, $y_k = g_k - g_z, . \, g_z = \nabla f(z)$

And Compute $a_k = \alpha_k g_k^T d_k$, $b_k = -\alpha_k y_k^T d_k$, if $b_k \ne 0$, then compute:

$\varphi_k = -\dfrac{a_k}{b_k}$ and update the variables as $x_{k+1} = x_k + \varphi_k \alpha_k d_k$; otherwise update the variables as $x_{k+1} = x_k + \alpha_k d_k$.

Step4: **(Finding the direction).** Compute the new search direction $d_{k+1} = -g_{k+1} + \beta_k^{new} d_k$, where the scalar parameters β_k^{new} are known in (5), (6) and (7).

Step5: **(Restart procedure).** If $\left|g_{k+1}^T g_k\right| \geq \psi \|g_{k+1}\|^2$, then go to **Step (1)** else continue.(this is **Powell restart**).[14]

Step6: (Loop). Let k = k + 1 and go to **Step (2)**.

3. Convergence Analysis.

Now, we have to prove the global convergence property of these three new CG-algorithms under the condition that the following assumption is hold.

Assumption (H**)**

(i) The level set $S = \{x : x \in R^n, f(x) \leq f(x_0)\}$ is bounded, where x_0 is the starting point.

(ii) In a neighborhood Ω of S, f is continuously differentiable and its gradient g is Lipschitz continuously, namely, there exists a constant $L \geq 0$ such that

$$\| g(x) - g(x_k) \| \leq L \| x - x_k \|, \forall x, x_k \in \Omega \qquad \ldots(8)$$

Obviously, from the Assumption (H, i) there exists a positive constant D such that:

$$D = \max\{\|x - x_k\|, \forall x, x_k \in S\} \qquad \ldots(9)$$

where D is the diameter of Ω. From Assumption (H, ii), we also know that there exists a constant $\Gamma \geq 0$, such that:

$$\|g(x)\| \leq \gamma, \forall x \in S \qquad \ldots(10)$$

On some studies of the CG-methods, the sufficient descent or descent condition plays an important role, but unfortunately some times, this condition is hard to hold.[16]

3.1. Theorem Suppose that Assumption (H) holds and satisfies the SWC (4). consider any CG-method (2)-(3) with scalar parameter β_k is defined in (5)-(7) respectively are satisfies the **Sufficient Descent** condition with:

$$c_1 = \{(1 - \frac{2v\sigma\psi + u\sigma + u\psi}{w})\}$$

$$c_2 = \{(1 - \frac{c(u\sigma + u\psi + 2v\sigma\psi)}{w})\}$$

$$c_3 = \{(1 - \frac{u\sigma - u\psi + 2v\psi\sigma}{w\sigma})\}$$

$$0 < u, v, \sigma, \psi < w \leq 1$$

Proof

Case (1) we have to prove that the CG-method d_{k+1} from (3) with (5) and multiplying by g_{k+1}, also put value of β_k^{new1}, we get:

$$d_{k+1}{}^{T} g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}^T g_k}{-wd_k{}^T g_k}\right] d_k^T g_{k+1} \qquad \ldots(11)$$

We obtain from (4)

$$\sigma d_k{}^T g_k \le d_k{}^T g_{k+1} \le -\sigma d_k{}^T g_k \qquad \ldots(12)$$

Since the Powell restarting criterion is defined as follows:

$$\left|g_{k+1}^T g_k\right| \ge \psi \|g_{k+1}\|^2 \qquad \ldots(13)$$

Then we get:

$$g_{k+1}^T g_k \le -\psi \|g_{k+1}\|^2 \qquad \ldots(14)$$

Using (12) and (14) in the inequality (11) become:

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 - \frac{u\|g_{k+1}\|^2}{wd_k{}^T g_k} d_k{}^T g_{k+1} - \frac{u\|g_k\|^2}{wd_k{}^T g_k} d_k{}^T g_{k+1} + \frac{2vg_{k+1}{}^T g_k}{wd_k{}^T g_k} d_k{}^T g_{k+1}$$

$$= -\|g_{k+1}\|^2 - \frac{ud_k{}^T g_{k+1}}{wd_k{}^T g_k}\|g_{k+1}\|^2 - \frac{ud_k{}^T g_k}{wd_k{}^T g_k} g_{k+1}{}^T g_k + \frac{2vd_k{}^T g_{k+1}}{wd_k{}^T g_k} g_{k+1}{}^T g_k$$

$$\le -\|g_{k+1}\|^2 + \frac{u\sigma d_k{}^T g_k}{wd_k{}^T g_k}\|g_{k+1}\|^2 + \frac{u\psi}{w}\|g_{k+1}\|^2 + \frac{2v\sigma\psi\, d_k{}^T g_k}{wd_k{}^T g_k}\|g_{k+1}\|^2$$

$$\le (-1 + \frac{2v\sigma\psi}{w} + \frac{u\sigma}{w} + \frac{u\psi}{w})\|g_{k+1}\|^2$$

Dividing by $\|g_{k+1}\|^2$, we get:

$$\frac{d_{k+1}{}^T g_{k+1}}{\|g_{k+1}\|^2} \le -\{(1 - \frac{2v\sigma\psi + u\sigma + u\psi}{w})\} = -c_1$$

Hence the sufficient descent condition hold, i.e.

$$d_{k+1}{}^T g_{k+1} \le -c_1\|g_{k+1}\|^2,\ c_1 \ge 0 \qquad \ldots(15)$$

$$0 < u, v, \sigma, \psi < w \le 1 \ .$$

Case (2) take d_{k+1} from (3) with (6) proceed by induction. For $k = 1$ we have:

$$d_1 = -g_1$$

and

$$d_1^T g_1 = -g_1{}^T g_1 = -\|g_1\|^2 \le 0$$

suppose that:

$$d_k^T g_k \le -c\|g_k\|^2 \qquad \ldots(16)$$

Multiplying the new search direction by g_{k+1} and put value of β_k^{new2}, we get:

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

$$= -\|g_{k+1}\|^2 + \frac{u\|g_{k+1}\|^2}{w\|g_k\|^2} d_k{}^T g_{k+1} + \frac{u\|g_k\|^2}{w\|g_k\|^2} d_k{}^T g_{k+1} - \frac{2vg_{k+1}{}^T g_k}{w\|g_k\|^2} d_k{}^T g_{k+1}$$

$$= -\|g_{k+1}\|^2 + \frac{ud_k{}^T g_{k+1}}{w\|g_k\|^2}\|g_{k+1}\|^2 + \frac{ud_k{}^T g_k}{w\|g_k\|^2} g_{k+1}{}^T g_k - \frac{2vd_k{}^T g_{k+1}}{w\|g_k\|^2} g_{k+1}{}^T g_k$$

