Journal of Education and Science
Vol. 27, No.3, 2018, pp. 143-162
ISSN 1812-125X
http://edusj.mosuljournals.com

A Global Convergence of Spectral Conjugate Gradient Method for Large Scale Optimization

Ghada M. Al-Naemi

Department of Mathematics, College of Computer Science and Mathematics, University of Mosul , IRAQ.

تاريخ الاستلام 2013/12/03 تاريخ القبول 2014/02/16

المستخلص

في هذا البحث, تم الاهتمام بطريقة التدرج المترافق في مسائل الامثلية غير المقيدة, وذلك لســـهولة التعامل معها ولأنها لاتحتاج لخزن أية مصـــفوفة., اقترحنا طريقتين جديدتين مطورتين طيفيتين لطريقة (CD). الطريقتان المقترحتان تكونان اتجاه خطوط بحث منحدرة دائما لدالة الهدف ,وفي كل تكرار اســـتخدمنا خط البحث القوي المعتمد على خط بحث غير تام. وكذلك تحققان صفة التقارب الشمولي للدوال العامة غير التربيعية, النتائج العددية تظهر كفاءة الطريقتين المقترحتين.

Abstract

In this paper, we are concerned with the conjugate gradient method for solving unconstrained optimization problems due to its simplicity and don’t store any matrices. We proposed two spectral modifications to the conjugate descent (CD). These two proposed methods produces sufficient descent directions for the objective function at every iteration with strong Wolfe line searches and with inexact line search, and also they are globally convergent for general non-convex functions can be guaranteed. Numerical results show the efficiency of these two proposed methods.

Introduction.

Let $f: R^n \to R$ be continuously differentiable function. Consider the unconstrained nonlinear optimization problem:

Minimize f(x), $x \in R^n$. (1)

We use g(x) to denote to the gradient of f at x. Due to need less computer memory especially, conjugate gradient method is very appealing for solving (1) when the number of variables is large. A conjugate gradient (CG) method generates a sequence of iterates by letting

$$x_k = x_{k-1} + \alpha_{k-1} d_{k-1}, \text{ k=0,1,2,...} \quad (2)$$

where the step-length α_k is obtained by carrying out some line search, and the search direction d_k is defined by

$$d_k = \begin{cases} -g_k & , \ if\ k=0 \\ -g_k + \beta_k d_{k-1}, & if\ k \geq 1 \end{cases}, \quad (3)$$

where β_k is scalar which determines the different CG methods [11]. There are many well- known formula for β_k, such as the Fletcher-Reeves(FR) [7], Polak-Ribirere-Polyak (PRP) [13] and [14], Hesteness-Stiefel (HS) [10], conjugate descent (CD) [8], Liu-Story (LS) [12], and Dai-Yuan (DY) [5]. In survey paper Hager and Zhang in [9] reviewed the development of different various of nonlinear gradient methods, with especial attention given to global convergence properties.

The standard CD method proposed by Fletcher [8], specifies the β_k^{CD} by

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \quad (4)$$

where $\|.\|$ denotes the Euclidean norm of vectors. An important of the CD method is that the method will produce a descent direction under the strong Wolfe line search [18]

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (5)$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k, \quad (6)$$

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

Another popular method to solving (1) is spectral CG method, which was developed originally by Barzilai and Browein [2]. Raydan in [17] further introduced the spectral CG method for potentially large-scale unconstrained optimization problems. Birgin and Marti'nez [3] proposed a spectral CG method by combining CG method and spectral gradient method [17], by multiplying the gradient g_k in the second part of equation (3) by the parameter η_k in the following manner:

$$d_k = \begin{cases} -g_k & , \ if\ k=0 \\ -\eta_k g_k + \beta_k d_{k-1}, & if\ k \geq 1 \end{cases}, \tag{7}$$

