SQU Journal for Science, 2022, 90-99 DOI: 10.53539/squjs.vol27iss2pp90-99
Sultan Qaboos University

Solving unconstrained optimization problems by a new conjugate gradient method with sufficient descent property

Farzad Rahpeymaii[1*], Majid Rostami[2]

[1] Department of Mathematics, Technical and Vocational University (TVU), Tehran, Iran,
[2] Department of Mathematics, Technical and Vocational University (TVU), Tehran, Iran,
****Email: rahpeyma_83@yahoo.com.***

ABSTRACT: There have been some conjugate gradient methods with strong convergence but numerical instability and conversely. Improving these methods is an interesting idea to produce new methods with both strong convergence and numerical stability. In this paper, a new hybrid conjugate gradient method is introduced based on the Fletcher formula (CD) with strong convergence and the Liu and Storey formula (LS) with good numerical results. New directions satisfy the sufficient descent property, independent of line search. Under some mild assumptions, the global convergence of new hybrid method is proved. Numerical results on unconstrained CUTEst test problems show that the new algorithm is very robust and efficient.

Keywords: Conjugate gradient methods; Unconstrained optimization; Global convergence; Sufficient descent property; Numerical results.

حل وتحسين مسائل غيرالمقيّدة من خلال طريقة الميول المترافقة خاصية الإنحدارالكافيه

فرزاد راهپيمايى، مجيد رستمى

الملخص: هناك بعض من طرق الإنحدار المزدوج، التى توافقت بقوة، لكن غير مستقرّة وغيرثابتة عدديا و عكسيا. تحسين هذه الطرق فكرة مثيرة لإنتاج طرق جديدة بالتقارب العددى القوى و الثابت. فى هذه المقالة، قد عرّفت و قدّمت طريقة الإنحدار المزدوج المركّب الجديد على أساس صيغة فلجر (CD) بالتقارب القوى و صيغة ليو واستورى (LS) بنتائج رقمية جيدة. تحت الشروط القياسية وبالنتائج العددية الجيدة. الجهات الجديدة المستقلة عن البحث الخطي، تم اثبات التقارب الجماعي للطريقة المركبة الجديدة. النتائج العددية على اختبار القضايا والمسائل غير المقيدة CUTEst تدلّ على أنّ الخوارزمية الجديدة، قويّة جدّا.

الكلمات المفتاحية: طرق الإنحدار المزدوج- التحسين غير المقيّد- التقارب الجماعى، شرط التقليل الكافى- النتائج العددية.

1. Introduction

Consider the unconstrained optimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1}$$

where $f:\mathbb{R}^n \to \mathbb{R}$ is a smooth nonlinear function whose gradient at x is available $g := g(x) = \nabla f(x)$. There are many iterative methods to solve unconstrained optimization problem including the Newton methods, the quasi-Newton methods, trust-region methods [21].

1.1 Conjugate gradient method

The conjugate gradient (CG) methods are famous iterative methods for solving large-scale unconstrained optimization problems whose iterative scheme is

$$x_{k+1} := x_k + \alpha_k d_k, \qquad x_0 \in \mathbb{R}^n. \tag{2}$$

Here $x_0 \in \mathbb{R}^n$ is an initial point, $\alpha > 0$ is a step size, which is obtained by an exact or inexact line search methods and d_k is a search direction computed by

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k g_{k-1}, & k \geq 1, \end{cases} \tag{3}$$

in which β_k is a scalar, called the CG parameter and $g_k := g(x_k)$. Different choices for CG parameter are available, some of which are as follows

$$\beta_k^{FR} := \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \qquad \text{Fletcher \& Reeves (FR) [8]} \tag{4}$$

$$\beta_k^{HS} := \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \qquad \text{Hestenes \& Stiefel (HS) [15]} \tag{5}$$

$$\beta_k^{CD} := -\frac{\|g_k\|^2}{g_{k-1}^T d_{k-1}}, \qquad \text{Fletcher (CD) [7]} \tag{6}$$

$$\beta_k^{PRP} := \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \qquad \text{Polak \& Ribire − Polyak (PRP) [23,24]} \tag{7}$$

$$\beta_k^{DY} := \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}, \qquad \text{Dai \& Yuan (DY) [4]} \tag{8}$$

$$\beta_k^{LS} := -\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}}, \qquad \text{Liu \& Storey (LS) [20]} \tag{9}$$

$$\beta_k^{HZ} := \left(y_{k-1} - 2d_{k-1}\frac{\|y_{k-1}\|^2}{d_{k-1}^T y_{k-1}}\right)^T \frac{g_k}{d_{k-1}^T y_{k-1}}, \qquad \text{Hager \& Zhang (HZ) [13]} \tag{10}$$

in which $\|\,.\,\|$ is the Euclidean norm and $y_{k-1} := g_k - g_{k-1}$. These methods require low memory.

