New Scale Dai-Yaun Conjugate Gradient Method for Unconstrained Optimization

Hamsa Th. Chilmerane
hamsathrot@uomosul.edu.iq
Department of Mathematics,
College of Computers Sciences and Mathematics
University of Mosul, IRAQ

Received on: 18/10/2011 **Accepted on: 15/02/2012**

ABSTRACT

Conjugate gradient algorithm is widely used for solving large-scale unconstrained optimization problems, because they do not need the storage of matrices. In this paper, we suggest a modified Dai-Yuan conjucay coefficient of conjugate gradient algorithm and propose new spectral form three-term conjugate gradient algorithm. These algorithms are used inexact line searches and Wolf line search conditions. These algorithms satisfied sufficient descent condition and the converge globally are provided under some assumptions. The numerical results indicate that the proposed algorithm is very effective and the new spectral algorithm is of very robust results depending on iterations and the number of known functions.

Keywords: conjugate gradient, inexact line search, conjugacy coefficient.

مقياس جديد لطريقة Dai-Yaun للتدرج المترافق في الأمثلية غير المقيدة

همسة ثروت جلميران
كلية علوم الحاسوب والرياضيات
جامعة الموصل، الموصل، العراق

تاريخ استلام البحث: 2011/10/18 **تاريخ قبول البحث: 2012/02/15**

الملخص

طريقة التدرج المترافق دائما تستخدم لحل المسائل الأمثلية غير المقيدة العالية الإبعاد لأنها لا تحتاج لخزن لأية مصفوفة. تم في هذا البحث اقتراح تحسين لمعامل ترافق Dai-Yuan لخوارزمية التدرج المترافق كما تم اشتقاق ذو قياس جديد من طريقة التدرج المترافق لذات المقادير الثلاثة باستخدام خط بحث غير مضبوط ، كل من الخوارزميتين تم برهانهما بأنهما يحققان شرط الانحدار الكافي التقارب الشامل تحت بعض الشروط . دلت النتائج العددية على أن الطريقة الجديدة تحقق نتائج عددية ذات كفاءة عالية بالاعتماد على عدد التكرارات وعدد حسابات الدالة.

الكلمات المفتاحية: التدرج المترافق، خط بحث غير تام، معامل الترافق.

1. Introduction:

In this paper, let us consider the nonlinear unconstrained optimization problem

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

where $f : R^n \to R$ is a continuously differentiable function, bounded from below. As we know, for solving this problem, let $x_0 \in R^n$ be the initial guess of the solution of problem (1). A non-linear conjugate gradient method is usually generates a sequence $\{x_k\}$ as form

$$x_{k+1} = x_k + \lambda_k d_k \qquad \text{……….. (2)}$$

where $\lambda_k > 0$ is obtained by line search, and the directions d_k are generated as

$$d_{k+1} = -g_{k+1} + \beta_k d_k \ , \ d_0 = -g_0 \qquad \text{……..…(3)}$$

$\beta_k \in R$ is known as a conjugate gradient coefficient , some well known formulas are given as follows:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}$$ (Hestenes-Stiefel, [5],(1952)) ……….. (4)

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}$$ (Fletcher-Reeves (FR), [3] ,(1964**))** ……….. (5)

…….. (6))) Ribière (PR),[8],(1969 -Polak($$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$$

(7) ……….. Assady ,[1],(1986)) -Bayati & Al-(Al $$\beta_k^{BA} = \frac{-y_k^T y_k}{d_k^T g_k}$$

$$\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}$$ (Fletcher (CD),[4] ,(1987)) ……….. (8)

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}$$ (Liu-Storey (LS),[6],(1991)) ……….. (9)

(10)………..))Yuan (DY),[2],(1999-Dai($$\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k}$$

where $g_k = \nabla f(x_k)$ and $y_k = g_{k+1} - g_k$.

