Iraqi Journal of Statistical Science (21) 2012
pp[15- 33]

New Six and Seven-Parameter Family for CG-Methods

Baan A. Hasen[1] **Nada F. Haasan[2]**

Abstract

In this paper, we have investigated two new conjugacy coefficients for Conjugate Gradient (CG) methods for solving nonlinear unconstrained optimization. First six-parameter family which depends on six parameters $(\lambda, \mu, \omega, \delta, \gamma, \varphi)$, second seven-parameter family which depends on seven parameters $(\lambda, \mu, \omega, \delta, \gamma, \varphi, \psi)$. The clear feature of these two families is obtaining eleven or twelve different forms of original conjugacy coefficient from six and seven-parameter family, respectively. The global convergence analysis of these two families had been proved. Numerical results for these two families on some well-known test functions were reported and compared with the result of some well known formula for β_k in CG-methods.

معاملان جديدان للعائلة السادسة والسابعة لطرائق الاتجاه المترافق

الملخص

تم في هذا البحث التقصي عن معاملين جديدين للاتجاه المترافق β_k لحل مسائل الأمثلية اللاخطية غير المقيدة. الأول معامل العائلة السادسة (β^{six}) وذلك من خلال اعتماده على ست قيم هي $(\lambda, \delta, \gamma, \mu, \omega, \varphi)$, والثاني معامل العائلة السابعة β^{seven} وذلك من خلال اعتماده على سبع قيم هي $(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$. الخاصية البارزة في كلا المعاملين هو انه من خلال التحكم بهذه القيم يمكن الحصول على أحد عشر معاملا أو اثني عشر معاملاً من معاملات الاتجاه المترافق الأصلية والتي تم اكتشافها من علماء معروفين على التوالي. تم إثبات التقارب الشامل لكلا الصيغتين ومقارنة النتائج العددية مع نتائج بعض الصيغ المشهورة لمعامل الاتجاه المترافق β_k لطرائق الاتجاه المترافق باستخدام عدد من دوال الاختبار.

---

[1]Prof/College of Computers Sciences and Math.\University of Mosul
[2]Lecturer / College of Computers Sciences and Math.\University of Mosul
Received:30/6 /2011 ____________________Accepted: 23 /10 / 2011

1. Introduction

We consider the following unconstrained minimization problem:

Minimize f(x) …………………(1)

Where $f:R^n \rightarrow R$ is smooth and its gradient $g(x)=\nabla f(x)$ is available. Conjugate gradient methods are very efficient for solving large-scale unconstrained optimization problems (1). The iterates of conjugate gradient methods are obtained by

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

$$d_k = \begin{cases} -g_k & for \quad k=1 \\ -g_k + \beta_k d_k & for \quad k \geq 2 \end{cases}$$

(3)

Where step size α_k is positive, which is computed by carrying out some line search, usually obtained by the strong Wolfe line search

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \quad (4)$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k \quad (5)$$

Where $0 < \delta < \sigma < 1$

Since 1952, there have been many formulas for the scalar β_k (conjugancey coefficient)

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}$$ (Hestenes-Stiefel, 1952)

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$ (Fletcher and Reeves,1964)

$$\beta_k^{PR} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}$$ (Polak-Ribiere ,1969)

$$\beta_k^{DX} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}$$ (Dixon ,1975)

$$\beta_k^{Perry} = \frac{g_k^T(y_{k-1} - v_{k-1})}{d_{k-1}^T y_{k-1}} \qquad \text{(Perry,1978)}$$

$$\beta_k^{PRP+} = \max\{0, \frac{y_k^T g_{k+1}}{g_k^T g_k}\} \qquad \text{(Powell, 1984)}$$

$$\beta_k^{LS} = -\frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}} \qquad \text{(Liu and Storey, 1991)}$$

Where $\|.\|$ means the Euclidean norm and , $y_{k-1}=g_k - g_{k-1}$, $v_{k-1}=x_k-x_{k-1}$.

Dai and Yuan in (Dai and Yuan,2001a) proposed one-parameter family of the CG method: i.e.

$$\beta_k^{one} = \frac{\|g_k\|^2}{\lambda\|g_{k-1}\|^2 + (1-\lambda)d_{k-1}^T y_{k-1}} \qquad \text{………………………… (6)}$$

Where $\lambda \in [0,1]$ is a parameter.

Nazareth in (Nazareth ,1999) proposed two-parameter family of the CG method: i.e.

$$\beta_k^{Two} = \frac{\lambda_k\|g_k\|^2 + (1-\lambda_k)g_k^T y_{k-1}}{\mu_k\|g_{k-1}\|^2 + (1-\mu_k)d_{k-1}^T y_{k-1}} \qquad \text{…………………… (7)}$$

where λ_k, $\mu_k \in [0,1]$ are parameters.

Also Dai and Yuan in (Dai and Yuan,2001) proposed a three-parameter family of nonlinear CG methods defined by:

$$\boldsymbol{\beta_k^{Three}} = \frac{\mathbf{(1-\lambda_k)\|g_k\|^2 + \lambda_k g_k^T y_{k-1}}}{\mathbf{(1-\mu_k-\omega_k)\|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}}} \qquad \text{…….. (8)}$$

Where λ_k, $\mu_k \in [0,1]$, $\omega_k \in [0,1-\mu_k]$ **.**

AL-Bayati and Ahmed,(2005) proposed a four-parameter family of nonlinear CG methods defined by:

$$\beta_k^{Four} = \frac{(1-\lambda_k-\delta_k)\|g_k\|^2 + \lambda_k g_k^T y_{k-1} + \delta_k\|y_{k-1}\|^2}{(1-\mu_k-\omega_k)\|g_{k-1}\|^2 + \mu_k d_{k-1}^T y_{k-1} - \omega_k d_{k-1}^T g_{k-1}} \qquad \text{………..(9)}$$

Where λ_k, $\mu_k \in [0,1]$, $\omega_k \in [0,1-\mu_k]$ and $\delta_k \in [0,1-\lambda_k]$ are parameters.