In order to guarantee the global convergence of CG methods, the search direction d_k must satisfy the sufficient descent condition [1]

$$g_k^T d_k \leq -c\|g_k\|^2, \qquad for\ all\ k \geq 0, \tag{11}$$

in which c is a positive constant. In CG methods to solve (1), after determining the descent search direction satisfying $g_k^T d_k \leq 0$ for all $k \geq 0$, the step size α_k needs to be found, which can be computed by inexact line search such as Armijo, Goldstein and Wolfe conditions. For a given constant $\rho \in (0,1)$, the Armijo line search is

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \rho \alpha_k g_k^T d_k. \tag{12}$$

The Armijo condition (12) along with

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k, \qquad 0 < \rho < \sigma < 1, \tag{13}$$

is called the Wolfe line search [21]. In addition, in a strong Wolfe line search, (13) is changed as

$$|g_{k+1}^T d_k| \leq -\sigma g_k^T d_k. \tag{14}$$

1.2 Applications of CG method

Conjugate gradient methods play an important role in solving large-scale unconstrained optimization problems which arise in economics, engineering, sciences and so on. Currently, the unconstrained optimization problems in impulse noise removal and image restoration have been solved by CG methods [2,17].

1.3 Contribution

Although LS has a good performance in practice, it is generally not a strong convergence. On the other hand, in CD with strong convergence the numerical results are not efficient. The purpose of this paper is to overcome these drawbacks. We improve and combine LS and CD to obtain a new method with a good performance in practice and strong convergence properties. The new method always produces a sufficient descent direction which the global convergence of it is established under some suitable assumptions. Also, we give some preliminary numerical experiments to illustrate the efficiency of new method.

In Section 2, we describe our proposed method and give its algorithm. The sufficient descent condition of new direction and the global convergence of the new algorithm are established in Section 3. In Section 4, we report some numerical results to show the efficiency of new method. Finally, we give our conclusion in Section 5.

2. Motivation and proposed algorithm

In this section, we describe the new method to solve unconstrained optimization problem (1). The LS proposed by Liu and Story [20]. The global convergence of LS with Grippo & Lucidi line search is established in [18]. This method has good numerical results but its global convergence properties are not strong. Many researchers have proposed several modifications of LS, see [19,26,27]. Fletcher in [7] proposed CD for a general objective function with strong convergence properties, but in numerical performance is weak. We improve and combine CD and LS to obtain a new method with both the strong convergence and the good numerical results. Based on LS and CD parameters, we obtain

$$\beta_k^1 := -\frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}} \qquad and \qquad \beta_k^2 := -\frac{\|y_{k-1}\|^2}{g_{k-1}^T d_{k-1}}. \tag{15}$$

Here, $\beta_k^1 := \beta_k^{LS}$ and in the numerator of β_k^2 term $\|y_{k-1}\|^2$ has replaced $\|g_k\|^2$ in parameter β_k^{CD}. Therefore, the new conjugate gradient parameter is obtained by

$$\beta_k^{new} := t_k \beta_k^2 - \beta_k^1, \tag{16}$$

with

$$t_k := 2\frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}. \tag{17}$$

The parameters β_k^1, β_k^2 and t_k guarantee the sufficient descent property and the global convergence. Furthermore, these parameters improve numerical results of new method in compared to CD and LS methods. Now, by substituting (15) and (17) in (16), we get

$$\beta_k^{new} := -2\frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}} \frac{\|y_{k-1}\|^2}{g_{k-1}^T d_{k-1}} + \frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}}. \tag{18}$$

Finally, the new search direction d_k is computed by

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k^{new} d_{k-1}, & k \geq 1. \end{cases} \tag{19}$$

Algorithm 1 solves the smooth unconstrained optimization problem (1). It takes the initial point $x_0 \in \mathbb{R}^n$ as input and uses the following parameters: $\varepsilon > 0$ (minimum threshold for the stopping test), $\kappa \in (0,1)$ and $0 < \rho < \sigma < 1$ (Line search parameters), k_{max} (maximum number of iterations), and $0 < \alpha_{min} < \alpha_{max} < \infty$ (minimum and maximum values for α_k). It returns $x^* := x_k$ and $f^* := f_k$ as an optimum and its function value. Note that α_{min} and α_{max} prevent the production of too small and large step size, respectively.

Algorithm 1 A new conjugate gradient method

(S0) Compute the initial function value $f_0 := f(x_0)$, the initial gradient vector $g_0 := g(x_0)$ and set $d_0 := -g_0$.

(S1) If $\|g_k\| \leq \varepsilon$ or $k > k_{max}$, stop.

(S2) Find α_k satisfying (12) and (14) and restrict $\alpha_k = max\{\alpha_{min}, min\{\alpha_k, \alpha_{max}\}\}$. Then, compute $x_{k+1} := x_k + \alpha_k d_k$, $f_{k+1} := f(x_{k+1})$ and $g_{k+1} := g(x_{k+1})$.

