A New Separable Logarithmic Algorithm for Non-Linear Optimization

Abbas Y. Al-Bayati and Hawraz N. Jabbar
1Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Iraq
2Department of Mathematics, College of Education, Suleim University, Iraq

Abstract: Problem statement: The idea of this study stemmed from the fact that most of the currently used optimization algorithms use a local quadratic representation of the objective function. It also arisen from the fact that the objective function may not be represented adequately by quadratic functions and the global minimizer may be obtained for objective functions. So, in this study, we generalized the field of quadratic model into the field of the non-quadratic model. Approach: A new non-quadratic model was suggested for solving unconstrained optimization problems, which modified the classical Conjugate Gradient (CG) algorithm by scaling the standard quadratic model. Results: The new algorithm was derived and evaluated theoretically and numerically for some standard well-known and effective test functions. The results, in general, indicated that the new algorithm had improvements on different well-known algorithms used in this study. Conclusion: The new proposed algorithm would be generic and easy to implement in all gradient based optimization process. Its simulation results showed that it was robust and had a potential significantly enhance the computational efficiency of iterations and function evaluations.

Key words: Unconstrained minimization, conjugate-gradient algorithm, rational models, global convergence, exact line searches

INTRODUCTION

This study considers the calculation of a local minimizer \( x^* \) say, for the problem

\[
\text{Min} f(x), \quad x \in \mathbb{R}^n, \quad f \in C^2
\]  

It is assumed that expressions for the elements of the gradient vector \( g(x) = \text{gradient of } f(x) \) are available are calculated but the Hessian matrix is not available. Standard algorithms for solving this problem include CG-algorithms which are iterative algorithms and generate a sequence of approximations to minimize a function \( f(x) \). This type of algorithms requires \( 4n \) locations of computer storage to implement. However, this study considers a more general model than the usual quadratic function

\[
F(x) = \frac{1}{2} x^T G x + b^T x
\]  

Definition 1: A set of vectors \( d_i \) in \( \mathbb{R}^n \) is said to be conjugate with respect to the symmetric positive definite matrix \( G \) if and only if \( d_i^T G d_j = 0 \) for \( i \neq j \).

Hestenes and Stiefel (1952) published the first CG-algorithm, for solving a system of linear algebraic equations. Fletcher and Reeves, (1964) were the first, among other scholars, to use this technique to minimize a non-linear function of several variables. Since then this algorithm has been used successfully to tackle many problems.

We can list outline of the CG-algorithm as follows:

Outline of the standard CG-algorithm: for given TOL and \( x_0 \in \mathbb{R}^n \) an initial estimate of the minimizer \( x^* \):

Step 1: Set \( d_0 = -g_0 \)

Step 2: For \( i = 1, 2, \ldots \), compute \( x_i = x_{i-1} + \lambda_i d_{i-1} \)

where \( \lambda_i \) is the optimal step-size obtained by a line search procedure.

Step 3: If \( \|s_i\| \leq \text{Tol} \) stop; otherwise continue.

Step 4: Calculate the new direction \( d_i = -g_i + \beta_i d_{i-1} \)

where \( \beta_i \) is the conjugacy coefficient.

It is considered as one of the following formulas:

Corresponding Author: Abbas Y. Al Bayati, Department of Mathematics, College of Computers Sciences and Mathematics, Mosul University, Iraq

498
\[ \beta_i = \frac{g_i^T (g_i - g_{i-1})}{g_{i-1}^T (g_i - g_{i-1})} \]  \hspace{1cm} (3a)

Called: Hestenes and Stiefel (HS):

\[ \beta_i = \frac{\|g_i\|^2}{\|g_{i-1}\|^2} \]  \hspace{1cm} (3b)

Called: Fletcher and Reeves (FR):

\[ \beta_i = \frac{g_i^T (g_i - g_{i-1})}{(g_i^T g_{i-1})} \]  \hspace{1cm} (3c)

Called: Polak and Ribiere (PR):

\[ \beta_i = \frac{\|g_i\|^2}{\|d_{i-1}^T g_i\|^2} \]  \hspace{1cm} (3d)

Called: Descent Dixon (DD):

\[ \beta_i = -\frac{g_i^T (g_i - g_{i-1})}{d_{i-1}^T g_{i-1}} \]  \hspace{1cm} (3e)

Called: Al-Bayati and Al-Assady (BA).