AL-Bayati and Mitras ,(2008) proposed five-parameter family of CG methods, defined by:

$$\beta_k^{Five} = \frac{(1-\lambda_k-\delta_k-\gamma_k)g_k^T(y_{k-1}-v_{k-1}) + \lambda_k\|g_k\|^2 + \delta_k g_k^T y_{k-1} + \gamma_k\|y_{k-1}\|^2}{\mu_k\|g_{k-1}\|^2 + \omega_k d_{k-1}^T y_{k-1} - (1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}} \qquad \text{.. (10)}$$

where λ_k, $\mu_k \in [0,1]$, $\omega_k \in [0,1-\mu_k]$, $\delta_k \in [0,1-\lambda_k]$ and $\gamma_k \in [0,1-\delta_k]$ are parameters.

2. New CG- Methods with Six and Seven Parameters

2.1 Six–Parameter Family

We notice that Dia and Yuan,(2003) presented a nonlinear conjugate ,which has the form (1)with

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} \quad , \quad \beta_k^{DY} = \frac{g_k^T d_k}{d_{k-1}^T g_{k-1}}$$

(We called the first formula DY1and the second formula DY2).

We observed that the formulae FR, HS, DX, PR, LS, NEW1, NEW2, NEW3, and PERRY,DY1 and DY2 (which is already included in the previous families) are sharing nominators and denominators. According to this, we proposed six-parameter family defined by:

$$\beta_k^{six} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k) g_k^T (y_{k-1}-v_{k-1}) + \lambda_k \|g_k\|^2 + \delta_k g_k^T y_{k-1} + \gamma_k \|y_{k-1}\|^2 - \varphi_k^T g_k^T d_k}{\mu_k \|g_{k-1}\|^2 + \omega_k d_{k-1}^T y_{k-1} - (1-\mu_k-\omega_k) d_{k-1}^T g_{k-1}} \qquad \text{…………(11)}$$

Where λ_k, $\mu_k \in [0,1]$,$\omega_k \in [0,1-\mu_k]$, $\delta_k \in [0,1-\lambda_k]$ and $\gamma_k \in [0,1-\delta_k]$, $\varphi \in [0,1-\gamma_k]$ parameters.[i.e. $(\lambda_k, \delta_k, \gamma_k, \varphi)$ are impossible to be equal to one at the same time; the same thing is also correct for μ_k and ω_k]. Then we have $2^6=64$, 53 cases failed, and 11 cases are succeed which are

(0,0,1,1,0,0),(1,1,0,0,0,0),(0,1,0,1,0,0),(1,0,0,0,0,0),(1,0,1,0,0,0),(0,0,0 ,1,0,0)

(0,1,0,0,1,0),(0,1,0,0,1,0),(0,0,1,0,1,0),(0,0,1,0,0,0),(0,0,0,0,0,1). i.e. when

$\lambda=0, \mu=0, \omega=1, \delta=1, \gamma=0, \varphi=0$	β^{six} reduces to β^{HS}
$\lambda=1, \mu=1, \omega=0, \delta=0, \gamma=0, \varphi=0$	β^{six} reduces to β^{FR}
$\lambda=0, \mu=1, \omega=0, \delta=1, \gamma=0, \varphi=0$	β^{six} reduces to β^{PR}
$\lambda=1, \mu=0, \omega=0, \delta=0, \gamma=0, \varphi=0$	β^{six} reduces to β^{DX}
$\lambda=1, \mu=0, \omega=1, \delta=0, \gamma=0, \varphi=0$	β^{six} reduces to β^{DY1}
$\lambda=0, \mu=0, \omega=0, \delta=1, \gamma=0, \varphi=0$	β^{six} reduces to β^{LS}

λ=0, μ=1, ω=0, δ=0, γ=1 ,φ=0 β^{six} reduces to β^{NEW1}
λ=0, μ=0, ω=0, δ=0, γ=1 ,φ=0 β^{six} reduces to β^{NEW2}
λ=0, μ=0, ω=1, δ=0, γ=1 ,φ=0 β^{six} reduces to β^{NEW3}
λ=0, μ=0, ω=1, δ=0, γ=0 ,φ=0 β^{six} reduces to β^{Perry}
λ=0, μ=0, ω=0, δ=0, γ=0 ,φ=1 β^{six} reduces to β^{DY2}

Which represent subfamilies of a six-parameter family

2.2 Seven-parameter Family:

Hager and Zhang, (2005 proposed a new formula for the scalar β_k defined by

$$\beta_k^N = (y_k - 2d_k \frac{\|y_k\|^2}{d_k^T g_k})^T \frac{g_{k+1}}{y_k^T d_k} \qquad \text{……………………(12)}$$

We observed that the formulas FR, HS, DX, PR, LS, NEW1, NEW2, NEW3, PERRY, Dy1, Dy2 and β^N are sharing nominators and denominators. We now list β_k^{seven} for the new seven – parameter family :

$$\beta_k^{seven} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k-\psi_k)g_k^T(y_{k-1}-v_{k-1})+\lambda_k\|g_k\|^2+\delta_k g_k^T y_{k-1}+\gamma_k\|y_{k-1}\|^2-\varphi_k^T g_k d_k+\psi_k y_1^T g_k}{\mu_k\|g_{k-1}\|^2+\omega_k d_{k-1}^T y_{k-1}-(1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}} \qquad \text{………………. (13)}$$

Where

$$y_1 = (y_k - 2d_k \frac{\|y_k\|^2}{d_k^T g_k})$$

Where λ_k, $\mu_k \in [0,1]$, $\omega_k \in [0,1- \mu_k]$, $\delta_k \in [0,1- \lambda_k]$ and $\gamma_k \in [0,1-\delta_k]$, $\varphi \in [0,1-\gamma_k]$ are parameters, i.e. $(\lambda_{k,}\delta_{k,}\gamma_{k,}\varphi,\psi)$ are impossible to be equal to one at the same time; the same thing is also correct for μ_k and ω_k].