(S3) Calculate β_{k+1}^1 and β_{k+1}^2 by (15), obtain the parameter t_{k+1} by (17), determine the parameter β_{k+1}^{new} by (16) and $d_{k+1} := -g_{k+1} + \beta_{k+1}^{new} d_k$.

(S4) Replace k by $k+1$ and go to (S1).

3. Convergence analysis

In this section, the sufficient descent property and the global convergence of the Algorithm 1 are established. To do so, we make some assumptions on the objective function.

(H1) The level set $L(x_0) = \{x \in \mathbb{R}^n \,|\, f(x) \leq f(x_0)\}$ is bounded, i.e., there exists a constant $B > 0$ such that

$$\|x\| \leq B, \qquad for\ all\ x \in L(x_0). \tag{20}$$

(H2) In some neighborhood $\Omega \subseteq L(x_0)$, the gradient of the objective function f is Lipschitz continuous, i.e., there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \qquad for\ all\ x \in \Omega. \tag{21}$$

From **(H1)** and **(H2)**, there exists a positive constant γ such that

$$\|g(x)\| \leq \gamma, \tag{22}$$

see [21]. We now show that the generated directions by Algorithm 1 satisfy the sufficient descent condition (11) with $c = \frac{7}{8}$ independent of line search type.

Lemma 1 Suppose that the direction d_k is generated by Algorithm 1. Then, we have

$$g_k^T d_k \leq -\frac{7}{8}\|g_k\|^2. \tag{23}$$

Proof By multiplying (19) in g_k^T and using (18), we obtain

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

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

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

Take

$$\upsilon_k := 2(g_k^T d_{k-1}) y_{k-1} \qquad \tilde{\upsilon}_k := \frac{1}{2}\,(g_{k-1}^T d_{k-1}) g_k.$$

Using $\upsilon_k^T \tilde{\upsilon}_k \le \frac{1}{2}(\|\upsilon_k\|^2 + \|\tilde{\upsilon}_k\|^2)$, we get

$$g_k^T d_k \le \frac{1}{\Theta_1^2}\left(-\|g_k\|^2 \Theta_1^2 + 2\Theta_2^2 \|y_{k-1}\|^2 + \frac{1}{8}\Theta_1^2 \|g_k\|^2 - 2\|y_{k-1}\|^2 \Theta_2^2\right) = -\frac{7}{8}\|g_k\|^2,$$

where

$$\Theta_1 := g_{k-1}^T d_{k-1} \qquad and \qquad \Theta_2 := g_k^T d_{k-1}.$$

Therefore, the search direction d_k always satisfies the sufficient descent condition.

■

Lemma 2 Suppose that **(H2)** holds and the tuning parameter α_{max} is given. Then, there exists a constant $\omega_1 > 0$ such that

$$|\beta_k^1| \le \omega_1 \frac{\|d_{k-1}\|}{\|g_{k-1}\|^2}. \tag{24}$$

Proof From Lemma 1, we have

$$g_{k-1}^T d_{k-1} \le -\frac{7}{8}\|g_{k-1}\|^2,$$

so that

$$\frac{1}{|g_{k-1}^T d_{k-1}|} \le \frac{8}{7\|g_{k-1}\|^2}. \tag{25}$$

By the definition of β_k^1, we obtain

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

Then, from the Cauchy-Schwarz inequality, **(H2)**, (22) and (25), we get

$$|\beta_k^1| \le \frac{\|g_k\|\|y_{k-1}\|}{|g_{k-1}^T d_{k-1}|} \le \frac{8\gamma L \alpha_{k-1}\|d_{k-1}\|}{7\|g_{k-1}\|^2} = \omega_1 \frac{\|d_{k-1}\|}{\|g_{k-1}\|^2},$$

with $\omega_1 := \frac{8\gamma L \alpha_{max}}{7}$.

■

Lemma 3 Suppose that **(H2)** holds and the tuning parameter α_{max} is given. Then, there exists a constant $\omega_2 > 0$ such that

$$|\beta_k^2| \le \omega_2 \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^2}. \tag{26}$$

Proof **(H2)**, (15) and (25) result in

$$|\beta_k^2| = \left|-\frac{\|y_{k-1}\|^2}{g_{k-1}^T d_{k-1}}\right| = \frac{\|y_{k-1}\|^2}{|g_{k-1}^T d_{k-1}|} \le \frac{8L^2 \alpha_{k-1}^2 \|d_{k-1}\|^2}{7\|g_{k-1}\|^2} = \omega_2 \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^2},$$

where $\omega_2 := \frac{8L^2 \alpha_{max}^2}{7}$.

