A New Extended PR Conjugate Gradient Method for Solving Smooth Minimization Problems

Prof. Dr. Abbas Y. AL-Bayati **Rana Z. AL-Kawaz**
Department Mathematics, College of Computer Sciences and Mathematics, Mosul University, Mosul, Iraq.

Abstract

In this paper, we have discussed and investigated an extended PR-CG method which uses function and gradient values. The new method involves the extended CG-methods and have the sufficient descent and globally convergence properties under certain conditions. We have got some important numerical results by improving an standard computer program compared with Wu and Chen (2010) method in this field.

طريقة PR الموسعة في التدرج المترافق لحل المسائل التصغيرية الناعمة

المستخلص

في هذا البحث تطرقنا إلى تقصي و اشتقاق نظري لطريقة PR الموسعة والتي تستخدم قيم الدالة والمشتقة. الطريقة الجديدة تشمل حقل طرائق التدرج المترافق الموسع وتمتلك خاصيتي الانحدار الحاد وخاصية التقارب الشامل تحت شروط معينة. تم الحصول على نتائج عددية متميزة عبر تطوير برامج قياسية في هذا المجال مقارنة مع طريقة Wu and Chen (2010) المتماثلة في نفس المجال.

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

1. Introduction.

Our problem is to minimize a function of n variables:

$$Min\, f(x), \text{ where } f : \mathbb{R}^n \to \mathbb{R} \quad \text{.....(1)}$$

is a smooth nonlinear function and its gradient $\nabla f(x)$ is available. At the current iterative point x_k, the Conjugate Gradient (CG) method has the following form:

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

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

where α_k is a step-length; d_k is a search direction; $g_k = \nabla f(x_k)$ and β_k is a parameter. The CG-method has played a special role in solving large-scale nonlinear optimization due to the simplicity of their iterations and their very low memory requirements, for example. Some well-known formulas for β_k are the Fletcher-Reeves (FR), Polak-Ribiére (PR), Hestenes-Stiefel (HS) methods which are given, respectively, by:

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \qquad \text{(3a)}$$

$$\beta_k^{PR} = \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}, \qquad \text{(3b)}$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \qquad \text{(3c)}$$

where

$$y_{k-1} = g_k - g_{k-1} \qquad \text{(3d)}$$

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:

$$f(x_{k-1} + \alpha_k d_k) - f(x_{k-1}) \le \delta \alpha_k g_{k-1}^T d_{k-1}, \qquad \text{(4a)}$$

$$g_k^T d_{k-1} \ge \sigma\, g_{k-1}^T d_{k-1}, \qquad \text{(4b)}$$

$$f(x_{k-1} + \alpha_k d_k) - f(x_{k-1}) \le \delta \alpha_k g_{k-1}^T d_{k-1}, \qquad \text{(5a)}$$

$$\left| g_k^T d_{k-1} \right| \le -\sigma\, g_{k-1}^T d_{k-1} \qquad \text{(5b)}$$

$$0 < \delta < 0.5 \le \sigma < 1 \qquad \text{(5c)}$$

Equations [(4a)-(4b)] and [(5a)-(5b)] are called the "Standard Wolfe" and "Strong Wolfe" conditions, respectively. It has been shown by Dai and Yuan [32] that for the FR scheme, the strong Wolfe-Powell conditions may not yield a direction of descent unless $\sigma \le 1/2$. In typical implementations of the Wolfe-Powell conditions, it is often most efficient to choose σ close to one. Hence, the constraint $\sigma \le 1/2$, needed to ensure descent, represents a significant restriction in the choice of the line search parameters. For the PR scheme, the strong Wolfe-Powell conditions may not yield a direction of descent for any choice of $\sigma \in (0,1)$. Although all these methods are equivalent in the linear case, their behaviors for general objective functions may be far different. In the PR method, if a bad direction and a tiny step from x_{k-1} to x_k are generated, the next direction d_k and the next step α_k are also likely to be poor unless a restart along the gradient direction is performed. For general functions, [19] 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. Both methods are preferred to the FR method in its numerical performance, because the methods essentially perform a restart after it encounters a bad direction. Nevertheless, [25] showed that the PR and the HS methods can cycle infinitely without approaching a solution, which implies that they do not have globally convergence.

Therefore, over the past few years, much effort has been put to find out new formulae for CG-methods such that they have not only global convergence property for general functions but also good numerical performance [21] and [26]. New kinds of nonlinear CG-methods are developed by using new conjugacy condition, such as [31]; [20]; [18] and [35]. Recently, [2] proposed a new three term preconditioned gradient memory method. Their method subsumes some other families of nonlinear preconditioned gradient memory methods as its subfamilies with Powell's restart criterion and inexact Armijo line searches. Their search direction was defined by:

$$d_k^{B\,\&\,L} = \begin{cases} -H_k g_k & if\ \ k=1 \\ -g_{k-1}+\beta_k H_{k-1}d_{k-1}-\alpha_k H_{k-2}d_{k-2} & if\ \ k>1 \end{cases} \qquad \text{.................................(6)}$$

where α_k is a step-size defined by inexact Armijo line search procedure and β_k is the conjugacy parameter. [11] introduced two versions CG-algorithm. Their search directions are defined by:

$$d_k^1 = \begin{cases} -g_k, & if\ k=0 \\ -g_k+\beta_k^{v1}d_{k-1}, & if\ \ k>0 \end{cases} \text{and}\ \beta_k^{v1}=(1-\frac{s_k^T y_k}{y_k^T y_k})\frac{(g_{k+1}^T y_k)}{s_k^T y_k} \qquad \text{..................(7)}$$

$$d_k^2 = \begin{cases} -g_k, & if\ k=0 \\ -g_k+\beta_k^{v2}d_{k-1}, & if\ \ k>0 \end{cases} \text{and}\ \beta_k^{v2}=(1-\frac{s_k^T y_k}{y_k^T y_k})\frac{(g_{k+1}^T y_k)}{d_k^T y_k}+\frac{s_k^T g_{k+1}}{d_k^T y_k} \qquad \text{........(8)}$$