Using strong Wolfe condition (4) and the Powell restarting condition (14) in the above inequality we obtain:

$$d_{k+1}^{\ \ T} g_{k+1} \le -\|g_{k+1}\|^2 - \frac{u\sigma d_k^{\ T} g_k}{w\|g_k\|^2}\|g_{k+1}\|^2 - \frac{u\psi d_k^{\ T} g_k}{w\|g_k\|^2}\|g_{k+1}\|^2 - \frac{2v\sigma\psi d_k^{\ T} g_k}{w\|g_k\|^2}\|g_{k+1}\|^2$$

$$\le -\|g_{k+1}\|^2 - \left(\frac{u\sigma + u\psi + 2v\sigma\psi}{w}\right)\frac{\|g_{k+1}\|^2}{\|g_k\|^2} d_k^{\ T} g_k$$

$$\frac{d_{k+1}^{\ \ T} g_{k+1}}{\|g_{k+1}\|^2} \le -(1 - \frac{c(u\sigma + u\psi + 2v\sigma\psi)}{w}) = -c_2$$

Hence the sufficient descent condition hold, i.e.

$$d_{k+1}^{\ \ T} g_{k+1} \le -c_2\|g_{k+1}\|^2,\ c_2 \ge 0 \qquad \ldots(17)$$

$0 < u, v, \sigma, \psi < w \le 1$.

Case (3) Also, take d_{k+1} from (3) with (7) and multiplying by g_{k+1} with the value of β_k^{new3} to get:

$$d_{k+1}^{\ \ T} g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2v g_k^T g_k}{w d_k^{\ T} y_k}\right] d_k^{\ T} g_{k+1} \qquad \ldots(18)$$

But:

$$d_k^{\ T} y_k = d_k^{\ T} g_{k+1} - d_k^{\ T} g_k \ge d_k^{\ T} g_{k+1}$$

$$\Rightarrow d_k^{\ T} y_k \ge d_k^{\ T} g_{k+1}$$

$$\Rightarrow \frac{1}{w d_k^{\ T} y_k} \le \frac{1}{w d_k^{\ T} g_{k+1}} \qquad \ldots(19)$$

Putting (19) in (18) yields:

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + \left[\frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2v g_{k+1}^T g_k}{w d_k^{\ T} g_{k+1}}\right] d_k^{\ T} g_{k+1}$$

$$\le -\|g_{k+1}\|^2 + \frac{u\|g_{k+1}\|^2}{w d_k^{\ T} g_{k+1}} d_k^{\ T} g_{k+1} + \frac{u\|g_k\|^2}{w d_k^{\ T} g_{k+1}} d_k^{\ T} g_{k+1} - \frac{2v g_{k+1}^{\ \ T} g_k}{w d_k^{\ T} g_{k+1}} d_k^{\ T} g_{k+1}$$

$$\le -\|g_{k+1}\|^2 + \frac{u d_k^{\ T} g_{k+1}}{w d_k^{\ T} g_{k+1}}\|g_{k+1}\|^2 + \frac{u d_k^{\ T} g_k}{w d_k^{\ T} g_{k+1}} g_{k+1}^{\ \ T} g_k - \frac{2v d_k^{\ T} g_{k+1}}{w d_k^{\ T} g_{k+1}} g_{k+1}^T g_k$$

Using (4) and (14) :

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + \frac{u}{w}\|g_{k+1}\|^2 + \frac{u d_k^{\ T} g_{k+1}}{w\sigma d_k^{\ T} g_{k+1}}(-\psi\|g_{k+1}\|^2) - \frac{2v d_k^{\ T} g_{k+1}}{w d_k^{\ T} g_{k+1}}(-\psi\|g_{k+1}\|^2)$$

$$\le -\|g_{k+1}\|^2 + \frac{u}{w}\|g_{k+1}\|^2 - \frac{u\psi}{w\sigma}\|g_{k+1}\|^2 + \frac{2v\psi}{w}\|g_{k+1}\|^2$$

Dividing this inequality by $\|g_{k+1}\|^2$ yields:

$$\frac{d_{k+1}^{\ \ T} g_{k+1}}{\|g_{k+1}\|^2} \le -(1 - \frac{u\sigma - u\psi + 2v\psi\sigma}{w\sigma}) = -c_3 \qquad \ldots(20)$$

Hence the sufficient descent hold i.e.

$$d_{k+1}{}^T g_{k+1} \le -c_3 \|g_{k+1}\|^2, \quad c_3 \ge 0 \qquad \ldots(21)$$

$0 < \{u, v,\ w,\ \sigma,\ \psi\} \le 1$.

3.2. Property Consider a general CG-method and suppose that[12]:

$$0 < \varsigma \le \|g_k\| \le \gamma, \quad \forall k \ge 0 \qquad \ldots(22)$$

we say that a CG-method has the Property (3.2), if there exists two constants b> 1 and $\lambda > 0$ such that for all k,

$$\left|\beta_k^{new}\right| \le b \qquad \ldots(23)$$

If $\|s_k\| \le \lambda$ then $\left|\beta_k^{new}\right| \le \dfrac{1}{2b}$ for all $\lambda > 0$ …(24)

3.3. Lemma Suppose that Assumption (H) hold. If there exists a constant $\varsigma > 0$ such that $\|g_k\| \ge \varsigma$, for all positive k, then the following holds . If d_k is satisfies the sufficient descent condition ($g_k^T d_k \le -c\|g_k\|^2, \forall k \ge 0$) and α_k is obtained by (4). The parameters ($\beta_k^{new1}, \beta_k^{new2}, \beta_k^{new3}$) in our CG-methods satisfy Property (3.2).

Proof

First we can prove this property for the first algorithm with the parameter(5):

$$\beta_k^{new1} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2v g_{k+1}^T g_k}{-w d_k{}^T g_k} \qquad 0 < u, v, w \le 1$$

$$\left|\beta_k^{new1}\right| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{w\left|-d_k{}^T g_k\right|} \qquad \ldots(25)$$

From (15) we have**:**

$$\frac{1}{d_k^T g_k} \ge \frac{1}{-c\|g_k\|^2}$$

$$\Rightarrow \frac{1}{-d_k^T g_k} \le \frac{1}{c\|g_k\|^2}$$

$$\Rightarrow \frac{1}{\left|-d_k^T g_k\right|} \le \frac{1}{c\|g_k\|^2} \qquad \ldots(26)$$

After putting (26) in (25) we get:

$$\left|\beta_k^{new1}\right| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{wc\|g_k\|^2}$$

$$\le \frac{u(\gamma^2 + \gamma^2) + 2v\gamma^2}{wc\gamma^2}$$

$$= \frac{2(u+v)}{wc} = b_1 = \frac{2u\gamma^2 + 2v\gamma^2}{wc\gamma^2}$$

Now, let us define:

$$\lambda_1 = \frac{8(u+v)^2 \alpha\gamma}{w^2 c} \quad \text{and} \quad \|s_k\| \le \lambda_1,\ \lambda_1 \ge 0 \qquad \ldots(27)$$

