New Scale Dai-Yau 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 conjugacy 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.

1. Introduction:

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

\[ \min \{ f(x); x \in \mathbb{R}^n \} \]  

(1)
where \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is a continuously differentiable function, bounded from below. As we know, for solving this problem, let \( x_0 \in \mathbb{R}^n \) 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

\[
x_{k+1} = x_k + \lambda_k d_k
\]

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

\[
d_{k+1} = -g_{k+1} + \beta_k d_k , \quad d_0 = -g_0
\]

\( \beta_k \in \mathbb{R} \) is known as a conjugate gradient coefficient, some well known formulas are given as follows:

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

\[
\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{d_k^T g_k} \quad \text{(Fletcher-Reeves (FR), [3],(1964))} \tag{5}
\]

\[
\beta_k^{RB} = \frac{-y_k^T y_k}{d_k^T g_k} \quad \text{(Ribière (PR),[8],(1969) -Polak(} \tag{6)
\]

\[
\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{d_k^T g_k} \quad \text{(Fletcher (CD),[4],(1987))} \tag{8}
\]

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

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

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

The line search in the conjugate gradient algorithms for \( \lambda_k \) computation is often based on the standard Wolfe conditions

The weak Wolfe-conditions:

\[
f(x_k + \lambda_k d_k) - f(x_k) \leq \delta \lambda_k g_k^T d_k \tag{11}
\]

\[
g(x_k + \lambda_k d_k)^T d_k \geq \sigma g_k^T d_k \tag{12}
\]

The strong Wolfe-conditions:

\[
f(x_k + \lambda_k d_k) - f(x_k) \leq \delta \lambda_k g_k^T d_k \tag{13}
\]

\[
|g(x_k + \lambda_k d_k)^T d_k| \leq -\sigma g_k^T d_k \tag{14}
\]

Eq.(14) can be written as

\[
\sigma g_k^T d_k \leq g_{k+1}^T d_k \leq -\sigma g_k^T d_k \tag{15}
\]

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

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
\[
\beta_{MFR}^{\mu} = \frac{\mu_1 \|s_{k+1}\|^2}{\mu_2 \|g_{k+1}^T d_k\| + \mu_3 \|g_k\|^2}, \quad \ldots \quad (16)
\]

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

\[
\beta_{MFR}^{\mu} = \begin{cases} 
\mu_1 g_{k+1}^T d_k + \|g_k\|^2, & \|s_{k+1}\|^2 > \|g_{k+1}^T g_k\| \\
0, & \text{otherwise}
\end{cases} \quad \ldots \quad (17)
\]

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:

\[
\beta_{MDY}^{\mu} = \begin{cases} 
\left\|g_{k+1}^T\right\|^2 - t g_{k+1}^T s_k \\
0, & \left\|g_{k+1}^T\right\|^2 > t g_{k+1}^T s_k
\end{cases} \quad \text{if} \quad \left\|g_{k+1}^T\right\|^2 > t g_{k+1}^T s_k \quad \ldots \quad (18)
\]

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 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 \( \Omega = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\} \).

(ii) In some neighborhood \( \Omega_0 \) of \( \Omega \), \( f \) is differentiable and it is gradient \( g(x) \) is Lipschitz continuous, namely, there exists a constant \( L > 0 \) such that \( \|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in \Omega_0 \). We can get from the assumption A that there exists positive constant \( \Gamma > 0 \), such that:

\[
\|g(x)\| \leq \Gamma, \quad \forall x \in \Omega
\]

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

\[
g_{k+1}^T d_{k+1} \leq -\zeta \|s_{k+1}\|^2
\]

where \( \zeta \) 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 \( d_k^T g_k < 0 \quad \forall k \)

Now, we prove if \( k = k + 1 \) then:

\[
d_{k+1} = -g_{k+1} + \beta_{MDY}^{\mu} d_k \quad \ldots \quad (19)
\]
Multiply sides both of eq. (19) by \( g_{k+1}^T \) we get:
\[
g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_k^{MDY} g_{k+1}^T d_k
\]

(20)

Dividing sides both of eq.(20) by \( \|g_{k+1}\|^2 \) we get:
\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \beta_k^{MDY} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}
\]

Since \( \beta_k^{MDY} \) is defined in (18), then we get
\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \left( \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{1}{\|g_{k+1}\|^2} \right) \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}
\]

(21)

Since \( s_k = \lambda_k d_k \) and \( y_k = g_{k+1} - g_k \) we get:
\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} + 1 = \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{\lambda_k g_{k+1}^T d_k}{\|g_{k+1}\|^2} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(22)

