

Two New Extended PR Conjugate Gradient Methods for Solving Nonlinear Minimization Problems

*Abbas Y. AL-Bayati

**Rana Z. AL-Kawaz

*College of Basic Education Telafer, Mosul University, Iraq.

**College of Computers Sciences and Mathematics, Mosul
University, Iraq

تاريخ القبول

تاريخ الاستلام 2012/03/26

2012/08/13

المستخلص

في هذا البحث تطرقنا إلى تقصي و اشتقاق نظري لنموذجين أكثر عمومية من النموذج التربيعي لطريقة PR الموسعة والتي تستخدم قيم الدالة والمشتقة. الطريقتان الجديدتان تشملان حقل طرائق التدرج المترافق الموسع وتمتلكان خاصيتي الانحدار الحاد و التقارب الشامل تحت شروط معينة. تم الحصول على نتائج عددية متميزة مقارنة مع طريقة Wu and Chen PRCG- (2010) المتماثلة في نفس المجال.

Abstract

In this paper, we have discussed and investigated two nonlinear extended PR-CG method which use function and gradient values. The two new methods involve the standard CG-methods and have the sufficient descent and globally convergence properties under certain conditions. We have got some important numerical results by comparing the new method with Wu and Chen PRCG-(2010) method in this field.

Key Words: *Extended Conjugate Gradient Method, Minimization Problems, Non-Quadratic Models, Sufficient Descent, Global Convergence.*

Introduction

This paper considers the calculation of a local minimizer x^* say, for the problem:

$$\text{Min } f(x); \text{ where } f: \mathbb{R}^n \rightarrow \mathbb{R} \quad \dots\dots(1)$$

is a smooth nonlinear function (of n variables) and its gradient vector $g_k = \nabla f(x_k)$ is available are calculated but the Hessian matrix is not available. At the current iterative point x_k , the Conjugate Gradient (CG) method has the following form:

$$x_{k+1} = x_k + \alpha_k d_k \quad \dots\dots(2a)$$

$$d_k = \begin{cases} -g_k, & k = 0 \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \end{cases} \quad \dots\dots(2b)$$

where α_k is a step-length; d_k is a search direction; β_k is a parameter. Standard algorithms for solving this problem include CG-algorithms which are iterative algorithms and generate a sequence of approximations to minimize a function $F(x)$ and their very low memory requirements. However, this paper considers a more general model than the usual quadratic function

$$F(x) = \frac{1}{2} x^T G x + b^T x + c \quad \dots\dots(3)$$

Some well-known formulas for β_k are the Fletcher-Reeves (FR), Polak-Ribière (PR), Hestenes-Stiefel (HS), Descent-Dixon (DD) methods which are given, respectively, by:

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}; \beta_k^{PR} = \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}; \beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}; \beta_k^{DD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} \quad \dots\dots(4)$$

where $y_{k-1} = g_k - g_{k-1}$

Another important issue related to the performance of CG-methods is the line search, which requires sufficient accuracy to ensure that the search directions yield descent. Common criteria for line search accuracy are the Wolfe-Powell conditions:

$$f(x_{k-1} + \alpha_k d_k) - f(x_{k-1}) \leq \delta \alpha_k g_{k-1}^T d_{k-1}, \quad \dots\dots(5a)$$

$$g_k^T d_{k-1} \geq \sigma g_{k-1}^T d_{k-1}, \quad \dots\dots(5b)$$

$$f(x_{k-1} + \alpha_k d_k) - f(x_{k-1}) \leq \delta \alpha_k g_{k-1}^T d_{k-1}, \quad \dots(6a)$$

$$|g_k^T d_{k-1}| \leq -\sigma g_{k-1}^T d_{k-1} \quad \dots(6b)$$

where $0 < \delta < 0.5 \leq \sigma < 1$

Equations (5) and (6) are called the ‘‘Standard Wolfe’’ and ‘‘Strong Wolfe’’ conditions, respectively. When ,quadratic functions and exact line searches are used, all the above formulas in (4) are equivalent. However, these formulas vary according to general functions. For general functions, [22] proved the global convergence of PR method with exact line search. On the other hand, the PR and HS methods perform similarly in terms of theoretical property. Nevertheless, [16] showed that the PR and the HS methods can cycle infinitely without approaching a solution, which implies that they do not have globally convergence.

In this paper, we have proposed, two new special formulas $\rho_k^{SB}, \rho_k^{STS}$ for ρ_k^B, ρ_k^{TS} , respectively, from this formulas we can conclude the $\rho_k^{New1}, \rho_k^{New2}$ and two new formulas $\beta_k^{New1}, \beta_k^{New2}$ for β_k applying the rational non-quadratic model .Moreover, $\beta_k^{New1}, \beta_k^{New2}$ keeps the property of PR method, namely, if a very small step is generated the next search direction tends to the Steepest Descent (SD) direction, preventing a sequence of tiny steps from happening. Furthermore, finite quadratic termination is retained for the new methods. Since the sufficient descent condition is a property of great importance for the global convergence analysis of any CG-method, we have modified the conjugacy parameter of [21] to implement the non-quadratic rational model which satisfies the sufficient descent property and the modified Wolfe-Powel conditions introduced by Andrei [6] we illustrate this condition in section 4 . In addition, the global convergence property of the new proposed CG-method is discussed and a set of numerical results presented show that the new proposed method is efficient.

Materials and Methods

2.1 Extended CG-Methods For Non-Quadratic Models.

Over years, various authors have published works in this area, In this paper, a more general model than quadratic one is suggested as a basis for a CG-algorithm. If $q(x)$ is a quadratic function, then a function $F(q(x))$ is defined as a non-linear scaling of $q(x)$ if the following invariancy condition holds:

$$\text{Minimize } f(x) = F[q(x)]; \text{ where } \frac{df}{dq} = f' > 0 \text{ and } q > 0 \quad \dots\dots(7)$$

where, x^* is the minimizer of $q(x)$ with respect to x for more details see [19] and f is monotonic increasing, may be better to represent the objective and thus it gives an advantage to method based on this model. In order to obtain better global rate of convergence for minimization methods when applied to more general functions than the quadratic.

The following properties for $f(x)$ are immediately derived from the above condition:

- Every contour line of $q(x)$ is a contour line of $f(x)$
- If x^* is a minimizer of $q(x)$, then it is a minimizer of $f(x)$
- If x^* is a local minimizer of $q(x)$, then it is a local minimizer of $f(x)$

Boland et al. (1979) was the first who observed that $q(x)$ and $F[q(x)]$ have determined the same search directions so that the finite termination property for their algorithm was satisfied. Many authors have proposed special models as follows:

a) A conjugate method which minimizer the function $F[q(x)] = (q(x))^p; p > 0$ and $x \in R^n$

in at most n step has been described by Fried [11].

b) The special polynomial case, $F[q(x)] = \varepsilon_1 q(x) + \frac{1}{2} \varepsilon_2 (q(x))^2$

where $\varepsilon_2, \varepsilon_1$ are scalars, has been investigated by Boland et al, [9].

c) A rational model has been developed by (Tassopoulos and Storey) [20]

where:
$$F[q(x)] = \frac{\varepsilon_1 q(x) + 1}{\varepsilon_2 q(x)}; \varepsilon_2 < 0; \text{ and } q(x) > 0$$

d) Al-Bayati introduced several non-quadratic rational models; see for example [2]; [1]; [4] and [3].