Then we have 2^7 =128, 116 cases failed, 12 casesed succeed, which are

(0,0,1,1,0,0,0),(1,1,0,0,0,0,0),(0,1,0,1,0,0,0),(1,0,0,0,0,0,0),(1,0,1,0,0,0,0,),(0,0,0,1,0,0,0).(0,1,0,0,1,0,0),(0,0,0,0,1,0,0),(0,0,1,0,1,0,0),(0,0,1,0,0,0,0),(0,0,0,0,0,1,0), (0,0,0,0,0,0,1) i.e. when:

λ=0, μ=0, ω=1, δ=1, γ=0 ,φ=0, ψ=0 β^{seven} reduces to β^{HS}

λ=1, μ=1, ω=0, δ=0, γ=0 ,φ=0, ψ=0 β^{seven} reduces to β^{FR}

λ=0, μ=1, ω=0, δ=1, γ=0 ,φ=0, ψ=0 β^{seven} reduces to β^{PR}

λ=1, μ=0, ω=0, δ=0, γ=0 ,φ=0, ψ =0 β^{seven} reduces to β^{DX}

λ=1, μ=0, ω=1, δ=0, γ=0 ,φ=0, ψ=0 β^{seven} reduces to β^{DY1}

λ=0, μ=0, ω=0, δ=1, γ=0 ,φ=0, ψ =0 β^{seven} reduces to β^{LS}

λ=0, μ=1, ω=0, δ=0, γ=1 ,φ=0, ψ=0 β^{seven} reduces to β^{NEW1}

λ=0, μ=0, ω=0, δ=0, γ=1 ,φ=0, ψ=0 β^{seven} reduces to β^{NEW2}

λ=0, μ=0, ω=1, δ=0, γ=1 ,φ=0, ψ=0 β^{seven} reduces to β^{NEW3}

λ=0, μ=0, ω=1, δ=0, γ=0 ,φ=0, ψ=0 β^{seven} reduces to β^{Perry}

λ=0, μ=0, ω=0, δ=0, γ=0 ,φ=1, ψ=0 β^{seven} reduces to β^{DY2}

λ=0, μ=0, ω=1, δ=0, γ=0 ,φ=0, ψ=1 β^{seven} reduces to β^{N}

which represent subfamilies of seven-parameter family.

3. Convergence properties

In this section, we study the global convergence properties of the new family of nonlinear CG-methods.

Theorem 3.1

Consider any CG method in the form (2), (3) and (11) with $\lambda_k, \mu_k \in [0,1]$, $\omega_k \in [0,1-\mu_k]$ and $\delta_k \in [0,1-\lambda_k]$, $\varphi \in [0,1-\delta_k]$ with the search condition $[\left|g(x_k+\alpha_k d_k)^T d_k\right| \le -\sigma g_k^T d_k]$ and with Powell's restarting criterion (AL-Bayati and Mitras , 2008). If the parameter ζ and σ satisfy

$$(1+\zeta)\sigma \le \frac{1}{2}$$

then we have for all $k \ge 1$,

$$0 < r_k < \frac{1}{1-(1+\zeta)\sigma}$$

For the proof of this theorem see(AL-Bayati and Mitras ,2008).

Assumption 3.2

Assume that the level set $L = \{x \in R^n, f(x) \le f(x_1)\}$ is bounded in some neighborhood N in L; f is continuously differentiable and its gradient g is Lipchitz continuous, i.e. there exists a constant λ>0 such that:

$\|g(x)-g(y)\| \le \lambda \|x-y\|$ for all x, y∈N ……………………….(14)

The above assumption implies that there exists a positive constant γ s.t.

$\|g(x)\| \le \gamma$ for all x∈L ………………………………..(15)

Lemma 3.3

Suppose that x_1 is a starting point for which the above assumption holds. Consider any method in the form (2), (3) with the following three properties: (i) $\beta_k \ge 0$; (ii) the strong Wolfe condition holds, and the sufficient descent condition holds for all k and some positive constant c; (iii) property (*) holds, namely, there exists constant $b > 1$ and $\hat{\lambda} > 0$ such that for all k:

$|\beta_k| \le b$ and if $g_k^T d_k \le -c\|g_k\|^2$ $\|x_k - x_{k-1}\| \le \hat{\lambda}$, $|\beta_k| \le (2b)^{-1}$, then the method converges (Bunday,1984) in the sense that:

$\lim_{k\to\infty} \inf \|g_k\| = 0$ ………………………………(16)

by this lemma we can prove the following general result for the six-parameter family of linear CG- methods.

Theorem 3.4

Suppose that x_1 is a starting point for which the above assumption holds. Consider any method in the form(2), (3) and (11)where

$\lambda_k \in [0,1], \mu_k \in [0,1], \omega_k \in [0,1-\mu_k], \delta_k \in [0,1-\lambda_k], \gamma_k \in [0,1-\delta_k]$,and $\varphi_k \in [0,1-\gamma_k]$ and the step-size satisfied the strong Wolfe condition with the restarting criterion (powell,1977)

$-\zeta\|g_k\|^2 \le g_k^T g_{k-1} \le \zeta\|g_k\|^2.$ ………………………(17)

Moreover, if

$$(1+\zeta)\sigma < \frac{1}{2} \qquad \text{………………………(18)}$$

$$\lambda_k + \delta_k + \gamma_k + \varphi_k \geq 1 - c_1\|x_k - x_{k-1}\| \qquad \text{……………………(19)}$$

where ($c_1 > 0$ is constant) are held. The new method converges in the sense that(16) holds.

