Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 2, November 2022, pp. 1184~1191
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i2.pp1184-1191

Global convergence of a modified RMIL+ nonlinear conjugate gradient method with strong wolfe

Abdelrhaman Abashar[1], Osman Omer Osman Yousif[2], Awad Abdelrahman Abdalla Mohammed[2], Mohammed A. Saleh[3]
[1]Department of Basic Sciences, Faculty of Engineering, Red Sea University (RSU), Portsudan, Sudan
[2]Department of Mathematics, Faculty of Mathematical and Computer Sciences, University of Gezira (U of G), Wad Medani, Sudan
[3]Department of Computer, College of Science and Arts in Ar Rass, Qassim University, Saudi Arabia

Article Info

Article history:

Received Jul 3, 2022
Revised Aug 20, 2022
Accepted Aug 30, 2022

Keywords:

Conjugate gradient method
Global convergence
Strong wolfe line search
Sufficient descent property
Unconstrained optimization

ABSTRACT

Nonlinear conjugate gradient (CG) methods are extensively used as an important technique for addressing large-scale unconstrained optimization problems which are arise in many aspects of science, engineering, and economics. That is due to their simplicity, convergence properties, and low memory requirements. To generate a new approximation solution in each iteration, the CG methods usually implement under the strong Wolfe line search. For good performance, many studies have been carried out to modify well-known CG methods. In this paper, we did some modifications on one of CG method called RMIL+ in order to obtain a new CG method possesses the sufficient descent property and the global convergence under strong Wolfe line search. The numerical results demonstrate that the suggested method outperforms other CG methods.

Corresponding Author:

Abdelrhaman Abashar
Department of Basic Sciences, Faculty of Engineering, Red Sea University
Red Sea State, Portsudan, Sudan
Email: abashar3373@gmail.com

1. INTRODUCTION

Considering the next unconstrained optimization problem,

$$\min f(x)\,,\quad x \in R^n, \tag{1}$$

where $f: R^n \rightarrow R$ is a continuous and differentiable. The conjugate gradient (CG) method considered as one of the choicest for solving (1), particularly for the case n is large. The nonlinear conjugate gradient method's iterative formula is given by,

$$x_{k+1} = x_k + \alpha_k d_k\,, \tag{2}$$

where x_k is present iterate point and α_k is step length, which is calculated by performing a line search, and d_k is the search direction, which is defined by,

$$d_{k+1} = \begin{cases} g_{k+1}\,, & if \quad k=0 \\ -g_{k+1} + \beta_{k+1} d_k, & if \quad k \geq 1, \end{cases} \tag{3}$$

where β_k is paramter. The classical conjugate gradient method includes the Hestenes and Stiefel [1]. The Fletcher and Reeves [2]. The Polak [3], method Polyak and Ribiere [4]. The conjugate descent method [5]. The Liu and Storey method, [6] and the Dai and Yuan method [7], the parameters β_k of these methods are as follow:

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, \qquad \beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \qquad \beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}},$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}, \qquad \beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}, \qquad \beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}.$$

usually, in the conjugate gradient methods convergence analysis and implementation, the step length α_k is requireed to satisfy some line search to be imprecise line searches [8]-[11], such as an Armijo line search or a strong Wolfe line search. The strong Wolfe line search is utilized to find α_k such as:

$$f(x_k + \alpha_k d_k) - f(x_k) \le \delta \alpha_k g_k^T d_k,$$
$$|g(x_k + \alpha_k d_k)^T d_k| \le -\sigma g_k^T d_k \qquad (4)$$

where $\delta \in \left(0, \frac{1}{2}\right), and\ \sigma \in (\delta, 1)$.

The nonlinear conjugate gradient method's sufficient descent condition is as (5).

$$g_k^T d_k \le -c\|g_k\|^2, \forall k \ge 1, c \in (0,1). \qquad (5)$$

The sufficient decent property and the global convergence have been studied by many researchers such as Baali [12] who established the global convergence of the FR method under strong Wolfe line search, Liu *et al.* [13] and Dai and Yuan [14] extended the results to $\sigma = \frac{1}{2}$, Gilbert and Nocedal [15], established the global convergence property of the PRP^+ method, the PRP^+ indicated that is a non-negative parameter. For more studies [16]-[18].