And from this relation, we have:

$$d_k^T g_k = \frac{1}{\alpha} s_k^T g_k \qquad \ldots(28)$$

By using (25) and (28) with the value of λ , we get:

$$|\beta_k^{new1}| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha}\|s_k\|\|g_k\|}$$

$$\le \frac{2(u+v)\alpha\gamma^2}{w\gamma\lambda_1} = \frac{2(u+v)\alpha\gamma}{w\lambda_1} = \frac{1}{2b_1}$$

Hence

$$|\beta_k^{new1}| \le \frac{1}{2b_1} \quad \text{when} \quad \|s_k\| \le \lambda_1 \ .$$

Second similarity, as in the first proof, we will deal with the new second algorithm as defined in (6):

$$\beta_k^{new2} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2v g_{k+1}^T g_k}{w\|g_k\|^2} \qquad 0 < u, v, w \le 1$$

$$|\beta_k^{new2}| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{w\|g_k\|^2}$$

$$\le \frac{u\gamma^2 + u\gamma^2 + 2v\gamma^2}{w\gamma^2} = \frac{2(u+v)}{w} = b_2$$

Now, again let us define:

$$\lambda_2 = \frac{8(u+v)^2 \alpha c \gamma^2}{w^2} \quad \text{and} \quad \|s_k\| \le \lambda_2, (\lambda_2 > 0) \qquad \ldots(29)$$

Now from the descent property (20) we have:

$$d_k^T g_k \le -c\|g_k\|^2$$

$$\Rightarrow |d_k^T g_k| \le c\|g_k\|^2$$

$$\Rightarrow \frac{1}{c\|g_k\|^2} \le \frac{1}{|d_k^T g_k|} \le \frac{1}{\|d_k\|\|g_k\|}$$

$$|\beta_k^{new2}| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha c}\|s_k\|\|g_k\|}$$

$$\le \frac{2\alpha c(u+v)\gamma^2}{w\lambda_2\gamma} = \frac{2\alpha c(u+v)\gamma}{w\lambda_2} = \frac{1}{2b_2}$$

Hence:

$$|\beta_k^{new1}| \le \frac{1}{2b_2} \quad ; \quad \text{when} \quad \|s_k\| \le \lambda_2 .$$

Third let us again try as in the last proof with the third algorithm (7) as:

$$\beta_k^{new3} = \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_k^T g_k}{wd_k^{\ T} y_k} \qquad \ldots(30)$$

by utilize from (4) to get:

$$\sigma d_k^{\ T} g_k \le d_k^{\ T} g_{k+1} \le -\sigma d_k^{\ T} g$$

$$\Rightarrow d_k^{\ T} g_{k+1} - d_k^{\ T} g_k \ge \sigma d_k^{\ T} g_k - d_k^{\ T} g_k$$

$$\Rightarrow d_k^{\ T} y_k \ge -(1-\sigma) d_k^{\ T} g_k \qquad \ldots(31)$$

By adding to $-d_k^T g_k \ge c\|g_k\|^2$ then (30) becomes:

$$\Rightarrow d_k^{\ T} y_k \ge c(1-\sigma)\|g_k\|^2 \qquad \ldots(32)$$

Taking the absolute values of (30) and since $d_k^T y_k = \frac{1}{\alpha} s_k^T y_k$ then:

$$\left|\beta_k^{new3}\right| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha}\left|s_k^{\ T} y_k\right|}$$

Using inequality (32) the above inequality yields:

$$\left|\beta_k^{new3}\right| \le \frac{u(\|g_{k+1}^T\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{wc(1-\sigma)\|g_k\|^2}$$

$$\le \frac{2(u+v)\gamma^2}{w(1-\sigma)\gamma^2} = \frac{2(u+v)}{w(1-\sigma)} = b_3$$

Now, also let us define:

$$\lambda_3 = \frac{8\alpha(u+v)^2\gamma}{w^2(1-\sigma)^2} \text{ and} \|s_k\| \le \lambda_3 \text{ , } (\lambda_3 > 0) \qquad \ldots(33)$$

$$\left|\beta_k^{new3}\right| \le \frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) + 2v\|g_{k+1}\|\|g_k\|}{\left|\frac{w}{\alpha}(-(1-\sigma)\right|\|s_k\|\|g_k\|)}$$

$$\le \frac{2\alpha(u+v)\gamma^2}{w(1-\sigma)\lambda_3\gamma} = \frac{2\alpha(u+v)\gamma}{w(1-\sigma)\lambda_3} = \frac{1`}{2b_3}$$

Hence

$$\left|\beta_k^{new3}\right| \le \frac{1}{2b_3} \quad ; \text{ when } \|s_k\| \le \lambda_3$$

3.4. Lemma Assume that d_{k+1} is a descent direction and g_k satisfies the Lipschitz condition $\|g(x) - g(x_k)\| \le L\|x - x_k\|$ for all x on the line segment connecting x and x_k, where L is constant If the line search direction satisfy (4), then[6]:

$$\alpha_k \ge \frac{(1-\sigma)\left|d_k^{\ T} g_k\right|}{L\|d_k\|^2} \qquad \ldots(34)$$

Proof Using curvature inequality in (4)

$$\sigma d_k^{\ T} g_k \le d_k^{\ T} g_{k+1} \le -\sigma d_k^{\ T} g_k$$

$\Rightarrow \sigma d_k^T g_k \le d_k^T g_{k+1}$ …(35)

Subtracting $d_k^T g_k$ from both sides of (35) and using Lipschitz condition yields: $(1-\sigma) d_k^T g_k \le d_k^T (g_{k+1} - g_k) \le L\alpha_k \|d_k\|^2$ …(36)

Since d_k is descent direction and $\sigma \le 1$, then (34) holds:

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

The conclusion of the following Lemma, often called the Zoutendijk condition is used to prove the global convergence of any nonlinear CG-method. It was originally given by Zoutendijk [18] under the Strong Wolfe line search (4). In following Lemma, we will prove this condition.