More recently, [5] introduced a new three-term CG-method. An attractive property of their proposed method is that the generated directions are always descending. Besides, this property is independent of line search used and the convexity of objective function. A remarkable property of the method is that it produces a descent direction at each iteration. Motivated by the nice descent property. In order to ensure the global convergence for general functions, Dai and Liao restrict β_k to be positive, that is:

$$\beta_k^{DL+} = \max\left\{\frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, 0\right\} - t\frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}}, \quad t \ge 0. \qquad \text{...(9)}$$

The search direction of their method was given by:

$$d_{k_k}^{Bayati\,\&\,Tae} = \begin{cases} -g_0, & if\ \ k=0, \\ -g_k+\beta_k^{DL+}d_{k-1}-\mu_k\left(y_{k-1}-(2\frac{\|y_{k-1}\|^2}{s_{k-1}^T y_{k-1}})s_{k-1}\right), & if\ \ k\ge 1, \end{cases} \qquad \text{...........(10)}$$

where β_k^{DL+} is defined in (1.9), and $\mu_k = g_k^T d_{k-1} / d_{k-1}^T y_{k-1}$ and t is defined by:

$$t = 2\frac{\|y_{k-1}\|^2}{s_{k-1}^T y_{k-1}} \qquad \text{...(11)}$$

Also, [7] proposed several extended CG-methods which are combine both quadratic and non-quadratic models. Their extended search directions are defined as:

$$d_k^{one} = \begin{cases} -g_k, & if\ k=0, \\ -\theta_k^{EPR}g_k+\beta_k^{PR}d_{k-1}, & if\ \ k>0, \end{cases} \qquad \text{...(12)}$$

$$\beta_k^{PR} = \frac{g_k^T z_{k-1}}{g_{k-1}^T g_{k-1}},\ \theta_k^{EPR} = 1+\rho_k\beta_k^{PR}\frac{g_k^T d_{k-1}}{\|g_k\|^2}-\psi\frac{g_k^T d_{k-1}}{g_{k-1}^T g_{k-1}},\ z_{k-1}=y_{k-1}+\varepsilon_1 s_{k-1} \qquad \text{........(13)}$$

$$d_k^{Two} = \begin{cases} -g_k & if\ \ k=0 \\ -g_k+\beta_k^{PR}d_{k-1}-\theta_k^{(1)}y_{k-1} & if\ \ k>0 \end{cases} \qquad \text{..(14)}$$

$$\beta_k^{PR} = \frac{g_k^T z_{k-1}}{g_{k-1}^T g_{k-1}},\ \theta_k^{(1)} = 1+\rho_k\beta_k^{PR}\frac{g_k^T d_{k-1}}{g_k^T g_k}-\psi\frac{g_k^T d_{k-1}}{g_{k-1}^T g_{k-1}},\ z_{k-1}=y_{k-1}+\varepsilon_1 s_{k-1} \qquad \text{......(15)}$$

$$d_k^{Three} = \begin{cases} -g_k & if \;\; k=0 \\ -g_k + \beta_k^{PR} d_{k-1} - \theta_k^{(2)} y_{k-1} & if \;\; k>0 \end{cases} \quad \ldots\ldots (16)$$

$$\beta_k^{PR} = \frac{g_k^T z_{k-1}}{g_{k-1}^T g_{k-1}}, \theta_k^{(2)} = \psi \rho_k \frac{\|g_k\|^2}{g_k^T y_{k-1}} \frac{g_k^T d_{k-1}}{g_{k-1}^T g_{k-1}} - \frac{g_k^T d_{k-1}}{g_{k-1}^T g_{k-1}},\; z_{k-1} = y_{k-1} + \varepsilon_1 s_{k-1} \ldots\ldots (17a)$$

$$\rho_k = \frac{(s_{k-1}^T g_{k-1}/2)^2}{(f_k - f_{k-1})^2} \quad \ldots\ldots (17b)$$

Finally, [3] considered a modified three term CG-method defined as :

$$d_{k+1} = -g_{k+1} + \beta_k d_k - \gamma_k y_k \quad \ldots\ldots (18)$$

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k} \text{ and } \quad \gamma_k^{\mathrm{mod}ified} = \frac{g_{k+1}^T d_k}{(1-u) g_k^T g_k + u\,|d_k^T g_k|}, \quad \ldots\ldots (19)$$

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

where $u \in (0,1]$ is a constant. Obviously, $\gamma_k^{\mathrm{mod}ified} = \gamma_k^{sc1}$ for u approaches 0, and $\gamma_k^{\mathrm{mod}ified} = \gamma_k^{sc2}$ for $u = 1$. The search direction generated by this method at each iteration satisfies the descent condition. The optimal value of the parameter u is given in (1.20).