This paper is organised into four sections. In section 2, a new parameter for the coefficient β_k is proposed followed by an algorithm. The sufficient descent condition and the global convergence analysis under strong Wolfe line search is presented in subsection 2.1. In section 3, the numerical performance of the new formula versus other well-known conjugate gradient methods are presented. Finally, in section 4, the conclusion is presented.

2. THE MODIFICATION METHOD

Recently, Rivaie *et al.* [19], [20], proposed two new formulas as (6), (7).

$$\beta_k^{RMIL} = \frac{g_k^T (g_k - g_{k-1})}{\|d_{k-1}\|^2}, \qquad (6)$$

$$\beta_k^{RMIL+} = \frac{g_k^T (g_k - g_{k-1} - d_{k-1})}{\|d_{k-1}\|^2}, \qquad (7)$$

Zhang [21] presented an improved formula called β_k^{NPRP} to Wei-Yao-Liu which is given by,

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{\|g_{k-1}\|^2}, \qquad (8)$$

the NPRP method satisfies the descent property given in condition (5).

Motivated by the in (7) and (8), we propose a modified formula of RMIL+ as:

$$\beta_k^{AO} = \begin{cases} \dfrac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)}, & if\ \|g_k\|^2 \ge \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}| \\ 0 & , otherwise. \end{cases} \qquad (9)$$

where AO denotes Abashar and Osman.

By defining the formula β_k^{AO}, we have a new CG method which can be described in Algorithm 2.1.

Algorithm 2.1
Step 0. Initialization, given $x_0 \in R^n$, $\varepsilon \geq 0, set\ d_0 = -g_0, k = 0$.
Step 1. If $\|g_k\| \leq \varepsilon$, then exit.
Step 2. Find α_k using (4).
Step 3. Set $x_{k+1} = x_k + \alpha_k d_k, g_{k+1} = g(x_{k+1}), if\ \|g_{k+1}\| \leq \varepsilon, then\ stop.$
Step 4. compute β_k by the (6, 7, 8, 9 and FR method), generated d_k by (3).
Step 5. Put $k = k + 1$ and go to Step 2.

2.1. Analysis of convergence

In this subsection, the analysis of Algorithm 2.1 is presented. We proved that the algorithm satisfies condition (5) and the properties of global convergence. The next lemma is required to simplify the new β_k^{AO}.

Lemma. 2.1.1

β_k^{AO} satisfies, $\beta_k^{AO} \leq \frac{\|g_k\|^2}{\|d_{k-1}\|^2}$, $\beta_k^{AO} \geq 0$.

From the definition (9), we get,

$$\beta_k^{AO} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)} \leq \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \tag{10}$$

using Cauchy- Schwarz inequality, we get (11).

$$\beta_k^{AO} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k||g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)} = 0$$
$$\beta_k^{AO} \geq 0. \tag{11}$$

Lemma. 2.1.2

Suppose that $\{g_k\}$ and $\{d_k\}$ are generated by the Algorithm 2.1 for $\sigma < \frac{1}{2}$, then,

$$\frac{\|g_k\|}{\|d_k\|} \leq 1, \quad \forall k \geq 0. \tag{12}$$

Proof. The proof is by induction. For $k = 0$, $\frac{\|g_0\|}{\|d_0\|} = 1 \leq 1$, hence (12) holds for $k = 0$.