3.5. Lemma Suppose Assumption (H) holds. Consider the iteration process of the form (2)-(3), where d_{k+1} satisfies the descent condition ($d_k^T g_k \le 0$) for all $k \ge 1$ and α_k satisfies (4). Then

$$\sum_{k\ge1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \qquad \text{…(37)}$$

Proof From the first inequality in (4) we can get:

$$f_{k+1} - f_k \le \delta \alpha_k g_k^T d_k$$

Combining this with the results in Lemma **(3.4)**, yields

$$f_{k+1} - f_k \le \frac{\delta(1-\sigma)}{L} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \qquad \text{…(38)}$$

Using the bound-ness of function f in Assumption (H), hence

$$\sum_{k\ge1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \qquad \text{…(39)}$$

3.6. Global Convergence Property For Uniformly Convex Functions

Under Assumption (H) on f , there exists a constant $\gamma \ge 0$, such that $\|\nabla f(x)\| \le \gamma$, for all $x \in S$, then for any CG-method with Strong Wolfe line search, the following general result holds.[8]

3.6.1. Theorem Let Assumption (H) holds and consider any CG-method (2)-(3), where d_{k+1} is a descent direction and α_k is obtained by (4) line search , if

$$\sum_{k\ge1} \frac{1}{\|d_k\|^2} = \infty \qquad \text{…(40)}$$

Then

$$\liminf_{k\to\infty} \|g_k\| = 0 \qquad \text{…(41)}$$

For uniformly convex function which satisfied the above assumptions we can prove that the norm of d_{k+1} given by (3) with (5), (6) and (7) is bounded above. Assume that the function f is uniformly convex function, i.e. there exists a constant $\mu \geq 0$ such that for all $x,\ x_k \in S$

$$(g(x)-g(x_k))^T(x-x_k) \geq \mu\|x-x_k\|^2 \qquad \ldots(42)$$

and the step-length α_k is obtained by the Strong Wolfe line search (4), Now try to prove the following result:

3.6.2. Theorem Suppose that Assumption (H) hold. Consider the algorithm (2.1), where $0<u,v,w\leq 1$, for $\gamma>0$, let $\|g_k\|\leq\gamma$, $\|g_{k+1}\|\leq\gamma$ and α_k is obtained by (SWC) line search. If $\|s_k\|$ tends to zero and there exists non-negative constants η_1 and η_2 such that[6]:

$$\|g_k\|^2 \geq \eta_1\|s_k\|^2 \ ; \ \|g_{k+1}\|^2 \leq \eta_2\|s_k\| \qquad \ldots(43)$$

and if f is a uniformly convex function, then:

$$\lim_{k\to\infty} g_k = 0 \qquad \ldots(44)$$

Proof

Case(1) we have from (5) and if we taking the absolute value:

$$\left|\beta_{k+1}^{new1}\right| \leq \frac{u(\|g_{k+1}\|^2+\|g_k\|^2)+2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha_k}\|s_k\|\|g_k\|}$$

$$\leq \frac{u\eta_2\|s_k\|+u\eta_1\|s_k\|^2+2v\gamma}{\frac{w}{\alpha_k}\xi\|s_k\|}$$

But $\|s_k\|=\|x-x_k\|$ and since $D=\max\{\|x-x_k\|,\forall x,x_k\in S\}$ is diameter of the level set S (9), then

$$\left|\beta_k^{new1}\right| \leq \frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{\frac{w}{\alpha_k}\xi D}$$

Taking the norm and square both sides of the new search direction, we get:

$$\|d_{k+1}\|^2 = \|-g_{k+1}+\beta_k^{new1}d_k\|^2$$

$$\leq \|g_{k+1}\|^2+2\left|\beta_k^{new1}\right|\|g_{k+1}\|\|d_k\|+(\beta_k^{new1})^2\|d_k\|^2$$

$$\leq \eta_2\|s_k\|+2\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{\frac{w}{\alpha_k}\xi D}\right)\frac{1}{\alpha_k}\gamma\|s_k\|+\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{\frac{w}{\alpha_k}\xi D}\right)^2\frac{1}{\alpha_k^2}\|s_k\|^2$$

$$\leq \eta_2\|s_k\|+2\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{\frac{w}{\alpha_k}\xi D}\right)\frac{1}{\alpha_k}\gamma\|s_k\|+\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{\frac{w}{\alpha_k}\xi D}\right)^2\frac{1}{\alpha_k^2}\|s_k\|^2$$

$$\leq \eta_2\|s_k\|+2\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{w\xi D}\right)\gamma\|s_k\|+\left(\frac{u(\eta_2+\eta_1 D)D+2v\gamma^2}{w\xi D}\right)^2\|s_k\|^2$$

$$\le \eta_2 D + 2(\frac{u(\eta_2+\eta_1 D)D + 2v\gamma^2}{w\xi})\gamma + (\frac{u(\eta_2+\eta_1 D)D + 2v\gamma^2}{w\xi})^2 = \psi_1^{\ 2}$$

Thus $\|d_{k+1}\|^2 \le \psi_1^{\ 2}$

$$\Rightarrow \frac{1}{\|d_{k+1}\|^2} \ge \frac{1}{\psi_1^{\ 2}}$$

Summation this inequality for all $k \ge 1$

$$\sum_{k\ge1}\frac{1}{\|d_{k+1}\|^2} \ge \sum_{k\ge1}\frac{1}{\psi_1^{\ 2}} = \infty$$

Using Theorem (3.2.1), hence $\lim_{k\to\infty}\inf\|g_k\| = 0$

But f is uniformly convex (3.37) therefore satisfies $\lim_{k\to\infty} g_k = 0$.

Case (2) similarly we begin taking the absolute to the second scalar parameter β_k^{new2}

$$\left|\beta_k^{new2}\right| \le \frac{u\|g_{k+1}\|^2 + u\|g_k\|^2 + 2v\|g_{k+1}\|\|g_k\|}{w\|g_k\|^2}$$

$$\le \frac{u\eta_2\|s_k\| + u\eta_1\|s_k\|^2 + 2v\gamma\gamma}{w\eta_1\|s_k\|^2}$$

$$\le \frac{u(\eta_2+\eta_1 D)D + 2v\gamma^2}{w\eta_1 D^2}$$

where D is a diameter of the level set S and the new direction can be evaluated from (3) with (6). Now since:

$$d_{k+1} = -g_{k+1} + \beta_k^{new2} d_k$$

$$\|d_{k+1}\|^2 = \left\|-g_{k+1} + \beta_k^{new2} d_k\right\|^2$$

$$\le \|g_{k+1}\|^2 + 2\left|\beta_k^{new2}\right|\|g_{k+1}\|\|d_k\| + (\beta_k^{new2})^2\|d_k\|^2$$

$$\le \eta_2\|s_k\| + 2(\frac{u(\eta_2+\eta_1\|s_k\|)\|s_k\| + 2v\gamma^2}{w\eta_1\|s_k\|^2})\frac{\gamma\|s_k\|}{\alpha_k} + (\frac{u(\eta_2+\eta_1\|s_k\|)\|s_k\| + 2v\gamma^2}{w\eta_1\|s_k\|^2})^2\frac{\|s_k\|^2}{\alpha_k^2}$$

$$\le \eta_2 D + 2(\frac{u(\eta_2+\eta_1 D)D + 2v\gamma^2}{w\alpha_k\eta_1 D})\gamma + (\frac{u(\eta_2+\eta_1 D)D + 2v\gamma^2}{w\alpha_k\eta_1 D})^2 = \psi_2^2$$

Thus $\|d_{k+1}\|^2 \le \psi_2^2$

$$\Rightarrow \frac{1}{\|d_{k+1}\|^2} \ge \frac{1}{\psi_2^{\ 2}}$$

Summation this inequality for all $k \ge 1$

$$\sum_{k\ge1}\frac{1}{\|d_{k+1}\|} \ge \sum_{k\ge1}\frac{1}{\psi_2^{\ 2}} = \infty$$

Using Theorem (3.7.1), hence $\lim_{k\to\infty}\inf\|g_k\| = 0$

But f is uniformly convex (3.37) therefore $\lim_{k\to\infty} g_k = 0$.