The line search in the conjugate gradient algorithms for λ_k computation is often based on the standard Wolfe conditions

The weak Wolfe-conditions:

$$f(x_k + \lambda_k d_k) - f(x_k) \le \delta \lambda_k g_k^T d_k \qquad \text{……….. (11)}$$

$$g(x_k + \lambda_k d_k)^T d_k \ge \sigma g_k^T d_k \qquad \text{……….. (12)}$$

The strong Wolfe-conditions:

$$f(x_k + \lambda_k d_k) - f(x_k) \le \delta \lambda_k g_k^T d_k \qquad \text{……….. (13)}$$

$$\left| g(x_k + \lambda_k d_k)^T d_k \right| \le -\sigma g_k^T d_k \qquad \text{……….. (14)}$$

Eq.(14) can be written as

$$\sigma g_k^T d_k \le g_{k+1}^T d_k \le -\sigma g_k^T d_k \qquad \text{……….. (15)}$$

Where $\delta \in (0,1)$ and $\sigma \in (\delta, \frac{1}{2})$

2. The Modified Dai-Yuan Conjucay Coefficient of Conjugate Gradient Algorithm

Zhang and Weijun omit. (2006) Page No. [9] proposed a modified of an FR (MFR) formula such as

$$\beta^{MFR}(\mu)=\frac{\mu_1\|g_{k+1}\|^2}{\mu_2|g_{k+1}^T d_k|+\mu_3\|g_k\|^2} \quad \text{……………………………….(16)}$$

Where $\mu_1\in(0,+\infty),\mu_2\in[\mu_1+\varepsilon,+\infty),\mu_3(0,+\infty)$ and ε is a small number. Zhang and Weijun (2008) Page No. [11] drive another modified an FR formula which is defined as follows:

$$\beta^{MFR}(\mu)=\begin{cases}\dfrac{\|g_{k+1}\|^2-|g_{k+1}^T g_k|}{\mu_1|g_{k+1}^T d_k|+\|g_k\|^2} & \|g_{k+1}\|^2\succ|g_{k+1}^T g_k| \\ 0 & \text{otherwise}\end{cases} \quad \text{……(17)}$$

In this paper, we are going to study the developed modified Dai-Yuan CG-method based on nonlinear function taking the idea of inexact line searches as follows:

$$\beta_k^{MDY}=\begin{cases}\dfrac{\|g_{k+1}^T\|^2-t\,g_{k+1}^T s_k}{d_k^T y_k+\mu d_k^T g_{k+1}} & \text{if } \|g_{k+1}^T\|^2\succ t\,g_{k+1}^T s_k \\ 0 & \end{cases} \quad \text{………(18)}$$

Clearly, this formula will reduce the DY formula if *f* is a strictly convex quadratic and the line search is exact.

3. Global Convergence Properties for Modified Dai-Yuan Conjucay Conjugate Gradient Algorithm:

In this section, the convergence properties of the modified Dai-Yuan conjugate gradient algorithm with the inexact line search analyze and in order to ensure the sufficient descent condition, using Wolf condition line search the global convergence analysis of the iterative methods, the following assumption is often needed:

Assumption (A):

(i) f is bounded below, on the level set $\Omega=\{x\in R^n : f(x)\le f(x_0)\}$.

(ii) In some neighborhood Ω_0 of Ω, f is differentiable and it is gradient $g(x)$ is Lipschitz continuous, namely, there exists a constant $L\succ 0$ such that

$\|g(x)-g(y)\|\le L\|x-y\|$, $\forall\, x,y\in\Omega_0$

We can get from the assumption A that there exists positive constant $\Gamma>0$, such that:

$$\|g(x)\|\le\Gamma \qquad \forall\, x\in\Omega$$

Lemma 1: Assume the assumption A holds, and the modified DY method generates sequence $\{x_k\}$, then we have

$$g_{k+1}^T d_{k+1}\le-\zeta\|g_{k+1}\|^2$$

where ζ is a positive constant

Proof:

We prove the theorem with strong Wolfe conditions, by induction, for initial direction $(k=0)$ we have

$d_0=-g_0 \;\rightarrow\; d_0^T g_0=-\|g_0\|^2<0$

Suppose $d_k^T g_k<0 \qquad \forall\, k$

Now, we prove if $k=k+1$ then:

$$d_{k+1}=-g_{k+1}+\beta_k^{MDY}d_k \quad \text{………………….(19)}$$

Multiply sides both of eq. (19) by g_{k+1}^T we get :