Proof:

We proceed by contradiction. Assume that

$$\lim_{k\to\infty} \inf\|g_k\| \neq 0 \quad \text{………………………………(20)}$$

then there exists a positive constant γ such that

$$\|g_k\| \geq \gamma, \text{ for all } k \geq 1 \qquad \text{……………………..(21)}$$

we can see from that for all $k \geq 1$, (AL-Bayati and Mitras ,2008)

$$r_k \geq 1 - \frac{(1+\zeta)\,\sigma}{[1-(1+\zeta)\,\sigma]} \quad \text{…………………………..(22a)}$$

which implies that

$$r_k \geq \frac{1-2(1+\zeta)\,\sigma}{1-(1+\zeta)\sigma} \cong c_2 \quad \text{………………………..(22b)}$$

where r_k is defined as $r_k = -\dfrac{g_k d_k}{\|g_k\|^2}$

Thus the sufficient descent condition defined by $g_k^T d_k \leq -c\|g_k\|^2$ holds.

We can show from condition (19)

$$\lambda_k + \delta_k + \gamma_k + \varphi_k - 1 \geq -c_1\|x_k - x_{k-1}\| \quad \text{……………….(23)}$$

multiply both sides of (23) by (-1)

$$1 - \lambda_k - \delta_k - \varphi_k \leq c_1\|x_k - x_{k-1}\| \quad \text{…………………………(24)}$$

Now

$$\beta_k = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k)g_k^T(y_{k-1}-v_{k-1}) + \lambda_k\|g_k\|^2 + \delta_k g_k^T y_{k-1} + \gamma_k\|y_{k-1}\|^2 - \varphi_k^T g_k^T d_k}{\mu_k\|g_{k-1}\|^2 + \omega_k d_{k-1}^T y_{k-1} - (1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}}$$

where $\|y_{k-1}\| = \|g(x_k) - g(x_{k-1})\| \leq \tau\|x_k - x_{k-1}\|$.

Now from (5), definition of r_k, the two conditions of Assumption, (22b) and substituting (24) ,we have

$$|\beta_k| \le \frac{c_1\|x_k - x_{k-1}\|\|g_k\|^2 + \lambda_k\|g_k\|^2 + \delta_k\|g_k\|\tau\|x_k - x_{k-1}\| + \gamma_k\|y_{k-1}\|^2 - \varphi_k^T g_k^T d_k}{\|g_{k-1}\|^2[\mu_k + \omega_k(1+\sigma)c_2 + (1-\mu_k-\omega_k)c_2]}$$

since $\lambda_k \in [0,1], \mu_k \in [0,1]$, $\omega_k \in [0, 1-\mu_k]$ and $\delta_k \in [0, 1-\lambda_k], \gamma_k \in [0, 1-\delta_k]$, $\varphi_k \in [0, 1-\gamma_k]$

(i.e. $\lambda_k = 0, \delta_k = 1, \mu_k = 0, \omega_k = 0, \gamma_k = 0, \varphi_k = 0)$) then we have :

$$|\beta_k| \le \frac{c_1\|x_k - x_{k-1}\|\|g_k\|^2 + \|g_k\|\tau\|x_k - x_{k-1}\|}{\|g_{k-1}\|^2 c_2} \quad \ldots\ldots\ldots\ldots\ldots\ldots(25)$$

$$\le \frac{(c_1\|g_k\|^2 + \|g_k\|\tau).\|x_k - x_{k-1}\|}{\|g_{k-1}\|^2 c_2}$$

$$\le c_3\|x_k - x_{k-1}\| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(26)$$

where $c_3 = \dfrac{c_1\gamma^2 + \tau\gamma}{c_2\gamma^2}$ ……………………………….(27)

Since the Assumption holds implying there exists a positive ψ constant such that

$$\|x\| \le \psi \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(28)$$

for $b = 2c_3\psi$ and $\hat{\lambda} = (4c_3^2\psi)^{-1}$,we have from (26) and (28) that

$$|\beta_k| \le b \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(29)$$

and if $\|x_k - x_{k-1}\| \le \hat{\lambda}$

$$|\beta_k| \le (2b)^{-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(30)$$

Thus property (iii) holds. In addition from(17) and(22a) imply that $\beta_k \ge 0$.Therefore the conditions of lemma are all satisfied and hence (16) holds. #

For more details see(AL-Bayati and Mitras ,2008).

Note:

To show that the new CG-method with β_k^{seven} is also a globally convergent method we have to show an equivalent relation between β_k^{six} and β_k^{seven} using the main properties of CG-method

(i) ELS i.e, $d_k^T g^* = 0$.

(ii) Orthogonality properties $g^{*T} y = 0$.

As follows: Assuming that $d_k = -g_k$ and let us consider:

$$\beta_k^{six} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k)g_k^T(y_{k-1}-v_{k-1})+\lambda_k\|g_k\|^2+\delta_k g_k^T y_{k-1}+\gamma_k\|y_{k-1}\|^2-\varphi_k^T g_k^T d_k}{\mu_k\|g_{k-1}\|^2+\omega_k d_{k-1}^T y_{k-1}-(1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}}$$

using the properties (i) and (ii) we get:

$$\beta_k^{six} = \frac{g_k^{\ T} g_k+\gamma_k g_{k-1}^T g_{k-1}}{g_{k-1}^T g_{k-1}} \quad \text{……………………………………..(31)}$$

And also

$$\beta_k^{seven} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k-\psi_k)g_k^T(y_{k-1}-v_{k-1})+\lambda_k\|g_k\|^2+\delta_k g_k^T y_{k-1}+\gamma_k\|y_{k-1}\|^2-\varphi_k^T g_k d_k+\psi_k y_1^T g_k}{\mu_k\|g_{k-1}\|^2+\omega_k d_{k-1}^T y_{k-1}-(1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}}$$

using the properties (i) and (ii) we get:

$$\beta_k^{seven} = \frac{g_k^T g_k+\gamma_k g_{k-1}^T g_{k-1}}{g_{k-1}^T g_{k-1}} \quad \text{…………………………………(32)}$$

Clearly (31),(32) are the same.