Case (3) from equation (7) we have also:

$$\left|\beta_{k+1}^{new3}\right| \le \frac{u\|g_{k+1}\|^2 + u\|g_k\|^2 + 2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha}\|s_k\|\|y_k\|}$$

But the gradient g is Lipschitz and f is uniformly convex then we get

$$\mu\|s_k\| \le \|y_k\| \le L\|s_k\|$$

Therefore

$$\left|\beta_k^{new3}\right| \le \frac{u\|g_{k+1}\|^2 + u\|g_k\|^2 + 2v\|g_{k+1}\|\|g_k\|}{\frac{w}{\alpha}\mu\|s_k\|^2}$$

$$\le \frac{u(\eta_2 + \eta_1\|s_k\|)\|s_k\| + 2v\gamma\gamma}{\frac{w}{\alpha}\mu D^2} = \frac{\alpha_k u(\eta_2 + \eta_1 D)D + 2v\alpha_k\gamma^2}{w\mu D^2}$$

Then

$$\left|\beta_k^{new3}\right| \le \frac{\alpha_k u(\eta_2 + \eta_1 D)D + 2v\alpha_k\gamma^2}{w\mu D^2}$$

Again we know that

$$d_{k+1} = -g_{k+1} + \beta_k^{new3} d_k$$

$$\|d_{k+1}\|^2 = \left\|-g_{k+1} + \beta_k^{new3} d_k\right\|^2$$

$$\le \|g_{k+1}\|^2 + 2\left|\beta_k^{new3}\right|\|g_{k+1}\|\|d_k\| + (\beta_k^{new3})^2\|d_k\|^2$$

$$\le \eta_2\|s_k\| + 2\left(\frac{\alpha_k u(\eta_2 + \eta_1 D)D + 2v\alpha_k\gamma^2}{w\mu D^2}\right)\frac{\gamma}{\alpha_k}\|s_k\| + \left(\frac{\alpha_k u(\eta_2 + \eta_1 D)D + 2v\alpha_k\gamma^2}{w\mu D^2}\right)^2\frac{1}{\alpha_k^2}\|s_k\|^2$$

$$\le \eta_2 D + 2\left(\frac{u(\eta_2 + \eta_1 D)D + 2v\gamma^2}{w\mu D}\right)\gamma + \left(\frac{u(\eta_2 + \eta_1 D)D + 2v\gamma^2}{w\mu D}\right)^2 = \psi_3^2$$

Thus $\|d_{k+1}\|^2 \le {\psi_3}^2$

$$\Rightarrow \frac{1}{\|d_{k+1}\|^2} \ge \frac{1}{{\psi_3}^2}$$

Summation this inequality for all $k \ge 1$

$$\sum_{k\ge1}\frac{1}{\|d_{k+1}\|^2} \ge \sum_{k\ge1}\frac{1}{{\psi_3}^2} = \infty$$

Using theorem (3.2.1) , hence $\liminf_{k\to\infty}\|g_k\| = 0$

But f is uniformly convex (3.37), therefore $\lim_{k\to\infty} g_k = 0$

3.7. Global Convergence Property For General Nonlinear Functions

For general nonlinear functions, the convergence analysis of our algorithms exploits insights developed by Gilbert and Nocedal [9]; Dai and Liao [8] and Hager and Zhang [10]. The global convergence proof of (New1, New2, New3) CG-algorithms is based on the Zoutendijk condition as showed in Lemma **(3.5)**. combined with the analysis showing that the descent property holds and $\|d_k\|$ is bounded. Suppose that the level set S is bounded and the function f is bounded from below.

3.7.1 Theorem Let Assumption (H) hold, d_{k+1} is descent direction and α_k is obtained by line search satisfies strong wolf line search, $c > 0$, $0 < u, v, w \le 1$ also constants $\gamma > 0$ such that $\|g_k\| \le \gamma$ as in (10) . Then the algorithm (2.1) satisfies, either $g_k = 0$ for some k or(44) s. t. $\lim_{k\to\infty}\inf\|g_k\| = 0$

Proof We will prove this theorem by using contradiction , then assume that the result is not true , So there exists a constant s. t.

$$\|g_k\| \ge \xi \quad ,\forall k \ge 1 \qquad \ldots(45)$$

Case (1) we begin with direction of CG-method contains the parameter β_k^{new1}

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

$$\le \|g_{k+1}\| + \frac{u\|g_{k+1}\|^2 + u\|g_k\|^2 + 2v\|g_{k+1}\|\|g_k\|}{w\|d_k\|\|g_k\|}\|d_k\|$$

$$\le \gamma + \frac{u\gamma^2 + u\gamma^2 + 2v\gamma^2}{w\gamma} = \zeta_1$$

Thus:

$$\|d_{k+1}\| \le \zeta_1 \qquad \ldots(46)$$

Since the level set S is bounded and the function f is bounded below using {Lemma (3.3) and Lemma (3.4)}, we get:

$$0 \le \sum_{k\ge1} \frac{(g_k{}^T d_k)^2}{\|d_k\|^2} < \infty$$

Combining with sufficient descent condition (15) yields:

$$\sum_{k\ge1} \frac{\xi^4}{\|d_k\|^2} \le \sum_{k\ge1} \frac{\|g_k\|^4}{\|d_k\|^2} \le \frac{1}{c_1} \sum_{k\ge1} \frac{(g_k{}^T d_k)^2}{\|d_k\|^2} < +\infty$$

Using the above inequality with (3.46) which contradiction to (40). Hence (41) holds s.t. $\lim_{k\to\infty}\inf\|g_k\| = 0$

Case (2) with direction of CG-method contains the parameter β_k^{new2}

$$\|d_{k+1}\| = \left\| -g_{k+1} + \left(\frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}{}^T g_k}{w\|g_k\|^2}\right) \frac{1}{\alpha_k} s_k \right\|$$

$$\le \|g_{k+1}\| + \frac{u\|g_{k+1}\|^2}{\alpha_k w\|g_k\|^2}\|s_k\| + \frac{u\|g_k\|^2}{\alpha_k w\|g_k\|^2}\|s_k\| + \frac{2v\|g_{k+1}\|\|g_k\|}{\alpha_k w\|g_k\|^2}\|s_k\|$$

$$\leq \gamma + \frac{u\gamma^2}{\alpha_k w\gamma^2}\|s_k\| + \frac{u}{\alpha_k w}\|s_k\| + \frac{2v\gamma\gamma}{\alpha_k w\gamma^2}\|s_k\|$$

$$= \gamma + (\frac{2(u+v)}{\alpha_k w})\|s_k\|$$

$$\leq \gamma + 2D(\frac{u+v}{\alpha_k w}) = \zeta_2$$

Thus:

$$\|d_{k+1}\| \leq \zeta_2 \qquad \ldots(47)$$

Since the level set S is bounded and the function f is bounded below using lemmas (3.2) and (3.3), we get:

$$0 \leq \sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

Combining with sufficient descent condition (17), s. t.