In this paper, we have proposed a new formula β_k^{New} for β_k applying the rational non-quadratic model and Perry's conjugacy condition [1].

$$d_k^T y_{k-1} = -(H_k g_k)^T y_{k-1} = -g_k^T s_{k-1} \quad \ldots\ldots (21)$$

where H_k is an approximation to the inverse Hessian and $s_{k-1} = x_k - x_{k-1}$. They respectively can be seen as the modifications of the method HS and PR. In comparison with classic CG methods, the decrease of the objective function value is contained in the two new formulae. Moreover, β_k^{New} 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 [14] to implement the non-quadratic rational model which satisfies the sufficient descent property and the standard Wolfe-Powel conditions. 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.

2. Materials and Methods.

2.1 Extended CG-Methods For Non-Quadratic Models.

Many attempts have been made to investigate more general function than the quadratic one as a basis for the CG-methods. Over years, various authors have published works in this area, and a large variety of methods have been derived to solve this problem for many sorts of objective functions. The CG-methods discussed so far assume a local quadratic representation of the objective function. However, quadratic models may not always be adequate to incorporate all the information which might be needed to represent the objective function successfully. Also in problems where the quadratic representations is not good. When we are remote from such a region, a non-quadratic model may better represent the objective function and that leads speculation on a better way to choose a type of a non-quadratic model.

2.2 Extended Rational CG-Method. [8]

The CG-methods so far discussed is a local quadratic representation of the objective function. In problems when the quadratic representation is not good, or when we are remote from such a region, quadratic function $f(q(x))$, where 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. In this paper, Al-Bayati's 1993 extended CG-method which is invariant to nonlinear scaling of quadratic rational functions is proposed and combined with the standard conjugacy condition of [14] to increase the efficiency of this type of CG-methods. There is some precedent for this approach, if $q(x)$ is quadratic function then a function f is defined as nonlinear scaling of $q(x)$ if the invariancy property to nonlinear scaling by [17] holds:

$$\min f(x) = f(q(x)) \quad \ldots\ldots(22)$$

$$\text{where } \frac{df}{dq} = f > 0 \text{ and } q > 0 \quad \ldots\ldots(23)$$

has been considered by [15]. Al-Bayati introduced several non-quadratic rational models; see for example Boland theorem [30]; [8]; [4]; [10] and [9]. Al-Bayati's, 1993 non-quadratic model to be investigated here, 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:

$$f[q(x)] = \frac{\varepsilon_1 q(x)}{1-\varepsilon_2 q(x)};\ \varepsilon_2 < 0, \quad \varepsilon_1 > 0 \quad \ldots\ldots(24)$$

Where

$$q(x) = \frac{1}{2}(x - x_{\min})^T Q(x - x_{\min}) \quad \ldots\ldots(25)$$

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 (23).

2.3 Outline of Al-Bayati's Extended Rational CG-Model.

Step 1: Compute a, b and c using $a = \frac{s^T g}{2}$; $b = w - a$ and $c = wa - (w-a)f$.

Step 2: If $|b| \le \delta$ *or* $|c| \le \delta$; set $\rho = 1$ and go to **Step 4**.

Step 3: Compute $\rho_k = \frac{(s_{k-1}^T g_{k-1}/2)^2}{(f_k - f_{k-1})^2}$ ………………………………………….....(26)

Step 4: Compute $d_k = -g_k + \frac{g_k^T(\rho_k g_k - g_{k-1})}{\|g_{k-1}\|^2} d_{k-1}$ …………………………….....(27)

Where δ is a suitable tolerance value; say $\delta = 1\times 10^{-11}$ .This direction d_k is then used instead of the direction used in the standard CG-formula and since the model satisfies conditions (23), the resulting algorithm has finite convergence on model (24). Recently, [6] introduced a new extended CG-method for which their search directions are defined by:

$$d_k^{B\&A} = \begin{cases} -g_k, & if\ k = 0, \\ -\theta_k^Z g_k + \overline{\beta}_k^{PR} d_{k-1}, & if\ k > 0, \end{cases} \quad \text{……...……………………...(28)}$$

$$\theta_k^Z = 1 + \rho_k \beta_k^Z \frac{g_k^T d_{k-1}}{\|g_k\|^2} \quad \text{and} \quad \overline{\beta}_k = \frac{g_k^T(\rho_k g_k - g_{k-1})}{\|g_{k-1}\|^2} \quad \text{…………………..…...(29)}$$

$$\beta_k^Z = \max\ (0, \min\ \beta^{PR}, \beta^{FR}\ ) \quad \text{…..…………...………………..……...(30)}$$

ρ_k is a scalar defined in (26).

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; $\beta_k^1,\ \beta_k^2,\ \beta_k^3\ and\ \beta_k^4$ respectively by making use of the Powell's restarting criterion and the Armijo-type line search defined by:

$$\beta_k^1 = \beta^{HS} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{d_{k-1}^T y_{k-1}} \quad \text{……………………...………………..…...…….(31)}$$

$$\beta_k^2 = \beta^{PR} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{\|g_{k-1}\|^2} \quad \text{……………….................……….....……....…...(32)}$$

$$\beta_k^3 = \max\{0, \beta^{PR}\} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{\|g_{k-1}\|^2} \quad \text{……………………..…………….…...…...(33)}$$

and for $\delta \le \mu < 1-\delta$ *and* $t > 0$: where the two constants are defined by :

$$A_k = \mu\|g_k\|^2 - (g_k)^T g_{k-1} + 2(f_{k-1} - f_k) + (s_{k-1})^T g_{k-1}, \qquad (34)$$

$$B_k = \mu\|g_k\|^2 - (g_k)^T g_{k-1} + 2(f_{k-1} - f_k) + t(s_{k-1})^T g_{k-1}, \qquad (35)$$

such that:

$$2(f_{k-1} - f_k) + t(s_{k-1})^T g_{k-1} \le 0, \qquad (36)$$

hold when $A_k \ge 0$ *and* $B_k \ge 0$

$$\beta_k^4 = \begin{cases} \dfrac{B_k}{|(g_k)^T d_{k-1}| + \|g_{k-1}\|^2}, & A_k \ge 0, \\ \dfrac{\mu\|g_k\|^2}{|(g_k)^T d_{k-1}| + \|g_{k-1}\|^2}, & A_k < 0 \end{cases} \qquad (37)$$

They proved that all the above CG-methods satisfy the sufficient descent condition and have the global convergence property.