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{MDY} g_{k+1}^T d_k \qquad \text{……………….(20)}$$

Dividing sides both of eq.(20) by $\|g_{k+1}\|^2$ we get :

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \beta_k^{MDY} \; g_{k+1}^T \, d_k$$

Since β_k^{MDY} is defined in (18), then we get

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \left(\frac{\|g_{k+1}\|^2 - t\, g_{k+1}^T s_k}{d_k^T y_k + \mu d_{k+1}^T g_k} \right) \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \qquad \text{……………….(21)}$$

Since $s_k = \lambda_k \, d_k$ and $y_k = g_{k+1} - g_k$ we get:

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \frac{\|g_{k+1}\|^2 - t\, \lambda_k g_{k+1}^T d_k}{d_k^T g_{k+1} - d_k^T g_k + \mu d_k g_{k+1}} \; \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \qquad \text{……………….(22)}$$

Since $d_k = -g_k$ and from eq.(15) we get:

$$\frac{g_{k+1}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \frac{\|g_{k+1}\|^2 - t\sigma\lambda_k \|g_k\|^2}{-\sigma\|g_k\|^2 + \|g_k\| - \sigma\mu\|g_k\|^2} \; \frac{\sigma\|g_k\|^2}{\|g_{k+1}\|^2} \qquad \text{……………(23)}$$

$$\frac{g_{k+1}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \left(\frac{\|g_{k+1}\|^2 - t\sigma\lambda_k \|g_k\|^2}{(1-\sigma-\sigma\mu)\|g_k\|^2} \right) \frac{\sigma\|g_k\|^2}{\|g_{k+1}\|^2} \qquad \text{……………….(24)}$$

Let $c = 1 - \sigma - \sigma\mu$, where c is positive constant ………….(25)

$$\frac{g_{k+1}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \left(\frac{\|g_{k+1}\|^2 - t\sigma\lambda_k \|g_k\|^2}{c\|g_k\|^2} \right) \frac{\sigma\|g_k\|^2}{\|g_{k+1}\|^2} \qquad \text{………….(26)}$$

$$\frac{g_{k+}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \frac{\sigma}{c} - \frac{t\sigma^2 \lambda_k \|g_k\|^2}{c\|g_{k+1}\|^2} \qquad \text{……………….(27)}$$

Since, $\|g_k\|^2 = \vartheta_1, \|g_{k+1}\|^2 = \vartheta_2$ and $\vartheta_1 / \vartheta_2 = \vartheta$, where $\vartheta, \vartheta_1, \vartheta_2$ are positive constants

$$\eta = \frac{\sigma}{c} - t\sigma^2 \lambda_k \vartheta \qquad \text{……………….(28)}$$

Now, we will summit eq.(28) in eq. (27),we get

$$\frac{g_{k+}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \eta$$

Now, we have two cases

1) If η is positive, we get

$$\frac{g_{k+}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le \eta \qquad \text{………….. (29)}$$

$$g_{k+1}^T d_{k+1} \le -\xi \|g_{k+1}\|^2 \text{, where } \xi = 1 - \eta \qquad \text{………… (30)}$$

2) If η is negative, we get

$$\frac{g_{k+}^T d_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \le -\eta \qquad \text{…………….(31)}$$

$$g_{k+1}^T d_{k+1} \le -\xi \|g_{k+1}\|^2 \text{, where } \xi = 1 + \eta \qquad \text{…...………. (32)}$$

Lemma (2): Suppose that the assumption (A) holds and considers any conjugate gradient methods (2) and (3), where d_k is a descent direction and λ_k is obtained by the strong Wolfe line search

If $\sum_{k\ge1} \frac{1}{\|d_k\|^2} = \infty$ Then, $\lim_{k\to\infty} \inf \|g_k\| = 0$.