■

Lemma 4 If **(H2)** holds, then there exists constant $\omega_3 > 0$ such that

$$|t_k| \le \omega_3 \frac{\|d_{k-1}\|}{\|g_{k-1}\|^2}. \tag{27}$$

Proof The Cauchy-Schwarz inequality, (17), (22) and (25) imply

$$|t_k| = \left|2\frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}\right| \leq 2\frac{\|g_k\|\|d_{k-1}\|}{|g_{k-1}^T d_{k-1}|} \leq \frac{16\gamma\|d_{k-1}\|}{7\|g_{k-1}\|^2} = \omega_3\frac{\|d_{k-1}\|}{\|g_{k-1}\|^2},$$

in which $\omega_3 := \frac{16\gamma}{7}$.

■

Lemma 5 Suppose that **(H1)** and **(H2)** hold. If the sequence $\{d_k\}$ be generated by Algorithm 1, then there exists a constant $M > 0$ such that

$$\|d_k\| \leq M, \qquad for\ all\ k \geq 0. \tag{28}$$

Proof

We use induction to prove this lemma. First (22) implies that

$$\|d_0\| = \|g_0\| \leq \gamma.$$

From the assumption of induction $\|d_{k-1}\|$ is bounded. Hence, the exists constant $M^* > 0$ such that

$$\|d_{k-1}\| \leq M^*. \tag{29}$$

Now, the definitions of d_k and β_k^{new}, give

$$\|d_k\| = \left\|-g_k + \beta_k^{new} d_{k-1}\right\| \leq \left\|g_k\right\| + \left|\beta_k^{new}\right|\|d_{k-1}\|$$

$$\leq \left\|g_k\right\| + \left(\left|\beta_k^1\right| + |t_k|\left|\beta_k^2\right|\right)\|d_{k-1}\|.$$

We get from (22), (29) and Lemmas 2-4 that

$$\|d_k\| \leq \gamma + \left(\omega_1\frac{\|d_{k-1}\|}{\left\|g_{k-1}\right\|^2} + \omega_2\omega_3\frac{\|d_{k-1}\|^3}{\left\|g_{k-1}\right\|^4}\right)\|d_{k-1}\|$$

$$\leq \gamma + \left(\frac{\omega_1}{\varepsilon^2}\|d_{k-1}\| + \frac{\omega_2\omega_3}{\varepsilon^4}\|d_{k-1}\|^3\right)\|d_{k-1}\|.$$

$$\leq \gamma + \left(\frac{\omega_1}{\varepsilon^2}M^* + \frac{\omega_2\omega_3}{\varepsilon^4}(M^*)^3\right)M^* := M.$$

■

The following result is a theorem in [29].

Lemma 6 Let d_k be a sufficient descent direction and assume the step size α_k satisfies the strong Wolfe line search (12) and (14). Then, based on (H1) and (H2), we have

$$\sum_{k=0}^{\infty}\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty. \tag{30}$$

Theorem 1 Let d_k be a sufficient descent direction and $\{x_k\}$ be the generated sequence by Algorithm 1. Also, (H1) and (H2) hold. Then

$$\lim_{n\to\infty}\inf\|g_k\| = 0. \tag{31}$$

Proof From Lemma 5, we obtain

$$\sum_{k=0}^{\infty}\frac{1}{\|d_k\|^2} \geq \sum_{k=0}^{\infty}\frac{1}{M^2} = +\infty. \tag{32}$$

We use contradiction to prove this theorem. Hence, there exists a constant $\epsilon > 0$ such that

$$\|g_k\| \geq \epsilon. \tag{33}$$

Let

$$\Xi_k = \frac{g_k^T d_k}{\|g_k\|\|d_k\|}. \tag{34}$$

Using Lemma 1, we obtain

$$\Xi_k = \frac{g_k^T d_k}{\|g_k\|\|d_k\|} \leq -\frac{7}{8}\frac{\|g_k\|^2}{\|g_k\|\|d_k\|} = -\frac{7}{8}\frac{\|g_k\|}{\|d_k\|}, \tag{35}$$

so that

$$\Xi_k^2 \geq \frac{49}{64}\frac{\|g_k\|^2}{\|d_k\|^2}. \tag{36}$$

From (33), (34) and (36), we can obtain

$$\frac{49}{64}\frac{\epsilon^2}{\|d_k\|^2} \leq \frac{49}{64}\frac{\|g_k\|^2}{\|d_k\|^2} \leq \Xi_k^2 = \frac{\left(g_k^T d_k\right)^2}{\|g_k\|^2\|d_k\|^2} \leq \frac{\left(g_k^T d_k\right)^2}{\epsilon^2\|d_k\|^2}. \tag{37}$$

By taking sums from both sides (37) and using Lemma 6, we have

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty,$$

which contradicts with (32). Hence, the proof of the desired result is completed.