4. A New Extended CG-Method.

Consider the following quadratic model we proceed as in [14]:

$$f(\mathrm{x}) = \frac{1}{2} x^T A x + b^T x + c \qquad (38)$$

where $A \in R^{n \times n}$ is a symmetric positive definite matrix, $b \in R^n$ and $c \in R$. Then $y_{k-1} = g_k - g_{k-1} = A s_{k-1}$. Substituting $\mathrm{x}_k = \mathrm{x}_{k-1} + \mathrm{s}_{k-1}$ into (38), we obtain:

$$f_k = \frac{1}{2} x_k^T A x_k + b^T x_k + c$$

$$f_k = \frac{1}{2}(x_{k-1} + s_{k-1})^T A (x_{k-1} + s_{k-1}) + b^T (x_{k-1} + s_{k-1}) + c$$

$$= \frac{1}{2} x_{k-1}^T A x_{k-1} + \frac{1}{2} s_{k-1}^T A s_{k-1} + b^T x_{k-1} + b^T s_{k-1} + c$$

$$f_k = f_{k-1} + \frac{1}{2} s_{k-1}^T A s_{k-1} + b^T s_{k-1} \qquad (39)$$

From Taylor series b=g we get:

$$f_k = f_{k-1} + \frac{1}{2} s_{k-1}^T A s_{k-1} + (g_{k-1})^T s_{k-1}$$

$$-g_{k-1}^T s_{k-1} = f_{k-1} - f_k + \frac{1}{2} s_{k-1}^T A s_{k-1}$$

Since $A s_{k-1} = g_k - g_{k-1}$, we have

$$-g_{k-1}^T s_{k-1} = f_{k-1} - f_k + \frac{1}{2} s_{k-1}^T (g_k - g_{k-1})$$

$$-\frac{1}{2} g_{k-1}^T s_{k-1} = f_{k-1} - f_k + \frac{1}{2} s_{k-1}^T g_k$$

Multiplying both sides by 2

$$-g_{k-1}^T s_{k-1} = 2(f_{k-1} - f_k) + s_{k-1}^T g_k \qquad (40)$$

It follows from Perry's conjugacy conditions (21) and (40) that

$$-g_{k-1}^T s_{k-1} = 2(f_{k-1} - f_k) - d_k^T y_{k-1} \quad \text{.....(41)}$$

Additionally, $d_k = -g_k + \beta_k d_{k-1}$ and (41) imply that:

$$-g_{k-1}^T s_{k-1} = 2(f_{k-1} - f_k) + g_k^T y_{k-1} - \beta_k d_{k-1}^T y_{k-1}$$

Which yields:

$$\beta_k = \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1} + g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}} \quad \text{.....(42a)}$$

$$= \beta^{HS} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{d_{k-1}^T y_{k-1}} \quad \text{.....(42b)}$$

If exact line search., i.e. $g_k^T d_{k-1} = 0$ and $d_{k-1} = -g_{k-1}$ is used in (42a) yields:

$$\beta_k = \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1} + g_k^T y_{k-1}}{\|g_{k-1}\|^2} \quad \text{.....(43a)}$$

$$= \beta^{PR} + \frac{2(f_{k-1} - f_k) + g_{k-1}^T s_{k-1}}{\|g_{k-1}\|^2} \quad \text{.....(43b)}$$

For more details see [14].

From Section (2) we can get ρ_k using (26) to use in the new extended CG method whose conjugacy parameter is defined by β_k^{New} such that:

$$\beta_k^{New} = \beta^{PR} + \frac{2(f_{k-1} - f_k) + \rho_k g_{k-1}^T s_{k-1}}{\|g_{k-1}\|^2} \quad \text{.....(44)}$$

Note that the scalar ρ_k may be rewritten as:

$$\rho_k = \frac{(s_{k-1}^T g_{k-1})^2}{4(f_{k-1} - f_k)^2} \quad \text{.....(45)}$$

By using (45), equation (44) becomes:

$$\beta_k^{New} = \beta_k^{PR} + \frac{2(f_{k-1} - f_k) + \dfrac{(s_{k-1}^T g_{k-1})^2}{4(f_{k-1} - f_k)^2}(g_{k-1}^T s_{k-1})}{\|g_{k-1}\|^2} \quad \text{.....(46)}$$

$$\beta_k^{New} = \beta_k^{PR} + \frac{8(f_{k-1} - f_k)^3 + (g_{k-1}^T s_{k-1})^3}{4(f_{k-1} - f_k)^2 \|g_{k-1}\|^2} \quad \text{.....(47)}$$

4.1 Outline of The 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: Set $x_{k-1} = x_k + \alpha_k d_k$**;** α_k is obtained by WP-procedure.

Step 4: 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 $d_{k+1} = -g_{k+1} + \beta_k^{New} d_k$ (β_k^{New} is defined in (47)), go to **Step 2.**

Step 5: If $\|g_{k+1}\| < \varepsilon$, stop else set k=k+1 go to **Step 3**.