Zhang in [19] take FR formula and $\eta_k = d_{k-1}^T y_{k-1} / \|g_{k-1}\|^2$ they proved that this method can guarantee to generate descent directions and is globally convergent. Matonoha and et al in [15] proposed a modified CD method by

$$\beta_k = \frac{g_k^T g_k}{\left|d_{k-1}^T g_{k-1}\right|}, and\ \eta_k = \frac{y_{k-1}^T d_{k-1}}{\left|d_{k-1}^T g_{k-1}\right|}. \tag{8}$$

Zhong in [20] they proposed the spectral PRP method by using the standard PRP formula with η_k defined by

$$\eta_k = 1 - \frac{g_k^T g_{k-1} . g_k^T d_{k-1}}{\|g_k\|^2 . d_{k-1}^T (g_k - g_{k-1})}. \tag{9}$$

Du and Liu in [6] they proposed the spectral HS method by using the standard HS formula with η_k defined in (9). Liu and Jiang [11] proposed success spectral gradient method by combining the CD method and spectral gradient method by the following manner

$$\beta_k = \begin{cases} \beta_k^{CD}, & if\ d_{k-1}^T g_k \leq 0 \\ 0 & , else \end{cases}, \tag{10}$$

and $$\eta_k = 1 - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}. \tag{11}$$

In this paper we proposed two spectral CG method, they based to the modification to the standard CD in (4), and then proposed a suitable η_k for each one to get a good spectral CD-CG methods.

The rest of this paper is organized as follows. In the next section, a new modified spectral CD-CG method is proposed by combining modification to CD method with η_k defined in (11) will be denoted by MCD1, and we give its algorithm. Section 3 will be devoted to prove the global convergence. In section 4, new proposed spectral CD-CG method is proposed by combining modification to CD method with η_k will be defined next in this section, will be denoted by MCD2, and we give its algorithm. In section 5 will be devoted to prove the global convergence. Finally in section 6, some numerical experiments will be done to test the efficiency of the two proposed methods.

2- Modified Spectral CD Conjugate Gradient and its Algorithm (MCD1).

In this section, we present a new modified CD method which is specified by

$$\beta_k^{MCD1} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} - \frac{\|g_k\|^2 g_k^T d_{k-1}}{\left(d_{k-1}^T g_{k-1}\right)^2}, \tag{12}$$

If exact line search is used, then β_k^{MCD1} will reduce to standard β_k^{CD}, and η_k in (11) equal one. However, we used inexact line search in our work. We put (12) with η_k defined in (11) in (7), we will get the direction of our proposed method

$$d_k^{MCD1} = -\left(1 - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}\right) g_k + \left(-\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} - \frac{\|g_k\|^2 . g_k^T d_{k-1}}{\left(d_{k-1}^T g_{k-1}\right)^2}\right) . d_{k-1} \tag{13}$$

Algorithm (2.1) (MCD1)

Step 0: Given $x_0 \in R^n, \varepsilon = 1*10^{-5}, k = 0,\ \delta \in (0,0.5), and\, \sigma \in (\delta,1)$.

Step 1: Set $d_k = -g_k$

Step 2: If $\|g_k\| \le \varepsilon$ then stop; else continue.

Step 3: Determine the steplength α_k by using the strong Wolfe line search conditions (5) and (6).

Step 4: Calculate new point x_k by (2).

Step 5: Compute d_k^{MCD1} by (13).

Step 6: If k=n, or $\left|g_k^T g_{k-1}\right| \geq 0.2\|g_k\|^2$, set k=1, and go to step 1; else, set

k=k+1, and go to step 2.

The following theorem shows that algorithm (2.1) possesses the sufficient descent condition with strong Wolfe line search (5) and (6).

Theorem 2.1

Let $\{x_k\}$ *and* $\{d_k\}$ be generated by algorithm (2.1), then we have

$$d_k^T g_k \leq -c\|g_k\|^2, \tag{14}$$

where $c = \left(1-\sigma^2\right)$.