2.2 ERCG-Method (Tassopoulos and Storey ,1984). [20]

Tassopoulos and Storey,1984 non-quadratic model is defined as:

$$F[q(x)] = \frac{\varepsilon_1 q(x) + 1}{\varepsilon_2 q(x)}; \quad \varepsilon_2 < 0; \text{ and } q(x) > 0 \quad \dots\dots(8a)$$

From Boland theorem [10] they get the parameter:

$$\rho_k^{TS} = \frac{(n)^2}{(2n + w)^2}; \text{ where } n = -(\lambda_{k-1} g_{k-1}^T d_{k-1})/2 \text{ and } w = f_k - f_{k-1}$$

Where they defined the direction as:

$$d_{k+1} = -g_{k+1} + \rho_k^{TS} \beta_k^{FR} d_k; \quad k = 1, 2, \dots$$

2.3 ERCG-Method (Al-Bayati,1993). [2]

Al-Bayati's, 1993 non-quadratic model is defined as the quotient of two quadratic functions and so belongs also to the class of rational functions , Al-Bayati's rational function model was considered by:

$$F[q(x)] = \frac{\varepsilon_1 q(x)}{1 - \varepsilon_2 q(x)}; \quad \varepsilon_2 < 0, \quad \varepsilon_1 > 0 \quad \dots\dots(8b)$$

$$\text{Where } q(x) = \frac{1}{2} (x - x_{\min})^T Q (x - x_{\min}) \quad \dots\dots(9)$$

is the quadratic function then it determines the solutions x_{\min} in a finite number of iterations not exceeding (n), and $F[q(x)]$ satisfy the property (7). He get the parameter:

$$\rho_k^B = \frac{(s_{k-1}^T g_{k-1})^2}{4(f_{k-1} - f_k)^2} \quad \dots\dots(10)$$

Where they defined the direction as:

$$d_{k+1} = -g_{k+1} + \rho_k^B \beta_k^{FR} d_k; \quad k = 1, 2, \dots$$

2.4 The Special Cases.

In this paper ,we introduced the two special cases of AL-Bayati's (1993) and Tassopoulos and Storey (1984) extended CG-method which are invariant to nonlinear scaling of quadratic rational functions are proposed. The first investigated model is defined as the quotient of two quadratic functions and so belongs also to the class of rational functions a special of AL-Bayati's rational function model was considered by:

$$F[q(x)] = \frac{q(x)}{1 - q(x)}; \quad \dots\dots(11)$$

From (7) We can rewritten (11) as:

$$f(x) = \frac{q(x)}{1-q(x)};$$

Where the function f is defined as nonlinear scaling of $q(x)$ and the invariancy property to nonlinear scaling (7) holds and $q(x)$ is defined in (9) is the quadratic function then it determines the solutions x_{\min} in a finite number of iterations not exceeding (n). It is shown the one-dimensional problem ($\min_{\alpha} f(x_k + \alpha d_k)$) and d_k is a search direction that the following updating process by the Boland theorem [10] to convert the quadratic model to a non-quadratic model in (12) :

$$d_1 = -g_1 \quad \dots\dots (12a)$$

$$d_{k+1} = -g_{k+1} + \rho_k \beta_k d_k; \quad k = 1, 2, \dots \quad \dots\dots(12b)$$

$$\rho_k = f'_k / f'_{k+1} \quad \dots\dots(12c)$$

$$\beta_k = g_{k+1}^T (g_{k+1} - g_k) / g_k^T g_k \quad \dots\dots(12d)$$

$$f' = \frac{df}{dq} \quad \dots\dots(12e)$$

we can write:

$$f = \frac{q}{1-q} \Rightarrow f - f q = q \Rightarrow f = (f+1)q$$

$$\Rightarrow q = \frac{f}{f+1}; \quad \dots\dots(13)$$

Since from the (12c) we have:

$$\rho_k = \frac{\frac{1-q_{k-1} + q_{k-1}}{(1-q_{k-1})^2}}{\frac{1-q_k + q_k}{(1-q_k)^2}} \Rightarrow \rho_k = \left(\frac{1-q_k}{1-q_{k-1}} \right)^2 \quad \dots\dots(14)$$

If substituting (13) in (14) we get:

$$\rho_k = \left(\frac{1 - \frac{f_k}{f_k+1}}{\frac{f_k}{f_k+1} + 1} \right)^2 = \left(\frac{f_k + 1 - f_k}{f_{k-1} + 1 - f_{k-1}} \cdot \frac{f_{k-1} + 1}{f_k + 1} \right)^2$$

$$\Rightarrow \rho_k^{SB} = \left(\frac{f_{k-1} + 1}{f_k + 1} \right)^2 \quad \dots\dots(15)$$

Similarly the special case of the rational function of Tassopoulos and Storey [20] in (8a). In this paper we shall state the rational function of AL-Assady and Shakory [18] considered by:

$$F[q(x)] = \frac{q(x)+1}{q(x)} \Rightarrow f = \frac{q+1}{q} \Rightarrow f q = q+1; \Rightarrow f q - q = 1$$

$$\Rightarrow q = \frac{1}{f-1}; \quad \dots\dots(16)$$

Since from the (12) we have:

$$\rho_k = \frac{\frac{q_{k-1} - 1 - q_{k-1}}{(q_{k-1})^2}}{\frac{q_k - 1 - q_k}{(q_k)^2}} \Rightarrow \rho_k = \left(\frac{q_k}{q_{k-1}} \right)^2 \quad \dots\dots(17)$$

If substituting (16) in (17) we get:

$$\rho_k = \left(\frac{\frac{1}{f_k - 1}}{\frac{1}{f_{k-1} - 1}} \right)^2 \Rightarrow \rho_k^{STS} = \left(\frac{f_{k-1} - 1}{f_k - 1} \right)^2 \quad \dots\dots(18)$$

The (18) and (15) are called the special case of rational function model of Tassopoulos and Storey (1984) and AL-Bayati's (1993) ,respectively.

2.5 Two New Combined of Rational Functions.

A. We introduce the combined of two rational function as a convex combination of the special case of Tassopoulos and Storey in the equation (18) and the rational function AL-Bayati's (1993) in the equation (10) non-quadratic model to be investigated here is considered by :

$$\rho_k^{New1} = \lambda \rho_k^B + (1 - \lambda) \rho_k^{STS}$$

$$\rho_k^{New1} = \lambda \cdot \left(\frac{s_{k-1}^T g_{k-1}}{2(f_k - f_{k-1})} \right)^2 + (1 - \lambda) \cdot \left(\frac{f_{k-1} - 1}{f_k - 1} \right)^2 \quad \dots\dots(19)$$

Where $0 < \lambda < 1$, we take the suitable value in the numerical result; $\lambda = \frac{1}{2}$.