4. 2 Theoretical Properties For The New Extended CG-Method.

In this section, we focus on the convergence behavior on the β_k^{New} method with exact line searches. Hence, we make the following basic assumptions on the objective function.

4. 3 Assumption.

f is bounded below in the level set $L_{x_0} = \{x \in R^n | f(x) \le f(x_0)\}$; 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:

$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$ for all x, y $\in L_{x_0}$ ……………..……………...……….(48)

4.4 Lemma

Consider a general CG-method, and suppose that $0 < \gamma \le \|g_k\| \le \bar{\gamma}$ holds. We call a method has **Lemma 4.4** if there exist two constants b>1 and p>0 such that for all k, $|\beta_k| \le b$ and

$\|s_k\| \le p \Rightarrow |\beta_k| \le \dfrac{1}{2b}$ …………………………..…………………………………(49)

4.5 Lemma (Zoutendijk condition).

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 (4 and 5). Then we have that:

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

4.6 Theorem

Suppose that **Assumption 4.3** holds. Consider the new extended CG-method defined in (47) with β_k^{New}, if α_k is obtained by an exact line search and then: $\liminf_{k\to\infty} \|g_k\| = 0$

Proof:

We now prove the theorem by contradiction and assume that there exists some constants γ > 0 such that $\|g_k\| \ge \gamma$ for all $k \ge 0$. The compactness of the level set L_{x_0} implies that there exists a constant $\bar{\gamma}$>0 such that $\|g_k\| \le \bar{\gamma}$. Since $\|s_k\| \to 0$, we know that there is a $\bar{k}$, for all $k > \bar{k}$ *such that* $k < \|s_k\| \le p$, where p is the same as in **Lemma 4.4**. Then, for all $k > \bar{k}$, we have:

$\|d_k\| \le \|g_k\| + |\beta_k^{New}| \, \|d_{k-1}\|$ …………………………...………………………….....(50)

$$\leq \bar{\gamma} + \frac{1}{2b}(\bar{\gamma} + \frac{1}{2b}\|d_{k-2}\|)$$

$$= (1 + \frac{1}{2b})\bar{\gamma} + \frac{1}{2b^2}(\|d_{k-2}\|)$$

$$\leq \ldots \leq (\frac{1}{1 - \frac{1}{2b}})\ \bar{\gamma} + \frac{1}{2b^{k-\bar{k}}}(\|d_{\bar{k}}\|)$$

$$\leq (\frac{2b}{2b-1})\bar{\gamma} + \|d_{\bar{k}}\| \equiv \bar{\eta} \quad \ldots\ldots(51)$$

Furthermore, we know

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{\gamma^4}{\|d_k\|^2} \quad \ldots\ldots(52)$$

we know using **Lemma 4.5** together with (52) ,yields

$$\sum_{k=0}^{\infty} \frac{\gamma^4}{\|d_k\|^2} < \infty \quad \ldots\ldots(53)$$

Which contradictions (51).
Therefore, we conclude the truth of the theorem.

4.7 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 Wolfe-Powell conditions (4) and (5) and d_k satisfies the new β_k^{New} CG-method, then the new extended method has sufficient descent directions i.e.,

$$d_k^T g_k \leq -c\|g_k\|^2; \quad c > 0 \quad \ldots\ldots(54)$$

Proof:

For initial direction we have:

$$d_1 = -g_1 \Rightarrow d_1^T g_1 = -\|g_1\|^2 \leq 0 \quad \ldots\ldots(55)$$

which satisfies (54). Now let the theorem be true for all $k-1$, i.e.

$$d_{k-1} = -g_{k-1} \Rightarrow d_{k-1}^T g_{k-1} = -\|g_{k-1}\|^2 \leq 0 \quad \ldots\ldots(56)$$

Multiplying the search direction of (47) by g_k^T yields:

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

Using Wolfe-Powell conditions (4) and (5) we have:

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

$$d_k^T g_k \leq -\|g_k\|^2 + (\frac{s_{k-1}^T y_{k-1}}{\|g_{k-1}\|^2})(\|g_k\|^2) + (\frac{(1-8\delta^3)}{4\delta^2 \|g_{k-1}\|^2})(s_{k-1}^T g_k)(g_{k-1}^T s_{k-1}) \quad \ldots\ldots(57)$$

If exact line searches are used then (57) becomes using (56):

$$d_k^T g_k \le -\|g_k\|^2 - (\frac{s_{k-1}^T g_{k-1}}{\|g_{k-1}\|^2})(\|g_k\|^2)$$
$$= -\|g_k\|^2 + \alpha_{k-1}\|g_k\|^2$$
$$= -(1-\alpha_{k-1})\|g_k\|^2$$
$$= -c\|g_k\|^2 \quad \ldots\ldots(58)$$

Hence, for ELS, the search directions are sufficiently descent since $c = (1-\alpha_{k-1}) > 0$.