(See [2])

Theorem (1)

Suppose that the assumption is (A) holds. Consider theْ iteration method in (2) where d_k is defined by (19) and β_k is defined in (18). Then, the new algorithm either stops at stationary point i.e. $\|g_k\| = 0$ or $\underset{k\to\infty}{Lim} \inf \|g_k\| = 0$

Proof

By Taking the absolute value of sides both of eq.(18) , we get:

$$\left|\beta_k^{MDY}\right| = \frac{\left|\|g_{k+1}\|^2 - t g_{k+1}^T s_k\right|}{\left|d_k^T y_k + \mu g_{k+1}^T d_k\right|} \le \frac{\|g_{k+1}\|^2}{\left|d_k^T y_k + \mu g_{k+1}^T d_k\right|} + \frac{\left|t g_{k+1}^T s_k\right|}{\left|d_k^T y_k + \mu g_{k+1}^T d_k\right|} \qquad \text{……….(33)}$$

Since $y_k = g_{k+1} - g_k$ and $s_k = \lambda_k d_k$

$$\therefore \left|\beta_k^{MDY}\right| \le \frac{\|g_{k+1}\|^2}{\left|d_k^T g_{k+1} - d_k^T g_k + \mu g_{k+1}^T d_k\right|} + \frac{t \lambda_k g_{k+1}^T d_k}{\left|d_k^T g_{k+1} - d_k^T g_k + \mu g_{k+1}^T d_k\right|} \qquad \text{…....(34)}$$

Since $d_k = -g_k$ and from strong Wolfe condition , we set:

$$\left|\beta_k^{MDY}\right| \le \frac{\|g_{k+1}\|^2}{\left|(1-\sigma-\sigma\mu)\|g_k\|^2\right|} + \frac{t \lambda_k \sigma \|g_k\|^2}{\left|(1-\sigma-\sigma\mu)\|g_k\|^2\right|} \qquad \text{..…(35)}$$

Summit eq.25, $\rho = t\lambda_k \sigma$ and $\vartheta_1 / \vartheta_2 = \vartheta$ in eq.35 ,we get

$$\left|\beta_k^{MDY}\right| \le \frac{1}{c\vartheta} + \frac{\rho}{\tau} \qquad \text{……(36)}$$

Let $c_1 = \frac{1}{c\vartheta} + \frac{\rho}{c}$

$$\left|\beta_k^{MDY}\right| \le c_1 \qquad \text{……(37)}$$

By taking the absolute value of sides both in eq. (19), we get

$$\|d_{k+1}\| = \|g_{k+1}\| + |\beta_k| \, \|d_k\| \qquad \text{.......(38)}$$

Since $s_k = \lambda_k d_k$, $\|g_{k+1}\| \prec \Gamma$ and $D = \{\|x - y\|, x, y \in S, where S \subseteq R\}$ we get

$$\|d_{k+1}\| \le \|g_{k+1}\| + \left|\beta_k^{New2}\right| \|s_k\| \le \Gamma + c_1 D \qquad \text{...... (39)}$$

$$\sum_{k \ge 1} \frac{1}{\|d_k\|^2} = \infty$$

by using lemma(2) we get $\lim_{k \to \infty} \|g_k\| = 0$.

4- New Scalar Modified DY Conjugate Gradient Method

In order to eliminate the truncation and rounding errors, the new scalar parameter is added to make the sequence and efficiency, as problem dimension, increase. Zhang (2006,2008) et al. [10, 12] made a scalar for modification FR method and proposed a descent FR method; in this paper, we define the new scalar DY CG algorithm as follows From three term CG algorithm, we get

$$d_{k+1} = -g_{k+1} + \beta_k^{MDY} d_k - \frac{g_{k+1}^T d_k}{d_k^T y_k + \mu \left| g_{k+1}^T d_k \right|} g_{k+1} \qquad \text{.........(40)}$$

Since $\beta_k^{MDY} \dfrac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k} = \dfrac{g_{k+1}^T d_k}{d_k^T y_k + \left|\mu g_{k+1}^T d_k\right|}$

$$\therefore d_{k+1} = -(1 + \beta_k^{MDY} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k}) g_{k+1} + \beta_k^{MDY} d_k \qquad \text{............ (41)}$$

Lemma (3):

Suppose the assumption (A) holds, let the sequence { x_k } generated by (2) and the step length λ_k satisfies Wolf conditions, then

$$g_{k+1}^T d_{k+1} \le -\delta \|g_{k+1}\|^2$$

where δ is a positive constant.