Proof: We can prove the conclusion by induction. From $\|g_0\|^2 = -g_0^T d_0$, the conclusion (13) holds for k=0. Now we assume that the conclusion is true for (k-1) and $g_{k-1} \neq 0$, i. e. $g_{k-1}^T d_{k-1} < 0$. We need to prove that the conclusion holds for k. Multiply both sides of (13) by g_k, we have

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

$$= -\|g_k\|^2 - \left(\frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}\right)^2 . \|g_k\|^2 . \tag{15}$$

Now, using (6) in (15), we obtain

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

and since $0 < \sigma < 1$, we will get (14).

3- The global convergence of MCD1 method.

In order to establish the global convergence result for the MCD1, we will impose the following assumptions for f, which have been used often in the literature to analyze the global convergence of CG methods with inexact line search.

Assumption (I): Let

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

(ii) In some neighborhood N of Ω, f is continuously differentiable and its gradient g satisfying Lipschitz conditions, namely, there exist a constant L>0, such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in \Omega. \qquad (16)$$

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

$$B = \{\|x - y\|, \forall x, y \in \Omega\}, \qquad (17)$$

where B is the diameter of Ω. From Assumption (I, ii), we can also find if there exist a constant $\omega > 0$, such that

$$\|g_k\| \le \omega, \forall x \in \Omega \qquad (18)$$

To prove global convergence by contradiction we assume that there is a positive constant ϖ such that

$$\|g_k\| \ge \varpi, \; for all\, k \ge 0 \;. \qquad (19)$$

We are going to prove that $\beta_k^{MCD1} \ge 0$. Using (6) and (14) in (12) we will get

$$\beta_k^{MCD1} \ge (1 + \sigma).\frac{1}{c} > 0. \qquad (20)$$

Theorem 3.1

Suppose that the Assumption (I) holds and consider any CG methods (2) and (7). The parameter β_k^{MCD1} defined by (12), and η_k defined by (11), the direction d_k^{MCD1} is descent direction and determined α_k by using (5) and (6), if

$$\sum_{k\geq 0}\frac{1}{\|d_k\|^2}=\infty$$

then

$$\lim_{k\to\infty}\inf\|g_k\|=0 \tag{21}$$

Proof: From (12) , (13) and (6) we get

$$\left|\beta_k^{MCD1}\right|=\left|-\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}-\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}g_k^T d_{k-1}\right|$$

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

$$\leq\left|\frac{\|g_k\|^2}{c\|g_{k-1}\|^2}+\sigma\frac{\|g_k\|^2}{c\|g_{k-1}\|^2}\right|,$$

Suppose that there exists a positive constants $\omega_1,\varpi_1,\omega_2 and \varpi_2>0$, such that $\varpi_1\leq\|g_k\|\leq\omega_1$, $\varpi_2\leq\|g_{k-1}\|\leq\omega_2$, so we get

$$\Rightarrow\left|\beta_k^{MCD1}\right|\leq\frac{|1+\sigma|}{c\varpi_2}\omega_1 \tag{22}$$

also,

$$|\eta_k|=\left|1-\frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}\right|\leq|1+\sigma| \tag{23}$$

Take the norm of the both sides of (7) with (22) and (23), its yield

$$\therefore\|d_{k+1}\|\leq|\eta_k|\|g_k\|+\left|\beta_k^{MCD1}\right|\|d_{k-1}\|$$

$$\leq(1+\sigma)\omega_1\left(1+\frac{B}{c\varpi_2}\right)=A$$

This relation implies

$$\sum_{k\geq 1}\frac{1}{\|d_k\|^2} > \frac{1}{A^2}\sum_{k\geq 1} 1 = +\infty \, .$$

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

Now to prove that the new algorithm is global convergence for general function, we establish a bounded for the change $(w_{k+1} - w_k)$ in the normalized direction $w_k = d_k / \|d_k\|$, which we will use to conclude, by contradiction, that the gradients cannot be bounded away from zero [16].

Lemma 3.2

Suppose that Assumption (I) hold and consider the CG algorithm (2.1), the direction d_k given by (13) satisfies the sufficient descent condition (14), and the line search satisfying the Zoutendijk condition [21], then $d_k \neq 0$ and

$$\sum_{k=1}^{\infty}\|w_k - w_{k-1}\|^2 < \infty \, , \tag{24}$$

where $w_k = d_k / \|d_k\|$.