Since \( d_k = -g_k \) and from eq.(15) we get:
\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{\lambda_k}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(23)

\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \left( \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{\lambda_k}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \right) \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(24)

Let \( c = 1 - \sigma - \sigma \mu \), where \( c \) is positive constant

\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \left( \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{\lambda_k}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \right) \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(25)

\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \left( \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} - \frac{\lambda_k}{\|g_{k+1}\|^2} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \right) \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(26)

\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \frac{\sigma}{c} - \frac{\lambda_k}{c} \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2}
\]

(27)

Since, \( \|g_{k+1}\|^2 = \theta_1 \), \( \|g_{k+1}\|^2 = \theta_2 \) and \( \theta_1 / \theta_2 = \theta \), where \( \theta, \theta_1, \theta_2 \) are positive constants
\[
\eta = \frac{\sigma}{c} - \frac{\lambda_k}{c} \theta
\]

(28)

Now, we will summit eq.(28) in eq. (27),we get
\[
\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \eta
\]

Now, we have two cases
1) If $\eta$ is positive, we get

$$g_k^T d_{k+1} + \|g_{k+1}\|^2 \leq \eta$$

$$\left| g_k^T d_{k+1} \right| \leq -\xi \|g_{k+1}\|^2$$, where $\xi = 1 - \eta$ \hfill \ldots \ldots \ldots (29)

2) If $\eta$ is negative, we get

$$g_k^T d_{k+1} + \|g_{k+1}\|^2 \leq -\eta$$

$$\left| g_k^T d_{k+1} \right| \leq -\xi \|g_{k+1}\|^2$$, where $\xi = 1 + \eta$ \hfill \ldots \ldots \ldots (30)

**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 $\lambda_k$ is obtained by the strong Wolfe line search

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

(See [2])

**Theorem (1):** Suppose that the assumption is (A) holds. Consider the iteration method in (2) where $d_k$ is defined by (19) and $\beta_k$ is defined in (18). Then, the new algorithm either stops at stationary point i.e. $\|g_k\| = 0$ or $\liminf_{k \to \infty} \|g_k\| = 0$

**Proof:**

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

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

$$= \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

$$\leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

$$\leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

$$\leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

$$\leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

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

$$\therefore \beta_k^{MDY} \leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

$$\leq \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

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

$$\beta_k^{MDY} \leq \frac{\left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k}{\left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k} \right| + \left| g_{k+1} \right| d_k y_k + \mu g_{k+1} d_k$$

Summit eq.25, $\rho = \tau \lambda_k \sigma$ and $\beta \lambda = \beta$ in eq.35,we get

$$\beta_k^{MDY} \leq \frac{1}{c \beta} + \frac{\beta}{\tau}$$

Let $c_1 = \frac{1}{c \beta} + \frac{\beta}{\tau}$

$$\beta_k^{MDY} \leq c_1$$
By taking the absolute value of sides both in eq. (19), we get
\[ \|d_{k+1}\| = \|g_{k+1}\| + \|\beta_k\|\|d_k\| \] ……(38)

Since \( s_k = \lambda_k d_k \cdot \|g_{k+1}\| \propto \Gamma \) and \( D = \|x - y\|, x, y \in S, \text{where} S \subset \mathbb{R} \), we get
\[ \|d_{k+1}\| \leq \|g_{k+1}\| + \|\beta_k\|^2 \|s_k\| \leq \Gamma + c_1 \|D\| \] …… (39)
\[
\sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} = \infty
\]

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
\[ d_{k+1} = -g_{k+1} + \beta_k^{MDY} d_k - \frac{g_{k+1}^T d_k}{d_k^T y_k + \mu g_{k+1}^T d_k} g_{k+1} \]

Since \( \beta_k^{MDY} = \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2 - \lambda_k g_{k+1}^T d_k} = \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - \lambda_k g_{k+1}^T d_k} \)
\[ \therefore d_{k+1} = -(1 + \beta_k^{MDY}) \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - \lambda_k g_{k+1}^T d_k} g_{k+1} + \beta_k^{MDY} d_k \]

Lemma (3):
Suppose the assumption (A) holds, let the sequence \( \{x_k\} \) generated by (2) and the step length \( \lambda_k \) satisfies Wolf conditions, then
\[ g_{k+1}^T d_{k+1} \leq -\delta \|g_{k+1}\|^2 \]

where \( \delta \) is a positive constant.