■

4. Numerical experiments

This section gives numerical results of some algorithms on a set of the nonlinear unconstrained optimization test problems from CUTEst collection [11], given in Table 1. The dimensions of test problems are from 2 to 12005 while the initial points are standard ones proposed in CUTEst. We apply the following algorithms to solve these test problems:

- **M1:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^1 d_{k-1}$.
- **M2:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^2 d_{k-1}$.
- **M3:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^{new} d_{k-1}$.
- **M4:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^{new+} d_{k-1}$ and $\beta_k^{new+} \coloneqq max\{0, \beta_k^{new}\}$.
- **DY:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^{DY} d_{k-1}$.
- **HZ:** Conjugate gradient method with $d_k \coloneqq -g_k + \beta_k^{HZ} d_{k-1}$.

All algorithms are implemented in Matlab 2011 programming environment on a 2.3Hz Intel core i3 processor laptop and 4GB of RAM with the double precision data type in Linux operations system. All algorithms are terminated whenever the inequality $\|g_k\| < 10^{-6}$ holds or the maximum number of iterations exceeds 10000. The tuning strong Wolfe line search parameters are taken as $\rho = 0.0001$, $\sigma = 0.9$, $\alpha_{min} = 10^{-8}$ and $\alpha_{max} = 10^8$.

Here, we use the performance profiles of Dolan & More [5] to compare **M1**, **M2**, **M3**, **M4**, **DY** and **HZ** algorithms on the test problems. We consider P as designates the percentage of problems which are solved within a factor τ of the best solver. The horizontal axis of the figure gives the percentage of the test problems for which an algorithm is the fastest (efficiency), while the vertical axis gives the percentage of the test problems that were successfully solved by each algorithm (robustness).

Figures 1-3 show that **M4** is the best in terms of the total number of iterations, the total number of function evaluations and time in seconds in comparison with others.

Table 1. Test functions taken from CUTEst collection.

No.	Test function	Dim	No.	Test function	Dim	No.	Test function	Dim
1	3PK	3	49	DQDRTIC	10000	97	NONDIA	1000
2	AIRCRFTB	8	50	DQRTIC	5000	98	NONDQUAR	3000
3	ALLINIT	4	51	EDENSCH	100	99	OSCIPANE	5000
4	ALLINITU	4	52	EG2	1000	100	OSCIPATH	2
5	ARGLINA	500	53	EG3	10000	101	OSLBQP	8
6	ARGLINB	200	54	EIGENA	2550	102	PALMER1C	8
7	ARWHEAD	5000	55	ENGVAL1	100	103	PALMER1D	7
8	BARD	3	56	ENGVAL2	3	104	PALMER2C	8
9	BDQRTIC	100	57	ERRINROS	50	105	PALMER3C	8
10	BEALE	2	58	EXPFIT	2	106	PALMER4C	8
11	BIGGS6	6	59	EXTROSNB	1000	107	PALMER5C	6
12	BIGGSB1	100	60	FLETCBV2	10000	108	PALMER6C	8
13	BOX2	3	61	FLETCHCR	500	109	PALMER7C	8
14	BOX3	3	62	FMINSRF2	5625	110	PALMER8A	6
15	BRKMCC	2	63	FMINSURF	5625	111	PALMER8C	8
16	BROWNDEN	4	64	FREUROTH	2	112	PENALTY1	100
17	BROYDN3D	5000	65	GENHUMPS	500	113	PENALTY2	50
18	BROYDN7D	500	66	GENROSE	500	114	POWELLBC	1000
19	BROYDNBD	5000	67	GROWTHLS	3	115	POWELLSG	5000
20	BRYBND	500	68	GULF	3	116	QR3DLS	610
21	CHAINWOO	1000	69	HAIRY	2	117	QUARTC	25
22	CHNROSNB	50	70	HATFLDD	3	118	ROSENBR	2
23	CLIFF	2	71	HATFLDF	3	119	S308	2

24	COSINE	1000	72	HATFLDFL	3	120	SCHMVETT	100
25	CRAGGLVY	1000	73	HEART6LS	6	121	SENSORS	2
26	CUBE	2	74	HEART8LS	8	122	SINEVAL	2
27	CUBENE	2	75	HELIX	3	123	SINVALNE	2
28	DALLASM	196	76	HILBERTA	10	124	SISSER	2
29	DALLASS	46	77	HILBERTB	10	125	SNAIL	2
30	DECONVU	63	78	HIMMELBA	2	126	SPARSINE	1000
31	DENSCHNA	2	79	HIMMELBC	2	127	SPARSQUR	10000
32	DENSCHNB	2	80	HIMMELBF	4	128	SPMSRTLS	4999
33	DENSCHNC	2	81	HIMMELBG	2	129	SROSENBR	1000
34	DENSCHNF	2	82	HIMMELBH	2	130	TAME	2
35	DIXMAANA	9000	83	HUMPS	2	131	TESTQUAD	100
36	DIXMAANB	3000	84	JENSMP	2	132	TOINTGOR	50
37	DIXMAANC	3000	85	KOWOSB	4	133	TOINTGSS	10000
38	DIXMAAND	3000	86	LIARWHD	5000	134	TOINTPSP	50
39	DIXMAANE	3000	87	LOGHAIRY	2	135	TOINTQOR	50
40	DIXMAANF	3000	88	MANCINO	100	136	TQUARTIC	10
41	DIXMAANG	3000	89	MATRIX2	6	137	TRIDIA	5000
42	DIXMAANH	3000	90	METHANOL	12005	138	VAREIGVL	500
43	DIXMAANI	3000	91	MODBEALE	2	139	VIBRBEAM	8
44	DIXMAANJ	3000	92	MOREBV	5000	140	WATSON	12
45	DIXMAANK	3000	93	MSQRTALS	1024	141	WEEDS	3
46	DIXMAANL	3000	94	MSQRTBLS	1024	142	WOODS	100
47	DIXON3DQ	1000	95	MINE5D	10733	143	YFITU	3
48	DJTL	2	96	NONCVXU2	1000	144	ZANGWIL2	2

