Australian Journal of Basic and Applied Sciences, 7(7): 302-309, 2013
ISSN 1991-8178

A Global Convergent Spectral Conjugate Gradient Method

[1]Abbas Y. Al-Bayati and [2]Hawraz N. Al-Khayat

[1]College of Basic Education Telafer, Mosul University, Iraq.
[2]College of Computer Sciences and Mathematics, Mosul University, Mosul, Iraq.

Abstract: In this paper, we are concerned with the Conjugate Gradient (CG) methods for solving unconstrained optimization problems. It is well-known that the direction generated by a CG-method may not be a descent direction of the objective function. In this paper, we have done a little modification to the Conjugate Descent (CD) method such that the direction generated by the modified method provides a descent direction for the objective function. This property depends neither on the line search used, nor on the convexity of the objective function. Moreover, the modified method reduces to the standard CD method if line search is exact. Under mild conditions, we prove that the modified method with strong Wolfe line search is globally convergent even if the objective function is non convex. We also present some numerical results to show the efficiency of the proposed method.

Key words: Spectral Conjugate Gradient, Global Convergence, Unconstrained Optimization, Descent Direction, Line Search.

INTRODUCTION

Our aim in this paper is to study the global convergence properties and practical computational performance of a new nonlinear spectral CG-method for unconstrained optimization with Powell restarting criterion and with appropriate conditions. We consider the following unconstrained optimization problem:

$$\min\{f(x) | x \in R^n\}, \tag{1.1}$$

where $f : R^n \to R$ is a continuously differentiable function. Nonlinear CG-methods are efficient for solving (1.1). The nonlinear CG-methods generate iterates by letting :

$$x_{k+1} = x_k + \alpha_k d_k , \quad k = 0,1,... \tag{1.2}$$

with

$$d_k = \begin{cases} -g_k, & if \;\; k = 0 \\ -g_k + \beta_k d_{k-1}, & if \;\; k \geq 1 \end{cases} \tag{1.3}$$

where x_k is the current iteration, $\alpha_k > 0$ is the step-length which is determined by some line search, d_k is the new search direction, $g_k = g(x_k)$ denotes the gradient of f at x_k, and β_k is a suitable parameter. There are many well-known formulas for β_k, such as the Fletcher-Reeves (FR) (Raydan, M., 1997), Polak-Ribiere (PR) (Andrei, N., 2008), Hestenes-Stiefel (HS) (Hestenes, M.R. and E. Stiefel, 1952), and Conjugate-Descent (CD) (Wolfe, P., 1969). The CG-method is a powerful line search method for solving optimization problems, and it remains very popular for engineers and mathematicians who are interested in solving large-scale problems. This method can avoid, like Steepest Descent (SD) method, the computation and storage of some matrices associated with the Hessian of objective functions. The original CD-method proposed by Fletcher (Wolfe, P., 1969), in which β_k^{CD} is defined by the following:

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} \tag{1.4}$$

where $\|.\|$ denotes the Euclidean norm of vectors. An important property of the CD- method is that the method will produce a descent direction under the strong Wolfe line search. In the strong Wolfe line search, the step-length α_k is required to satisfy the following:

$$\begin{aligned} & f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \\ & \left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k \end{aligned} \tag{1.5}$$

Corresponding Author: Abbas Y. Al-Bayati, College of Basic Education Telafer, Mosul University, Iraq.
E-mail: profabbasalbayati@yahoo.com

where $0 < \delta < \sigma < 1$.

Another popular method to solving problem (1.1) is the Spectral Gradient (SG) method, which was developed originally by Barzilai and Borwein (Birgin, E.G. and J.M. Martinez, 2001). In 1988, Raydan (Liu, J. and Y. Jiang, 2012) further introduced the SG-method for potentially large-scale unconstrained optimization problems. The main feature of this method is that only gradient directions are used at each line search whereas a non-monotone strategy guarantees global convergence. As well as, this method outperforms sophisticated CG-method in many problems. Birgin and Martinez (Al-Baali, M., 1985) proposed three kinds of spectral CG-methods. The direction d_k is given by the following way