B. We introduce another combined of two rational function as a convex combination of the special case of AL-Bayati's in (15) and the and AL-Bayati's in (10) non-quadratic model to be investigated here is considered by :

$$\rho_k^{New2} = \lambda \rho_k^B + (1-\lambda) \rho_k^{SB}$$

$$\rho_k^{New2} = \lambda \cdot \left(\frac{s_{k-1}^T g_{k-1}}{2(f_k - f_{k-1})} \right)^2 + (1-\lambda) \cdot \left(\frac{f_{k-1} + 1}{f_k + 1} \right)^2 \quad \dots\dots(20)$$

Where $0 < \lambda < 1$, we take the suitable value in the numerical result; $\lambda = \frac{1}{2}$

3. Wu and Chen (2010) CG-Method.

In this section, we are going to present the recent work of the two well-known Scientist Wu and Chen in (2010). They introduced several well-known CG-formulas. The conjugacy parameters of these CG-methods are given by; β_k^1 , β_k^2 , respectively by making use of the **Powell's restarting** criterion and the **Armijo**-type line search defined by:

$$\beta_k^1 = \beta_k^{HS} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{d_{k-1}^T y_{k-1}} \quad \dots\dots(21)$$

$$\beta_k^2 = \beta_k^{PR} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{\|g_{k-1}\|^2} \quad \dots\dots(22)$$

They proved that all the above CG-methods satisfy the sufficient descent condition and have the global convergence property for more details see [21].

4. A New Extended CG-Method.

4.1 Transform Quadratic model to Non-quadratic.

Consider the following quadratic model we proceed as in [21]:

$$f(x) = \frac{1}{2} x^T A x + b^T x + c \quad \dots\dots(23)$$

where $A \in R^{n \times n}$ is a symmetric positive definite matrix, $c \in R$ and $b \in R^n$ Wu and Chen in (2010) obtain the new formula of conjugate gradient method in (22) which based on the quadratic model, in this paper we

expand the conjugate coefficient (22) by add ρ_k using (19)-(20) to get two new extended CG method whose conjugacy parameters are defined by:

$$\beta_k^{New1} = \beta_k^{PR} + \frac{2(f_{k-1} - f_k) + \left(\lambda_1 \cdot \left(\frac{s_{k-1}^T g_{k-1}}{2(f_k - f_{k-1})} \right)^2 + \lambda_2 \cdot \left(\frac{f_{k-1} - 1}{f_k - 1} \right)^2 \right) (g_{k-1}^T s_{k-1})}{\|g_{k-1}\|^2} \dots(24)$$

$$\beta_k^{New1} = \beta_k^{PR} + \frac{8(f_{k-1} - f_k)^3 + \lambda_1 \cdot (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} + \lambda_2 \cdot \frac{(f_{k-1} - 1)^2 (g_{k-1}^T s_{k-1})}{(f_k - 1)^2 \|g_{k-1}\|^2} \dots(25)$$

$$\beta_k^{New2} = \beta_k^{PR} + \frac{2(f_{k-1} - f_k) + \left(\lambda_1 \cdot \left(\frac{s_{k-1}^T g_{k-1}}{2(f_k - f_{k-1})} \right)^2 + \lambda_2 \cdot \left(\frac{f_{k-1} + 1}{f_k + 1} \right)^2 \right) (g_{k-1}^T s_{k-1})}{\|g_{k-1}\|^2} \dots(26)$$

$$\beta_k^{New2} = \beta_k^{PR} + \frac{8(f_{k-1} - f_k)^3 + \lambda_1 \cdot (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} + \lambda_2 \cdot \frac{(f_{k-1} + 1)^2 (g_{k-1}^T s_{k-1})}{(f_k + 1)^2 \|g_{k-1}\|^2} \dots(27)$$

where $1 > \lambda_1, \lambda_2 > 0$.

4.1.1 An Acceleration Scheme of the Line Search Parameter.

In the CG-methods the search directions tend to be poorly scaled and as a consequence the line search must perform more function evaluations in order to obtain a suitable step-length α_k . In order to improve the performances of the CG-methods the efforts were directed to design procedures for direction computation based on the second order information. Jorge Nocedal [14] pointed out that in CG methods the step lengths may differ from 1 in a very unpredictable manner. They can be larger or smaller than 1 depending on how the problem is scaled. Numerical comparisons between CG methods and the limited memory QN method, by Liu and Nocedal [13], show that the latter is more successful [8]. Here, we have pointed out Andrei's [7] acceleration scheme; basically, this modifies the step length in a multiplicative manner to improve the reduction of the function values along the iterations [5, 6].

4.1.2 Outline of The Two New Extended CG-Method.

Step 1: Given $x_1 \in R^n$; ($\varepsilon > 0$); (k) is an index of the algorithm

Step 2: Set $k=1$; $d_k = -g_k$

Step 3: Using the modification WP line search conditions which fully described by Andrei (2009) determine the step length α_k , such that, compute: $z_{k+1} = x_k + \alpha_k d_k$, $y_k = g_k - g_z$, $g_z = \nabla f(z)$.

Acceleration scheme, compute , $a_k = \alpha_k g_k^T d_k, b_k = -\alpha_k y_k^T d_k$. If $b_k \neq 0$, then compute $\lambda_k = -\frac{a_k}{b_k}$ and update the variables as $x_{k+1} = x_k + \lambda_k \alpha_k d_k$, otherwise update the variables as $x_{k+1} = x_k + \alpha_k d_k$.

Step 4: Compute f_{k+1} , g_{k+1} , use $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$.

Step 5: If Powell restarting, $|g_k^T g_{k-1}| > 0.2 \|g_k\|^2$, satisfied then set:

$d_{k+1} = -g_{k+1}$, else set $d_{k+1} = -g_{k+1} + \beta_k^{New} d_k$ ($\beta_k^{New1}, \beta_k^{New2}$ is defined in (25), (27)), go to **Step 2**.

Step 6: If $\|g_{k+1}\| < \varepsilon$, stop else set $k=k+1$ go to **Step 3**.

4.2 Theoretical Properties for the Two New Extended CG-Method.

In this section, we focus on the convergence behavior on the $\beta_k^{New1}, \beta_k^{New2}$ methods with inexact line searches. Hence, we make the following basic assumptions on the objective function.