From (3), multiplying by g_{k+1}^T , we get,

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k$$
$$\|g_{k+1}\|^2 = \beta_{k+1} g_{k+1}^T d_k - g_{k+1}^T d_{k+1} \tag{13}$$

from (4) and absolute value of (13), obained,

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

by applying (12), and substitute (10), we get,

$$\|g_{k+1}\|^2 \leq -\sigma \frac{\|g_{k+1}\|^2}{\|d_k\|^2}\|g_k\|\|d_k\| + \|g_{k+1}\|\|d_{k+1}\| \tag{14}$$
$$\|g_{k+1}\|^2 \leq -\sigma\|g_{k+1}\|^2 \frac{\|g_k\|}{\|d_k\|} + \|g_{k+1}\|\|d_{k+1}\|$$
$$\|g_{k+1}\|^2 \leq -2\sigma\|g_{k+1}\|^2 + \|g_{k+1}\|\|d_{k+1}\|$$

$$\|g_{k+1}\|^2 + 2\sigma\|g_{k+1}\|^2 \leq \|g_{k+1}\|\|d_{k+1}\| \tag{15}$$

Divided by $\|g_{k+1}\|\|d_{k+1}\|$, we get,

$$\frac{\|g_{k+1}\|}{\|d_{k+1}\|} \leq \frac{1}{1+2\sigma} \tag{16}$$

hence this holds true for $k + 1$.

Theorem 2.1.1

Assume that g_k and d_k are produced by the methods (2) and (3), respectively, and that the step size α_k is calculated by (4), if $\sigma < \frac{1}{2}$, then the relation,

$$\frac{-1+2\sigma}{1+2\sigma} \leq \frac{g_k^T d_k}{\|g_k\|^2} \leq \frac{-1-2\sigma}{1+2\sigma}, \tag{17}$$

holds. Henceforth, condition (5) holds as $g_k \neq 0$.

Proof. By induction, true if $k = 0$, assume (17) is true if $k \geq 0$, from (3), we have:

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k \tag{18}$$

Dividing both sides by $\|g_{k+1}\|^2$

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

using strong Wolfe condition (4) we have (19).

$$|\beta_{k+1} g_{k+1}^T d_k| \leq -\sigma |\beta_k|\, g_k^T d_k$$
$$-1 + \sigma\, \beta_{k+1} \frac{g_k^T d_k}{\|g_{k+1}\|^2} \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq -1 - \sigma\, \beta_{k+1} \frac{g_k^T d_k}{\|g_{k+1}\|^2} \tag{19}$$

By using (10) and Cauchy inequality, we have,

$$-1 + \sigma \frac{\|g_{k+1}\|^2}{\|d_k\|^2} \frac{\|g_k\|\|d_k\|}{\|g_{k+1}\|^2} \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq -1 - \sigma\, \frac{\|g_{k+1}\|^2}{\|d_k\|^2} \frac{\|g_k\|\|d_k\|}{\|g_{k+1}\|^2}$$
$$-1 + \sigma \frac{\|g_k\|}{\|d_k\|} \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq -1 - \sigma\, \frac{\|g_k\|}{\|d_k\|} \tag{20}$$

from the induction hypothesis (16), we obtain (21).

$$\frac{-1+2\sigma}{1+2\sigma} \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq \frac{-1-2\sigma}{1+2\sigma} \tag{21}$$

We conclude that (17), holds for $k + 1$.

Assumption 2.1.1

(i) The set $\Omega = \{x \in R^n\, f(x) \leq f(x_0)\}$ is bounded with an initial guess x_0.

(ii) f is continuously differentiable and its gradient is Lipschitz continuous in some neighborhood N of Ω, that is, there exists a constant $l > 0$ such that $\|g(x) - g(y)\| \leq l\, \|x - y\|, \forall\, x, y \in N)$.

Theorem 2.1.2

Suppose that Assumption 2.1.1 holds. Let $\{g_k\}$ be obtained by Algorithm 2.1, then $\lim_{k\to\infty} \|g_k\| = 0$.

Proof. We use contradiction, that is, there is a scalar $\varepsilon > 0$, such that (22).

$$\|g_k\| \geq \varepsilon, \tag{22}$$

From (4), we have (23).