Proof:

The proof by induction is also

$$d_{k+1} = -(1 + \beta_k^{MDY} \frac{g_{k+1}^T d_k}{(\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k)}) g_{k+1}^T + \beta_k^{MDY} d_k \qquad \text{................(42)}$$

Multiplying both sides of the equation (42) by g_{k+1}^T and divide both sides by $\|g_{k+1}^T\|^2$,we get:

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = -\beta_k^{MDY} \frac{g_{k+1}^T d_k}{\|g_{k+1}^T\|^2 (\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k)} \|g_{k+1}^T\|^2 + \frac{\beta_k^{MDY} g_{k+1}^T d_k}{\|g_{k+1}^T\|^2} \qquad \text{.... (43)}$$

By using (15) we get:

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 \le -\beta_k^{MDY} \left(\frac{\sigma_k \|g_k\|^2}{\|g_{k+1}\|^2 + \sigma_k t\lambda_k \|g_k\|^2} + \frac{\sigma_k \|g_k\|^2}{\|g_{k+1}^T\|^2} \right) \qquad \text{...................(44)}$$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2}+1 \le -\beta_k^{MDY}\left(\frac{\sigma_k\|g_k\|^2\|g_{k+1}\|^2+\sigma_k\|g_k\|^2(\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2)}{\|g_{k+1}\|^2(\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2)}\right) \quad \text{..... (45)}$$

since $\beta_k^{MDY} = \frac{\|g_{k+1}\|^2 - tg_{k+1}^T s_k}{d_k^T y_k + \mu g_{k+1}^T d_k} = \frac{\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k}{d_k^T g_{k+1} - d_k^T g_k + \mu\, g_{k+1}^T d_k}$

……..(46) Summit eq.15 in eq.46 ,we get

$$\beta_k^{MDY} \le \frac{\|g_{k+1}\|^2 - t\lambda_k\sigma\|g_k\|^2}{c\|g_k\|^2} = \frac{\|g_{k+1}\|^2}{c\|g_k\|^2} - \frac{t\lambda_k\sigma\|g_k\|^2}{c\|g_k\|^2} \le \frac{\|g_{k+1}\|^2}{c\|g_k\|^2} \quad \text{………..(47)}$$

$$\therefore\ \beta_k^{MDY} \le \frac{\|g_{k+1}\|^2}{c\|g_k\|^2} \quad \text{………..(48)}$$

Summit (48) in eq. (45) ,we get

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

$$\frac{g_{k+1}^T d_{k+1}+1}{\|g_{k+1}\|^2} \le -\left(\frac{2\sigma_k\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2}{c(\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2)}\right) \quad \text{……………(49)}$$

Let $\rho = \left(\frac{2\sigma_k\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2)}{c(\|g_{k+1}\|^2+\sigma t\lambda_k\|g_k\|^2)}\right)$, where ρ is positive constant

$$\frac{g_{k+1}^T d_{k+1}+1}{\|g_{k+1}\|^2} \le -\rho \quad \text{…………………….(50)}$$

$\therefore\ g_{k+1}^T d_{k+1} \le -\delta\|g_{k+1}\|^2$, where δ is positive define as $\delta = 1+\rho$

Theorem (2)

consider the iteration method (1) where d_k Suppose the assumption A holds. Then, the new algorithm either stops at stationary point i.e.

$\|g_k\| = 0$ or $\lim\inf_{k\to\infty}\|g_k\| = 0$

Proof:

The proof is by contradiction i.e. if theorem is not true, then $\|g_k\| \ne 0$ then there exists a positive scalar ε such that:

$$\|g_k\| \ge \varepsilon\ , \qquad \forall\ k \quad \text{...(51)}$$

Let $\psi = (1+\beta_k^{MDY}\frac{g_{k+1}^T d_k}{(\|g_{k+1}\|^2 - t\lambda_k g_{k+1}^T d_k)})$

Therefore, eq.(41) becomes

$$f(x) = \sum_{i=1}^{n/4}\left(\exp(x_{4i-3})+10x_{4i-2}\right)^2 + 100(x_{4i-2}+x_{4i-1})^6 + \left(\tan(x_{4i-1}-x_{4i})\right)^4 + (x_{4i-3})^8 + (x_{4i}-1)^2,\ x_0 = (1,2,2,2,...)^T. \quad \text{.......(52)}$$