4.3 Assumption.[21]

f is bounded below in the level set $L_{x_0} = \{x \in R^n | f(x) \leq f(x_0)\}$; in some neighborhood U of the level set L_{x_0} , f is continuously differentiable and its gradient ∇f is Lipschitz continuous in the level set L_{x_0} , namely, there exists a constant $L > 0$ such that:

$$\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\| \quad \text{for all } x, y \in L_{x_0} \quad \dots\dots(28)$$

4.4 Assumption.[21] The level set L_{x_0} is compact.

4.5 Lemma. [12]

Consider a **general CG-method**, and suppose that $0 < \gamma \leq \|g_k\| \leq \bar{\gamma}$ holds. We call a method has **Lemma 4.5** if there exist two constants $b > 1$ and $p > 0$ such that for all k , $|\beta_k| \leq b$ and

$$\|s_k\| \leq p \Rightarrow |\beta_k| \leq \frac{1}{2b} \quad \dots\dots(29)$$

4.6 Lemma. (Zoutendijk Condition).[22]

Suppose that **Assumption 4.3** holds. Consider any CG-type method in the form of $x_{k+1} = x_k + \alpha_k d_k$ where d_k is a descent direction and α_k satisfies the Wolfe-Powell line search conditions in **(5)-(6)** . Then we have that:

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

4.7 Lemma.

Suppose that the **Assumption 4.3 and 4.4** hold. Consider the CG-algorithm 4.1.2, and α_k is obtained by the strong Wolfe-Powell (6) and for all $k \geq 0$, there exists the positive constant γ such that $\gamma \leq \|g_k\|$, then the new algorithm **(4.1.2)** has the **Lemma 4.5**.

Proof.

It follows from the definition of β_k^{New1} in (25), and by the compactness of the level set L_{x_0} , there exist three constants $M_1 > 0, M_2 > 0, \bar{\gamma} > 0$ such that:

$$\|x_k\| \leq M_1, |f(x_k)| \leq M_2, \|g_k\| \leq \bar{\gamma}, x_k \in L_{x_0}$$

Under **Assumption 4.3 and 4.4**, we know there is a constant $L_1 > 0$ satisfying :

$$|f(x_k) - f(x_{k-1})| \leq L_1 \|x_k - x_{k-1}\|, x_k, x_{k-1} \in L_{x_0}$$

$$|\beta_k^{New1}| = \left| \beta_k^{PR} + \frac{8(f_{k-1} - f_k)^3 + \lambda_1 \cdot (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} + \lambda_2 \cdot \frac{(f_{k-1} - 1)^2 (g_{k-1}^T s_{k-1})}{(f_k - 1)^2 \|g_{k-1}\|^2} \right|$$

$$|\beta_k^{New1}| \leq \left| \beta_k^{PR} \right| + \left| \frac{8(f_{k-1} - f_k)^3 + \lambda_1 \cdot (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} \right| + \left| \lambda_2 \cdot \frac{(f_{k-1} - 1)^2 (g_{k-1}^T s_{k-1})}{(f_k - 1)^2 \|g_{k-1}\|^2} \right| \dots\dots(30)$$

The third term from the last equation can be simplified as:

$$\left| \frac{(f_{k-1} - 1)(f_{k-1} - 1)}{(f_k - 1)(f_k - 1)} \right| \leq \frac{(|f_{k-1}| + 1)^2}{(|f_k| + 1)^2} = \frac{(M_2 + 1)^2}{(M_2 + 1)^2} = 1$$

$$|\beta_k^{New1}| \leq \frac{2\bar{\gamma}^2}{\gamma^2} + \frac{64M_2^3 + 8\lambda_1 \cdot M_1^3 \bar{\gamma}^3}{16M_2^2 \gamma^2} + \frac{2\lambda_2 \bar{\gamma} M_1}{\gamma^2}$$

$$|\beta_k^{New1}| \leq \frac{2(4M_2^2 \bar{\gamma}^2 + 8M_2^3 + \lambda_1 \cdot M_1^3 \bar{\gamma}^3 + 4\lambda_2 M_1 M_2^2 \bar{\gamma})}{4M_2^2 \gamma^2} \equiv b \dots\dots(31)$$

Define

$$p = \frac{4M_2^2 \gamma^2}{2b(4M_2^2 \bar{\gamma}^2 + 8M_2^3 + \lambda_1 \cdot M_1^3 \bar{\gamma}^3 + 4\lambda_2 M_1 M_2^2 \bar{\gamma})} \dots\dots(32)$$

we obtain from the Lipschitz continuous that if $\|s_k\| \leq p$, for all $k \geq 0$ then

$$|\beta_k^{New1}| \leq \frac{\bar{\gamma} L \|s_k\|}{\gamma^2} + \frac{8L_1^3 \|s_k\|^3 + 8\lambda_1 \cdot \bar{\gamma}^3 \|s_k\|^3}{4L_1^2 \|s_k\|^2 \gamma^2} + \frac{\lambda_2 \bar{\gamma} \|s_k\|}{\gamma^2}$$

$$|\beta_k^{New1}| \leq \frac{p(4\bar{\gamma} L L_1^2 + 8L_1^3 + \lambda_1 \cdot \bar{\gamma}^3 + 4\lambda_2 L_1^2 \bar{\gamma})}{4L_1^2 \gamma^2} = \frac{1}{2b} \dots\dots(33)$$

Therefore, for b and p defined in (31) and (32) respectively, it follows that the relation (29) hold, and similar if we want to proof that β_k^{New2} has the **Lemma 4.5**

4.8 Theorem.

Suppose that **Assumption 4.3** and **4.4** holds. Consider the new extended CG-method defined in (25) , (27) with β_k^{New1} , β_k^{New2} if α_k is obtained by an exact line search. If $\|s_k\| \rightarrow 0$, then: $\liminf_{k \rightarrow \infty} \|g_k\| = 0$

Proof:

We now prove the theorem by contradiction and assume that there exists some constants $\gamma > 0$ such that $\|g_k\| \geq \gamma$ for all $k \geq 0$. The compactness of the level set L_{x_0} implies that there exists a constant $\bar{\gamma} > 0$ such that $\|g_k\| \leq \bar{\gamma}$. Since $\|s_k\| \rightarrow 0$, we know that there is a \bar{k} , for all $k > \bar{k}$ such that $k < \|s_k\| \leq p$, where p is the same as in **Lemma 4.6**. Then, for all $k > \bar{k}$, we have:

$$\begin{aligned} \|d_k\| &\leq \|g_k\| + |\beta_k^{New}| \|d_{k-1}\| && \dots\dots(34) \\ &\leq \bar{\gamma} + \frac{1}{2b}(\bar{\gamma} + \frac{1}{2b}\|d_{k-2}\|) \\ &= (1 + \frac{1}{2b})\bar{\gamma} + \frac{1}{(2b)^2}(\|d_{k-2}\|) \\ &\leq \dots \leq (\frac{1}{1 - \frac{1}{2b}})\bar{\gamma} + \frac{1}{(2b)^{k-\bar{k}}}(\|d_{\bar{k}}\|) \end{aligned}$$

Such that $\frac{1}{1 - \frac{1}{2b}}$ is a geometric and the $\lim_{k \rightarrow \infty} \frac{1}{(2b)^{k-\bar{k}}} = 1$

$$\leq (\frac{2b}{2b-1})\bar{\gamma} + \|d_{\bar{k}}\| \equiv \bar{\eta} \quad \dots\dots(35)$$

Furthermore, we know

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{\gamma^4}{\|d_k\|^2} \quad \dots\dots(36)$$

we know using **Lemma 4.6** together with (36), yields

$$\sum_{k=0}^{\infty} \frac{\gamma^4}{\|d_k\|^2} < \infty \quad \dots\dots(37)$$

Which contradictions (35). Therefore, we conclude the truth of the theorem.