$$d_k = -\theta_k g_k + \beta_k d_{k-1} \tag{1.6}$$

where the parameter β_k is computed the following

$$\beta_k^1 = \frac{(\theta_k y_{k-1} - s_{k-1})^T g_k}{s_{k-1}^T y_{k-1}},\ \beta_k^2 = \frac{\theta_k y_{k-1}^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}^T g_{k-1}},\ \beta_k^3 = \frac{\theta_k g_k^T g_k}{\alpha_{k-1}\theta_{k-1} g_{k-1}^T g_{k-1}} \tag{1.7}$$

respectively, and θ_k is taken to be the spectral gradient and computed by the following

$$\theta_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \tag{1.8}$$

where $y_{k-1} = g_k - g_{k-1}$, $s_{k-1} = x_k - x_{k-1}$. The numerical results show that these methods are very effective. Unfortunately, they cannot guarantee to generate descent directions. More spectral CG-methods have also been reported in (Gilbert, J.C. and J. Nocedal, 1992; Liu, J., et al., 2012; Polak, E. and G. Ribiere, 1969). In the case where Armijo-type line search or Wolfe-type line search is used, the descent property of d_k determined by (1.3) is in general not guaranteed. In order to ensure descent property, Dixon (Gilbert, J.C. and J. Nocedal, 1992) and Al-Baali (Hager, W.W. and H. Zhang, 2005) suggested to use the SD-direction $-g_k$ instead of d_k determined by (1.3) in the case where d_k is not a descent direction. By the use of this hybrid technique, Dixon (Gilbert, J.C. and J. Nocedal, 1992) and Al-Baali (Hager, W.W. and H. Zhang, 2005) obtained the global convergence of the CG-methods with some inexact line searches. Quite recently, it is noticed that there are many modified CG-methods studied. Liu et al. (2012) take modification to the CD-method such that the direction generated is always a descent direction and d_k is defined by the following:

$$d_k = \begin{cases} -g_k, & if\ \ k = 1 \\ -\theta_k g_k + \beta_k d_{k-1}, & if\ \ k \ge 2 \end{cases} \tag{1.9}$$

where β_k is specified by the following

$$\beta_k = \begin{cases} \beta_k^{CD}, & if\ g_k^T d_{k-1} \le 0, \\ 0, & else, \end{cases} \tag{1.10}$$

and

$$\theta_k = 1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}} \tag{1.11}$$

They prove that this method can guarantee to generate descent directions and is globally convergent. The paper has the following structure; in the next section, a new spectral CG-method is proposed. Section 3 will be devoted to prove the global convergence of the new proposed method. In Section 4, some numerical experiments will been reported to test the efficiency, especially in comparison with the existing other methods. Some concluding remarks will be given in the last section.

2. New Spectral Conjugate Descent Method:

In this section we have, first, to investigate how to determine a descent direction of objective function. Let x_k be the current iterate. Let d_k be defined by

$$d_k = \begin{cases} -g_k, & \text{if} \quad k=0 \\ -\theta_k g_k + \beta_k^{CD} d_{k-1}, & \text{if} \quad k \ge 1 \end{cases} \tag{2.1}$$

where β_k^{CD} is specified by (1.4) and let us consider the following new parameter:

$$\theta_k^{New} = -\frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T g_{k-1}} - \frac{d_{k-1}^T g_k g_k^T g_{k-1}}{\|g_k\|^2 d_{k-1}^T g_{k-1}} \tag{2.2}$$

The new method reduces to the standard CD method if the line search is exact. But generally we refer to use the inexact line search (s.t. Wolfe line search). We first prove that d_k is a sufficiently descent direction.

2.1 Lemma:

Suppose that d_k is given by (2.1) and (2.2). Furthermore assume that α_k satisfies strong Wolfe condition (1.5) with $\sigma_k < 0.5$ and if Powell restart is used (i.e. $\left|g_k^T g_{k-1}\right| < 0.2\, g_k^T g_k$). Then, the following result

$$g_k^T d_k \le -c_1 \|g_k\|^2 \tag{2.3}$$

holds for any $k \ge 0$.

Proof.