By taking the square of both sides of eq.(52),we get

$$\|d_{k+1}\|^2 = (\beta_k)^2\|d_k\|^2 - 2\psi\beta_k^{MDY} g_{k+1}^T d_k + \psi^2\|g_{k+1}\|^2 \qquad \text{.......... ..(53)}$$

$$\|d_{k+1}\|^2 = (\beta_k)^2\|d_k\|^2 - 2\psi\beta_k^{MDY} g_{k+1}^T (\frac{d_{k+1}+\psi g_{k+1}}{\beta_k^{MDY}}) d_k + \psi^2\|g_{k+1}\|^2 \qquad \text{.........(54)}$$

$$\|d_{k+1}\|^2 = (\beta_k)^2\|d_k\|^2 - 2\psi_k g_{k+1}^T d_{k+1} - 2\psi^2\|g_{k+1}^T\|^2 + \psi^2\|g_{k+1}\|^2 \qquad \text{.......... ....(55)}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} = (\beta_k)^2 \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{2\psi g_{k+1}^T d_{k+1}}{(g_{k+1}^T d_{k+1})^2} - \frac{\psi^2\|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \qquad \text{.......... ....(56)}$$

From Lemma (3) and (52) we get:

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|g_{k+1}\|^4}{c\|g_k\|^4}\frac{\|d_k\|^2}{\|g_{k+1}\|^4} + \frac{2\psi}{\|g_{k+1}\|^2} - \frac{\psi^2}{\|g_{k+1}\|^2} \qquad \text{.......... ....(57)}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|d_k\|^2}{c\|g_k\|^4} + \frac{2\psi}{\|g_{k+1}\|^4} - \frac{\psi^2}{\|g_{k+1}\|^2}. \qquad \text{.......... ....(58)}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|d_k\|^2}{c\|g_k\|^4} - \frac{1}{\|g_{k+1}\|^4}(\psi_k^2 - 2\psi_k + 1 - 1) \qquad \text{.......... ....(59)}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|d_k\|^2}{c\|g_k\|^4} - \frac{1}{\|g_{k+1}\|^4}(\psi_k - 1)^2 + \frac{1}{\|g_{k+1}\|^4} \qquad \text{.......... ....(60)}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|d_k\|^2}{c\|g_k\|^4} + \frac{1}{\|g_{k+1}\|^4} \qquad \text{.......... ....(61)}$$

$$\sum_{k\ge1} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \le \gamma \sum_{k\ge1} \frac{1}{\|g_{k+1}\|^2} \le \frac{k}{\varepsilon^2}$$

$$\sum_{k\ge1} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \gamma \sum_{k\ge1} \frac{1}{k} = \infty$$

This contradiction (51), the proof is complete

3.1 The New Algorithm :

Step 1: For the initial point, $x_0 \in R^n$, ε, Set $d_0 = -g_0$,k = 1 , if $\|g_0\| \le \varepsilon$, then stop.

Step 2: Set $d_k = -g_k$

Step 3: Find $\lambda_k > 0$ satisfying the Wolf conditions.

Step 4: Let $x_{k+1} = x_k + \lambda_k d_k$ and If $\|g_{k+1}\| \le \varepsilon$ then stop.

Step 5: Compute β_k by the formula (21) , then generate d_{k+1} by (43)

Step 6: If $k = n \ or \ \frac{\left| g_k^T g_{k+1} \right|}{\left\| g_{k+1} \right\|^2} \geq 0.2$,then go to step 2.

Step 7: Set k: = k+1, go to Step3.

5. Numerical Results

In order to assess the performance of the proposed algorithm; we tested suggestion algorithm over (10) non-linear unconstrained test functions (see appendix). All the results are obtained by using (Pentium 4 computer). All programs are written in FORTRAN 90 language and for all cases the stopping criterion taken to be: $\left\| g_{k+1} \right\| \leq 10^{-5}$

All the algorithms in this paper use the same ILS strategy($\sigma =$ 0.9)The comparative performance for all of these algorithms is evaluated by considering number of function Evaluation (*NOF*) and number of iteration (*NOI*).