4.9 Theorem.

Suppose that **Assumption 4.3** holds. If there exists a constant $\gamma > 0$ such that $\|g_k\| \geq \gamma$, for all $k \geq 0$. If α_k is obtained by modified Wolfe-Powell conditions and d_k is defined in (2b) of the new $\beta_k^{New1}, \beta_k^{New2}$ CG-method and $0 < \lambda_k < 1$, then the new extended method has sufficient descent directions i.e.,

$$d_k^T g_k \leq -c \|g_k\|^2; \quad c > 0 \quad \dots\dots(38)$$

Proof:

For initial direction we have:

$$d_1 = -g_1 \Rightarrow d_1^T g_1 = -\|g_1\|^2 \leq 0 \quad \dots\dots(39)$$

which satisfies (35). Now let the theorem be true for all $k-1$, i.e.

$$d_{k-1} = -g_{k-1} \Rightarrow d_{k-1}^T g_{k-1} = -\|g_{k-1}\|^2 \leq 0 \quad \dots\dots(40)$$

Since our function f is uniformly **convex** function either in the quadratic or in the non-quadratic regions, then there exist a Lipschitz constant $L > 0$ and a constant, $\eta > 0$ such that:

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq \eta \|x - y\|^2 \quad \text{for all } x, y \in L_{x_0} \quad \dots\dots(41)$$

Or equivalently:

$$y_{k-1}^T s_{k-1} \geq \eta \|s_{k-1}\|^2 \quad \text{and} \quad \eta \|s_{k-1}\|^2 \leq y_{k-1}^T s_{k-1} \leq L \|s_{k-1}\|^2 \quad \dots\dots(42)$$

A. Multiplying the search direction of (25) by g_k^T yields:

$$\begin{aligned} d_k^T g_k = & -\|g_k\|^2 + \left(\frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}} \right) (s_{k-1}^T g_k) + \left(\frac{8(f_{k-1} - f_k)^3 + \lambda_1 (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} \right) (s_{k-1}^T g_k) \\ & + \lambda_2 \frac{(1 - f_{k-1})(s_{k-1}^T g_{k-1})}{(1 - f_k) \|g_{k-1}\|^2} (s_{k-1}^T g_k) \end{aligned}$$

For inexact line search using Wolfe-Powell conditions (5) and (6) we have:

$$d_k^T g_k \leq -\|g_k\|^2 + \left(\frac{s_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2}\right)(\|g_k\|^2) + \left(\frac{-8\delta^3 (g_{k-1}^T s_{k-1})^3 + \lambda_1 (g_{k-1}^T s_{k-1})^3}{4\delta^2 (g_{k-1}^T s_{k-1})^2 \|g_{k-1}\|^2}\right)(s_{k-1}^T g_k) \\ + \lambda_2 \zeta_1 \frac{(s_{k-1}^T g_{k-1})}{\|g_{k-1}\|^2} (s_{k-1}^T g_k)$$

$$d_k^T g_k \leq -\|g_k\|^2 + \left(\frac{s_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2}\right)(\|g_k\|^2) + \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2 \|g_{k-1}\|^2}\right)(s_{k-1}^T g_k)(g_{k-1}^T s_{k-1}) \\ + \lambda_2 \zeta_1 \frac{(s_{k-1}^T g_{k-1})}{\|g_{k-1}\|^2} (s_{k-1}^T g_k) \quad \dots\dots(43)$$

$$d_k^T g_k \leq -\|g_k\|^2 + L(\alpha_{k-1})^2 \|g_k\|^2 - \alpha_{k-1} \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right)(s_{k-1}^T g_k) - \alpha_{k-1} \lambda_2 \zeta_1 (s_{k-1}^T g_k) \dots(44)$$

Powell and criteria which are defined as:

$$g_k^T g_{k-1} > \psi \|g_k\|^2; \quad \psi \in (0,1) \quad \dots\dots(45)$$

From [17] Powell restarting criterion (45) we have:

$$g_k^T s_{k-1} = \alpha_{k-1} g_k^T d_{k-1} = -\alpha_{k-1} g_k^T g_{k-1} \quad \dots\dots(46)$$

Using (45) and (46) in (44):

$$d_k^T g_k \leq -\|g_k\|^2 + L(\alpha_{k-1})^2 \|g_k\|^2 + \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) \|g_k\|^2 \\ + \psi \lambda_2 \zeta_1 (\alpha_{k-1})^2 \|g_k\|^2 \quad \dots\dots(47)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq -1 + L(\alpha_{k-1})^2 + \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) + \psi \lambda_2 \zeta_1 (\alpha_{k-1})^2 \quad \dots\dots(48)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq - (1 - \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) - L(\alpha_{k-1})^2 - \psi \lambda_2 \zeta_1 (\alpha_{k-1})^2) \quad \dots\dots(49)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq -c, \quad c > 0 \quad \text{for} \quad \dots\dots(50) \\ 0 < \delta < 0.5; \quad 0 < \alpha, \psi, L, \lambda < 1$$

B. Multiplying the search direction of (27) by g_k^T yields:

$$d_k^T g_k = -\|g_k\|^2 + \left(\frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}\right)(s_{k-1}^T g_k) + \left(\frac{8(f_{k-1} - f_k)^3 + \lambda_1 (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2}\right)(s_{k-1}^T g_k) \\ + \lambda_2 \frac{(1 + f_{k-1})(s_{k-1}^T g_{k-1})}{(1 + f_k) \|g_{k-1}\|^2} (s_{k-1}^T g_k)$$

For inexact line search using Wolfe-Powell conditions (5) and (6) we have:

$$d_k^T g_k \leq -\|g_k\|^2 + \left(\frac{s_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2}\right)(\|g_k\|^2) + \left(\frac{-8\delta^3 (g_{k-1}^T s_{k-1})^3 + \lambda_1 (g_{k-1}^T s_{k-1})^3}{4\delta^2 (g_{k-1}^T s_{k-1})^2 \|g_{k-1}\|^2}\right)(s_{k-1}^T g_k) \\ + \lambda_2 \zeta_2 \frac{(s_{k-1}^T g_{k-1})}{\|g_{k-1}\|^2} (s_{k-1}^T g_k)$$

$$d_k^T g_k \leq -\|g_k\|^2 + \left(\frac{s_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2}\right)(\|g_k\|^2) + \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2 \|g_{k-1}\|^2}\right)(s_{k-1}^T g_k)(g_{k-1}^T s_{k-1}) \\ + \lambda_2 \zeta_2 \frac{(s_{k-1}^T g_{k-1})}{\|g_{k-1}\|^2} (s_{k-1}^T g_k) \quad \dots\dots(51)$$

$$d_k^T g_k \leq -\|g_k\|^2 + L(\alpha_{k-1})^2 \|g_k\|^2 - \alpha_{k-1} \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right)(s_{k-1}^T g_k) - \alpha_{k-1} \lambda_2 \zeta_2 (s_{k-1}^T g_k) \dots\dots(52)$$