Proof: Obviously, we have $d_k \neq 0$. Therefore, w_k is well defined.

Now, from (19) and theorem 3.1 it follows that

$$\sum_{k\geq 0}\frac{1}{\|d_k\|^2} < \infty \, ,$$

otherwise (21) holds, contradicting (19). Define

$$u_k = \frac{v_k}{\|d_k\|} \;\; and \;\; r_k = \beta_k^{MCD1}\frac{\|d_{k-1}\|}{\|d_k\|} \geq 0 \, .$$

Therefore, we have

$$w_k = \frac{d_k}{\|d_k\|} = \frac{-\eta_k g_k + \beta_k^{MCD1} d_{k-1}}{\|d_k\|}$$

$$= \frac{-\eta_k g_k}{\|d_k\|} + \beta_k^{MCD1}\frac{\|d_{k-1}\|}{\|d_k\|}.\frac{d_{k-1}}{\|d_{k-1}\|}$$

$$= u_k + r_k \, w_{k-1} .$$

Using the identity $\|w_k\| = \|w_{k-1}\| = 1$, therefore

$$\|u_k\| = \|w_k - r_k w_{k-1}\| = \|r_k w_k - w_{k-1}\| \tag{25}$$

(the last equality can be verified by squaring both sides). Using the condition $r_k \geq 0$, the triangle inequality, and (25), we obtain

$$\|w_k - w_{k-1}\| \leq \|(1 + r_k)(w_k - w_{k-1})\|$$

$$\leq \|w_k - r_k w_{k-1}\| + \|r_k w_k - w_{k-1}\|$$

$$= 2\|u_k\|.$$

From the definition of v_k, and using (6) we get

$$\|v_k\| = \left\| -\left(1 - \frac{g_k^T d_{k-1}}{d_{k-1}^T g_{k-1}}\right) . g_k \right\| \leq \|-(1+\sigma) g_k\|$$

$$\leq (1+\sigma)\omega_1 = E$$

With the above estimates we get

$$\sum_{k\geq 1} \|w_k - w_{k-1}\|^2 = \sum_{k\geq 1} 2\|u_k\|^2 = 4 \sum_{k\geq 1} \frac{\|v_k\|^2}{\|d_k\|^2}$$

$$= 4E^2 \sum_{k\geq 1} \frac{1}{\|d_k\|^2} < \infty .$$

Therefore (24) holds, which complete the proof.

4- Modified Spectral CD Conjugate Gradient and its Algorithm (MCD2).

In this section, we present a new modified CD method which is specified by

$$\beta_k^{MCD2} = \frac{\|g_k\|^2}{\left|d_{k-1}^T g_{k-1}\right|}, \tag{26}$$

and let us consider the new parameter η_k by:

$$\eta_k = \frac{\|g_k\|^2 + \left|g_k^T d_{k-1}\right|}{\left|d_{k-1}^T g_{k-1}\right|}. \tag{27}$$

If exact line search is used, then β_k^{MCD2} will reduce to standard β_k^{CD}, and η_k equal one. However, we used inexact line search in our work. We put (26) and (27) in (7), we will get new direction

$$d_k^{MCD2} = -\left(\frac{\|g_k\|^2 + |g_k^T d_{k-1}|}{|d_{k-1}^T g_{k-1}|}\right) g_k + \left(\frac{\|g_k\|^2}{|d_{k-1}^T g_{k-1}|}\right) . d_{k-1}. \quad (28)$$

Algorithm (4.1) (MCD2)

Step 0: Given $x_0 \in R^n, \varepsilon = 1*10^{-5}, k = 0,\ \delta \in (0,0.5), and\, \sigma \in (\delta,1)$.

Step 1: Set $d_k = -g_k$

Step 2: If $\|g_k\| \le \varepsilon$ then stop; else continue.

Step 3: Determine the steplength α_k by using the strong Wolfe line search conditions (5) and (6).