For the details of these formulas (Fletcher, 1987; Al-Bayati and Al-Assady, 1994).

When, quadratic functions and exact line searches are used, all the above formulas for \( \beta_i \) s are equivalent. However, these formulas vary according to general functions.

Several algorithms have been suggested as alternative ways of modifying the classical CG-algorithms (Al-Bayati, 1993; 2001; 2007; Al-Bayati and Al-Assady, 1994; 1997; Al-Bayati and Al-Naemi, 1995; Al-Bayati and Rassam, 2009).

CG-algorithms have, in general, the following:

**BASIC properties:**

- The conjugacy condition
- The orthogonality condition
- The descent direction
- The quadratic termination condition with Exact Line Search (ELS)

**MATERIALS AND METHODS**

**Extended CG-algorithms for non-quadratic models:**

In this study, a more general model than quadratic one is suggested as a basis for a CG-algorithm. If \( q(x) \) is a quadratic function, then a function \( F(q(x)) \) is defined as a non-linear scaling of \( q(x) \) if the following condition holds:

\[ f(x) = F(q(x)), \quad \frac{df}{dq} = f' > 0 \quad \text{and} \quad q(x) > 0 \]  \hspace{1cm} (4)

where, \( x^* \) is the minimizer of \( q(x) \) with respect to \( x \), (Spedicato, 1976).

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

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

Boland et al. (1979) was the first who observed that \( q(x) \) and \( f(q(x)) \) have determined the same search directions so that the finite termination property for their algorithm was satisfied. A CG-algorithm which minimizes the function

\[ f(x) = q(q(x))^p, \quad p > 0 \quad \text{and} \quad x \in \mathbb{R}^n \]  \hspace{1cm} (5a)

in at most \( n \) step was described by (Fried, 1971).

The special case:

\[ F(q(x)) = \epsilon_1 q(x) + \frac{1}{2} \epsilon_2 q'(x) \]  \hspace{1cm} (5b)

where, \( \epsilon_1 \) and \( \epsilon_2 \) are scalars, was investigated by (Boland et al., 1979).

(Tassopoulos and Storey, 1984) proposed two different rational models defined by:

\[ F(q(x)) = \frac{\epsilon_1 q(x) + 1}{\epsilon_2 q(x)}, \quad \epsilon_2 < 0 \quad \text{and} \quad q(x) > 0 \]  \hspace{1cm} (5c)

where, \( \epsilon_1 \) and \( \epsilon_2 \) are scalars and:

\[ F(q(x)) = \frac{\epsilon q(x)}{1 + q(x)}, \quad \epsilon > 0 \quad \text{and} \quad q(x) > 0 \]  \hspace{1cm} (5d)

Al-Bayati (1993) proposed another rational model defined by:

\[ F(q(x)) = \frac{\epsilon q(x)}{1 - \epsilon q_1(x)}, \quad \epsilon_1 > 0 \quad \text{and} \quad \epsilon_2 < 0 \]  \hspace{1cm} (5e)

In this study, a new logarithmic model is investigated and tested on a set of several standard test functions, on the assumption that condition (1) holds. An extended CG-algorithm is developed which is based on this new model and scales \( q(x) \) by the natural log function for the rational \( q(x) \) functions:
\[ F(q(x)) = \log \left( \frac{e_{q(x)}}{1 - e_{q(x)}} \right), \quad e_{\cdot} < 0 \quad (6) \]

We first observe that \( q(x) \) and \( F(q(x)) \) given by (6) have identical contours, though with different function values and they have the same unique minimum point denoted by \( x^* \).

**New ECG-algorithm for non-quadratic model:** For any \( f \) satisfying the condition (4), it is shown in (Boland et al., 1979) that the updating process given below generates identical conjugate directions and the same sequence of approximations \( x_i \) to the minimizer \( x^* \).