For inexact line searches we have:

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:

$$(\nabla f(x) - \nabla f(y))^T (x-y) \ge \eta \ \|x-y\|^2 \text{ for all x, y} \in L_{x_0} \quad \ldots\ldots(59)$$

Or equivalently:

$$y_{k-1}^T s_{k-1} \ge \eta\|s_{k-1}\|^2 \quad and \quad \eta\|s_{k-1}\|^2 \le y_{k-1}^T s_{k-1} \le L\|s_{k-1}\|^2 \quad \ldots\ldots(60)$$

$$d_k^T g_k \le -\|g_k\|^2 + L(\alpha_{k-1})^2\|g_k\|^2 - \alpha_{k-1}(\frac{(1-8\delta^3)}{4\delta^2})(s_{k-1}^T g_k) \quad \ldots\ldots(61)$$

From Powel restarting criterion we have:

$$g_k^T g_{k-1} > \psi\|g_k\|^2; \quad \psi \in (0,1) \quad \ldots\ldots(62)$$

$$g_k^T s_{k-1} = \alpha_{k-1} \ g_k^T d_{k-1} = -\alpha_{k-1} \ g_k^T g_{k-1} \quad \ldots\ldots(63)$$

Using (62) and (63) in (61):

$$d_k^T g_k \le -\|g_k\|^2 + L(\alpha_{k-1})^2\|g_k\|^2 + \psi(\alpha_{k-1})^2 (\frac{(1-8\delta^3)}{4\delta^2})\|g_k\|^2 \quad \ldots\ldots(64)$$

$$(d_k^T g_k)/(\|g_k\|^2) \le -1 + L(\alpha_{k-1})^2 + \psi(\alpha_{k-1})^2 (\frac{(1-8\delta^3)}{4\delta^2}) \quad \ldots\ldots(65)$$

$$(d_k^T g_k)/(\|g_k\|^2) \le - \ (1 - \psi(\alpha_{k-1})^2 (\frac{(1-8\delta^3)}{4\delta^2}) - L(\alpha_{k-1})^2) \quad \ldots\ldots(66)$$

$$(d_k^T g_k)/(\|g_k\|^2) \le - \ c, \quad c > 0 \quad for \quad 0 < \delta < 0.5; \quad 0 < \alpha, \psi, L < 1 \quad \ldots\ldots(67)$$

Thus our new proposed extended CG-method has also sufficient descent directions using inexact line searches under the condition that Powell restarting condition must be used. Therefore, the method has a global convergent property by satisfying the conditions of Zoutendijk theorem [19].

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 35 nonlinear unconstrained problems that have second derivatives available. These test problems are contributed in CUTE and their details are given in the

Appendix. Our numerical results are divided in three branches according to the numerical experiments with their number of variables:

1- 10 numerical experiments with n = 100, 200, ..., 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 [14] using the same set of test problems. All these methods terminate when the following stopping criterion is met:

$\|g_k\| \le 10^{-6}$ ………………………………………………………………………….(68)

Tables 5.1, 5.2 and 5.3 compare some numerical results for the modified PRCG method due to Wu & Chen and the 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.

Tables 5.4, 5.5 and 5.6 compare the percentage performance of the 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.

It is clear from **Table (5.4)** that taking, over all, the tools as a 100% for the Wu & Chen PRCG method, the New Extended PRCG method has an improvement, in about (12.3%) NOI; (11.5%) NOFG and (2.5%) TIME, also from **Table (5.5)** that taking, over all, the tools for PRCG method has an improvement, in about (6.1%) NOI; (5.3%) NOFG and (3.4%) TIME. It is clear from **Table (5.6)** that taking, over all, the tools for PRCG method has an improvement, in about (12.3%) NOI; (11.3%) NOFG and (2.2%) TIME. These results indicate that new extended PRCG method is in general is the best.

Table (5.1)
COMPARISON BETWEEN THE NEW AND (WU & CHEN) METHODS FOR THE TOTAL OF (35) PROBLEMS WITH n= 100, 200, … ,1000

Prob.	**Wu & Chen/2010** NOI/NOFG/TIME	**New Extended PRCG** NOI/NOFG/TIME
1	1709/2017/1.30	1700/2008/1.36
2	219/412/0.03	219/412/0.03
3	75/96/0.03	75/96/0.02
4	1592/1724/0.99	1592/1724/1.05
5	1044/1137/0.16	1044/1137/0.15
6	331/358/0.28	349/377/0.29
7	10506/10624/1.10	6831/6938/0.70
8	143/182/0.20	161/196/0.21
9	319/453/0.04	319/453/0.03
10	205/282/0.11	205/282/0.09
11	561/677/0.14	558/674/0.17
12	205/317/0.03	205/317/0.04
13	32/64/0.01	32/64/0.02
14	1314/1400/0.18	1314/1400/0.19
15	4179/4259/0.76	4179/4259/0.74
16	126/147/0.03	126/147/0.04
17	90/118/0.03	90/118/0.03
18	109/133/0.03	109/133/0.04
19	1279/1368/0.21	1279/1368/0.21
20	75/96/0.04	75/96/0.01
21	947/1109/0.10	947/1080/0.10
22	645/678/0.26	645/678/0.25
23	1190/1326/0.49	1182/1318/0.50
24	137/211/0.01	137/211/0.00
25	251/330/0.04	251/330/0.03
26	860/934/0.27	875/949/0.30
27	144/194/0.00	149/191/0.03
28	80/160/0.05	80/160/0.06
29	85/105/0.07	85/105/0.07
30	44/76/0.01	44/76/0.04
31	206/258/0.10	206/258/0.10
32	1144/1248/0.14	1089/1199/0.19
33	27/77/0.00	27/77/0.00
34	80/110/0.03	80/110/0.02
35	211/307/0.02	200/280/0.00
Total	**30164/32987/7.29**	**26459/29221/7.11**

Table (5.2)