Step 4: Calculate new point x_k by (2).

Step 5: Compute d_k^{MCD1} by (28).

Step 6: If k=n, or $|g_k^T g_{k-1}| \ge 0.2\|g_k\|^2$, set k=1, and go to step 1; else, set k=k+1, and go to step 2.

The following theorem shows that algorithm (4.1) possesses the sufficient descent condition with strong Wolfe line search (5) and (6).

Theorem 4.1

Let $\{x_k\} and \{d_k\}$ be generated by algorithm (2.1), then we have

$$d_k^T g_k \le -c\|g_k\|^2, \quad (29)$$

Proof: We can prove the conclusion by induction. From $\|g_0\|^2 = -g_0^T d_0$, the conclusion (29) holds for k=0. Now we assume that the conclusion is true for (k-1) and $g_{k-1} \neq 0$, i. e. $g_{k-1}^T d_{k-1} < 0$. We need to prove that the conclusion holds for k. Multiply both sides of (28) by g_k, we have

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

$$= -\frac{\|g_{k-1}\|^2}{|d_{k-1}^T g_{k-1}|} \|g_k\|^2 < 0,$$

let $c = \frac{\|g_{k-1}\|^2}{|d_{k-1}^T g_{k-1}|} \leq 1, \; if \; \|g_{k-1}\|^2 \leq |d_{k-1}^T g_{k-1}|$.

W e will get (29), so the proof is complete.

5- The global convergence of MCD2 method.

In this section we are going to prove the global convergence of the proposed method MCD2.

Theorem 5.1

Suppose that the Assumption (I) holds and consider any CG methods (2) and (7). The parameter β_k^{MCD2} defined by (28), and η_k defined by (27), the direction d_k^{MCD2} is descent direction and determined α_k by using (5) and (6), if

$$\sum_{k \geq 0} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty, \tag{30}$$

then

$$\underset{k \to \infty}{Lim} \inf \|g_k\| = 0. \tag{31}$$

Proof: we can rewrite (7) as follows

$$d_k + \eta_k g_k = \beta_k d_{k-1},$$

and squaring both side of the above equation, we get

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

Dividing the above equation by $\|g_k\|^4$, we have

$$\frac{\|d_k\|^2}{\|g_k\|^4} = \frac{\|d_{k-1}\|^2}{\left(d_{k-1}^T g_{k-1}\right)^2} - \frac{\eta_k^2 \|g_k\|^2 + 2\eta_k g_k^T d_k}{\|g_k\|^4}$$

$$\leq \frac{\|d_{k-1}\|^2}{\left(d_{k-1}^T g_{k-1}\right)^2} - \frac{\eta_k^2 \|g_k\|^2 - 2\eta_k \|g_k\|^2}{\|g_k\|^4}$$

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

$$\leq \frac{\|d_{k-1}\|^2}{\left(d_{k-1}^T g_{k-1}\right)^2} + \frac{1}{\|g_k\|^2} . \qquad (32)$$

Noting that $\|d_0\|^2 = -g_0^T d_0 = \|g_0\|^2$, we get

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2} . \qquad (33)$$

Therefore, it follows from (33) and (19) that

$$\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \geq \frac{\varpi^2}{k+1} ,$$

which indicates

$$\sum_{k \geq 0} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k \geq 0} \frac{\varpi_1^2}{k+1} = +\infty . \qquad (34)$$

This contradicts the Zoutendijk condition [21]. Therefore the conclusion (31) holds, so the proof is complete.

The above theorem show that the new proposed method is independent to any line search is descent and global convergent.

Lemma 5.2

Suppose that Assumption (I) hold and consider the CG algorithm (4.1), the direction d_k given by (28) satisfies the sufficient descent condition (29), and the line search satisfying the Zoutendijk condition [21], then $d_k \neq 0$ and

$$\sum_{k=1}^{\infty} \|w_k - w_{k-1}\|^2 < \infty \tag{34}$$

Where $w_k = d_k / \|d_k\|$.

Proof: Obviously, we have $d_k \neq 0$. Therefore, w_k is well defined.