1. Define: \( n = -(\lambda_{i-1} s_{i-1} d_{i-1}) / 2, \quad w = F_i - F_{i-1} \), \( F_i = F(q(x_i)) \) and compute \( (n + w) \).
2. Compute \( d_i = -g_i + \rho_i \langle g_i, d_{i-1} \rangle d_i \).
3. Compute \( \rho_i = \min(n/(2n w), 1) \).
4. Calculate the new direction \( d_i = -g_i + \rho_i \langle g_i, d_{i-1} \rangle d_i \).
5. Calculate the new direction \( d_i = -g_i + \rho_i \langle g_i, d_{i-1} \rangle d_i \).

**Outline of the new ECG-algorithm:**

1. Set \( d_0 = -g_0 \).
2. For \( i = 1, 2, \ldots \)
   - Compute \( x_i = x_{i-1} + \lambda_{i-1} d_{i-1} \) where \( \lambda_{i-1} \) is the optimal step-size obtained by Armijo linear search procedure: set scalars \( s_k, \beta, \sigma \) with:
     \[ s_k = \frac{g_i^T d_k}{\|d_k\|^2}, \quad \beta \in (0, 1) \]
   - And \( \sigma \in (0, \frac{1}{2}) \) and we set \( \lambda_k = \beta^m s_k \), where \( m_k \) is the 1st non-negative integer \( m \) for which:
     \[ w = \exp(f_i) - \exp(f_{i-1}) \]
Step 3: Calculate $f$ which is a function value of an actual iterate as:
\[
\exp(f) = 1 + f + f^2/2! + \ldots
\]
and define:
\[
n = \lambda_{i+1} \left\| g_{i+1} \right\| / 2.
\]
\[
w = \exp(f_i) - \exp(f_{i+1}).
\]
\[
c = w - n \exp(f_{i+1}).
\]
Step 4: If $|w| < 0.1 \times 10^{-5}$ or $|c| < 0.1 \times 10^{-5}$, then set $i = 0, 1$ are tried successively until the inequality above is satisfied for $m = m_i$.

Step 5: Compute:
\[
\beta_i = \rho_i \left( \left\| g_i \right\|^2 / \left\| g_{i-1} \right\|^2 \right)
\]
where, $\rho_i$ is defined by different modified FR, HS, PR, DD and BA respectively. CG-algorithms are usually implemented by restarts in order to avoid an accumulation of errors affecting the search directions. It is therefore generally agreed that restarting is very helpful in practice, so we have used the following restarting criterion in our practical investigations. If the new direction satisfies:
\[
d_i \cdot g_i \geq -0.8 \left\| g_i \right\|^2
\]
then a restart is also initiated and the new direction is sufficiently downhill.

### The derivation of the new ECG-algorithm

The implementation of the extended CG-algorithm has been performed for general functions $F(q(x))$ of the form of (13). The unknown quantities $\rho_i$ were expressed in terms of available quantities of the algorithm (i.e., function and gradient values of the objective function).

It is first assumed that neither $\varepsilon_1$ nor $\varepsilon_2$ is zero in Eq. 6. Solving (6) for $q(x)$, yields:
\[
q(x) = \frac{\exp(f_i)}{\varepsilon_1 (\exp(f_i) + \varepsilon_i / \varepsilon_2)}
\]
and using the expression for $\rho_i$ (henceforth $\rho^{\text{new}}_i$)
\[
\rho^{\text{new}}_i = \frac{f'_i}{f'_i} \left( \frac{\exp(f_i)}{\exp(f_i) + \varepsilon_i / \varepsilon_2} \right)^2
\]
the quantity that has to be determined explicitly is $(\varepsilon_i / \varepsilon_2)$.