COMPARISON BETWEEN THE NEW AND (WU & CHEN) METHODS FOR THE TOTAL OF (35) PROBLEMS WITH n = 100, 300,500,700, 900

Prob.	**Wu & Chen/2010** NOI /NOFG/TIME	**New Extended PRCG** NOI/NOFG/TIME
1	939/1102/0.68	930/1093/0.69
2	113/208/0.02	113/208/0.01
3	37/47/0.02	37/47/0.00
4	775/848/0.45	775/848/0.46
5	513/559/0.06	513/559/0.08
6	170/184/0.13	170/184/0.14
7	4591/4644/0.38	3622/3735/0.23
8	81/98/0.09	81/98/0.09
9	168/228/0.02	168/228/0.03
10	93/134/0.04	93/134/0.05
11	223/293/0.05	223/293/0.06
12	89/140/0.00	89/140/0.01
13	18/35/0.00	18/35/0.00
14	663/701/0.11	663/701/0.09
15	3307/3345/0.53	3319/3357/0.53
16	65/75/0.03	65/75/0.06
17	45/60/0.00	45/60/0.00
18	54/65/0.02	54/65/0.02
19	639/683/0.10	639/683/0.09
20	37/47/0.01	37/47/0.02
21	455/494/0.06	455/494/0.05
22	373/403/0.14	374/404/0.14
23	432/495/0.19	424/487/0.19
24	68/87/0.00	68/87/0.00
25	119/174/0.02	119/174/0.03
26	456/487/0.13	510/549/0.14
27	69/100/0.00	74/99/0.00
28	40/81/0.03	40/80/0.01
29	43/53/0.03	43/53/0.03
30	22/38/0.02	22/38/0.02
31	95/122/0.05	95/122/0.05
32	623/682/0.08	595/657/0.06
33	12/35/0.00	14/42/0.00
34	39/55/0.02	39/55/0.02
35	104/152/0.02	100/140/0.01
Total	**15570/16954/3.53**	**14626/16067/3.41**

Table (5.3)
COMPARISON BETWEEN THE NEW AND (WU & CHEN) METHODS FOR THE TOTAL OF (35) PROBLEMS WITH n= 100, 400,700, 1000

Prob.	Wu & Chen/2010 NOI /NOFG/TIME	New Extended PRCG NOI/NOFG/TIME
1	895/1011/0.67	886/1002/0.65
2	87/164/0.02	87/164/0.02
3	29/38/0.01	29/38/0.01
4	631/685/0.44	631/685/0.43
5	406/443/0.06	406/443/0.06
6	135/146/0.13	135/146/0.11
7	4778/4809/0.55	3057/3110/0.36
8	53/68/0.07	66/79/0.06
9	133/191/0.01	141/183/0.02
10	84/119/0.03	84/119/0.05
11	186/230/0.04	227/279/0.08
12	98/138/0.02	98/138/0.01
13	14/28/0.00	14/28/0.00
14	513/541/0.09	513/541/0.07
15	2156/2188/0.40	2156/2188/0.46
16	53/61/0.03	53/61/0.03
17	36/48/0.02	36/48/0.01
18	43/52/0.01	43/52/0.03
19	510/545/0.07	510/545/0.08
20	29/38/0.01	29/38/0.01
21	379/412/0.04	379/412/0.07
22	334/355/0.12	334/355/0.12
23	386/428/0.18	382/424/0.15
24	56/84/0.02	56/84/0.00
25	98/124/0.01	98/124/0.02
26	265/281/0.08	327/351/0.12
27	61/84/0.01	61/84/0.02
28	32/64/0.01	32/64/0.03
29	35/43/0.03	35/43/0.03
30	18/30/0.00	18/30/0.01
31	76/95/0.03	76/95/0.03
32	393/437/0.05	397/441/0.06
33	9/27/0.00	9/27/0.00
34	32/44/0.01	32/44/0.00
35	82/120/0.02	80/112/0.01
Total	**13125/14171/3.29**	**11517/12577/3.22**

Table (5.4)
PERCENTAGE PERFORMANCE OF TABLE (5.1)

TOOLS	**WU & CHEN (2010)**	**NEW**
NOI	100%	87.7%
NOFG	100%	88.5%
TIME	100%	97.5%

Table (5.5)
PERCENTAGE PERFORMANCE OF TABLE (5.2)

TOOLS	**WU & CHEN (2010)**	**NEW**
NOI	100%	93.9%
NOFG	100%	94.7%
TIME	100%	96.6%

Table (56)
PERCENTAGE PERFORMANCE OF TABLE (5.3)

TOOLS	**WU & CHEN (2010)**	**NEW**
NOI	100%	87.7%
NOFG	100%	88.7%
TIME	100%	97.8%

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] A. Perry, (A modified conjugate gradient algorithm) Operations Research, **26**, 1073–1078, 1978.

[2] A. Y. Al-Bayati and I. S. Latif, (A new three-term preconditioned gradient memory algorithm for nonlinear optimization problems) American Journal of Applied Science, Science Publication, New York, **4**, 81-87, 2008.

[3] A. Y. Al-Bayati and B. A. Hassan, (A Modified three-term Conjugate Gradient Method) Journal of Association for the Advancement of Modeling and Simulation, (AMSE), Spain-France**, (Accepted for publication , July, 2011**).

[4] A. Y. Al-Bayati and G. M., Al- Naemi, (New extended CG-method for nonlinear optimization) J. of Mu'tah, Jordan, **10**, 69-87, 1995.

[5] A. Y. Al-Bayati and H. W. Altae, (A new three-term non-linear conjugate gradient method for unconstrained optimization) Canadian Journal on Science & Engineering Mathematics, Canada, **1**, 108-124, 2010.

[6] A. Y. Al-Bayati and R. M. Abdullah, (An Extended CG-Method with a New Non-Monotone Line Search Procedure for Constrained Optimization) Australian Journal of Basic and Applied Sciences, INSInet Publication Australia, **4**(12), 6069-6085, 2010.