4. Outlines of the new proposed family:

Step 1: Set k=1, d_k=-g_k, and the initial point x_1.

Step 2: Set $x_{k+1} = x_k + \lambda_k d_k$, where is a scalar chosen in such a way that $f_{k+1}<f_k$.

Step3: Check for convergence, i.e. if, $\|g(x)\| < \varepsilon$ where ε positive

tolerance, stop; otherwise continue.

Step 4: Compute the new search direction defined by:

$d_k = -g_k + \beta_k d_{k-1}$, where β_k is one of the following formulas:

$$\beta_k^{six} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k)g_k^T(y_{k-1}-v_{k-1})+\lambda_k\|g_k\|^2+\delta_k g_k^T y_{k-1}+\gamma_k\|y_{k-1}\|^2-\varphi_k^T g_k^T d_k}{\mu_k\|g_{k-1}\|^2+\omega_k d_{k-1}^T y_{k-1}-(1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}}$$

Or

$$\beta_k^{seven} = \frac{(1-\lambda_k-\delta_k-\gamma_k-\varphi_k-\psi_k)g_k^T(y_{k-1}-v_{k-1})+\lambda_k\|g_k\|^2+\delta_k g_k^T y_{k-1}+\gamma_k\|y_{k-1}\|^2-\varphi_k^T g_k d_k+\psi_k y_1^T g_k}{\mu_k\|g_{k-1}\|^2+\omega_k d_{k-1}^T y_{k-1}-(1-\mu_k-\omega_k)d_{k-1}^T g_{k-1}}$$

Step 5: If k=n or $\|g_k\|^2 < 0.1\|d_k\|^2$, go to step 1. Else, set k=k+1 and go to step 2.

5. Numerical Result and discussion

In order to assess the performance of the new two scalars, several CG methods are used with these two scalars and tested over five generalized selected well-known test functions with standard dimensions (N=1000).

These two scalars are assumed to have the convergence property when each element of the gradient vector is less than 1E-5,i.e. $\|g_{k+1}\| \le 1*10^{-5}$. A cubic fitting procedure which was described in detail by Bunday,(1984) ,was used as a line search procedure

All results are obtained using Pentium 4. All programs are written in FORTRAN language. The comparative performances for all these methods are evaluated by considering both Number Of Function Evaluations (NOF) and Number Of Iterations (NOI).

Tables (1-5) show numerical results of six-Parameter family CG Methods with different values of ($\lambda,\delta,\gamma,\mu,\omega,\varphi$) for the functions Powell, Wood, Cantral, Rosen, respectively.

Tables(6-10) show numerical results of seven-Parameter family CG Methods with different values of ($\lambda,\delta,\gamma,\mu,\omega,\varphi,\psi$) for the functions Powell, Wood, Sum, Cantral, respectively.

Note: In tables (1-10) (***) means number exceeds 2500.

Table (1), (Powell fn. with N=1000)
Numerical Results of six-Parameter Family CG Method with Different value of

				λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.3 δ=0.4
				NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.6	φ=0.2	664 287	536 248	290 135	257 126	226 111
γ=0.0	μ=0.5	ω=0.0	φ=0.1	207 101	319 139	470 213	122 59	504 239
γ=0.3	μ=0.2	ω=0.3	φ=0.4	2153 1021	244 117	685 301	349 167	2222 1035
γ=0.4	μ=0.0	ω=0.1	φ=0.6	236 116	1210 598	791 379	*** 2001	525 251
γ=0.2	μ=0.4	ω=0.0	φ=0.3	108 52	403 188	199 95	336 158	110 53
γ=0.0	μ=0.0	ω=0.0	φ=0.01	274 134	1065 505	141 68	1015 471	354 175

$(\lambda, \delta, \gamma, \mu, \omega, \varphi)$

In the above table we have implemented the new six- parameter family with a different value for the six different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(436 , 199) ,** β_k^{FR} **(2007 , 1001) ,** β_k^{PR} **(2011 , 1002) ,** β_k^{DX} **(2026, 1005) ,** β_k^{DY1} **(2006 , 1002) ,** β_k^{LS} **(209, 93) ,** β_k^{NEW1} **(*** , 2022),** β_k^{NEW2} **(*** , 2064),** β_k^{NEW3} **(***, 2001),** β_k^{PERRY} **(2006, 1002),** β_k^{DY2} **(***, ***).**

The best value at (λ=0.3,δ=0.2, γ=0.2, μ=0.4, ω=0.0, φ=0.3).