Now, from (19) and theorem 3.1 it follows that

$$\sum_{k \geq 0} \frac{1}{\|d_k\|^2} < \infty ,$$

otherwise (21) holds, contradicting (31). Define

$$u_k = \frac{v_k}{\|d_k\|} \ \ and \ \ r_k = \beta_k^{MCD2} \frac{\|d_{k-1}\|}{\|d_k\|} \geq 0 .$$

Therefore, we have

$$w_k = \frac{d_k}{\|d_k\|} = \frac{-\eta_k g_k + \beta_k^{MCD2} d_{k-1}}{\|d_k\|}$$

$$= \frac{-\eta_k g_k}{\|d_k\|} + \beta_k^{MCD2} \frac{\|d_{k-1}\|}{\|d_k\|} \cdot \frac{d_{k-1}}{\|d_{k-1}\|}$$

$$= u_k + r_k w_{k-1} .$$

Using the identity $\|w_k\| = \|w_{k-1}\| = 1$, therefore

$$\|u_k\| = \|w_k - r_k w_{k-1}\| = \|r_k w_k - w_{k-1}\| \tag{35}$$

Using the condition $r_k \geq 0$, the triangle inequality, and (41), we obtain

$$\|w_k - w_{k-1}\| \leq \|(1 + r_k)(w_k - w_{k-1})\|$$

$$\leq \|w_k - r_k w_{k-1}\| + \|r_k w_k - w_{k-1}\|$$

$$= 2\|u_k\| . \tag{36}$$

From the definition of v_k , and using (6) we get

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

$$\leq (1+\sigma)\omega_1 = Z$$

With the above estimates we get

$$\sum_{k\geq1} \|w_k - w_{k-1}\|^2 = \sum_{k\geq1} 4\|u_k\|^2 = 4\sum_{k\geq1} \frac{\|v_k\|^2}{\|d_k\|^2}$$

$$\leq \sum_{k\geq1} \frac{4Z^2}{\|g_k\|^4} \frac{\|g_k\|^4}{\|d_k\|^2}$$

$$\leq \frac{4Z^2}{\varpi_1} \sum_{k\geq1} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty .$$

Therefore (34) holds, which complete the proof.

6- Numerical results

In this section, we reported some numerical results that we obtained with the implementation of the two new methods MCD1 and MCD2 on a set of unconstrained test functions. The cod were written in Fortran 90 and in double precision arithmetic. Our experiments performed on a set of (35) large scale nonlinear unconstrained test functions. These test functions are contributed in CUTE (Bongratz [4] and Andrei [1]).

All these algorithms are implemented with strong Wolfe Powell line search conditions (5) and (6) with $\delta = 0.001$ *and* $\sigma = 0.9$. All these methods terminated when the following stopping criterion is satisfied:

$$\|g_{k+1}\| < 1\times10^{-5} \qquad (37)$$

We record the number of iterations denoted by (NOI), the number of function evaluations denoted by (NOF), for purpose of our comparisons. Table (1) and (2) gives a computational results of the two new methods (namely: MCD1 and MCD2) against the standard CD method with n= 100 and 10000, respectively. While Table (3) and (4) gives the percentage performance of these two proposed methods (MCD1 and MCD2) against the standard CD method taking over all the tools as 100%.

Table (1): Comparison between MCD1; MCD2 against Standard CD with different test problems with dimension n=100.

Test Functions	Standard CD Method		MCD1 Method		MCD2 Method	
	NOI	NOF	NOI	NOF	NOI	NOF
Wood	25	62	24	53	21	49
Dixon	465	1012	463	1008	471	1029
Powell-3	21	45	20	43	19	41
Strait	60	122	28	68	6	15
Sum	16	88	18	90	15	81
Shallow	13	32	13	33	10	25
Wolfe	65	131	45	91	45	91
Cosine	9	20	8	18	8	18
BDQRTIC	88	231	87	229	80	213
DENSCHNB	8	19	7	18	6	15
Dixmaana	6	14	6	14	5	12
Dixmaanb	5	13	5	13	5	13
Dixmaanc	6	15	5	13	5	13
Dixmaane	115	336	46	137	46	137
Dixmaang	51	154	46	139	50	152
Dixmaani	92	261	46	137	45	134
Diagonal-2	102	326	42	161	45	171
Diagonal-4	2	5	2	5	2	5
Diagonal-5	2	7	2	7	2	7
Diagonal-6	2	7	2	7	2	7