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

Thus from (3) and (10), we obtain,

$$\|d_k\|^2 \leq \|g_k\|^2 + 2|\beta_k||g_k^T d_k| + \beta_k^2 \|d_{k-1}\|^2, \tag{23}$$
$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{\sigma}{1+2\sigma} \|g_{k-1}\|^2 |\beta_k| + \beta_k^2 \|d_{k-1}\|^2,$$
$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{\sigma}{1+2\sigma} \|g_{k-1}\|^2 \frac{\|g_k\|^2}{\|d_{k-1}\|^2} + \beta_k^2 \|d_{k-1}\|^2,$$
$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{\sigma}{1+2\sigma} \|g_k\|^2 \left(\frac{\|g_{k-1}\|^2}{\|d_{k-1}\|^2}\right)$$

by applying (16), we get,

$$\|d_k\|^2 \le \|g_k\|^2 + \frac{\sigma}{1+2\sigma}\|g_k\|^2\left(\frac{1}{1+2\sigma}\right)^2$$
$$\|d_k\|^2 \le \|g_k\|^2 + \|g_k\|^2 M \quad , \quad M = \frac{\sigma}{(1+2\sigma)^2}$$
$$\|d_k\|^2 \le (1+M)\|g_k\|^2 \tag{24}$$

dividing both sides by $\|g_k\|^4$ to obtain (25).

$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \frac{(1+M)}{\|g_k\|^2}$$
$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \sum_{i=0}^{k} \frac{(1+M)}{\|g_i\|^2} \tag{25}$$

Therefor it follows from (22) and (25),

$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \frac{(1+M)k}{\varepsilon^2}$$
$$\frac{\|g_k\|^4}{\|d_k\|^2} \ge \frac{\varepsilon^2}{(1+cM^2)k} \tag{26}$$
$$\frac{\|g_k\|^4}{\|d_k\|^2} \ge \frac{c^2}{k}$$

this implies that,

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = \infty$$

this contradicts the condition of Zoutendijk [22]. Hence, the proof is come true. We now state that Algorithm 2.1 satisfies the property (*). The Property (*) states that: given a method of form (2), (3), and,

$$0 < \gamma \le \|g_k\| \le \bar{\gamma}, \tag{27}$$

where γ and $\bar{\gamma}$ are positive constant, a method is said to have property. (*), if for all $k \ge 1$, there a constant $b > 1$, $\lambda > 0$, such that $|\beta_k| \le b$ and $\|s_k\| \le \lambda$ implies $|\beta_k| \le \frac{1}{2b}$, where $s_k = x_{k+1} - x_k$.

Lemma 2.1.3
Let Assumption 2.1.1 be satisfied, then property (*) holds when Algorithm 2.1 applied.
Proof. Suppose that (27) holds, set $b = \frac{\overline{\gamma^2}}{\gamma^2} > 1, \lambda = \frac{\gamma^2}{4L\bar{\gamma}b} > 0$.

From the definition of β_k^{AO} that,

$$|\beta_k^{AO}| = \left|\frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)}\right| \le \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \le \frac{\overline{\gamma^2}}{\gamma^2} = b, \tag{28}$$

by assumption 2.2.1, and properties of norm, we can get that if $s_{k-1} \le \lambda$, then,

$$|\beta_k^{AO}| = \left|\frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|}{max(\|d_{k-1}\|^2, \|g_{k-1}\|^2)}\right| \le \frac{\left|\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T g_{k-1}|\right|}{\|g_{k-1}\|^2} \le \frac{\left|\|g_k\|(\|g_k\| - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|}\right|}{\|g_{k-1}\|^2},$$
$$\le \frac{\|g_k\|\left(\left\|g_k - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|}\right\|\right)}{\|g_{k-1}\|^2} \le \frac{\|g_k\|\left(\|g_k - g_{k-1}\| + \left\|g_{k-1} - \frac{|g_k^T g_{k-1}|}{\|g_{k-1}\|}\right\|\right)}{\|g_{k-1}\|^2} \le \frac{\|g_k\|\,(\|g_k - g_{k-1}\| + \|g_{k-1} - g_k\|)}{\|g_{k-1}\|^2},$$
$$\le \frac{\|g_k\|\,(\|g_k - g_{k-1}\| + \|g_k - g_{k-1}\|)}{\|g_{k-1}\|^2} \le \frac{2L\lambda\bar{\gamma}}{\gamma^2} = \frac{1}{2b}.$$