Table (2), (Wood fn. with N=1000)
Numerical Results of six-Parameter Family CG Method with value Different of $(\lambda, \delta, \gamma, \mu, \omega, \varphi)$

				λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.3 δ=0.4
				NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.6	φ=0.2	560 271	453 208	640 275	344 169	560 271
γ=0.0	μ=0.5	ω=0	φ=0.1	274 133	214 104	240 117	234 114	272 132
γ=0.3	μ=0.2	ω=0.3	φ=0.4	555 239	882 400	698 340	533 229	872 397
γ=0.4	μ=0.0	ω=0.1	φ=0.6	572 282	926 433	1022 444	444 220	784 388
γ=0.2	μ=0.4	ω=0.0	φ=0.3	940 411	656 276	951 432	333 163	803 346
γ=0.0	μ=0.0	ω=0.0	φ=0.01	168 80	214 104	201 95	172 83	196 94

In the above table we have implemented the new six- parameter family with a different value for the six different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} (184 , 90), β_k^{FR} (*,***), β_k^{PR} (169 , 82), β_k^{DX} (1874 , 913), β_k^{DY1} (***, 1263), β_k^{LS} (152 , 73) , β_k^{NEW1} (1034 , 514) , β_k^{NEW2} (1895, 798), β_k^{NEW3} (1374, 660) , β_k^{PERRY} (184 , 90) , β_k^{DY2} (***,***) .**

The best value at (λ=0.3,δ=0.2, γ=0.0, μ=0.0, ω=0.0, φ=0.01).

Table (3) ,(Cantral fn. with N=1000)
Numerical Results of six-Parameter Family CG Method with value Different of $(\lambda, \delta, \gamma, \mu, \omega, \varphi)$

				λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.3 δ=0.4
				NOF(NOI	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.6	φ=0.2	139 28	266 38	118 25	240 34	115 24
γ=0.0	μ=0.5	ω=0.0	φ=0.1	175 26	335 38	277 38	149 20	220 29
γ=0.3	μ=0.2	ω=0.3	φ=0.4	91 19	226 32	313 39	198 29	62 16
γ=0.4	μ=0.0	ω=0.1	φ=0.6	212 34	194 33	336 48	266 41	435 60
γ=0.2	μ=0.4	ω=0.0	φ=0.3	93 19	203 28	145 22	185 26	135 25
γ=0.0	μ=0.0	ω=0.0	φ=0.01	130 20	274 35	173 22	196 29	221 28

In the above table we have implemented the new six- parameter family with different value for the six a different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} (169 , 22), β_k^{FR} (250 , 34) , β_k^{PR} (118 , 19), β_k^{DX} (fail), β_k^{DY1} (176 , 22) , β_k^{LS} (135 , 19), β_k^{NEW1} (309 , 47) , β_k^{NEW2} (299 , 40) , β_k^{NEW3} (299 , 48), β_k^{PERRY} (224 , 31), β_k^{DY2} (2166 , 299) .

The best value at (λ=0.3,δ=0.4, γ=0.3, μ=0.2, ω=0.3, φ=0.4).

Table (4), (Rosenbrock fn. with N=1000)
Numerical Results of six-Parameter Family CG Method with Different value of $(\lambda, \delta, \gamma, \mu, \omega, \varphi)$

				λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.3 δ=0.4
				NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.6	φ=0.2	191 67	110 42	110 45	109 45	150 54
γ=0.0	μ=0.5	ω=0.0	φ=0.1	82 36	90 36	74 33	68 29	84 36
γ=0.3	μ=0.2	ω=0.3	φ=0.4	101 47	93 42	53 22	158 66	101 47
γ=0.4	μ=0	ω=0.1	φ=0.6	122 53	196 89	107 45	140 62	106 45
γ=0.2	μ=0.4	ω=0.0	φ=0.3	75 33	80 31	88 39	66 29	75 33
γ=0.0	μ=0.0	ω=0.0	φ=0.01	74 30	131 46	72 30	55 21	108 42

In the above table we have implemented the new six- parameter family with a different value for the six different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(58 , 22) ,** β_k^{FR} **(244, 84) ,** β_k^{PR} **(59 , 24) ,** β_k^{DX} **(***, 1073) ,** β_k^{DY1} **(401 , 162) ,** β_k^{LS} **(fail) ,** β_k^{NEW1} **(112, 45) ,** β_k^{NEW2} **(511 , 226) ,** β_k^{NEW3} **(477, 229) ,** β_k^{PERRY} **(57 , 22) ,** β_k^{DY2} **(***, ***) .**

The best value at (λ=0.0,δ=0.5, γ=0.0, μ=0.0, ω=0.0, φ=0.01).

Table (5), (Cubic fn. with N=1000)
Numerical Results of six-Parameter Family CG Method with Different value of $(\lambda,\delta,\gamma,\mu,\omega,\varphi)$

				λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.3 δ=0.4
				NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.6	φ=0.2	87 39	121 50	99 45	113 51	89 40
γ=0.0	μ=0.5	ω=0.0	φ=0.1	50 21	48 21	28 11	39 16	50 21
γ=0.3	μ=0.2	ω=0.3	φ=0.4	84 37	84 35	93 43	83 37	107 49
γ=0.4	μ=0	ω=0.1	φ=0.6	140 65	256 123	130 60	89 39	136 63
γ=0.2	μ=0.4	ω=0.0	φ=0.3	69 30	69 29	90 40	52 22	69 30
γ=0.0	μ=0	ω=0.0	φ=0.01	59 25	43 18	54 23	39 15	58 25

In the above table we have implemented the new six -parameter family with different value for the six a different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(35 , 13) ,** β_k^{FR} **(86 , 36) ,** β_k^{PR} **(38 , 16) ,** β_k^{DX} **(152, 71) ,** β_k^{DY1} **(49 , 20) ,** β_k^{LS} **(fail) ,** β_k^{NEW1} **(fail),** β_k^{NEW2} **(180 ,76) ,** β_k^{NEW3} **(266 , 101) ,** β_k^{PERRY} **(37, 14) ,** β_k^{DY2} **(***, ***) .**

The best value at (λ=0.25,δ=0.25, γ=0.0, μ=0.5, ω=0.0, φ=0.1).