$$-d_k^T g_k \geq c_2\|g_k\|^2 > 0$$

$$\sum_{k\geq 1} \frac{\xi^4}{\|d_k\|^2} \leq \sum_{k\geq 1} \frac{\|g_k\|^4}{\|d_k\|^2} \leq \frac{1}{c_2}\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty$$

Using the above inequality with (47) which contradiction to (40). Hence (3.34) holds s.t. $\liminf_{k\to\infty}\|g_k\| = 0$

Case(3) with direction of CG-method contains the parameter β_k^{new3}

$$\|d_{k+1}\| = \left\| -g_{k+1} + (\frac{u(\|g_{k+1}\|^2 + \|g_k\|^2) - 2vg_{k+1}^T g_k}{\frac{w}{\alpha_k} s_k^T y_k})\frac{1}{\alpha_k} s_k \right\|$$

$$\leq \|g_{k+1}\| + \frac{u\|g_{k+1}\|^2}{w\|s_k\|\|y_k\|}\|s_k\| + \frac{u\|g_k\|^2}{w\|s_k\|\|y_k\|}\|s_k\| + \frac{2v\|g_{k+1}\|\|g_k\|}{w\|s_k\|\|y_k\|}\|s_k\|$$

But: $\|y_k\| = \|g_{k+1} - g_k\| \leq \|g_{k+1}\| + \|g_k\| \leq 2\gamma$ and from equation (45), we obtain

$$\Rightarrow 2\xi \leq \|y_k\| \leq 2\gamma$$

$$\leq \|g_{k+1}\| + \frac{u\|g_{k+1}\|^2}{2w\xi\|s_k\|}\|s_k\| + \frac{u\|g_k\|^2}{2w\xi\|s_k\|}\|s_k\| + \frac{2v\|g_{k+1}\|\|g_k\|}{2w\xi\|s_k\|}\|s_k\|$$

$$\leq \|g_{k+1}\| + \frac{u\|g_{k+1}\|^2}{2w\xi} + \frac{u\|g_k\|^2}{2w\xi} + \frac{2v\|g_{k+1}\|\|g_k\|}{2w\xi}$$

Therefore

$$\|d_{k+1}\| \leq \gamma + \frac{u\gamma^2 + u + 2v\gamma\gamma}{2w\xi} = \zeta_3$$

Thus

$$\|d_{k+1}\| \leq \zeta_3 \qquad \ldots(48)$$

Since the level set S is bounded and the function f is bounded below using lemmas (3.2) and (3.3), we get:

$$0 \le \sum_{k\ge1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

Combining with sufficient descent condition (21), s. t.

$$\sum_{k\ge1} \frac{\xi^4}{\|d_k\|^2} \le \sum_{k\ge1} \frac{\|g_k\|^4}{\|d_k\|^2} \le \frac{1}{c_3} \sum_{k\ge1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty$$

Using the above inequality with (48) which contradiction to (40). Hence (41) holds s.t. $\liminf_{k\to\infty} \|g_k\| = 0$.

4. Numerical Results

The main work of this section is to report the performance of the new methods on a set of test problems. the codes were written in Fortran and in double precision arithmetic. All the tests were performed on a PC. Our experiments were performed on a set of 35-nonlinear unconstrained problems that have second derivatives available. These test problems are contributed in CUTE [7] and their details are given in the Appendix. for each test function we have considered 10 numerical experiments with number of variable n= 100,200,……,1000. In order to assess the reliability of our new proposed methods, we have tested them against FR and PR classical CG-methods using the same test problems. All these methods terminate when the following stopping criterion is met.

$$\|g_{k+1}\| \le 1\times10^{-5} \qquad \ldots(49)$$

We also force these routines stopped if the iterations exceed 1000 or the number of function evaluations reach 2000 without achieving the minimum. We use $\delta = 10^{-4}$, $\sigma = 0.1$ in the Wolfe line search routine. Tables (4.1); (4.2) and (4.3) compare some numerical result for (New1, New2 and New3) CG-methods against FR and PR CG-methods respectively, these tables indicate for (n) as a dimension of the problem;(NOI) number of iterations; (NOFG) number of function and gradient evaluation;(Time) the total time required to complete the evaluation process for each test problem. In Tables (4.4, 4.5, 4.6) we have compared the percentage performance of the new and FR&PR methods taking over all the tools as 100% . In order to summarize our numerical results , we have concerned only on the total of different dimensions n= 100, 200,……,10000, for all tools used in these comparisons.

Table 4.1 Comparison between new1 and classical FR and PR CG-methods for the total of n different dimensions n= 100, 200, ... ,1000 for each test problems with parameter (u =0.3, v=0.3, w=0.8; $\varepsilon=1*10^{-5}$) .

Table 4.2 Comparison between new2 and classical FR and PR CG-methods for the total of n different dimensions n= 100, 200, ... ,1000 for each test problems with parameter (u =0.3, v=0.7, w=0.8 ; $\varepsilon=1*10^{-5}$).

Table 4.3 Comparison between new3 and classical FR and PR CG-methods for the total of n different dimensions n= 100, 200, ... ,1000 for each test problems with parameter (u =0.4, v=0.4, w=0.1; $\varepsilon=1*10^{-5}$).

Table (4.1) Comparison between New1, FR and PR CG-methods for the total of n different dimensions n= 100, 200, ...,1000, for each test problem (u =0.4, v=0.6, w=0.8, ε=1*10^{-5}).