From Powell restarting criterion (45) , and (46) in (44):

$$d_k^T g_k \leq -\|g_k\|^2 + L(\alpha_{k-1})^2 \|g_k\|^2 + \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) \|g_k\|^2 \\ + \psi \lambda_2 \zeta_2 (\alpha_{k-1})^2 \|g_k\|^2 \quad \dots\dots(53)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq -1 + L(\alpha_{k-1})^2 + \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) + \psi \lambda_2 \zeta_2 (\alpha_{k-1})^2 \quad \dots\dots(54)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq - (1 - \psi(\alpha_{k-1})^2 \left(\frac{(\lambda_1 - 8\delta^3)}{4\delta^2}\right) - L(\alpha_{k-1})^2 - \psi \lambda_2 \zeta_2 (\alpha_{k-1})^2) \quad \dots\dots(55)$$

$$(d_k^T g_k) / (\|g_k\|^2) \leq -c ; \quad c > 0 \text{ for } 0 < \delta < 0.5; 0 < \alpha, \psi, L, \lambda < 1 \quad \dots\dots(56)$$

Thus our new proposed extended CG-method has sufficient descent directions using **inexact line searches** under the condition that Powell restarting condition must be used.

4.10 Theorem

Suppose that **Assumption 4.3** hold. Consider the method (2)-(3) with the following three properties:

- (i) $\beta_k \geq 0$ for all k;
- (ii) the line search satisfies the Zoutendijk condition, and the sufficient descent condition (38);
- (iii) **Lemma 4.4** holds.

Then $\liminf_{k \rightarrow \infty} \|g_k\| = 0$, for more details see [12].

Therefore, the method has a global convergent property by satisfying the conditions of Zoutendijk theorem and the line search satisfy the strong Wolfe condition then from Gilbert and Nocedal in [12] these method is global convergent.

5. Numerical Results

The main work of this section is to report the performance of the new method on a set of test problems. The codes are written in Fortran and in double precision arithmetic. All the tests are performed on a PC. Our experiments are performed on a set of case 35 nonlinear unconstrained problems that have second derivatives available. These test problems are contributed in CUTE and their details are given in the Appendix.

In **Case** are divided in **three** branches according to the numerical experiments with their number of variables:

- 1- 10 numerical experiments with $n = 100, 200, \dots, 1000$.
- 2- 5 numerical experiments with $n = 100, 300, 500, 700, 900$.
- 3- 4 numerical experiments with $n = 100, 400, 700, 1000$.

In order to assess the reliability of our new proposed method, we have tested it against the standard Wu & Chen's modified PRCG-method (Wu and Chen, 2010) using the same set of test problems. All these methods terminate when the following stopping criterion is met:

$$\|g_k\| \leq 10^{-6} \quad \dots\dots(57)$$

Tables 5.1, 5.3 and 5.5 compare some numerical results for the modified PRCG method of Wu & Chen and the two new extended PRCG method for 35 test functions. In all these tables (n) indicates for the dimension of the problem; (NOI) indicates for the number of iterations; (NOFG) indicates for the number of function and gradient evaluations; (TIME) indicates for the total time required to complete the evaluation process for each test problem.

In **Tables 5.2, 5.4 and 5.6** we have compared the percentage performance of the two new extended PRCG-methods against the standard Wu & Chen PRCG-method taking over all the tools as 100%. In order to summarize our numerical results, we have concerned only on the **Total** of n different dimensions for all tools used in these comparisons.

Table (5.1) Comparison between the new 1 &2 and (wu & chen) methods for the total of different dimensions of (35) problems with

n= 100, 200, ... ,1000

Pro b.	n	Wu & Chen/2010			New Extended 1 PRCG/2012			New Extended 2 PRCG/2012		
		NOI	NOFG	TIME	NOI	NOFG	TIME	NOI	NOFG	TIME
1	100	169	206	0.03	164	201	0.02	167	204	0.01
	800	317	347	0.3	278	327	0.11	278	327	0.25
	1000	182	212	0.21	127	157	0.07	127	157	0.14
2	200	19	39	0	19	39	0	19	39	0
3	100	7	9	0	7	9	0	7	9	0
4	300	110	124	0.04	110	124	0.02	110	124	0.03
5	100	95	103	0.02	95	103	0	95	103	0
6	100	35	38	0.01	34	37	0	34	37	0
	300	36	39	0.02	24	26	0	24	26	0
	400	36	39	0.01	35	38	0.01	35	38	0.03
	500	37	40	0.03	29	31	0	29	31	0.01
	800	37	40	0.05	36	39	0.01	36	39	0.05
	1000	37	40	0.07	36	39	0	36	39	0.07
7	100	524	544	0.02	15	18	0	15	18	0
	200	575	602	0.01	14	17	0	27	30	0
	300	226	228	0.02	9	12	0	15	18	0
	700	457	460	0.07	301	303	0.02	18	21	0.02
	800	536	539	0.09	41	44	0	91	94	0
	900	227	230	0.03	71	74	0	109	112	0.02
	1000	481	485	0.09	29	32	0	215	238	0.05
8	900	17	20	0.05	17	20	0	17	20	0.03
9	300	39	51	0	27	38	0	26	37	0
	900	27	39	0	26	38	0.02	26	38	0
10	100	19	27	0	16	23	0	16	23	0
	200	21	34	0.02	20	22	0	20	22	0

Two New Extended PR Conjugate Gradient Methods for Solving Nonlinear ...

	700	22	30	0.01	21	23	0	21	23	0
	800	21	23	0.02	19	27	0	19	27	0.01
11	200	51	62	0	49	60	0	34	43	0.02
	700	40	52	0.01	38	50	0.02	38	50	0.02
12	600	15	23	0.02	15	23	0	15	23	0
13	1000	3	6	0.02	3	6	0	3	6	0
14	500	149	157	0.02	149	157	0	149	157	0.01
15	900	490	498	0.12	490	498	0.03	490	498	0.11
16	1000	8	11	0	8	11	0.02	8	11	0
17	900	9	12	0.05	9	12	0	9	12	0.01
18	200	11	14	0.02	11	14	0	11	14	0
19	600	130	139	0.03	130	139	0.02	130	139	0.01
20	900	7	9	0.01	7	9	0	7	9	0
21	100	92	99	0.02	92	99	0	92	99	0
22	200	36	45	0	34	43	0	36	45	0
	300	43	45	0	38	47	0	43	45	0.02
	400	50	52	0	49	52	0	50	52	0.01
	700	95	97	0.05	53	55	0.01	95	97	0.05
	800	98	100	0.05	44	46	0	94	96	0.05
	900	94	96	0.04	70	72	0.02	94	96	0.04
	1000	90	92	0.07	47	49	0.02	71	73	0.05
23	200	152	164	0.03	73	80	0	136	149	0.01
	400	169	185	0.05	80	96	0	143	159	0.05
	800	148	165	0.09	124	144	0.01	92	122	0.06
	900	145	163	0.12	87	105	0.02	65	69	0.03
	1000	136	155	0.11	50	52	0.01	40	42	0.05
24	600	13	25	0.02	13	25	0	13	25	0.01
25	600	24	35	0	23	25	0	23	25	0
26	100	96	106	0.01	95	100	0.01	74	79	0.01
	600	123	145	0.04	107	129	0.01	49	71	0.02
	800	109	119	0.04	43	45	0	60	62	0.03

Two New Extended PR Conjugate Gradient Methods for Solving Nonlinear ...