3. NUMERICAL DISCUSSIONS

In this section, we ran some experiments to put the Algorithm 2.1 to the test; we referred to test problem addressed in Andrei [23]. In order to compare the performance of the proposed formula with those of the CG methods listed in (6) and (7). The comparisons were based on the amount of time spent on the CPU and the number of iterations. Considered $\varepsilon = 10^{-6}$ and $\|g_k\| \le \varepsilon$ as a stopping criterion as presented in Hillstrom [24]. All test problems in Table 1 are executed using MATLAB.

Table 1. Problem functions list

No.	Function	N	Initial point
1.	Six hump	2	(4,4),(10,10)
2.	Booth	2	(25,25),(100,100)
3.	Treccani	2	(5,5),(10,10)
4.	Three hump	2	(-5,5),(-3,12)
5.	Zettl	2	(15,15),(10,10)
6.	Leon	2	(20,20),(50,50)
7.	Matyas	2	(15,15),(6,6)
8.	Wood	4	(6,6,6,6),(13,13,13,13)
9.	Colville	4	(10,10,10,10),(30,30,30,30)
10.	Powell	4	(2,2,2,2),(20,20,20,20)
11.	Power	4	(2,2,2,2),(8,8,8,8)
		10,100	(2,2,…,2),(8,8,…,8)
12.	Extended Peanlty	2, 4	(5,5),(11,11),(5,5,5,5),(11,11,11,11)
		100,500	(5,5,…5),(11,11,…,11)
13.	Generalized Tridiagonal 1	2,4	(5,5),(20,20),(5,5,5,5),(20,20,20,20)
		10,100,500	(5,5,…,5),(20,20,…,20)
14.	Raydan 1	2,4	(3,3),(10,10),(3,3,3,3),(10,10,10,10)
		10,100	(3,3,…,3),(10,10,…,10)
15.	Dixon and Price	4	(80,80,80,80),(150,150,150,150)
		10,100	(80,80,…,80),(150,150,…,150)
16.	Hager	4	(2,2,2,2),(15,15,15,15)
		10,100	(2,2,…,2),(15,15,…,15)
17.	Flethcr	4	(40,40,40,40),(60,60,60,60)
		10,100,500	(40,40,…,40),(60,60,…,60)
18.	Nonscomp	2	(8,8),(-1,-1)
		10,100	(8,8,…,8),(-1,-1,…,-1)
19.	Extended Freudenstein and Roth	4	(-3,-3,-3,-3),(7,7,7,7)
		10,100,500	(-3,-3,…,-3),(7,7,…,7)
20.	Generalized Tridiagonal 2	4	(1,1,1,1),(4,4,4,4)
		10,100,500	(1,1,…,1),(4,4,…,4)
21.	Extended Quadratic Penalty QP2	4	(6,6,6,6),(14,14,14,14)
		10,100,500,1000	(6,6,…,6),(14,14,…,14)
22.	Extended Beale	4	(-1,-1,-1,-1),(2,2,2,2)
		10,100,500,1000,10000	(-1,-1,…,-1),(2,2,…,2)
23.	Diagonal 4	500,1000,10000	(7,7,…,7),(18,18,…,18)
24.	Extended Maratos	10,100,500,1000	(0.5,0.5,…,0.5),(1.5,1.5,…,1.5)
25.	Shallow	100,500,1000,10000	(-5,-5,…,-5),(-20,-20,…,-20)
26.	Extended Rosen Brock	100,500,1000,10000	(5,5,…,5),(10,10,…,10)
27.	Extended White and Holst	10,100,500,1000,10000	(-5,-5,…,-5),(-7,-7,…,-7)
28.	Quadratic QF2	10,100,500,1000	(2,2,…,2),(60,60,…,60)
29.	Extended Denschnb	10,100,500,1000,10000	(4,4,…,4),(16,16,…,16)
30.	Extended Himmelblau	10,100,500,1000,10000	(30,30,…,30),(200,200,…,200)