[7] A. Y. Al-Bayati and Sh. K. Hassan, (New extended Polak-Ribiere CG-Methods Non-Linear For Unconstrained Optimization) Canadian Journal on Science & Engineering Mathematics, SEM-1001-013, AM Publishers Corporation, Canada, **2**(1), 9-18, 2011.

[8] A. Y. Al-Bayati, (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.

[9] A. Y. Al-Bayati, (New Extended CG methods for combined non-quadratic models in nonlinear optimization) J. of Al-Yarmouk, Jordan, **16**, 347-366, 2007.

[10] A. Y. Al-Bayati, (New generalized CG-methods for the non-quadratic model in unconstrained optimizations) J. Al- Yarmouk, **10**, 9-25, 2001.

[11] A. Y. Al-Bayati; A. J. Salim. and K. K. Abbo, (Two versions of CG-algorithms based on conjugcy conditions for unconstrained optimization) American Journal of Economic and Business administration, Science Publication, USA, **1**(2), 97-104, 2009.

[12] C. Gilbert and J. Nocedal, (Global convergence properties of conjugate gradient methods for optimization). SIAM Journal on Optimization, **2** , 21–42, 1992.

[13] C. Y. Wu and G. Q. Chen, (A smoothing conjugate gradient algorithm for nonlinear complementarily problems) Journal of Systems Science and Systems Engineering, **17**, 460–472, 2008.

[14] C. Y. Wu and G. Q. Chen, (New type of conjugate gradient algorithm for unconstrained optimization problems) Journal of Systems Engineering and Electronics, **21**(6), 1000-1007, 2010.

[15] D. Goldfarb, (Variable-Metric and Conjugate-Direction Method in Unconstrained Optimization) Resent Developments, ACM, Proceedings, National Meeting, Boston, Massachusetts , 496-506,1972.

[16] D. H. Li; Y. Y Nie and J. P. Zeng, et al. (Conjugate gradient method for the linear complementarily problem with S-matrix) Mathematical and Computer Modeling, **48 ,** 918–928. 2008.

[17] E. Spedicato, (A Variable Metric method for function minimization derived from invariancy to nonlinear scaling) Journal of optimization theory and applications, **20** , 315-329.1976.

[18] G. Y. Li; C. M. Tang and Z. X. Wei, (New conjugacy condition and related new conjugate gradient methods for unconstrained optimization) Journal of Computational and Applied Mathematics, **202**, 523–539, 2007.
[19] G. Zoutendijk, (Nonlinear programming, computational methods in Integer and nonlinear programming) North-Holland Amsterdam , 37–86, 1970.
[20] H. Yabe and M. Takano, (Global convergence properties of nonlinear conjugate gradient methods with modified secant condition) Computational Optimization and Applications, **28** , 203–225, 2004.
[21] H. Yabe and N. Sakaiwa, (A new nonlinear conjugate gradient method for unconstrained optimization) Journal of the Operations Research, **48** , 284–296, 2005.
[22] L. Zhang , W. J. Zhou and D. H. Li, (Global convergence of a modified Fletcher-Reeves conjugate method with Armijo-type line search) Numerische Mathematik , **104** , 561–572, 2006.
[23] L. Zhang and W. J. Zhou, (Two descent hybrid conjugate gradient methods for optimization) Journal of Computational and Applied Mathematics, **216**, 251–264, 2008.
[24] L. Zhang; W. J. Zhou and D. H. Li, (A descent modified Polak-Ribiere Polyak conjugate gradient method and its global convergence) IMA Journal of Numerical. Analysis, **26**, 629–640, 2006.
[25] M. J. D. Powell, (Non-convex minimization calculations and the conjugate gradient method) Lecture Notes in Mathematics, 122–141, 1984.
[26] N. Andrei. and D. Yuan, (Conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization) Applied Mathematics Letters, **21**, 165–171. 2008.
[27] S. Burer., (Semi-definite programming in the space of partial positive Semi-definite matrices), SIAM Journal on Optimization, **14** , 139–172. 2003.
[28] S. Lalos and K. Berberidis, (An efficient conjugate gradient method in the frequency domain application to channel equalization) Signal Processing, **88** , 99–116, 2008.
[29] T. Johansson and L. Daniel, (A variational conjugate gradient method for determining the fluid velocity of a slow viscous flow) Applicable Analysis, **85**, 1327–1341, 2006.
[30] W. R. Boland; E. R. Kamgnia and J. S. Kowalik, (A conjugate gradient optimization method invariant to nonlinear scaling) Journal of optimization theory and applications, **27**, 221-230, 1979.
[31] Y. H. Dai and L. Z. Liao, (New conjugacy conditions and related nonlinear conjugate gradient methods) Applied Mathematics and Optimization, **43** , 87–101, 2001.
[32] Y. H. Dai and Y. Yuan, (Nonlinear Conjugate Gradient Methods) Shanghai Science and Technology Publisher, Shanghai , 2000.
[33] Y. Yuan, (Analysis on the conjugate gradient method) Optimization Methods and Software, **2** , 19–29.J, 1993.
[34] Z. Wei; G. Li and L. Qi, (New quasi-Newton methods for unconstrained optimization problems) Applied Mathematics and Computation, **175**, 1156–1188, 2006.
[35] Z. Wei; G. Yu and G. Yuan, et. al. (The super linear convergence of a modified BFGS-type method for unconstrained optimization) Computational Optimization and Application, **29**, 315–332 , 2004.