During every interaction $(\varepsilon_i / \varepsilon_2)$ must be evaluated as a function of known available quantities from the relation:
\[
g_{i-1}^T = f'_{i-1} G(x_{i-1} - x^*)
\]
where, $G$ is the Hessian Matrix and $x^*$ is the minimum point. Now, we have:
\[
g_{i-1}^T (x_i - x^*) = g_{i-1}^T (x_{i-1} + \lambda_{i-1} d_{i-1} - x^*)
\]
\[
= g_{i-1}^T (x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1}
\]
\[
g_{i}^T (x_{i+1} - x^*) = g_{i}^T (x_i - \lambda_{i+1} d_{i+1} - x^*) = g_{i}^T (x_i - x^*)
\]

Since $g_i^T d_{i+1} = 0$, therefore, we can express $\rho^{\text{new}}_i$ as follows:
\[
\rho^{\text{new}}_i = (g_{i-1}^T (x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1}) / (g_i^T (x_i - x^*))
\]
From (14) and (15), we get:
\[
\rho^{\text{new}}_i = \frac{f'_{i-1} (x_{i-1} - x^*)^T G(x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1}}{f'_{i} (x_i - x^*)^T G(x_i - x^*)}
\]
Therefore:
\[
\rho^{\text{new}}_i = (2f'_{i-1} q_{i-1} + \lambda_{i-1} g_{i-1}^T d_{i-1}) / 2f'_i q_i
\]
\[
= \rho^{\text{new}}_i (q_{i-1} / q_i) + \lambda_{i-1} g_{i-1}^T d_{i-1} / (2f'_i q_i)
\]
The quantities $(q_{i-1} / q_i)$ and $f'_i q_i$ can be rewritten as:
\[
(q_{i-1} / q_i) \left( (\sqrt{\rho^{\text{new}}_i}) / \sqrt{\exp(f_{i-1}) / \exp(f_i)} \right)
\]
\[
 f_{i+1} = \frac{\exp(f_i) + \varepsilon_i / \varepsilon_2}{\varepsilon_i / \varepsilon_2} 
\]  \hspace{1cm} (19)

Substituting (18) and (19) in (17), gives:

\[
\rho_i^{\text{NEW}} = \left(\frac{\rho_i}{\rho_i^{\text{NEW}}}ight) \left(\frac{\exp(f_{i+1})}{\exp(f_i)}\right) + 
\left(\frac{\lambda - \varepsilon_i / \varepsilon_2}{\exp(f_i) + \varepsilon_i / \varepsilon_2}\right) \left(\frac{\varepsilon_i / \varepsilon_2}{\exp(f_i) + \varepsilon_i / \varepsilon_2}\right)
\]  \hspace{1cm} (20)

Using the transformation:

\[
\lambda_{i-1}g_i^T d_{i-1} = 2n
\]  \hspace{1cm} (21)

and from RHS of (13) ; (20) and (21), we have:

\[
\frac{(\exp(f_{i+1})) - \exp(f_i) + \varepsilon_i / \varepsilon_2}{\exp(f_i) + \varepsilon_i / \varepsilon_2} = \frac{\exp(f_{i+1})}{\exp(f_i)}
\]  \hspace{1cm} (22)

Now we are prove the global convergence of the new extended CG-algorithm as follows: we first assume that:

**Assumption 1:** The function \( f \) has lower bound on \( L_0 = \{x \in \mathbb{R}^n \} \), where \( x_0 \in \mathbb{R}^n \) is available.

**Assumption 2:** The gradient \( g = \nabla f(x) \) is Lipschitz continuous in an open convex set \( B \) which contains \( L_0 \), i.e., \( \exists L > 0 \) such that \( \|g(x) - g(y)\| \leq L \|x - y\| \) for all \( x, y \in B \).

**Theorem 2:** If Assumptions 1; 2 and \( \frac{g_i^T d_k}{\|d_k\|^2} < 0 \) holds, then for Armijo line search procedure the new extended CG-algorithm generates an infinite sequence \( \{x_k\} \) that satisfy \( \lim_{k \to \infty} \frac{\|d_k\|^2}{\|d_k\|^2} = 0 \).