Table (1) Comparison suggestion algorithm which is defined in eq.(19) with standard DY CG-algorithm and scale DY CG- algorithm which is in eq. (41)

Table (1)

Scale algorithm		**new algorithm**		**old algorithm**		**Dim.**	**Test fun.**
NOF	**NOI**	**NOF**	**NOI**	**NOF**	**NOI**		
132	27	157	27	159	27	100	Central
76	35	78	35	78	35	100	Wood
139	50	97	36	98	36	100	Powell
79	14	80	15	80	15	100	Sum
80	39	82	40	87	43	100	Wolfe
1079	460	1090	465	1024	470	100	Dixon
42	14	22	8	22	8	100	Fred
76	30	76	30	76	30	100	Rosen
48	23	45	21	45	21	100	Powell3
177	49	177	50	178	56	100	Miele
22	9	25	10	22	9	100	Shallow
160	29	169	28	171	28	1000	Central
67	31	80	36	80	36	1000	Wood
142	25	97	36	98	36	1000	Powell
90	20	98	22	98	22	1000	Sum
97	48	90	46	99	49	1000	Wolf
1015	340	1053	335	1156	351	1000	Dixon
45	16	20	7	22	8	1000	Fred
76	30	76	30	76	30	1000	Rosen
44	21	45	21	45	21	1000	Powell3

Scale algorithm		new algorithm		old algorithm		Dim.	Test fun.
128	43	238	62	260	70	1000	Miele
22	9	27	11	22	9	1000	Shallow
244	32	254	36	297	37	10000	Central
77	35	80	36	80	36	10000	wood
102	35	113	41	114	41	10000	Powell
92	19	119	30	176	40	10000	Sum
252	125	268	133	268	133	10000	Wolfe
1041	460	1050	483	1007	460	10000	Dixon
45	16	20	7	22	8	10000	Fred
76	30	76	30	76	30	10000	Rosen
43	21	45	21	45	21	10000	Powell3
202	54	302	70	310	60	10000	Miele
22	10	23	9	22	10	10000	Shallow
6031	2199	6272	2249	6413	2286		Total

APPENDIX

1. Generalized Cantreal Function:

$$f(x)=\sum_{i=1}^{n/4}\begin{bmatrix}(\exp(x_{4i-3})-x_{4i-2})^4+100(x_{4i-2}-x_{4i-1})^6+\\ \arctan(x_{4i-1}-x_{4i}))^4+x_{4i-3}\end{bmatrix}, x_0=(1,2,2,2,...)^T.$$

2. Extended Wood Function

$$f(x)=\sum_{i=1}^{n/4}(100(x_{4i-3}^2-x_{4i-2})^2+(x_{4i-3}-1)^2+90(x_{4i-1}^2-x_{4i})^2+$$

$$(1-x_{4i-1})^2+10.1((x_{4i-2}-1)^2+(x_{4i}-1)^2)+19.8(x_{4i-2}-1)(x_{4i}-1)),$$

$$x_0=(-3,-1,-3,-1,...,-3,-1,-3,-1)^T.$$

3. Generalized Powell Function:

$$f(x)=\sum_{i=1}^{n/4}\begin{bmatrix}(x_{4i-3}+10x_{4i-2})^2+5(x_{4i-1}-x_{4i})^2+\\ (x_{4i-2}-2x_{4i-1})^4+10(x_{4i-3}-x_{4i})^4\end{bmatrix}, x_0=(3,-1,0,1,...,3,-1,0,1)^T.$$

4. Sum of Quatrics (SUM) Function:

$$f(x)=\sum_{i=1}^{n}(x_i-i)^4 \quad , x_0=(1,...)^T.$$

5. Wolfe Function:

$$f(x)=[-x_1(3-x_1/2)+2x_2-1]^2+\sum_{i=1}^{n-1}\begin{bmatrix}[x_{i-1}-x_i(3-x_i/2+2x_{i+1}-1)]^2\\ +[x_{n-1}-x_n(3-x_n/2)-1]^2\end{bmatrix} \quad x_0=(-1,...)^T$$