If $k=0$, then $d_k^T g_k = -\|g_k\|^2$. Then, from (1.4), (2.1), and (2.2), it is follows that:

$$g_k^T d_k = -\theta_k^{New} \|g_k\|^2 + \beta_k^{CD} g_k^T d_{k-1}$$

$$g_k^T d_k = -\left[-\frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T g_{k-1}} - \frac{d_{k-1}^T g_k g_k^T g_{k-1}}{\|g_k\|^2 d_{k-1}^T g_{k-1}} \right] \|g_k\|^2 - \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} g_k^T d_{k-1}$$

$$g_k^T d_k = \frac{1}{d_{k-1}^T g_{k-1}} \left[d_{k-1}^T y_{k-1}\, g_k^T g_k + d_{k-1}^T g_k\, g_k^T g_{k-1} - g_k^T d_{k-1}\, g_k^T g_k \right]$$

$$= \frac{1}{d_{k-1}^T g_{k-1}} \left[d_{k-1}^T g_k\, g_k^T g_k - d_{k-1}^T g_{k-1}\, g_k^T g_k + d_{k-1}^T g_k\, g_k^T g_{k-1} - g_k^T d_{k-1}\, g_k^T g_k \right]$$

$$= \frac{1}{d_{k-1}^T g_{k-1}} \left[-d_{k-1}^T g_{k-1}\, g_k^T g_k + d_{k-1}^T g_k\, g_k^T g_{k-1} \right]$$

$$g_k^T d_k = -\|g_k\|^2 + \frac{d_{k-1}^T g_k}{d_{k-1}^T g_{k-1}} g_k^T g_{k-1}$$

$$= -\|g_k\|^2 - \left(-\frac{d_{k-1}^T g_k}{d_{k-1}^T g_{k-1}}\right) g_k^T g_{k-1}$$

From second Wolfe (1.5)

$$d_{k-1}^T g_k \ge \sigma\, d_{k-1}^T g_{k-1} \Rightarrow -d_{k-1}^T g_k \le -d_{k-1}^T g_{k-1}$$

yields,

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

$$= -\|g_k\|^2 + \sigma\, g_k^T g_{k-1}$$

since the Powell restarting criterion (Liu, J., et al., 2012) is defined as follows:

$$\left|g_k^T g_{k-1}\right| \le -0.2 \|g_k\|^2 \tag{2.4}$$

$$g_k^T d_k \le -\|g_k\|^2 + \sigma(0.2)\|g_k\|^2$$

$$g_k^T d_k \le -\left[1 - \sigma(0.2)\right] \|g_k\|^2$$

where

$$c = [1 - \sigma(0.2)] > 0$$

and

$$g_k^T d_k \le -c\|g_k\|^2$$

we obtain the desired result. From *Lemma 2.1*, it is known that d_k is a descent direction of f at x_k Furthermore, if the exact line search is used, then:

$$\theta_k^{New} = -\frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T g_{k-1}} - \frac{d_{k-1}^T g_k g_k^T g_{k-1}}{\|g_k\|^2 d_{k-1}^T g_{k-1}} = 1$$

In this case, the proposed spectral CD-method reduces to the standard CD- method, However, it is often that the exact line search is time-consuming and sometimes is unnecessary. In the following, we are going to develop a new algorithm, where the search direction d_k is chosen by (2.1)-(2.2) and the step-length is determined by strong Wolfe-type inexact line search.

2.2 New Algorithm.

Step 1: Initialization: Take $x_0 \in R^n$ and the parameter $0 < \delta \le \sigma < 1$. Compute $f(x_0)$ and $g_0 = \nabla f(x_0)$ and set $d_0 = -g_0$ for $k = 0$.

Step 2: Computation of the Line Search: Compute α_k satisfying Wolfe conditions s.t:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k$$

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

where $0 < \delta \le \sigma < 1$ and then evaluate $x_{k+1} = x_k + \alpha_k d_k$

Step 3: Test for Convergence: If ($\|g_k\|_\infty \le 10^{-5}$ or $\left|\alpha_k g_k^T d_k\right| \le 10^{-10}|f_k|$) is satisfied then the iterations are stopped.

Step 4: Restarting Criterion: If Powell restarting criterion s.t.

$$\left| g_k^T g_{k-1} \right| \ge 0.2\|g_k\|^2$$

is satisfied then do a restart step by SD direction; otherwise continue.