Figures 1 and 2 display performance results that were established using the performance profile proposed by Dolan and More [25]. Based on their performance profile, we take $t_{p,s}$ to be the outcome when the solver s is used to solve problem, and $r_{p,s}$ to be the ratio $\frac{t_{p,s}}{\min\{t_{p,s}:s\in S\}}$, where S is the set of all solvers. Then we can order the values $r_{p,s}$ increasingly and plot them versus $p_s(t)$, where $p_s(t)$ is the ratio $\frac{The\ order\ the\ problem}{Total\ number\ of\ problems}$. Clearly the method of top curve is the winner. Overall, a solver with a high p(t) value or the curve that seems on top of the Figures is the most effective problem solver. As can be seen in Figures 1 and 2, our new proposed hast the choicest results whenever it could solve every test problem as in Table 2.

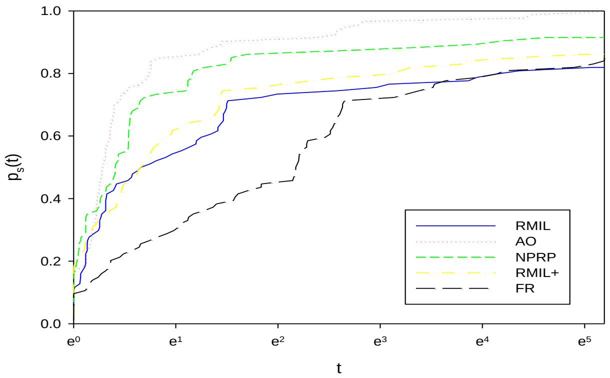

Figure 1. Performance results based on the number of iterations

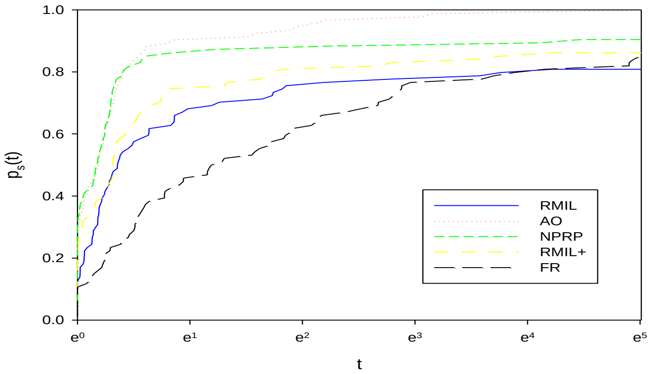

Figure 2. Performance results based on the CPU time

Table 2. Relative efficiency of the methods

AO	RMIL	RMIL+	NPRP	FR
1	0.7812	0.8143	0.8516	0.8094

4. CONCLUSION

In this paper, we presented a parameter for β_k that has better convergence. Numerical outcomes have reflected that proposed formula β_k highlighted better than FR, RMIL, RMIL+ and NPRP. In the future, the new formula can be applied under another line search.

ACKNOWLEDGEMENTS

The Authors are grateful to anonymous referees and the editor for their comments and suggestions that improved our paper.