**Proof:** For the new algorithm set \( K_1 = \{k | \lambda_k = \lambda_k \} \), \( K_2 = \{k | \lambda_k < \lambda_k \} \) by (9b) we have:

\[
f_k - f_{k-1} \geq -\lambda_k \sigma g_i^T d_k, \forall k \in K_1
\]  \hspace{1cm} (25)

\[
f_k - f_{k-1} \geq -\alpha_i \sigma g_i^T d_k, \forall k \in K_2
\]  \hspace{1cm} (26)

By line search rule (9), since \( \lambda_k / \beta \leq s_k, \forall k \in K_2 \), we have:

\[
f_k - f_{k-1} \geq -\lambda_k \sigma g_i^T d_k, \forall k \in K_2
\]  \hspace{1cm} (27)

Using the mean value theorem on the LHS of the above inequality, there exists \( \theta_k \in [0,1] \) such that:

\[
-\lambda_k \sigma g(x_k + \lambda_k \theta_k d_k / \beta) g_i^T d_k / \beta \geq -\lambda_k \sigma g_k^T d_k, \forall k \in K_2
\]  \hspace{1cm} (28)

Thus:

\[
\lambda_k \geq \frac{\beta(1-\sigma)g_i^T d_k}{L \|d_k\|^2}, \forall k \in K_2
\]  \hspace{1cm} (28)
It follows from (26) and (28):

$$f_k - f_{k+1} > \frac{\beta \sigma (1 - \sigma)}{L} \left( - \frac{g_k^T d_k}{\| d_k \|^2} \right)^2, \ k \in K_i$$  \hspace{1cm} (29)

By (25) and the definition of $s_k$, we have:

$$f_k - f_{k+1} \geq \sigma \left( - \frac{g_k^T d_k}{\| d_k \|^2} \right)^2, \ k \in K_i$$  \hspace{1cm} (30)

Set $\eta = \min \{ \sigma, \beta \sigma (1 - \sigma) / L \}$, by (29) and (30), we have:

$$f_k - f_{k+1} \geq \eta \left( - \frac{g_k^T d_k}{\| d_k \|^2} \right)^2$$  \hspace{1cm} (31)

By the first scaling property and $d_k^T g_k < 0$, it follows that:

$$\lim_{k \to \infty} \left( - \frac{g_k^T d_k}{\| d_k \|^2} \right)^2 = 0$$

**Theorem 3:** Suppose that $x_i$ is an arbitrary point for which Assumptions 1 and 2 hold. For the new extended CG-algorithm assume that the line search satisfies the conditions (9a) and (9b). Then if $g_k^T d_k < 0 \ \forall k \geq 1$ we have:

$$\lim_{k \to \infty} \| g_k \| = 0$$  \hspace{1cm} (32)

**Proof:** Suppose that the theorem is not true, then $\exists$ a constant $\alpha > 0$ such that:

$$\| g_k \| \geq \alpha \ \forall k \geq 1$$  \hspace{1cm} (33)

Now, for our new search direction defined in Step 6 of the new extended algorithm we have:

$$\left( \frac{\| g_k \|^2}{\| d_k \|^2} \right) \leq \left( \frac{\| d_{k-1} \|^2}{\| d_{k-1} \|^2} \right) + \frac{1}{\| g_k \|}$$  \hspace{1cm} (34)

Now, for any $k \geq 1$:

$$\left( \frac{\| g_k \|^2}{\| d_k \|^2} \right) \leq \left( \frac{\| d_{k-1} \|^2}{\| d_{k-1} \|^2} \right) + \frac{1}{\alpha^2}$$  \hspace{1cm} (35)

$$\Rightarrow \sum_{k=1}^{\infty} \left( \frac{g_k^T d_k}{\| d_k \|^2} \right)^2 = \infty$$  \hspace{1cm} (36)

which is a contradiction. Hence, the theorem must be true and the algorithm has super-linear convergence.

**RESULTS**

In order to test the effectiveness of the new algorithm which has been used to extend the standard CG-algorithm, the comparative tests involve several standard; effective and well-known test functions. These test functions have been chosen and solved numerically by utilizing the new and established algorithms.

Table 1-4 utilize the comparisons between our new extended CG-algorithm which is corresponding to the new non-quadratic model represented in (6), denoted by (NEW), the classical Fletcher and Reeves CG-algorithm, denoted by (FRCG), the rational model of Tassopoulos and Storey, denoted by (TS); Al-Bayati rational model, denoted by (B) for low and high dimensions. The cubic line search (Bunday, 1984), Armijo line search (Al-Bayati and Rassam, 2009) and the well-known convergence criterion $\| g_k \| < 1 \times 10^{-5}$ were used.