Step 5: Computation of the New Scalar Parameters: compute the following parameter β_k^{CD} and θ_k^{New} from:

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}$$

$$\theta_k^{New} = -\frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T g_{k-1}} - \frac{d_{k-1}^T g_k g_k^T g_{k-1}}{\|g_k\|^2 d_{k-1}^T g_{k-1}}$$

Step 6: Search Direction: Compute the new search direction d_k as

$$d_k = -\theta_k^{New} g_k + \beta_k^{CD} d_{k-1}$$

Step 7: Set $k=k+1$ and go to Step 2.

It is well known that, if f is bounded along the direction d_k, then there exists a step length α_k satisfying the Wolfe line search conditions (1.5). In our algorithm, when the Powell restarting condition (2.6) is satisfied, then we restart the algorithm with the negative gradient. More sophisticated reasons for restarting the algorithms have been proposed in the literature (Polak, E. and G. Ribiere, 1969; Powell, M.J.D., 1977), but we are interested in the performance of a CG-Algorithm that uses this restart criterion associated to a direction satisfying the conjugacy condition. Under reasonable assumptions, conditions (1.5) and (2.6) are sufficient to prove the global convergence of the algorithm.

3. Convergence Analysis:

In this section, we are in a position to study the global convergence of Algorithm (2.2).We first state the following mild assumptions, which will be used in the proof of global convergence property.

Assumption (H):

(i) The level set $S=\{x : x \in R^n, f(x) \le f(x_1)\}$ is bounded, where x_1 is the starting point.

(ii) In a neighborhood Ω of S, f is continuously differentiable and its gradient g is Lipschitz continuously, namely, there exists a constant $L \ge 0$ such that

$$\| g(x) - g(x_k) \| \le L \| x - x_k \|, \forall x, x_k \in \Omega \tag{3.1}$$

Obviously, from the Assumption (H, i) there exists a positive constant D such that:

$$D = \max\{\|x - x_k\|, \forall x, x_k \in S\} \tag{3.2}$$

where D is the diameter of Ω. From Assumption (H, ii), we also know that there exists a constant $\Gamma \ge 0$, such that:

$$\|g(x)\| \le \Gamma,\ \forall x \in S \tag{3.3}$$

On some studies of the CG-methods, the sufficient descent or descent condition plays an important role. Unfortunately, this condition is hard to hold.

3.1 Theorem.:

Under Assumptions (H, i) and (H, ii) , suppose that d_k is given by (2.1) and (2.2) where α_k satisfies strong Wolfe condition (1.5) with $\sigma_k < 0.5$ then it holds that

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{3.4}$$

Proof.

Suppose that there exists a positive constant $\varepsilon > 0$ such that

$$\|g_k\| \ge \varepsilon \tag{3.5}$$

For all k. Then, from (2.1), it follows that

$$\|d_k\|^2 = d_k^T d_k$$

$$= (-\theta_k^{New} g_k + \beta_k^{CD} d_{k-1})(-\theta_k^{New} g_k + \beta_k^{CD} d_{k-1})$$

$$= (\theta_k^{New})^2 \|g_k\|^2 - 2\theta_k^{New} \beta_k^{CD} d_{k-1} g_k + (\beta_k^{CD})^2 \|d_{k-1}\|^2$$

$$= (\theta_k^{New})^2 \|g_k\|^2 - 2\theta_k^{New} (d_k^T + \theta_k^{New} g_k^T) g_k + (\beta_k^{CD})^2 \|d_{k-1}\|^2$$

$$= (\theta_k^{New})^2 \|g_k\|^2 - 2\theta_k^{New} d_k^T g_k - 2(\theta_k^{New})^2 \|g_k\|^2 + (\beta_k^{CD})^2 \|d_{k-1}\|^2$$

$$= (\theta_k^{New})^2 \|d_{k-1}\|^2 - 2\theta_k^{New} d_k^T g_k - (\theta_k^{New})^2 \|g_k\|^2 \tag{3.6}$$

Dividing the both sides of the above equality by $(g_k^T d_k)^2$, then from (1.4), (2.3), (3.1), and (3.6) we obtain:

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} = \frac{(\beta_k^{CD})^2 \|d_{k-1}\|^2 - 2\theta_k^{New} d_k^T g_k - (\theta_k^{New})^2 \|g_k\|^2}{(g_k^T d_k)^2}$$

$$= \left[\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} \right]^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - (\theta_k^{New})^2 \frac{\|g_k\|^2}{(g_k^T d_k)^2} - 2\,\theta_k^{New} \frac{1}{(g_k^T d_k)}$$

$$\le \left[\frac{\|g_k\|^2}{c\|g_{k-1}\|^2} \right]^2 \frac{\|d_{k-1}\|^2}{c^2 \|g_k\|^4} - (\theta_k^{New})^2 \frac{\|g_k\|^2}{c^2 \|g_k\|^4} - 2\,\theta_k^{New} \frac{1}{c\|g_k\|^2}$$

$$\leq \left[\frac{\|g_k\|^2}{c\|g_{k-1}\|^2}\right]^2 \frac{\|d_{k-1}\|^2}{c^2\|g_k\|^4} - \left[\theta_k^{New}\frac{\|g_k\|}{c\|g_k\|^2} + \frac{1}{\|g_k\|}\right]^2 + \frac{1}{\|g_k\|^2}$$

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|d_{k-1}\|^2}{c^4 (g_{k-1}^T d_{k-1})^2} + \frac{1}{\|g_k\|^2} \leq \frac{\|d_{k-1}\|^2}{c^4 (g_{k-1}^T d_{k-1})^2} + \frac{1}{\varepsilon^2}$$

Since

$$d_1 = -g_1$$

So that

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{\|d_1\|^2}{(g_1^T d_1)^2} + \frac{k-1}{\varepsilon^2}$$

$$= \frac{1}{\|g_1\|^2} + \frac{k-1}{\varepsilon^2} < \frac{1}{\varepsilon^2} + \frac{k-1}{\varepsilon^2} = \frac{k}{\varepsilon^2}$$

Thus

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} > \sum_{k=1}^{\infty} \frac{\varepsilon^2}{k} = +\infty$$

Which is contrary to proof this theorem. Hence , the proof is complete.

In the last year, we interest to the general nonlinear functions, and the convergence analysis that often exploits insights developed by Gilbert and Nocedal (1992) by Hager and Zhang (2005). The global convergence proof of the new algorithm is based on the Zoutendijk condition (1970) combined with the analysis showing that the sufficient descent condition holds and d_k is bounded.

4. Numerical Experiments:

The main work of this section is to report the performance of the new methods on a set of test problems. The codes were written in Fortran and in double precision arithmetic. All the tests were performed on a PC. Our experiments were performed on a set of 35-nonlinear unconstrained problems that have second derivatives available. These test problems are contributed in CUTE (Fletcher, R., 1987) and their details are given in the Appendix. for each test function we have considered 10 numerical experiments with number of variable n= 100,200,……,1000. In order to assess the reliability of our new proposed methods, we have tested them against standard (CD & FR) classical CG-methods and (MFR) using the same test problems. All these methods terminate when the following stopping criterion is met.

$$(\|g_k\|_\infty \leq 10^{-5} \text{ or } |\alpha_k g_k^T d_k| \leq 10^{-10}|f_k|) \qquad (4.1)$$

we also force these routines stopped if the iterations exceed 1000 or the number of function evaluations reach 2000 without achieving the minimum. We use $\delta = 10^{-4}$, $\sigma = 0.1$ in the Wolfe line search routine. **Tables (4.1)** compares some numerical result for New spectral CG-methods against (CD & FR & MFR) CG-methods respectively, this table indicate for (n) as a dimension of the problem;(NOI) number of iterations; (NOFG) number of function and gradient evaluation;(Time) the total time required to complete the evaluation process for each test problem. In **Table (4.2**) we have compared the percentage performance of the new spectral CG-method and CD & FRCG and MFR methods taking over all the tools as 100% . In order to summarize our numerical results , we have concerned only on the total of different dimensions n= 100, 200,……,1000, for all tools used in these comparisons.

Table 4.1: Comparison between new spectral CG-method and CD, FRCG and MFR-CG methods for the total of n different dimensions n= 100, 200, ... ,1000 for each test problems.