REFERENCES

[1] M. R. Hestenes and E. L. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Natl. Bur. Stand. Sec.* vol. 49, pp. 409–432, 1952, doi: 10.6028/jres.049.044.
[2] F R. Fletcher and C. Reeves, "Function minimization by conjugate gradient," *J. Comput*, vol. 7, pp. 149-154, 1964, doi: 10.1093/comjnl/7.2.149.
[3] B. T. Polyak, "The conjugate gradient method in extreme problems," *USSR Comp. Math. Phys*, vol. 9, pp. 94-112, 1969, doi: 10.1016/0041-5553 (69)90035-4.
[4] E. Polakand and G. Ribie`re, "Note Sur la convergence de directions conjuge`es," *Rev. Francaise Informat Recherche Operationelle, 3e Anne`e* vol. 16, pp. 35-43, 1969, doi: 10.1051/m2an/196903R100351.
[5] R. Fletcher, "Practical method of optimization, second ed.," *Unconstrained Optimization*, vol. I, 1987, doi: 10.1002/9781118723203.
[6] Y. Liu, and C. Story, "Efficient generalized conjugate gradient algorithms, part 1: Theory," *J. Optim. Theor. Appl.* vol. 69, pp. 129-137, 1992, doi: 10.1007/BF00940464.
[7] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient with a strong global convergence properties," *SIAM J. Optim.*, vol. 10, pp. 177-182, 1999, doi: 10.1137/S1052623497318992.
[8] Z. Wei, S. Yao, and L. Liu, "The convergence properties of some new conjugate gradient methods," *Appl. Math. Comput*, vol. 183, pp. 1341–1350, 2006, doi: 10.1016/j.amc.2006.05.150.
[9] Z. F. Dai, "Two modified HS type conjugate gradient methods for unconstrained optimization problems," *Nonlinear Anal.*, vol. 74, pp. 927–936, 2011, doi: 10.1016/j.na.2010.09.046.
[10] W. W. Hager and H. C. Zhang, "A new conjugate gradient method with guaranteed descent and efficient line search," *SIAM J. Optim.*, vol. 16, pp.170–192, 2005, doi: 10.1137/030601880.
[11] H. Huang, Z. Wei, and S. Yao, "The proof of the sufficient descent condition of the Wei–Yao–Liu conjugate gradient method under the strong Wolfe–Powell line search," *Appl. Math. Comput.*, vol. 189, pp. 1241–1245, 2007, doi: 10.1016/j.amc.2006.12.006.
[12] M. Al-Baali, "Descent property and global convergence of fletcher-reeves method with inexact line search," *IMA J. Numer. Anal.*, vol. 5, pp. 121–124, 1985, doi: 10.1093/imanum/5.1.121.
[13] G. Liu, J. Han, and H. Yin, "Global convergence of the fletcher–reeves algorithm with inexact line search," *Appl. Math. J. Chin. Univ. Ser.*, vol. 10, pp. 5–82, 1995, doi: 10.1007/BF02663897.
[14] Y. H. Dai, Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM J. Optim.*, vol. 10, pp. 177–182, 2000, doi: 10.1137/S1052623497318992.
[15] N. J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM J. Optim.2*, pp. 21–42, 1992, doi: 10.1137/0802003.
[16] A. Abdelrahman, O. O. O. Yousif, M. Mhammed, and M. K. Elbashir. "Global convergence of nonlinear conjugate gradient coefficients with inexact line search," *The Scientific Journal of King Faisal University: Basic and Applied Sciences*, vol. 22, no. 2, pp. 86–91, 2021, doi: 10.37575/b/sci/210058.
[17] O. O. O. Yousif, "The convergence properties of RMIL+ conjugate gradient method under strong Wolfe line search," *Applied Mathematics and Computation*, vol. 367, p. 124777, 2020, doi: 10.1016/j.amc.2019.124777.

[18] O. O. O. Yousif, A. Abdelrahman, M. Mohammed, and M. A. Saleh. "A sufficient condition for the global convergence of conjugate gradient methods for solving unconstrained optimisation problems," *The Scientific Journal of King Faisal University: Basic and Applied Sciences*, vol. 23, no. 1, pp. 106–12, 2022, doi: 10.37575/b/sci/220013.

[19] M. Rivaie, M. Mamat, W. J. Leong, and I. Mod, "A new class of nonlinear conjugate gradient coefficient with global convergence properties," *Appl. Math. Comput.*, vol. 218, pp. 11323-11332, 2012, doi: 10.1016/j.amc.2012.05.030.

[20] M. Rivaie, M. Mamat, and A. Abashar, "A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches," *Appl. Math. Comput.*, vol. 268, pp. 1152-1163, 2015, doi: 10.1016/j.amc.2015.07.019.