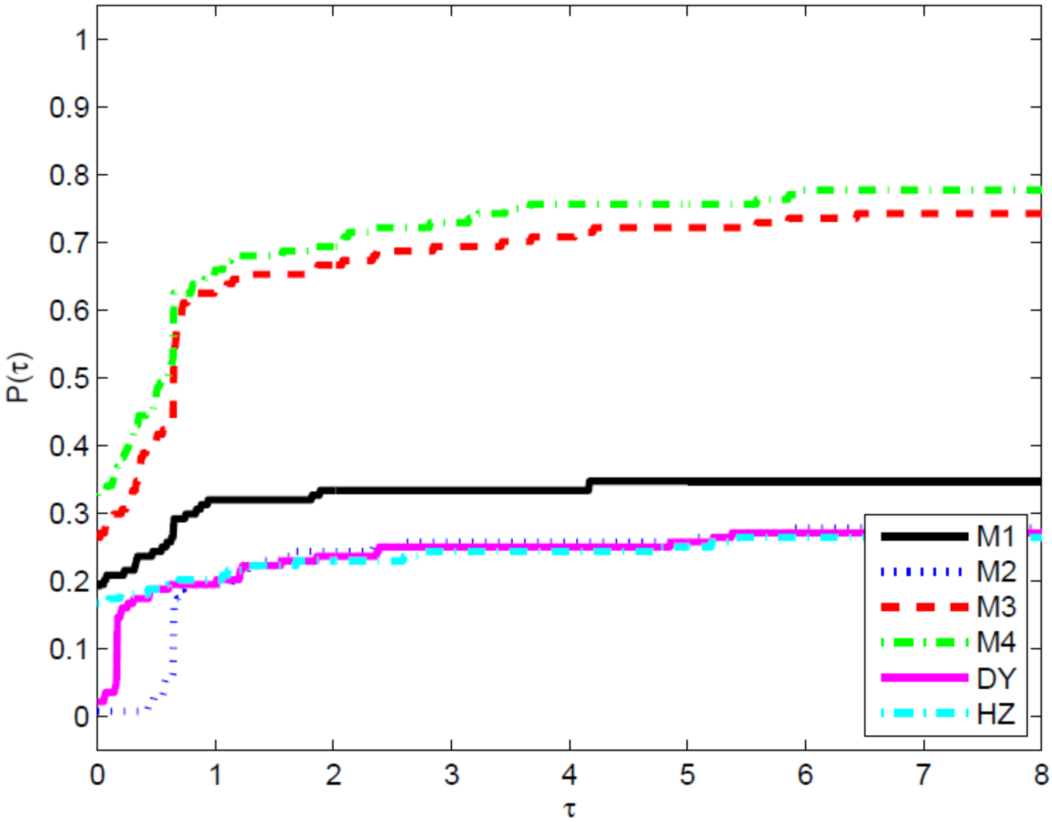

Figure 1. Dolan-More performance profiles for the total number of iterations.