Prob.	Classical CD method			Classical FR method			Modified FR method			New spectral CD method		
	NOI	NOFG	CPU	NOI	NOFG	CPU	NOI	NOFG	CPU	NOI	NOFG	CPU
1	116	249	0.11	126	258	0.09	123	256	0.1	114	242	0.1
2	63	191	0.02	63	191	0.01	63	191	0.02	63	191	0.01
3	93	262	0.05	98	296	0.05	93	262	0.06	93	262	0.05
4	405	835	0.20	395	772	0.22	285	584	0.12	135	298	0.05
5	206	441	0.03	199	401	0.01	176	376	0.04	157	338	0.03
6	249	646	0.25	184	368	0.14	157	340	0.12	45	95	0.02
7	2041	10969	0.25	146	473	0.02	124	417	0.00	77	252	0.01

8	63	232	0.09	68	255	0.08	63	232	0.09	63	232	0.09
9	138	395	0.02	140	360	0.02	137	358	0.02	124	328	0.01
10	141	421	0.07	141	337	0.08	157	352	0.07	79	183	0.02
11	293	737	0.09	247	519	0.08	187	446	0.05	84	229	0.03
12	120	323	0.02	114	301	0.02	117	325	0.02	117	325	0.01
13	20	167	0.01	22	209	0.03	20	167	0.01	20	167	0.01
14	313	623	0.04	314	632	0.05	237	500	0.03	131	287	0.02
15	138	422	0.04	201	537	0.04	147	459	0.02	96	371	0.01
16	110	288	0.03	110	288	0.03	110	288	0.03	110	288	0.03
17	128	291	0.03	128	291	0.01	128	291	0.03	128	291	0.03
18	113	296	0.03	113	296	0.05	113	296	0.03	113	296	0.03
19	203	464	0.03	193	417	0.03	178	390	0.03	129	286	0.02
20	93	262	0.05	100	304	0.06	93	262	0.04	93	262	0.04
21	190	423	0.01	181	413	0.02	168	375	0.03	161	356	0.02
22	214	491	0.08	213	390	0.06	168	355	0.04	116	234	0.03
23	418	1392	0.25	399	813	0.12	302	675	0.14	104	230	0.05
24	115	326	0.02	115	326	0.03	115	326	0.02	115	326	0.01
25	164	370	0.01	151	363	0.03	150	339	0.02	30	80	0.02
26	524	1824	0.27	405	801	0.13	294	659	0.11	140	312	0.06
27	94	272	0.01	94	270	0.02	93	269	0.00	89	262	0.00
28	93	279	0.07	97	300	0.06	93	279	0.06	93	279	0.06
29	84	262	0.08	90	277	0.08	84	262	0.08	84	262	0.08
30	56	245	0.05	56	245	0.04	56	245	0.04	56	245	0.04
31	141	421	0.07	141	337	0.06	157	352	0.07	79	183	0.01
32	352	960	0.03	342	692	0.02	274	603	0.01	119	267	0.02
33	10	30	0.00	10	30	0.00	10	30	0.00	10	30	000
34	70	90	0.01	80	100	0.03	70	90	0.01	70	90	0.01
35	211	933	0.03	150	360	0.03	150	360	0.02	150	360	0.02
Total	7782	26832	2.45	5626	13222	1.85	4892	12011	1.58	3387	8739	1.05

Percentage performance of the new algorithms against 100% (CD, FR, MFR) algorithms respectively, as follows in Tables (4.2), (4..3), and (4.4).

Table (4.2):

Tools	Classical CD Method	New Spectral CD Method
NOI	100%	43.5 %
NOFG	100%	32.6 %
CPU	100%	42.8 %

Clearly, from the above table, we have found that the new proposed algorithm beats classical CD algorithm in about (56.5%) NOI; (67.4%) NOFG and (57.2%)Time.

Table (4.3):

Tools	Classical FR Method	New Spectral CD Method
NOI	100%	60.2 %
NOFG	100%	66 %
CPU	100%	56.8 %

Clearly, from the above table, we have found that the new proposed algorithm beats classical FR algorithm in about (39.8%) NOI; (34%) NOFG and (43.2%)Time.

Table (4.4):

Tools	MFR Method	New Spectral CD Method
NOI	100%	69.2 %
NOFG	100%	72.7 %
CPU	100%	66.4 %

Clearly, from the above table, we have found that the new proposed algorithm beats modified FR algorithm in about (30.8%) NOI; (27.3%) NOFG and (33.6%)Time.

Appendix.