**Table 1:** Comparisons of different algorithms for non-quadratic models $2 \leq n \leq 10$ (low dimensions)

| Test | FRCG | TS | B | NEW |
|------|------|----|---|-----|
| fun | N | NOI | NOF | NOI | NOF | NOI | NOF | NOI | NOF |
| 1   | 2  | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 2   | 10 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 3   | 4  | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 4   | 20 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 5   | 40 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| Total | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |

**Table 2:** Comparison of different algorithms for non-quadratic models $20 \leq n \leq 80$ (low dimensions)

| Test | FRCG | TS | B | NEW |
|------|------|----|---|-----|
| fun | N | NOI | NOF | NOI | NOF | NOI | NOF | NOI | NOF |
| 1   | 20 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 2   | 20 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 3   | 40 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 4   | 80 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 5   | 80 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| Total | 386 | 836 | 284 | 643 | 210 | 535 | 229 | 528 |

**Table 3:** Comparisons of different algorithm for non-quadratic model $100 \leq n \leq 400$ (high dimensions)

| Test | FRCG | TS | B | NEW |
|------|------|----|---|-----|
| fun | N | NOI | NOF | NOI | NOF | NOI | NOF | NOI | NOF |
| 1   | 100 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 2   | 200 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 3   | 300 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 4   | 400 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| 5   | 500 | 188 | 472 | 169 | 430 | 264 | 775 | 155 | 389 |
| Total | 572 | 1248 | 425 | 974 | 242 | 686 | 288 | 644 |
Specifically, we record the Number Of Function calls (NOF), and the Number Of Iteration calls (NOI) for the purpose of our comparisons. For overcoming the overflow in our computer programs we have scaled the function values by a parameter equal to 1.E-10.

It is obvious that the new algorithm, for this set of low dimensionality test functions, improves:

- FRCG in (18%) NOI and (20%) NOF
- TS in (9%) NOI and (12%) NOF
- B in (40%) NOI and (50%) NOF

Also, the new algorithm, for the high dimensions, beats:

- FRCG in (41%) NOI and (37%) NOF
- TS in (20%) NOI and (18%) NOF
- B in (0%) NOI and (2%) NOF

In Table 4, taking the NOI and NOF 100%, we can determine the performance range of the new extended CG-algorithm according to others used in this study. Clearly it has (-15-98%) improvements in the NOI and it has about (6-93%) improvements in NOF.

We conclude that our new proposed rational logarithmic model is superior to some models in both quadratic and non-quadratic models.

The details of used test functions in this study are given in (Al-Bayati, 2007) and they are:

**Rosen-Brock function:**

\[
 f(x) = 100(x_2 - x_1^2)^2 + (1-x_1)^2, \quad x_0 = [-1.2, 1]^T
\]

**Dixon function:**

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

**Generalized Wood function:**

\[
 f(x) = \sum_{i=1}^{n/4} \left[ 100(x_{4i-3}^2 - x_{4i-2}^2)^2 + (1-x_{4i-3})^2 
+ 90(x_{i}^2 - x_{i-1}^2)^2 + (1-x_{i})^2 
+ 10.1(x_{4i-2} - 1)^2 + (x_{i} - 1)^2 
+ 19.8(x_{4i-2} - 1)(x_{i} - 1) \right], \quad x_0 = [-3, -1, -3, -1, \ldots, -3, -1, -3, -1]^T
\]

**Generalized Powell function:**

\[
 f(x) = \sum_{i=1}^{n} \left[ (x_{4i-3} - 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 
+ (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \right], \quad x_0 = [3, -1, 0, 1, \ldots, 3, -1, 0, 1]^T
\]

**Generalized Miele Function:**

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

**Non-diagonal function:**

\[
 f(x) = \sum_{i=1}^{n} \left[ 100(x_1 - x_i)^2 + (1 - x_i)^2 \right], \quad x_0 = [-1, \ldots, -1]^T
\]

**Oren and Spedicato OSP Function:**

\[
 f(x) = \left( \sum_{i=1}^{n} i(x_i)^2 \right)^2, \quad x_0 = [1, \ldots, 1]^T
\]

**DISCUSSION**

In this study, we have introduced a new non-quadratic CG-algorithm which is based on log of two rational functions. The new algorithm is compared with three well-known algorithms; namely FRCG; TS, (Tassopoulos and Storey, 1984) and B; (Al-Bayati, 1993) algorithms using seven well-known non-linear test functions with different dimensions taken from (Al-Bayati, 2007). Our numerical results indicate that the new technique has an improvements of about (-15-98%) in NOI against the others, while it saves about (6-93%) in NOF against the others.