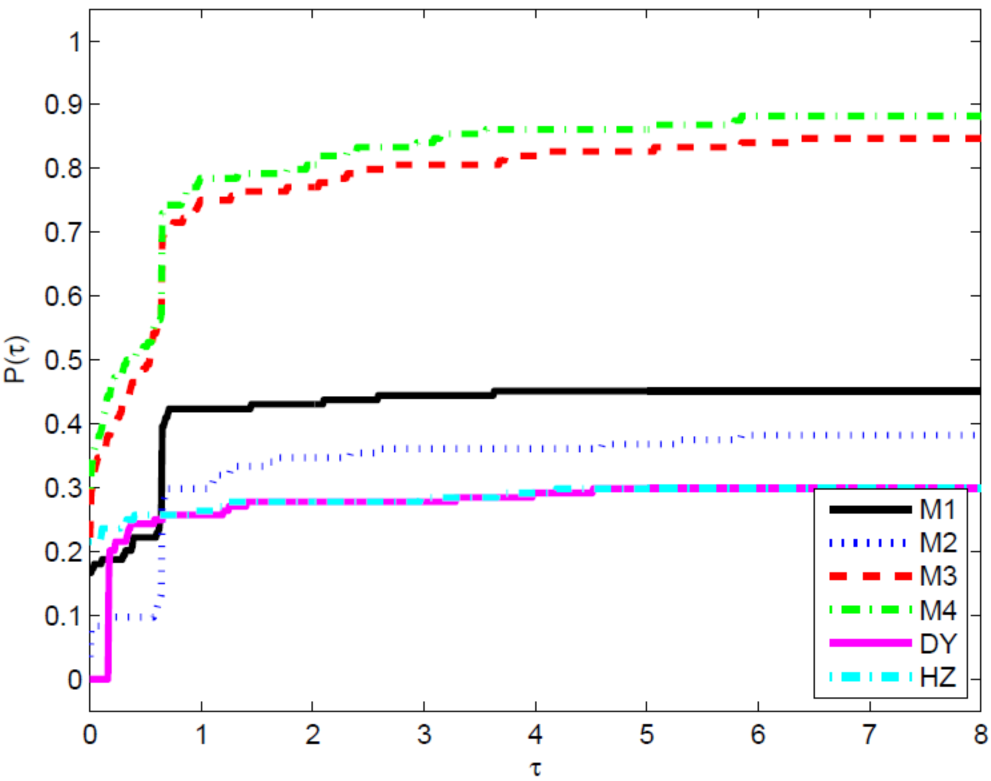

Figure 2. Dolan-More performance profiles for the total number of function evaluations.

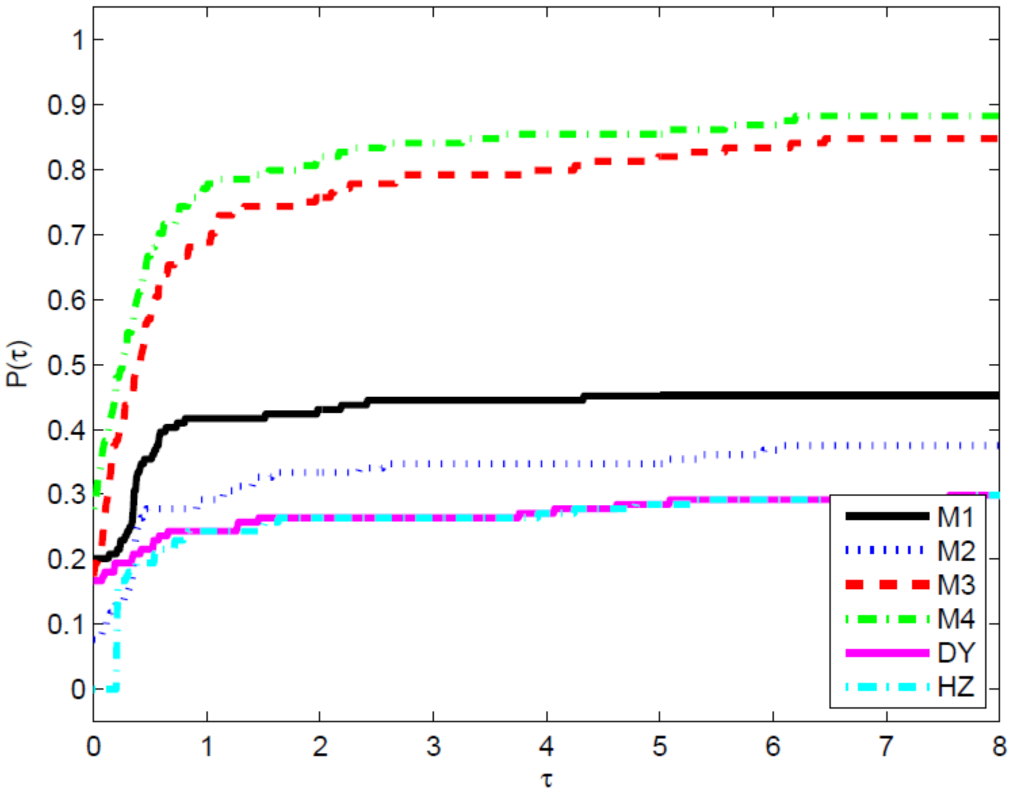

Figure 3. Dolan-More performance profiles for the time in seconds.

5. Conclusion

In this paper, we proposed a new conjugate gradient method for solving the unconstrained optimization problems with improving and combining LS and CD parameters. The new search directions of our algorithm satisfy the sufficient descent condition. It inherits the strong global convergence properties of CD and the numerical efficiency of LS. The global convergence under some mild assumptions is stablished. Our numerical experiments show that **M4** is better than other algorithms in terms of the total number of iterations, the number of function evaluations and time in seconds.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgment

The authors would like to thank the editor and reviewers for their valuable comments which helped to improve the manuscript.