[21] L. Zhang, "An improved Wei–Yao–Liu nonlinear conjugate gradient method for optimization computation," *Appl. Math. Comput.* vol. 215, pp. 2263–2274, 2009, doi: 10.1016/j.amc.2009.08.016.

[22] G. Zoutendijk, "*Nonlinear programming computational methods, in: J.Abadie (Ed.),*" Integer and Nonlinear Programming, North-Holland, Amsterdam, pp. 37–86, 1970.

[23] N. Andrei, "An unconstrained optimization test functions collection," *Adv. Model. Optim.*, vol. 10 pp. 147–161, 2008.

[24] K. E. Hilstrom, "A simulation test approach to the evaluation of nonlinear optimization algorithms," *ACM. Trans. Math. Softw.*, vol. 3, pp. 305–315, 1977, doi: 10.1145/355759.355760.

[25] E. Dolan and J. J. More, "Benchmarking optimization software with performance profile," *Math. Prog.*, vol. 91, pp. 201–213, 2002, doi: 10.1007/s101070100263.

BIOGRAPHIES OF AUTHORS

Abdelrhaman Abashar is Associate Professor at Faculty of Engineering, Red Sea University (RSU), Port sudan, Sudan. He Holds a PhD degree in Mathematical Sciences, School of Informatics and Applied Mathematics, University Malaysia Terengganu, Malaysia (UMT). His research areas are unconstrained optimizations 'Conjugate Gradient Methods'. He is Head of Department Basic Sciences in Dec. 2015. He is Dean of Faculty of Post Graduate Studies, Red Sea Univiersity, Sudan from October 2019- April 2022. He is Technical Program Committees, Co- Chair, First International Conference on Engineering and Applied Sciences (ICEAS2017, Marach 7-9), Sudan. His research interests include numerial analaysis, computation. He can be contacted at emails: abashar3373@gmail.com or abashar3373@rsu.edu.sd.

Osman Omer Osman Yousif is Assistant Professor at Faculty of Mathematical and Computer Sciences, University of Gezira, Sudan. He received the Ph.D. from Universiti Malaysia Terengganu (UMT) in 2015 with specialization in Optimization. He is the Head of Mathematics Department since 2019 up to now. He was also the postgraduate coordinator. He was a supervisor, a co-supervisor, and an examiner of more than 30 master students. His research interests include Optimization, Numerical Computation, and Operation Research. He has 19 publications in conference proceedings, ISI journals, and Scopus journals with 3 H-index and 51 citations. He can be contacted at emails: osman.omer@uofg.edu.sd or osman_om@hotmail.com.

Awad Abdelrahman Abdalla Mohammed is an assistant Professor in University of Gezira. He has received his Ph.D. degree in mathematical sciences from University of Malaysia Terengganu, Malaysia in 2017. Also, he is now a head of the documentation committee and a member of courses development committee. He interested in optimization and computational mathematics. He supervised to undergraduate and M.Sc. students. He has participated in conferences in Malaysia, Sudan and served as a co-chair of ICCCEEE18 conference which held in Khartoum in 2018. He can be contacted at emails: awad.abdalla26@yahoo.com or awadabdalla@uofg.edu.sd.

Mohammed A. Saleh was born in Riyadh, Kingdom of Saudi Arabia (KSA). He received the B.Sc. (Honor) in Mathematical and Computer Science, Faculty of Mathematical and Computer Science, University of Gezira, Sudan. He obtained the M.Sc. (First Class) in Information Security, Faculty of Computer Science and Information System, University of Technology (UTM), Malaysia, and the Ph.D. in Information Security (Computer Science), Faculty of Computing, University of Technology (UTM), Malaysia. From 2010 to 2014, he was a Security Engineer with the Sudanese Nation Information Center (NIC). As well, he was Lecturer at Gezira University. Currently, he is an Assistant Professor at Qassim University, KSA. His research interests include Malware Analysis and Artificial Intelligence in Cybersecurity. He is the author of a couple of journal articles. He can be contacted at email: m.saleh@qu.edu.sa.