Proof:
The proof by induction is also
\[ d_{k+1} = -(1 + \beta_k^{MDY}) \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - \lambda_k g_{k+1}^T d_k} g_{k+1} + \beta_k^{MDY} d_k \]

Multiplying both sides of the equation (42) by \( g_{k+1}^T \) and divide both sides by \( \|g_{k+1}\|^2 \), we get:
\[ \frac{g_{k+1}^T d_{k+1} + 1}{\|g_{k+1}\|^2} = -\beta_k^{MDY} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2 - \lambda_k g_{k+1}^T d_k} g_{k+1}^T + \beta_k^{MDY} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \]

By using (15) we get:
\[ \frac{g_{k+1}^T d_{k+1} + 1}{\|g_{k+1}\|^2} \leq -\beta_k^{MDY} \left( \frac{\sigma_k \|g_k\|^2}{\|g_{k+1}\|^2 + \sigma_k t \lambda_k \|g_k\|^2} + \frac{\sigma_k \|g_k\|^2}{\|g_{k+1}\|^2} \right) \]

………(44)
\[ g_k^T d_{k+1} \leq -\beta_k^{MDY} \left( \sigma_k \|g_k\|^2 + \sigma_k \|g_{k+1}\|^2 + \sigma \lambda_k \|g_k\|^2 \right) \] ..........................(45)

since

\[ \beta_k^{MDY} = \frac{\|g_{k+1}\|^2 - \sigma g_{k+1}^T s_k}{d_k^T g_{k+1} - \sigma g_{k+1}^T d_k} = \frac{\|g_{k+1}\|^2 - \sigma g_{k+1}^T d_k}{d_k^T g_{k+1} - \sigma g_{k+1}^T d_k} \]

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

\[ \beta_k^{MDY} \leq \frac{\|g_{k+1}\|^2}{c \|g_k\|^2} \] ..........................(47)

\[ \therefore \beta_k^{MDY} \leq \frac{\|g_{k+1}\|^2}{c \|g_k\|^2} \] ..........................(48)

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

\[ g_k^T d_{k+1} + 1 \leq \frac{\|g_{k+1}\|^2}{c \|g_k\|^2} \left( \|g_{k+1}\|^2 \sigma + \sigma \lambda_k \|g_k\|^2 \right) \]

\[ g_k^T d_{k+1} + 1 \leq \frac{\|g_{k+1}\|^2}{c \|g_k\|^2} \left( \|g_{k+1}\|^2 \sigma + \sigma \lambda_k \|g_k\|^2 \right) \]

Let \( \rho = \frac{2 \sigma \lambda_k \|g_k\|^2}{c \|g_k\|^2} \), where \( \rho \) is positive constant

\[ g_k^T d_{k+1} + 1 \leq -\rho \] .................................(50)

\[ \therefore g_k^T d_{k+1} \leq -\delta \|g_{k+1}\|^2 \], where \( \delta \) 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.

\[ \|g_k\| = 0 \text{ or } \lim_{k \to \infty} \inf \|g_k\| = 0 \]

Proof:

The proof is by contradiction i.e. if theorem is not true, then \( \|g_k\| \neq 0 \) then there exists a positive scalar \( \varepsilon \) such that:

\[ \|g_k\| \geq \varepsilon, \quad \forall \ k \] .................................(51)

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

Therefore, eq.(41) becomes

\[ f(x) = \sum_{i=4}^{41} \left( \exp(x_{i+3}) + 10x_{i+2} \right)^2 + 100(x_{i+2} + x_{i+1})^4 + (\tan(x_{i+1} - x_i))^4 + (x_{i+1})^2 \cdot x_0 = (1,2,2,2,\ldots)^T. \] .................................(52)
By taking the square of both sides of eq. (52), we get
\[
\|d_{k+1}\|^2 = (\beta_k)^2 \|d_k\|^2 - 2\psi \beta_k^{\text{MDY}} g_k^T g_{k+1}^T d_k + \psi^2 \|g_{k+1}\|^2
\]
\[
\|d_{k+1}\|^2 = (\beta_k)^2 \|d_k\|^2 - 2\psi \beta_k^{\text{MDY}} g_k^T (d_{k+1} + \psi g_{k+1}) d_k + \psi^2 \|g_{k+1}\|^2
\]
\[
\|d_{k+1}\|^2 = (\beta_k)^2 \|d_k\|^2 - 2\psi \beta_k^{\text{MDY}} g_k^T d_{k+1} - 2\psi^2 \|g_{k+1}\|^2 + \psi^2 \|g_{k+1}\|^2
\]
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} = (\beta_k)^2 \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - 2\psi \frac{g_k^T d_{k+1}}{(g_{k+1}^T d_{k+1})^2} - \psi^2 \frac{\|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}
\]
From Lemma (3) and (52) we get:
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{2\psi}{\|g_{k+1}\|^2} - \psi^2 \frac{\|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}
\]
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{2\psi}{\|g_{k+1}\|^2} - \psi^2 \frac{\|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}
\]
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{1}{\|g_{k+1}\|^4} (\psi_k^2 - 2\psi_k + 1 - 1)
\]
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{1}{\|g_{k+1}\|^4} (\psi_k - 1)^2 + \frac{1}{\|g_{k+1}\|^4}
\]
\[
\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^4}
\]
\[
\sum_{k=1}^n \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \leq \gamma \sum_{k=1}^n \frac{1}{\|g_{k+1}\|^2} \leq \frac{k}{\varepsilon^2}
\]
\[
\sum_{k=1}^n \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \geq \gamma \sum_{k=1}^n \frac{1}{k} = \infty
\]
This contradiction (51), the proof is complete.