6. Dixon Function

$$F(x)=(1-x_1)^2+(1-x_{10})^2+\sum_{i=2}^{n}(x_i^2-x_{i+1})^2, \; x_0=(-1;...)^T$$

7. Fred Function

$$f(x)=\sum_{i=1}^{n/2}(-13+x_{2i-1}+((5-x_{2i})x_{2i}-2)x_{2i})^2$$

$$+(-29+x_{2i-1}+((1+x_{2i})x_{2i}-14)x_{2i})^2, \; x_0=(30,3,...,30,3)^T.$$

8. Rosenborck Function:

$$f(x)=\sum_{i=1}^{n/2}100(x_{2i}-x_{2i-1}^2)^2+(1-x_{2i-1})^2 \quad x_0=(-1.2,1,...)^T.$$

9. Generalized Powell 3 Function:

$$f(x)=\sum_{i=1}^{n/3}\left\{3-\left[\frac{1}{1+(x_i-x_{2i})^2}\right]-\sin(\frac{\pi x_{2i}x_{3i}}{2})-\exp\left[-(\frac{x_i+x_{3i}}{x_{2i}}-2)^2\right]\right\},$$

$x_0=(0,1,2;...)^T$.

10.Generalized Shallow Function:

$$f(x) = \sum_{i=1}^{n/2} \left[x_{2i-1}^{2} - x_{2i} \right]^{2} + (1 - x_{2i-1})^{2}\ ,\ ,\ x_{0} = (-2, -2, \ldots)^{T}.$$

11. Generalized Miele Function:

$$f(x) = \sum_{i=1}^{n/4} \left(\exp\left(x_{4i-3}\right) + 10x_{4i-2}\right)^{2} + 100\left(x_{4i-2} + x_{4i-1}\right)^{6} + \left(\tan\left(x_{4i-1} - x_{4i}\right)\right)^{4}$$
$$+ \left(x_{4i-3}\right)^{8} + \left(x_{4i} - 1\right)^{2}, x_{0} = (1,2,2,2,\ldots)^{T}.$$

REFERENCES

[1] AL - Bayati, A.Y. and AL-Assady, N.H. (1986). "Conjugate gradient method" Technical Research report, No (1), school of computer studies, Leeds university.

[2] Dai,Y. and Yuan, Y. (1999), "A Nonlinear conjugate gradient method with a strong global convergence property",SIAM J. Optim.,10. ,177– 182.

[3] Fletcher, R., and Reeves,C., (1964) ," Function minimization by conjugate gradients", Comput. J. 7 ,149–154.

[4] Fletcher, R. (1987) ,"Practical Methods of Optimization", John Wiley & Sons, New York.

[5] Hestenes,M.R and Stiefel, E., (1952)," Method of conjugate gradient for solving linear equations", J. Res. Nat. Bur. Stand. 49, 409–436.

[6] Liu,Y. and Storey,C.,(1991) ," Efficient generalized Conjugate Gradient Algorithms", part1:theory,J. Optimizat .Theor conjugate Appl.69,129–137.

[7] Perry, A. (1978), "A modified conjugate gradient algorithms", Operations Research, Vol.(26), 1073–1078.

[8] Polak,E.,and Ribie`re, G., (1969) , "Note Sur la convergence de Directionsconjuge`es, Rev. Francaise Informat Recherche Operationelle" ,3e Anne`e 16 35–43.

[9] Zhang,L., Zhou,Li, , D.H. ,(2006),Nonlinear conjugate gradient methods for optimization problems, Ph.D. Thesis, College of Mathematics and Econometrics, Hunan University Changsha, China.

[10] Zhang,L., Zhou,Li, , D.H. ,(2006)," Global convergence of a modified Fletcher-Reeves conjugate Method with Armijo-type line search" Numer. Math. 104 ,561–572.

[11] Zhang, Li , Zhou ,Weijun,(2007) ," A descent nonlinear conjugate gradient methodfor large-scale unconstrained optimization Applied Mathematics and Computation 187 , 636–643.

[12] Zhang,L.,Weijun Zhou, (2008) , "Two descent hybrid conjugate gradient methods for optimization ",College of Mathematics and Computational Science, Journal of Computational and Applied Mathematics 216 ,251 –264.