References

1. Al-Baali, M. Descent property and global convergence of the Fletcher-Reeves method with inexact line search, *IMA Journal of Numerical Analysis*, 1985, **5**, 121-124.
2. Alhawarat, A., Salleh, Z. and Masmali, I.A. A convex combination between two di_erent search directions of conjugate gradient method and application in image restoration, *Mathematical Problems in Engineering*, 2021.
3. Conn, A. R., Gould, N.I.M. and Toint, P.L. Convergence of quasi-Newton matrices generated by the symmetric rank one update, *Mathematical programming*, 1991, **50**, 177-195.
4. Dai, Y.H. and Yuan, Y. A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Journal on Optimization*, 1999, **10**, 177-182.
5. Dolan, E.D. and More, J.J. Benchmarking optimization software with performance profiles, *Mathematical Programming*, 2002, **91**, 201-213.
6. Esmaeili, H. and Kimiaei, M. An improved adaptive trust-region method for unconstrained optimization, *Mathematical Modelling and Analysis*, 2014, **19(4)**, 469-490.
7. Fletcher, R. Practical Methods of Optimization, John Wiley & Sons, New York, 1987.
8. Fletcher, R. and Reeves, C. Function minimization by conjugate gradients, *Computer Journal*, 1964, **7(2)**, 149-154.
9. Gill, P. and Murray, W. Quasi-Newton methods for unconstrained optimization, *IMA Journal of Applied Mathematics*, 1972, **9(1)**, 91-108.
10. Gill, P. and Leonard, M. W. Reduced-Hessian quasi-Newton methods for unconstrained optimization, *SIAM Journal on Optimization*, 2001, **12(1)**, 209-237.
11. Gould, N.I.M., Orban, D. and Toint, P.L. CUTEst: A constrained and unconstrained testing environment with safe threads, *Technical Report*, RAL-TR-2013-005.
12. Griewank, A. The global convergence of Broyden-like methods with a suitable line search, *Journal of the Australian Mathematical Society Ser. B*, 1986, **28**, 75-92.
13. Hager, W.W. and Zhang, H. A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM Journal on optimization*, 2005, **16(1)**, 170-192.
14. Hager, W.W. and Zhang, H. The limited memory conjugate gradient method, *SIAM Journal on Optimization*, 2013, **23(4)**, 2150-2168.
15. Hestenes, M.R. and Stiefel, E.L. Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Bureau of Standards*, 1952, **49**, 409-436.
16. Kimiaei, M. and Ghaderi, S. A new restarting adaptive trust-region method for unconstrained optimization, *Journal of the Operations Research Society of China*, 2017, **5**, 487-507.
17. Kimiaei, M. and Rostami, M. Impulse noise removal based on new hybrid conjugate gradient approach, *Kybernetika*, 2016, **5**, 791-823.
18. Li, Z.F., Chen, J. and Deng, N.Y. A new conjugate gradient method and its global convergence properties, *Mathematical Programming*, 1997, **78**, 375-391.
19. Li, M. and Feng, H. A sufficient descent LS conjugate gradient method for unconstrained optimization problems, *Applied Mathematics and Computation*, 2011, **218**, 1577-1586.
20. Liu, Y.L. and Storey, C. Efficient generalized conjugate gradient algorithms, part1, *Journal of Optimization Theory and Applications*, 1991, **69(1)**, 129-137.
21. Nocedal, J. and Wright, S.J. Numerical Optimization, Springer series in Operation Research, Springer, New York,1999.

22. Ortega, J.M. and Rheinboldt, W.C. Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
23. Polak, E. and Ribiere, G. Note sur la convergence de directions conjugees, *Rev. Francaise Informat Recherche Opertionelle, 3e Annee*, 1969, **16**, 35-43.
24. Polyak, B.T. The conjugate gradient method in extreme problems, *USSR Computational Mathematics and Mathematical Physics*, 1969, **9**, 94-112.
25. Rahpeymaii, F., Kimiaei, M. and Bagheri, A. A limited memory quasi-Newton trust-region method for box constrained optimization, *Journal of Computational and Applied Mathematics,* 2016, **303**, 105-118.
26. Yang, X., Luo, Z. and Dai, X. A global convergence of LS-CD hybrid conjugate gradient method, *Advances in Numerical Analysis*, https://doi.org/10.1155/2013/517452.
27. Zhang, L. A new Liu-Storey type nonlinear conjugate gradient method for unconstrained optimization problems, *Journal of Computational and Applied Mathematics*, 2009, **225**, 146-157.
28. Zhou, Q., Zhou, F. and Cao, F. A nonmonotone trust region method based on simple conic models for Unconstrained optimization, *Applied Mathematics and Computation*, 2013, **225**, 295-305.
29. Zoutendijk, G. Nonlinear Programming, Computational Methods. *In: Abadie J. (ed) Integer and Nonlinear Programming. Amsterdam,* North-Holland, 1970, 37-86.

Received 13 July 2021
Accepted 16 June 2022