Rosen	30	78	30	78	29	76
Ex. Powell	309	622	66	176	45	119
Ex. Beal U63	11	27	11	27	11	27
Ex. Block Diagonal	24	50	22	46	20	42
Ex. Himmelbau	19	150	15	128	15	125
Ex. Penalty	9	21	6	15	6	15
Ex. PSC1	11	24	9	21	7	16
Ex. Quadratic Penalty	10	22	7	16	7	16
Ex. Three Exponential	32	67	9	21	7	17
OSP	49	160	49	160	49	160
Miele	46	152	46	153	45	144
Cubic	80	173	15	43	15	43
NONDIA (Shanno 78)	7	16	7	16	7	16
TRIDIA	90	182	89	179	86	173
Scaled Quadratic	57	116	58	117	57	115
Total	**1944**	**4760**	**1355**	**3480**	**1294**	**3332**

Table (2): Comparison between MCD1; MCD2 against Standard CD with different test problems with dimension n=10000.

Test Functions	Standard CD Method		MCD1 Method		MCD2 Method	
	NOI	NOF	NOI	NOF	NOI	NOF
Wood	31	76	29	69	26	59
Dixon	505	1097	486	1064	498	1078
Powell-3	22	47	21	45	20	43

Strait	66	134	56	115	6	15
Sum	34	161	29	123	23	102
Shallow	14	35	17	53	10	25
Wolfe	83	167	69	150	114	211
Cosine	12	29	9	20	9	20
BDQRTIC	123	356	124	359	119	346
DENSCHNB	9	21	8	23	6	15
Dixmaana	7	17	5	13	5	13
Dixmaanb	6	16	5	14	5	14
Dixmaanc	7	18	6	16	6	16
Dixmaane	401	1268	379	1143	377	1130
Dixmaang	406	1304	402	1263	394	1183
Dixmaani	398	1312	378	1139	377	1130
Diagonal-2	503	1638	346	1485	370	1518
Diagonal-4	3	7	3	7	2	6
Diagonal-5	2	7	2	7	2	7
Diagonal-6	2	7	2	7	2	7
Rosen	30	78	30	78	29	76
Ex. Powell	521	1051	132	405	53	154
Ex. Beal U63	12	29	12	29	12	29
Ex. Block Diagonal	27	56	24	50	20	42
Ex. Himmelbau	11	526	9	402	8	390
Ex. Penalty	6	22	6	22	6	22
Ex. PSC1	13	24	9	21	7	16

Ex. Quadratic Penalty	10	32	11	35	7	16
Ex. Three Exponential	35	81	12	31	8	21
OSP	621	2111	603	2002	551	1815
Miele	68	240	71	282	60	216
Cubic	94	199	24	70	16	45
NONDIA (Shanno 78)	5	15	5	15	5	15
TRIDIA	335	671	333	667	329	659
Scaled Quadratic	631	1266	632	1269	626	1253
Total	**5053**	**14124**	**4289**	**12499**	**4108**	**11707**

Table (3): Percentage performance of the MCD1 and MCD2 methods against the standard CD method with different test problems with dimension n=10000.

Measurement	Standard CD Method	MCD1 Method	MCD2 Method
NOI	100%	69.70%	66.56%
NOF	100%	73.11%	70.00%

Table (4): Percentage performance of the MCD1 and MCD2 methods against the standard CD method with different test problems with dimension n=10000.