Number of Problem	Classic FR NOI/NOFG/TIME	Classic PR NOI/NOFG/TIME	New 1 NOI/NOFG/TIME
1	366 697 0.24	369 710 0.24	422 706 0.34
2	269 674 0.03	270 676 0.02	188 381 0.04
3	111 337 0.05	111 337 0.03	70 99 0.01
4	687 1318 0.32	703 1342 0.35	481 629 0.26
5	308 612 0.05	311 602 0.04	278 363 0.04
6	255 494 0.18	254 500 0.21	168 192 0.13
7	309 841 0.05	383 996 0.04	333 361 0.03
8	79 291 0.09	79 291 0.09	135 169 0.17
9	215 527 0.00	218 534 0.01	194 329 0.02
10	187 452 0.10	187 452 0.10	172 247 0.09
11	344 694 0.09	346 701 0.08	233 353 0.06
12	185 476 0.03	185 476 0.01	167 287 0.03
13	57 300 0.01	57 300 0.03	27 60 0.00
14	786 1493 0.08	784 1480 0.11	547 642 0.09
15	519 1159 0.08	530 1201 0.10	458 529 0.06
16	138 370 0.05	138 370 0.04	125 145 0.07
17	145 358 0.04	145 358 0.05	87 109 0.03
18	140 369 0.03	140 369 0.02	100 130 0.02
19	322 675 0.03	326 680 0.05	347 432 0.04
20	111 337 0.05	111 337 0.03	69 90 0.03
21	321 699 0.03	320 687 0.03	330 444 0.05
22	328 585 0.10	312 562 0.09	215 255 0.07
23	643 1286 0.21	648 1289 0.22	284 360 0.10
24	148 410 0.04	148 410 0.03	112 174 0.04
25	236 562 0.04	240 568 0.05	230 291 0.02
26	664 1292 0.19	681 1303 0.20	297 393 0.11
27	124 348 0.01	124 348 0.02	118 191 0.02
28	117 359 0.08	117 359 0.02	118 191 0.02
29	102 314 0.08	102 314 0.08	75 95 0.03
30	116 435 0.07	116 435 0.10	41 73 0.03
31	187 452 0.08	187 452 0.08	172 247 0.07
32	573 1123 0.08	562 1097 0.06	420 514 0.05
33	10 30 0.00	10 30 0.02	14 49 0.00
34	80 100 0.02	80 100 0.02	80 100 0.03
35	184 431 0.03	184 431 0.01	218 466 0.03
Total	9266 20900 4.54	9478 20097 4.48	7325 10096 4.03

Table (4.2) Comparison between New2 , FR and PR CG-methods for the total of n different dimensions n= 100, 200, ... ,1000, for each test problem (u =0.3, v=0.4, w=0.5, $\varepsilon=1*10^{-5}$).

Number of Problem	Classic FR NOI/NOFG/TIME			Classic PR NOI/NOFG/TIME			New 2 NOI/NOFG/TIME		
1	366	697	0.24	369	710	0.24	409	710	0.36
2	269	674	0.03	270	676	0.02	188	381	0.03
3	111	337	0.05	111	337	0.03	69	90	0.03
4	687	1318	0.32	703	1342	0.35	505	633	0.26
5	308	612	0.05	311	602	0.04	244	331	0.05
6	255	494	0.18	254	500	0.21	279	305	0.22
7	309	841	0.05	383	996	0.04	416	444	0.05
8	79	291	0.09	79	291	0.09	135	169	0.15
9	215	527	0.00	218	534	0.01	196	329	0.01
10	187	452	0.10	187	452	0.10	163	226	0.08
11	344	694	0.09	346	701	0.08	258	379	0.08
12	185	476	0.03	185	476	0.01	165	258	0.03
13	57	300	0.01	57	300	0.03	26	58	0.00
14	786	1493	0.08	784	1480	0.11	563	661	0.08
15	519	1159	0.08	530	1201	0.10	376	447	0.05
16	138	370	0.05	138	370	0.04	101	122	0.03
17	145	358	0.04	145	358	0.05	86	108	0.01
18	140	369	0.03	140	369	0.02	100	130	0.02
19	322	675	0.03	326	680	0.05	309	423	0.07
20	111	337	0.05	111	337	0.03	69	90	0.03
21	321	699	0.03	320	687	0.03	305	419	0.03
22	328	585	0.10	312	562	0.09	241	282	0.08
23	643	1286	0.21	648	1289	0.22	298	442	0.12
24	148	410	0.04	148	410	0.03	103	165	0.00
25	236	562	0.04	240	568	0.05	219	311	0.04
26	664	1292	0.19	681	1303	0.20	399	445	0.11
27	124	348	0.01	124	348	0.02	111	183	0.01
28	117	359	0.08	117	359	0.02	70	151	0.05
29	102	314	0.08	102	314	0.08	75	95	0.04
30	116	435	0.07	116	435	0.10	41	73	0.04
31	187	452	0.08	187	452	0.08	162	224	0.06
32	573	1123	0.08	562	1097	0.06	407	521	0.05
33	10	30	0.00	10	30	0.02	24	77	0.01
34	80	100	0.02	80	100	0.02	70	100	0.03
35	184	431	0.03	184	431	0.01	169	270	0.02
Total	9266	20900	4.54	9478	20097	4.48	7351	10052	4.13

Table (4.3) Comparison between New3 , FR and PR CG-methods for the total of n different dimensions n= 100, 200, ... ,1000, for each test problem (u =0.4, v=0.6, w=0.8, ε=1*10^{-5}).

Number of Problem	Classic FR NOI/NOFG/TIME	Classic PR NOI/NOFG/TIME	New 3 NOI/NOFG/TIME
1	366 697 0.24	369 710 0.24	417 690 0.38
2	269 674 0.03	270 676 0.02	191 384 0.01
3	111 337 0.05	111 337 0.03	69 90 0.03
4	687 1318 0.32	703 1342 0.35	461 613 0.26
5	308 612 0.05	311 602 0.04	284 374 0.05
6	255 494 0.18	254 500 0.21	141 164 0.09
7	309 841 0.05	383 996 0.04	382 411 0.05
8	79 291 0.09	79 291 0.09	128 162 0.14
9	215 527 0.00	218 534 0.01	190 309 0.04
10	187 452 0.10	187 452 0.10	174 246 0.08
11	344 694 0.09	346 701 0.08	217 329 0.06
12	185 476 0.03	185 476 0.01	158 292 0.00
13	57 300 0.01	57 300 0.03	26 58 0.00
14	786 1493 0.08	784 1480 0.11	554 644 0.08
15	519 1159 0.08	530 1201 0.10	465 536 0.08
16	138 370 0.05	138 370 0.04	125 145 0.03
17	145 358 0.04	145 358 0.05	87 109 0.01
18	140 369 0.03	140 369 0.02	100 130 0.03
19	322 675 0.03	326 680 0.05	331 435 0.06
20	111 337 0.05	111 337 0.03	69 90 0.03
21	321 699 0.03	320 687 0.03	315 395 0.05
22	328 585 0.10	312 562 0.09	219 246 0.08
23	643 1286 0.21	648 1289 0.22	277 356 0.12
24	148 410 0.04	148 410 0.03	112 174 0.00
25	236 562 0.04	240 568 0.05	254 372 0.04
26	664 1292 0.19	681 1303 0.20	303 365 0.11
27	124 348 0.01	124 348 0.02	130 189 0.01
28	117 359 0.08	117 359 0.02	70 151 0.03
29	102 314 0.08	102 314 0.08	75 95 0.05
30	116 435 0.07	116 435 0.10	41 73 0.05
31	187 452 0.08	187 452 0.08	174 246 0.08
32	573 1123 0.08	562 1097 0.06	423 521 0.06
33	10 30 0.00	10 30 0.02	22 70 0.00
34	80 100 0.02	80 100 0.02	70 100 0.03
35	184 431 0.03	184 431 0.01	180 270 0.02
Total	9266 20900 4.54	9478 20097 4.48	7234 9834 4.02

Percentage Performance of each New algorithm against 100% of Fletcher-Reeves (FR), Polak- Ribiere (PR), algorithms respectively, as follows in Tables (4.4), (4.5), (4.6).