Table (6), (Powell fn. with N=1000)
Numerical Results of seven-Parameter Family CG Method with Different value of
$(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$

					λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.4 δ=0.0
					NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.3	φ=0.2	ψ=0.7	*** 878	630 171	192 90	637 248	7497 1086
γ=0.0	μ=0.5	ω=0.0	φ=0.1	ψ=0.5	2111 526	396 136	1625 448	714 218	*** 911
γ=0.3	μ=0.2	ω=0.7	φ=0.0	ψ=0.1	909 319	204 85	1495 422	491 183	168 81
γ=0.2	μ=0.4	ω=0.0	φ=0.3	ψ=0.4	2172 546	803 225	529 212	366 163	**** 723
γ=0.0	μ=0.0	ω=0.0	φ=0.0	Ψ=0.0	262 118	242 107	238 102	270 115	140 69

In the above table we have implemented the new seven- parameter family with a different value for the seven different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(436 , 199),** β_k^{FR} **(2007 , 1001),** β_k^{PR} **(2011 , 1002), (**β_k^{DX} **2026 , 1005),** β_k^{LS} **(209 , 93),** β_k^{DY1} **(2006, 1002) ,** β_k^{NEW1} **(*** , 2022) ,** β_k^{NEW2} **(*** , 2064),** β_k^{NEW3} **(*** , 2001) ,** β_k^{PERRY} **(2006 , 1002) ,** β_k^{DY2} **(***, ***),** β_k^{N} **(*** , 1002).**

The best value at (λ=0.4,δ=0.0, γ=0.3, μ=0.2, ω=0.7, φ=0.0,Ψ=0.1).

Table (7), (Wood fn. with N=1000)
Numerical Results of seven-Parameter Family CG Method with Different value of
$(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$

					λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.4 δ=0.0
					NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.31	φ=0.2	ψ=0.7	460 215	777 321	730 348	487 241	594 278
γ=0.0	μ=0.5	ω=0.0	φ=0.1	ψ=0.5	282 135	236 114	285 138	240 117	292 143
γ=0.3	μ=0.2	ω=0.7	φ=0.0	ψ=0.1	458 222	876 385	646 259	304 149	1284 495
γ=0.1	μ=0.0	ω=0.1	φ=0.4	ψ=0.6	295 138	242 114	261 125	359 176	315 149
γ=0.0	μ=0.0	ω=0.0	φ=0.0	Ψ=0.02	259 125	176 85	125 58	236 115	241 116

In the above table we have implemented the new seven- parameter family with a different value for the seven different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(184, 90) ,** β_k^{FR} **(***, 4410) ,** β_k^{PR} **(169, 82) ,** β_k^{DX} **(1874 , 913) ,** β_k^{DY1} **(*** , 1263) ,** β_k^{LS} **(152 , 73) ,** β_k^{NEW1} **(1034 , 514) ,** β_k^{NEW2} **(1895 ,798) ,** β_k^{NEW3} **(1374 , 660) ,** β_k^{PERRY} **(184 ,90) ,** β_k^{DY2} **(*** , ***) ,** β_k^{N} **(453 , 178).**

The best value at (λ=0.25,δ=0.25, γ=0.0, μ=0.0, ω=0.0, φ=0.0,Ψ=0.02).

Table(8), (Sum fn. with N=1000)

Numerical Results of seven-Parameter Family CG Method with Different value of

$(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$

					λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.4 δ=0.0
					NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.31	φ=0.2	ψ=0.7	241 57	218 49	257 56	260 59	234 55
γ=0.0	μ=0.5	ω=0.0	φ=0.1	ψ=0.5	105 21	112 22	87 21	101 22	143 27
γ=0.3	μ=0.2	ω=0.7	φ=0.0	ψ=0.1	173 42	173 39	179 41	159 40	171 41
γ=0.1	μ=0.0	ω=0.1	φ=0.4	ψ=0.6	192 39	300 29	131 27	111 28	125 30
γ=0.2	μ=0.4	ω=0.0	φ=0.3	ψ=0.4	147 35	189 39	171 38	153 34	194 40
γ=0.0	μ=0.0	ω=0.0	φ=0.0	Ψ=0.02	132 25	134 26	120 23	116 24	159 32

In the above table we have implemented the new seven- parameter family with a different value for the seven different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(127, 28) ,** β_k^{FR} **(147 , 26) ,** β_k^{PR} **(104, 22) ,** β_k^{DX} **(147 , 26) ,** β_k^{DY1} **(145, 26),** β_k^{LS} **(111 , 24) ,** β_k^{NEW1} **(207, 55),** β_k^{NEW2} **(202, 63) ,** β_k^{NEW3} **(213, 63) ,** β_k^{PERRY} **(107 , 23) ,** β_k^{DY2} **(134 , 24) ,** β_k^{N} **(121, 23).**

The best value at (λ=0.25,δ=0.25, γ=0.0, μ=0.5, ω=0.0, φ=0.1,Ψ=0.5).

Table(9), (Cantral fn. with N=1000)
Numerical Results of seven-Parameter Family CG Method with Different value of
$(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$

					λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.4 δ=0.0
					NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.31	φ=0.2	ψ=0.7	142 27	216 36	268 38	275 42	141 29
γ=0.0	μ=0.5	ω=0	φ=0.1	ψ=0.5	196 28	226 30	335 40	162 23	239 32
γ=0.3	μ=0.2	ω=0.7	φ=0.0	ψ=0.1	99 20	301 41	148 28	164 25	151 26
γ=0.1	μ=0.0	ω=0.1	φ=0.4	ψ=0.6	79 17	219 35	309 40	285 39	550 74
γ=0.2	μ=0.4	ω=0.0	φ=0.3	ψ=0.4	222 34	295 40	218 35	221 34	363 49
γ=0.0	μ=0.0	ω=0.0	φ=0.0	Ψ=0.02	250 33	202 28	183 25	159 25	261 36

In the above table we have implemented the new seven- parameter family with a different value for the seven different parameters .The best results are shown above in comparison with some classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} (169 , 22), β_k^{FR} (250 , 34), β_k^{PR} (118 , 19) , β_k^{DX} (fail), β_k^{DY1} (176, 22), β_k^{LS} (135, 19), β_k^{NEW1} (309 , 47), β_k^{NEW2} (299 , 40), β_k^{NEW3} 3(299, 48), β_k^{PERRY} (224, 31), β_k^{DY2} (2166 , 299), β_k^{N} (151, 21).