	900	65	67	0.05	58	60	0.01	45	47	0.03
27	500	13	23	0	12	21	0	12	21	0
28	1000	8	16	0.02	8	16	0	8	16	0
29	300	9	11	0.02	9	11	0	9	11	0
30	900	4	8	0.02	4	8	0	4	8	0
31	700	22	30	0.01	21	23	0	21	23	0.02
	800	21	23	0.02	19	27	0	19	27	0.01
32	100	106	116	0	104	114	0	104	114	0
	200	144	153	0	134	156	0	134	156	0
	300	128	141	0	128	143	0.02	80	92	0.02
33	100	1	4	0	1	4	0	1	4	0
34	900	8	11	0.02	8	11	0	8	11	0
35	100	18	27	0	16	25	0	15	24	0
	200	20	30	0	16	25	0	15	24	0
	300	20	29	0	16	25	0	15	24	0
	400	20	29	1	18	27	0	17	26	0
	500	20	29	0	18	28	0	17	27	0.02
	600	20	29	0	18	28	0	17	27	0
	700	20	29	0	18	28	0	17	27	0
	800	20	29	0	18	28	0	17	27	0
	900	20	29	0	18	28	2	18	28	0
	1000	20	29	0	18	28	0	18	28	0
Total	7662	8351	3.42	4331	4968	2.51	4407	5054	1.49	

Table (5.2) Percentage performance of table (5.1)

Methods	NOI	NOFG5	TIME
---------	-----	-------	------

Wu & Chen 2010	100 %	100 %	100 %
New algorithm 2012 (1)	56.5	59.5	73.4
New algorithm 2012 (2)	57.5	60.5	43.6

Table (5.3) Comparison between the new 1&2 and (wu & chen) methods for the total of (35) problems with

n= 100, 300, ... ,900

Prob.	Wu & Chen/2010			New Extended 1 PRCG/2012			New Extended 2 PRCG/2012		
	NOI	NOFG	TIME	NOI	NOFG	TIME	NOI	NOFG	TIME
1	936	1099	0.7	940	1106	0.7	943	1109	0.65
2	113	208	0.01	113	208	0.02	113	208	0.02
3	37	47	0.02	37	47	0	37	47	0.03
4	775	848	0.46	775	848	0.47	775	848	0.48
5	513	559	0.08	513	559	0.06	513	559	0.08
6	151	163	0.11	170	184	0.1	170	184	0.14
7	4244	4277	0.31	1098	1112	0.13	207	222	0.01
8	81	98	0.1	81	98	0.09	81	98	0.09
9	168	228	0.02	168	228	0.01	168	228	0.03
10	93	134	0.05	93	134	0.07	93	134	0.03
11	219	289	0.06	280	337	0.06	214	268	0.06
12	89	140	0.02	89	140	0	89	140	0
13	18	35	0.02	18	35	0.02	18	35	0.02
14	660	698	0.09	663	701	0.09	660	698	0.09
15	3311	3349	0.48	3315	3353	0.51	3311	3349	0.52
16	65	75	0.02	70	80	0.03	65	75	0.02
17	45	60	0.02	45	60	0.02	45	60	0.03
18	54	65	0.01	54	65	0.02	54	65	0.01
19	639	683	0.11	639	683	0.09	639	683	0.09
20	37	47	0.01	37	47	0	37	47	0

21	455	494	0.06	455	494	0.06	455	494	0.07
22	374	404	0.14	278	301	0.1	337	360	0.12
23	429	492	0.19	406	475	0.15	315	364	0.09
24	68	87	0.01	73	102	0.02	68	87	0
25	119	174	0.03	125	135	0.02	120	157	0.02
26	510	549	0.17	422	458	0.13	421	472	0.13
27	74	99	0.02	69	102	0.01	69	102	0.01
28	40	80	0.01	40	81	0.01	40	81	0.04
29	43	53	0.01	43	53	0.01	43	53	0.01
30	22	38	0.01	22	38	0	22	38	0.01
31	95	122	0.03	95	122	0.05	95	122	0.04
32	572	632	0.08	701	763	0.08	649	708	0.08
33	12	35	0	12	35	0	12	35	0
34	39	55	0.02	39	55	0.02	39	55	0.01
35	98	144	0.02	86	133	0	82	129	0
Total	15198	16560	3.5	12064	13372	3.15	10999	12314	3.03

Table (5.4) Percentage performance of table (5.3)

Methods	NOI	NOFG	TIME
Wu & Chen 2010	100 %	100 %	100 %
New algorithm 2012 (1)	79.3	80.7	90
New algorithm 2012 (2)	72.4	74.4	86.6

Table (5.5) Comparison between the new 1&2 and (wu & chen) methods for the total of (35) problems with

n= 100, 400,700 ,1000

Prob.	Wu & Chen/2010			New Extended 1 PRCG/2012			New Extended 2 PRCG/2012		
	NOI	NOFG	TIME	NOI	NOFG	TIME	NOI	NOFG	TIME

Two New Extended PR Conjugate Gradient Methods for Solving Nonlinear ...

1	892	1008	0.66	836	952	0.63	839	955	0.63
2	87	164	0	87	164	0	87	164	0.01
3	29	38	0.03	29	38	0.01	29	38	0.02
4	627	681	0.42	627	681	0.4	631	685	0.42
5	406	443	0.06	406	443	0.05	406	443	0.06
6	137	147	0.13	135	146	0.1	135	146	0.13
7	2518	2548	0.34	799	810	0.09	330	362	0.06
8	66	79	0.09	66	79	0.04	66	79	0.08
9	141	183	0.02	141	183	0.01	141	183	0
10	84	119	0.03	84	119	0.05	84	119	0.05
11	233	285	0.08	252	304	0.07	236	288	0.08
12	98	138	0.02	98	138	0	98	138	0.01
13	14	28	0	14	28	0.02	14	28	0
14	510	538	0.1	513	541	0.07	510	538	0.07
15	2156	2188	0.4	2160	2192	0.41	2156	2188	0.41
16	53	61	0.03	57	65	0.02	53	61	0.01
17	36	48	0	36	48	0	36	48	0
18	43	52	0	43	52	0.01	43	52	0.02
19	510	545	0.06	510	545	0.08	510	545	0.08
20	29	38	0	29	38	0.01	29	38	0.02
21	379	412	0.04	379	412	0.07	379	412	0.06
22	334	355	0.15	201	222	0.07	315	336	0.13
23	386	428	0.19	339	393	0.13	264	289	0.09
24	56	84	0.02	67	115	0.01	67	115	0.02
25	98	124	0	101	118	0.02	101	118	0.02
26	333	357	0.12	322	341	0.09	244	275	0.08
27	61	84	0	61	84	0.02	61	84	0
28	32	64	0.03	32	64	0.02	32	64	0.02
29	35	43	0.03	35	43	0.02	35	43	0.02
30	18	30	0.02	18	30	0.02	18	30	0.01