Measurement	Standard CD Method	MCD1 Method	MCD2 Method
NOI	100%	84.88%	81.30%
NOF	100%	88.49%	82.89%

From Table (3) we have obtained the following results: MCD1 saves (NOI 30.30%), (NOF 26.89%), and MCD2 saves (NOI 33.44%), (NOF 30.0%) compared with standard CD method. While from Table (4) we have obtained the following results: MCD1 saves (NOI 15.12%), (NOF 26.89%), and MCD2 saves (NOI 18.70%), (NOF 17.11%) compared with standard CD method

Reference

[1] Andrei, N., "An unconstrained optimization test functions collection". J. of Advance Modeling and Optimization, 10:147-161 (2008).

[2] Barzilai, J. and Borwein, I.M., "Two- point step size gradient methods", IMA J. Numerical Analysis, 8:141-148 (1988).

[3] Birgin, E. and Martinez, M., A spectral conjugate gradient method for unconstrained optimization, Applied Mathematics. and optimization, 43:117-128 (2001).

[4] Bongratz, I., Conn, A. R., Gould, N. I. M. and Toint, P. L., CUTE: "constrained and unconstrained testing environments", ACM Trans. Math. Softw., 21: 123-160 (1995) .

[5] Dai, Y. H. and Yuan, Y., "Convergence properties of the conjugate descent method". Adv. Math., 25:552-562 (1996).

[6] Du, X. and Liu, J., "Global convergence of a spectral HS conjugate gradient method", Advance in Control Engineering Information Science, 15:1487-1492 (2011). Doi:10.1016/j.proeng.2011.08.276.

[7] Fletcher, R. and Reeves, C. M., "Function minimization by conjugate gradients", Computer Journal, 7:149-154 (1964).

[8] Fletcher, R., "Practical Methods of Optimization", 2^{nd} ed., A Wiley-Interscience Publication, John Wiley & Sons, Inc., NY, USA (1987).

[9] Hager, W. and Zhang, H., "A survey of non-linear conjugate gradient methods", Pacific J. Optimization , 2:35-58 (2006).

[10] Hestenes, M. R., and Stiefel, E. L., "Methods of conjugate gradient for solving linear systems". J. Research Nat. Bur. Standards, 49:409-436 (1952).

[11] Liu, J., Jiang, Y., "Global Convergence of a Spectral Conjugate Gradient Method for Unconstrained Optimization". Abstract and Applied Analysis, Hindawi Publishing Cororation, 2012:1-12 (2012).

doi:10.1155/2012/758287.

[12] Liu, D. and Story, C., "Efficient generalized conjugate gradient algorithms., part l : Theory, J. Optimization Theory and Applications", 69: 129-137 (1991), .

[13] Polak, E., and Ribiere, G., "Not sur la convergence de directions conjugue'e". Rev. Franaise Informant Researche Operationelle, 3e Anne'e., 16:35-43 (1969a).

[14] Polyak, B. T., (1969), "The conjugate gradient method in extreme problems". URSS Comp. Math. Phys., 9: pp.94-112 (1969b).

[15] Matonoha, C., Luksan, L. and Vlcek, J., " Computational experience with conjugate gradient methods for unconstrained optimization", Technical Report, 1038:1-17 (2008).

[16] Nocedal, J. and Glibart, J., "Global convergence properties of conjugate gradient methods for optimization", SIAM J. Optimization, 2:21-42 (1992).

[17] Raydan, M., " The Barzilai, and Borwein gradient method for the large scale unconstrained minimization problem", SIAM J. Optimization, 7:26-33 (1997).

[18] Wolfe, P., Convergence conditions for ascent methods. SIAM Review, 11:226-235 (1969).

[19] Zhang, L., Zhou, W. and Li, D. H., "Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line-search", Numerical Mathematics, 104:561-572 (2011).

doi:10.1007/s0021-006-0028-z.

[20] Zhong, W., Zhan, L.Y. and Ya, L.W., " New spectral PRP conjugate gradient method for constrained optimization, 24:16-22 (2011).

[21] Zoutendijk, G., Nonlinear programming computational methods, In Integer and Nonlinear programming, J. Abadie (Ed.), North-Holland, Amsterdam (1970).