The best value at (λ=0.3,δ=0.2, γ=0.1, μ=0.0, ω=0.1, φ=0.4,Ψ=0.6).

Table(10), (Cubic fn. with N=1000)
Numerical Results of seven-Parameter Family CG Method with Different value of
$(\lambda, \delta, \gamma, \mu, \omega, \varphi, \psi)$

					λ=0.3 δ=0.2	λ=0.1 δ=04	λ=0.25 δ=0.25	λ=0.0 δ=0.5	λ=0.4 δ=0.0
					NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)	NOF(NOI)
γ=0.4	μ=0.1	ω=0.31	φ=0.2	ψ=0.7	129 59	59 25	188 90	110 49	167 79
γ=0.0	μ=0.5	ω=0	φ=0.1	Ψ=0.5	73 33	51 22	57 26	45 18	67 30
γ=0.3	μ=0.2	ω=0.7	φ=0.0	ψ=0.1	89 39	59 24	87 39	150 44	102 46
γ=0.1	μ=0	ω=0.1	φ=0.4	ψ=0.6	71 32	111 50	91 41	69 30	158 30
γ=0.2	μ=0.4	ω=0.0	φ=0.3	ψ=0.4	97 45	83 36	150 66	104 44	113 52
γ=0.0	μ=0.0	ω=0.0	φ=0.0	Ψ=0.02	46 19	41 17	48 20	32 12	50 22

In the above table we have implemented the new seven- parameter family with a different value for the seven different parameters .The best results are shown above in comparison with the classical (conjugancy coefficient) β_k and as follows:

β_k^{HS} **(35 , 13),** β_k^{FR} **(86 , 36) ,** β_k^{PR} **(38, 16) ,** β_k^{DX} **(152, 71) ,** β_k^{DY1} **(49 , 20),** β_k^{LS} **(fail) ,** β_k^{NEW1} **(fail),** β_k^{NEW2} **(180 ,76),** β_k^{NEW3} **(266, 101),** β_k^{PERRY} **(37, 14),** β_k^{DY2} **(***, ***),** β_k^{N} **(81, 36).**

The best value at (λ=0.0,δ=0.5, γ=0.0, μ=0.0, ω=0.0, φ=0.0,Ψ=0.02).

References

[1] Al-Bayati, A.Y.and Al-Assady, N.H. 1986, Conjugate gradient algorithms for constrained optimization, Technical Report, School of Computer Studies, Leeds University, U.K.

[2] Al-Bayati, A.Y. and Ahmed, H.I. 2005, "A New Four –parameter Family of Non-linear Conjugate Gradient Methods", Iraq, Journal of Statistical Science, University of Mosul, Vol.7, pp. 18-38.

[3] Al-Bayati, A.Y. and Mitras, B.A. 2008, "A New Five – parameter Family of Non-linear Conjugate Gradient Methods", (2008), 2nd international conference of mathematics, University of Aletho, Syria.

[4] Bunday, B., 1984, "Basic optimization method", Edward Arnold, Bedford Square, London, U.K.

[5] Dai, Y.H. and Yuan, Y., 2003, "A Class of globally convergent conjugate methods", Research Report ICM-98-030, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences**.**

[6] Dai, Y. H. and Yuan, Y., 2001a, "An efficient hybrid conjugate gradient method for unconstrained optimization", Annals of Operations Res., no. 103, pp.33-47 .

[7] Dai, Y. H. and Yuan, Y., 2001, "A Three- parameter family of non-linear conjugate gradient Methods ",Mathematics of Computation, no. 70,pp.1155-1167

[8] Dixon, L. G. W., 1972, "Nonlinear optimization", the English Universities Press Limited.

[9] Fletcher, R.and Reeves, C. M., 1964, "Function minimization by conjugate gradient", Computer Journal, vol. 7, pp.149-154.

[10] Hager, W. W., Zhang,H., 2005," A new conjugate gradient method guaranteed with decent and efficient line search", SIAM J. Optim., 16,170-192.
[11] Hestenes, M. R. And Stiefel, E. L., 1952, "Methods of conjugate gradient for solving linear systems", Journal of Research of the National Bureau of Standards, Vol. 49, pp. 409-436.
[12] Liu, Y. and Storey, C., 1991,"Efficient generalized conjugate gradient algorithms", Part 1: Theory, Journal of Optimization Theory and Applications, vol.69, pp. 129-137.
[13] Nazareth, L., 2001, " Conjugate gradient methods", Encyclopedia of Optimization (C. Floudas and P.Pardalos, eds.), Kluwer Academic Publishers, Boston,USA and Dordrecht, The Netherlands,I(A-D),PP. 319-323.
[14] Perry, A., 1969, "A Modified conjugate gradient algorithm, Operation Research", vol. 26, pp.1073-1078.
[15] Polak, E. and Ribiere, G., 1969, "Note sur la convergence des methods de directions conjugate", Rev. Fr Infr, Rech. Oper. vol.16, R1.
[16] Powell, M.J.D., 1977, "Restart procedures for the conjugate gradient method", Mathematical Programming, vol.12, pp.241- 254.
[17] Powell, M.J.D., 1984, "Nonconvex minimization calculations and the conjugate gradient method. Numerical Analysis(Dunde,1983), lecture Notes in mathematics, vol.1066, Springer Verlag, Berlin, pp.122-141.