1)Trigonometric 2)Penalty 3)Raydan 4)Hager 5)Generalized Tri-diagonal 6)Extended Three Exp-Terms 7)Diagonal4 8)Diagonal 9)Extended Himmelblau 10)Extended PSC1 11)Extended BD1 12)Extended Quadratic Penalty QP1 13)Extended EP1 14)Extended Tridiagonal-2 15)ARWHEAD (CUTE) 16)DIXMAANA (CUTE) 17)DIXMAANB (CUTE) 18)DIXMAANC (CUTE) 19)EDENSCH (CUTE) 20)DIAGONAL-6 21)ENGVAL1 (CUTE) 22)DENSCHNA (CUTE) 23)DENSCHNC (CUTE) 24)DENSCHNB (CUTE) 25)DENSCHNF (CUTE) 26)Extended Block-Diagonal BD2 27)Generalized quarticGQ1 28)DIAGONAL 7 29)DIAGONAL-8 30)Full Hessian 31)SINCOS 32)Generalized quartic GQ2 33)ARGLINB (CUTE) 34)HIMMELBG (CUTE) 35)HIMMELBH (CUTE).

REFERENCES

Al-Baali, M., 1985. "Descent Property and Global Convergence of The Fletcher Reeves Method with Inexact Line Search". IMA Journal Numer. Anal., 5: 121-124.

Andrei, N., 2008. "40 Conjugate Gradient Algorithms for Unconstrained Optimization, a Survey on their Definition". ICI Technical Report. No. 13/08, March 14.

Barzilai, J. and J.M. Borwein, 1988. "Two-Point Step Size Gradient Methods". IMA Journal of Numerical Analysis., 8(1): 141-148.

Birgin, E.G. and J.M. Martinez, 2001. "A Spectral Conjugate Gradient Method for Unconstrained Optimization". Applied Mathematics and Optimization, 43(2): 117-128.

Bongartz, K.E., A.R. Conn, N.I.M. Gould and P.L. Toint, 1995. "CUTE: Constrained and Unconstrained Testing Environments". ACM Trans, Math. Software., 21: 123-160.

Dixon, L.C.W., 1970. Nonlinear Optimization: a Survey of the State of the Art. In: Evans, D.J. (eds.) Software for Numerical Mathematics. Academic, New York, pp: 193-216.

Fletcher, R. and C.M. Reeves, 1964. "Function Minimization by Conjugate Gradients" The Computer Journal., 7: 149-154.

Fletcher, R., 1987. Practical Methods of Optimization: Unconstrained Optimization. John Wiley & Sons, New York, NY, USA.

Gilbert, J.C. and J. Nocedal, 1992. "Global Convergence Properties of Conjugate Gradient Methods for Optimization". SIAM Journal Optimization, 2: 21-42.

Hager, W.W. and H. Zhang, 2005. "A New Conjugate Gradient Method with Guaranteed Descent and an Efficient Line Search". SIAM Journal Optimization., 16: 170-192.

Hestenes, M.R. and E. Stiefel, 1952. "Methods of Conjugate Gradients for Solving Linear Systems". Journal of Research of the National Bureau of Standards., 49: 409-436.

Liu, J. and Y. Jiang, 2012. "Global Convergence of a Spectral Conjugate Gradient Method for Unconstrained Optimization". Hindawi Publishing Corporation, Abstract and Applied Analysis. Article ID 758287, doi:10.1155/2012/758287.

Liu, J., X. Du, and K. Wang, 2012. "A Mixed Spectral CD-DY Conjugate Gradient Method". Hindawi Publishing Corporation, Journal of Applied Mathematics. Article ID 569795, doi:10.1155/2012/569795.

Polak, E. and G. Ribiere, 1969. "Note Sur la Convergence de Methods de Directions Conjugates" 3(16): 35-43.

Powell, M.J.D., 1977. "Restart Procedures for the Conjugate Gradient Method". Mathematical Program, 12: 241-254.

Raydan, M., 1997. "The Barzilai and Borwein Gradient Method for the Large Scale Unconstrained Minimization Problem". SIAM Journal on Optimization., 7(1): 26-33.

Wolfe, P., 1969. "Convergence Conditions for Ascent Methods". SIAM Rev., 11: 226-235.

Zoutendijk, G., 1970. "Nonlinear Programming, Computational Methods in Integer and Nonlinear Programming". North-Holland Amsterdam. pp: 37-86.