31	76	95	0.03	76	95	0.05	76	95	0.06
32	393	437	0.05	415	459	0.04	402	446	0.04
33	9	27	0	9	27	0	9	27	0
34	32	44	0.02	32	44	0.01	32	44	0.02
35	78	115	0.01	70	108	0	68	106	0
Total	10989	12030	3.18	9079	10121	2.64	8536	9582	2.73

Table (5.6) Percentage performance of table (5.5)

Methods	NOI	NOFG	TIME
Wu & Chen 2010	100 %	100 %	100 %
New algorithm 2012 (1)	82.6	84.1	83
New algorithm 2012 (2)	77.7	79.7	85.8

Discussion

It is clear from **Table (5.2)** that taking, over all, the tools as a 100% for the Wu & Chen PRCG method, the two New Extended PRCG method has an improvement, in about New 1 Extended PR CG method (43.5%) NOI , (40.5%) NOFG and (26.6%) TIME, in about New 2 Extended PR CG method (42.5%) NOI , (39.5%) NOFG and (56.4%) TIME.

In from **Table (5.4)** that taking, over all, the tools for PRCG method has an improvement, in about New 1 Extended PR CG method (20.7%) NOI , (19.3%) NOFG and (10%) TIME , in about New 2 Extended PR CG method (27.6%) NOI , (25.6%) NOFG and (13.4%) TIME.

It is clear from **Table (5.6)** that taking, over all, the tools for PRCG method has an improvement, in about New 1 Extended PR CG method (17.4%) NOI , (15.9%) NOFG and (17%) TIME , in about New 2 Extended PR CG method (22.3%) NOI , (20.3%) NOFG and (14.2%) TIME. These results indicate that the two new extended PRCG method is in general is the best.

Appendix.

(1)Trigonometric (2)Penalty (3)Raydan (4)Hager (5)Generalized Tridiagonal (6)Extended Three Exp-Terms (7)Diagonal4 (8)Diagonal (9)Extended Himmelblau (10)Extended PSC1 (11)Extended BD1 (12)Extended Quadratic Penalty QP1 (13)Extended EP1 (14)Extended Tridiagonal-2 (15)ARWHEAD (CUTE) (16)DIXMAANA (CUTE) (17)DIXMAANB (CUTE) (18)DIXMAANC (CUTE) (19) EDENSCH (CUTE) (20)DIAGONAL-6 (21)ENGVAL1 (CUTE) (22)DENSCHNA (CUTE) (23)DENSCHNC (CUTE) (24)DENSCHNB (CUTE) (25)DENSCHNF (CUTE) (26)Extended Block-Diagonal BD2 (27)Generalized quarticGQ1 (28)DIAGONAL 7 (29)DIAGONAL-8 (30)Full Hessian (31)SINCOS (32)Generalized quartic GQ2 (33)ARGLINB (CUTE) (34)HIMMELBG (CUTE) (35)HIMMELBH (CUTE)

References

- [1] Al-Bayati A. Y. and Al- Naemi G. M., New extended CG-method for nonlinear optimization. *J. of Mu'tah, Jordan*, **10**:69-87 (1995)
- [2] Al-Bayati A. Y., A New Non-Quadratic Model for Unconstrained Non-linear Optimization Method. *Natural and applied series Mu'tah Journal for research and studies Mu'tah University, Jordan*, **8**(1):133-155 (1993)
- [3] Al-Bayati A. Y., New Extended CG methods for combined non-quadratic models in nonlinear optimization. *J. of Al-Yarmouk, Jordan*, **16**:347-366 (2007)
- [4] Al-Bayati A. Y., New generalized CG-methods for the non-quadratic model in unconstrained optimizations. *J. Al- Yarmouk*, **10**:9-25(2001)
- [5] Andrei N., Accelerated hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization. *ICI Technical Report*, (2008)
- [6] Andrei N., Acceleration of conjugate gradient algorithms for unconstrained optimization. *Applied Mathematics and Computation*, **213**:361-369 (2009)
- [7] Andrei N., Accelerated scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. *European Journal of Operational Research*, **204**:410–420 (2010)
- [8] Andrei N., Performance profiles of conjugate gradient algorithms for unconstrained optimization). In: Floudas, C.A., Pardalos, P.M. (eds.) *Encyclopedia of Optimization*, 2nd edition., p, 2938–2953. Springer, New York (2009)
- [9] Boland W. R. and kowalik J. S., Extended Conjugate Gradient method with Restarts. *Journal of Optimization Theory and Application*, **28**:1-9 (1979)
- [10] Boland W. R.; Kamgnia E. R. and Kowalik J. S., A conjugate gradient optimization method invariant to nonlinear scaling. *Journal of optimization theory and applications*, **27**:221-230 (1979)

- [11] Fried I., N-step Conjugate Gradient Minimization Scheme for Non-Quadratic Function. *AIAA Journal*, **19**:2286-2287 (1971)
- [12] Gilbert C. and Nocedal J., Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on Optimization*, **2**:21–42 (1992)
- [13] Nocedal J. and Liu D.C., On the limited memory BFGS method for large scale optimization). *Mathematical Program.* **45**:503–528 (1989)
- [14] Nocedal J., "Conjugate gradient methods and nonlinear optimization". In: L. Adams, J.L. Nazareth, (eds.) (*Linear and Nonlinear Conjugate Gradient—Related Methods*), 9–23. SIAM, Philadelphia ,1996.
- [15] Perry A., A modified conjugate gradient algorithm. *Operations Research*, **26**:1073–1078 (1978)
- [16] Powell M. J. D., Non-convex minimization calculations and the conjugate gradient method. *Lecture Notes in Mathematics*, 122–141 (1984)
- [17] Powell M. J. D., Restart procedure for the conjugate gradient method. *Mathematical Programming*, **2**:241-254 (1977)
- [18] Shakory O. S., M. Sc. Thesis, College of Computer Sciences And Mathematics, University of Mosul (2003). (In Arabic)
- [19] Spedicato E., A Variable Metric method for function minimization derived from invariancy to nonlinear scaling. *Journal of optimization theory and applications*, **20**: 315-329 (1976)
- [20] Tassopoulos A. and Storey C., Use of Non, Quadrate Model in A conjugate Gradient Model of Optimization with inexact line search. *JOTA Journal Theory and Applications*, **43**:353-730 (1984)
- [21] Wu C. Y. and Chen G. Q., New type of conjugate gradient algorithm for unconstrained optimization problems. *Journal of Systems Engineering and Electronics*, **21**(6):1000-1007 (2010)
- [22] Zoutendijk G., *Nonlinear programming, computational methods in Integer and nonlinear programming.* North-Holland Amsterdam , 37–86 (1970)