**CONCLUSION**

In this research, a new fast extended CG algorithm is introduced. The proposed algorithm improved the standard FRCG; TS and B algorithms by adaptively modifying the search direction. The new proposed algorithm is generic and easy to implement in all
The simulation results showed that it is robust and has a potential significantly enhance the computational efficiency of iterations and function evaluations.

REFERENCES

Al-Bayati, A.Y., 1993. A new non-quadratic model for unconstrained non-linear optimization. Mu’tah J. Res. Stud. Nat. Applied Sci. Ser., 8: 131-155.
Al-Bayati, A.Y., 2001. New generalized CG-methods for the non-quadratic model in unconstrained optimizations. J. Al-Yarmouk, 10: 9-25.
Al-Bayati, A.Y., 2007. New extended CG-methods for combined non-quadratic models in nonlinear optimization. J. Al-Yarmouk, 16: 347-366.
Al-Bayati, A.Y. and N.H. Al-Assady, 1994. Minimization of extended quadratic functions with inexact line searches. JOTA, 82: 139-147. DOI: 10.1007/BF02191784
Al-Bayati, A.Y. and N.H. Al-Assady, 1997. PCG-method for nonlinear unconstrained optimization. Mu’tah J. Res. Stud. Nat. Applied Science Ser., 12:13-29.
Al-Bayati, A.Y. and G.M. Al-Naemi, 1995. New extended CG-methods for non-linear optimization. Mu’tah J. Res. Stud. Nat. Applied Science Ser., 10: 69-87.
Al-Bayati A. Y. and S.Y. Rassam, 2009. A modified globally convergence CG-algorithm with a hybrid line search procedure. Aust. J. Basic Applied Sci., 3: 2188-2206. http://www.insipub.com/ajbas/2009/2188-2206.pdf
Boland, W.R. E.R. Kamgnia and J.S. Kowalik, 1979. A conjugate gradient method invariant to non-linear scaling. J. Optimiz. Theory Appli., 27: 221-230. DOI: 10.1007/BF00933228
Bunday, B., 1984. Basic Optimization Methods. 1st Edn., Edward Arnold, Bedford Square, London, ISBN: 0713135069, pp: 24-28.
Fletcher, R. and C.M. Reeves, 1964. Function minimization by conjugate gradients. Computer J., 7: 149-154. DOI: 10.1093/comjnl/7.2.149
Fletcher, R., 1987. Practical Methods of Optimization. 2nd Edn., John Wiley and Son, Toronto and Singapore, New York, ISBN: 10: 0471494631, pp: 80-87.
Fried, I., 1971. N-step conjugate gradient minimization scheme for non-quadratic functions. AIAA J., 9: 2286-2287. DOI: 10.2514/3.6507
Hestenes, M.R. and E. Stiefel, 1952. Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bureau of Stand., 5: 409-436. http://nvl.nist.gov/pub/nistpubs/jres/049/6/V49.N06.A08.pdf
Renpu, G., 1989. A Parallel global optimization algorithm for rational separable factorable function. J. Applied Math. Computer., 32: 61-72. DOI: 10.1016/0096-3003(89)90047-7
Spedicato, E., 1976. A variable metric method for function minimization derived from invariancy to non-linear scaling. J. Optimiz. Theory Appli., 20: 315-329. DOI: 10.1007/BF00933626
Tassopoulos, A. and C. Storey, 1984. A conjugate direction method based on a non-quadratic model. J. Optimiz. Theory Appli., 43: 371-381. DOI: 10.1007/BF00934461