Table (4.4) Performance of the New1 algorithm against 100% of Fletcher-Reeves (FR) and Polak- Ribiere (PR) algorithm, as followed in Table (4.1).

Tools	FR	New1		PR	New1
NOI	100%	79.052%		100%	77.284%
NOFG	100%	53.549%		100%	50.236%
Time	100%	57.202%		100%	89.955%

Table (4.5) Performance of the New2 algorithm against 100% of Fletcher-Reeves (FR) and Polak- Ribiere (PR) algorithm, as followed in Table (4.2).

Tools	FR	New2		PR	New2
NOI	100%	79.333%		100%	77.558%
NOFG	100%	48.059%		100%	50.017%
Time	100%	90.969%		100%	92.187%

Table (4.6) Performance of the New3 algorithm against 100% of Fletcher-Reeves (FR) and Polak-Ribiere (PR) algorithm, as followed in Table (4.3).

Tools	FR	New3		PR	New3
NOI	100%	78.070%		100%	76.324%
NOFG	100%	47.052%		100%	48.632%
Time	100%	88.546%		100%	89.732%

From the above tables we have concluded that the first new algorithm beats FR and PR CG-algorithms in all NOI; NOFG and Time in about (10-50)% percentages. However, the second new algorithm is also beats FR and PR CG-algorithms in all NOI; NOFG and Time in about (8-52)% percentages. Also the third new algorithm is also beats FR and PR CG-algorithms in all NOI; NOFG and Time in about (11-53)% percentages.

5. Conclusions

In this paper, by using scaling parameter idea, we have proposed three new scaled CG-methods with (2.1), (2.2) and (2.3) for β_k, under some assumptions. Our CG-methods have been shown to be globally convergent for uniformly convex and general functions respectively. Some numerical results have been reported against BA1; BA2; BA3; FR and PRCG-algorithms which showed the effectiveness of our new proposed CG-algorithms with the scalars u ,v and w.

6. Appendix

The details of the 35-test functions used are: 1-Extended Trigonometric Function. 2-Extended Penalty Function. 3-Raydan2 Function. 4-Hager Function. 5-Generalized Tridiagonal-1 Function. 6-Extended Three Exponential Function. 7-Diagonal 4 Function. 8-Diagonal5 Function. 9-Extended Himmelblau Function. 10-Generalized PSC1 Function. 11- Extended Block Diagonal BD1 Function. 12-Extended Quadratic Penalty QP1 Function. 13-Extended Quadratic QF2 Function. 14- Extended EP1 Function.15-Extended Tri-diagonal 2 Function. 16- DIXMAANA Function. 17- DIXMAANB Function. 18- DIXMAANC Function. 19-EDENSCH Function. 20- DIAGONAL 6 Function. 21-ENGVALI Function. 22-DENSCHNA Function. 23- DENSCHNC Function. 24-DENSCHNB Function. 25-DENSCHNF Function. 26- Extended Block–Diagonal BD2 Function. 27-Generalized quadratic GQ1 Function. 28- DIAGONAL 7 Function. 29- DIAGONAL 8 Function. 30- Full Hessian Function.
31-SINCOS Function. 32- Generalized quadratic GQ2 Function. 33-ARGLINB Function. 34-HIMMELBG Function. 35-HIMMELBH Function.

REFERENCES

[1] Al-Bayati Y., Jassim A. and Abbo K., (2010)," New scaled CG methods based on Newton direction for solving unconstrained optimization problems ". Australian J. of Basic and Applied Sciences, 1190-1202.

[2] Al-Bayati A. Y. and Ban A. Matras, (2008), "A New Five Parameter Family for CG-Methods ", Allepo Conf. 10 (1).

[3] Al-Bayati A.Y. and Al-Assady N. H., (1986)," Conjugate Gradient Algorithms for Unconstrained Optimization ", Technical Report, No. (1/86), School of Computer Studies, Leeds University, U.K.

[4] Al-Baali M., (1985), "Descent Property and Global Convergence of the Fletcher–Reeves Method with Inexact Line Search ", IMA Journal of Numerical Analysis, 5, pp.121-124..

[5] Andrei N. (2009), "Accelerated of conjugate gradient algorithm for unconstrained optimization", Applied Mathematical Computational, 213, pp.361–369.

[6] Andrei N., (2009)," Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization", J. Optim. Theory Appl. 141, pp. 249–264.

[7] Bongartz K. E., Conn A.R., Gould N.I.M. and Toint P.L., CUTE, (1995), "constrained and unconstrained testing environments". ACM Trans. Math. Software, 21, pp. 123-160.

[8] Dai Y.H. and Liao L.Z., (2001), "New conjugacy conditions and related nonlinear conjugate gradient methods", Applied. Mathematic. Optimization. 43, 87–101.

[9] Dai Y. H. and Yuan Y., (1998), "Nonlinear Conjugate Gradient Methods", Shanghai Scientific and Technical Publishers, Shanghai.

[10] Gilbert J.C. and Nocedal J. ,(1992), "Global convergence properties of conjugate gradient methods for optimization". SIAM J. Optimization. 2, pp. 21–42.

[11] Hager W.W. and Zhang H., (2005), "A new conjugate gradient method with guaranteed descent and an efficient line search", SIAM J. Optim. 16, pp.170–192.

[12] Hager W.W. and Zhang H., (2005), "A survey of nonlinear conjugate gradient methods", Department of Mathematics, University of Florida, Gainesville, FL, pp. 32611- 8105.

[13] Powell M. J. D., (1984), "Non-convex Minimization Calculations and the Conjugate Gradient Method", Lecture Notes in Mathematics, 1066, pp.121-141.

[14] Powell M.J.D., (1977), "Restart procedures for the conjugate gradient method", Mathematical Programming, 12, pp. 241–254.

[15] Sun W. and Yuan Y., (1999), "Theory and Methods of Optimization", Science Press of
China, Beijing.

[16] Yabe H. and Takano M., (2004), "Global Convergence Properties of Nonlinear Conjugate Gradient Methods with Modified Secant Condition", Computational Optimization and Applications, 28, 203–225.

[17] Zhang L. and Zhou W. J., (2008), "Two Descent Hybrid Conjugate Gradient Methods for Optimization", Journal of Computational and Applied Mathematics, 216, pp. 251-264.

[18] Zoutendijk G., (1970), "Nonlinear programming, computational methods in integer and nonlinear programming", North Holland Amsterdam, pp. 37-86.