3.1 The New Algorithm:
Step 1: For the initial point, \(x_0 \in R^n, \varepsilon\), Set \(d_0 = -g_0, k = 1\), if \(\|g_0\| \leq \varepsilon\), 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 \(\|g_{k+1}\| \leq \varepsilon\) then stop.
Step 5: Compute \(\beta_k\) by the formula (21), then generate \(d_{k+1}\) by (43)
New Scale Dai-Yaun Conjugate Gradient …

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

Step 7: Set \( k = k+1 \), go to Step 3.

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: \( \|g_{k+1}\| \leq 10^{-5} \)

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>
<thead>
<tr>
<th>Scale algorithm</th>
<th>new algorithm</th>
<th>old algorithm</th>
<th>Dim.</th>
<th>Test fun.</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOF</td>
<td>NOI</td>
<td>NOF</td>
<td>NOI</td>
<td>NOF</td>
</tr>
<tr>
<td>132</td>
<td>27</td>
<td>157</td>
<td>27</td>
<td>159</td>
</tr>
<tr>
<td>76</td>
<td>35</td>
<td>78</td>
<td>35</td>
<td>78</td>
</tr>
<tr>
<td>139</td>
<td>50</td>
<td>97</td>
<td>36</td>
<td>98</td>
</tr>
<tr>
<td>79</td>
<td>14</td>
<td>80</td>
<td>15</td>
<td>80</td>
</tr>
<tr>
<td>80</td>
<td>39</td>
<td>82</td>
<td>40</td>
<td>87</td>
</tr>
<tr>
<td>1079</td>
<td>460</td>
<td>1090</td>
<td>465</td>
<td>1024</td>
</tr>
<tr>
<td>42</td>
<td>14</td>
<td>22</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>76</td>
<td>30</td>
<td>76</td>
<td>30</td>
<td>76</td>
</tr>
<tr>
<td>48</td>
<td>23</td>
<td>45</td>
<td>21</td>
<td>45</td>
</tr>
<tr>
<td>177</td>
<td>49</td>
<td>177</td>
<td>50</td>
<td>178</td>
</tr>
<tr>
<td>22</td>
<td>9</td>
<td>25</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>160</td>
<td>29</td>
<td>169</td>
<td>28</td>
<td>171</td>
</tr>
<tr>
<td>67</td>
<td>31</td>
<td>80</td>
<td>36</td>
<td>80</td>
</tr>
<tr>
<td>142</td>
<td>25</td>
<td>97</td>
<td>36</td>
<td>98</td>
</tr>
<tr>
<td>90</td>
<td>20</td>
<td>98</td>
<td>22</td>
<td>98</td>
</tr>
<tr>
<td>97</td>
<td>48</td>
<td>90</td>
<td>46</td>
<td>99</td>
</tr>
<tr>
<td>1015</td>
<td>340</td>
<td>1053</td>
<td>335</td>
<td>1156</td>
</tr>
<tr>
<td>45</td>
<td>16</td>
<td>20</td>
<td>7</td>
<td>22</td>
</tr>
<tr>
<td>76</td>
<td>30</td>
<td>76</td>
<td>30</td>
<td>76</td>
</tr>
<tr>
<td>44</td>
<td>21</td>
<td>45</td>
<td>21</td>
<td>45</td>
</tr>
<tr>
<td>Scale</td>
<td>new_algorithm</td>
<td>old_algorithm</td>
<td>Dim.</td>
<td>Test fun.</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>---------------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td>128</td>
<td>43</td>
<td>238</td>
<td>62</td>
<td>1000</td>
</tr>
<tr>
<td>22</td>
<td>9</td>
<td>27</td>
<td>11</td>
<td>1000</td>
</tr>
<tr>
<td>244</td>
<td>32</td>
<td>254</td>
<td>36</td>
<td>10000</td>
</tr>
<tr>
<td>77</td>
<td>35</td>
<td>80</td>
<td>36</td>
<td>10000</td>
</tr>
<tr>
<td>102</td>
<td>35</td>
<td>113</td>
<td>41</td>
<td>10000</td>
</tr>
<tr>
<td>92</td>
<td>19</td>
<td>119</td>
<td>30</td>
<td>10000</td>
</tr>
<tr>
<td>252</td>
<td>125</td>
<td>268</td>
<td>133</td>
<td>10000</td>
</tr>
<tr>
<td>1041</td>
<td>460</td>
<td>1050</td>
<td>483</td>
<td>10000</td>
</tr>
<tr>
<td>45</td>
<td>16</td>
<td>20</td>
<td>7</td>
<td>10000</td>
</tr>
<tr>
<td>76</td>
<td>30</td>
<td>76</td>
<td>30</td>
<td>10000</td>
</tr>
<tr>
<td>43</td>
<td>21</td>
<td>45</td>
<td>21</td>
<td>10000</td>
</tr>
<tr>
<td>202</td>
<td>54</td>
<td>302</td>
<td>70</td>
<td>10000</td>
</tr>
<tr>
<td>22</td>
<td>10</td>
<td>23</td>
<td>9</td>
<td>10000</td>
</tr>
<tr>
<td>6031</td>
<td>2199</td>
<td>6272</td>
<td>2249</td>
<td>6413</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX

1. Generalized Cantreal Function:

\[
f(x) = \sum_{i=1}^{n} \left[ \left( \exp(x_{i-3}) - x_{4i-2} \right)^4 + 100(x_{4i-2} - x_{4i-1})^6 + \arctan(x_{4i-1} - x_{4i}) \right] + x_{4i-3},
\]

\[x_0 = (1,2,2,...)^T.\]

2. Extended Wood Function

\[f(x) = \sum_{i=1}^{n/4} \left( 100(x_{4i-3}^2 - x_{4i-2}^2) + (x_{4i-3} - 1)^2 + 90(x_{4i-1}^2 - x_{4i})^2 \right.
\]
\[+ (1 - x_{4i-1})^2 + 10.1((x_{4i-2} - 1)^2 + (x_{4i} - 1)^2) + 19.8(x_{4i-2} - 1)(x_{4i} - 1),
\]

\[x_0 = (-3,-1,-3,...,-3,-1)^T.\]

3. Generalized Powell Function:

\[
f(x) = \sum_{i=1}^{n} \left[ (x_{i-3} + 10x_{i-2})^2 + 5(x_{i-1} - x_{i})^2 + (x_{i-2} - 2x_{i-1})^4 + 10(x_{i-3} - x_{i})^4 \right],
\]

\[x_0 = (3,-1,0,1,...,3,-1,0,1)^T.\]

4. Sum of Quatrics (SUM) Function:

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

5. Wolfe Function:

\[
f(x) = \left[ -x_1 (3 - x_1/2) + 2x_2 - 1 \right]^2 + \sum_{i=1}^{n-1} \left[ \left( x_{i-1} - x_i \right) (3 - x_i/2 + 2x_{i+1} - 1) \right]^2
\]

\[x_0 = (-1, \ldots)^T.\]

6. Dixon Function

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

7. Fred Function

\[f(x) = \sum_{i=1}^{n/2} (-13 + x_{2i-1} + ((5 - x_{2i})x_{2i} - 2)x_{2i})^2
\]
\[+ (-29 + x_{2i-1} + ((1 + x_{2i})x_{2i} - 14)x_{2i})^2, \quad x_0 = (30,3,..,30,3)^T.\]

8. Rosenborck Function:

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

9. Generalized Powell 3 Function:

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

\[x_0 = (0, 1, 2, \ldots)^T.\]
10. Generalized Shallow Function:

\[ f(x) = \sum_{i=1}^{n/2} \left[ x_{2i-1}^2 - x_{2i} \right]^2 + (1 - x_{2i-1})^2 , \quad x_0 = (-2, -2, \ldots)^T. \]

11. Generalized Miele Function:

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


