JOURNAL OF COMPUTATIONAL INNOVATION AND ANALYTICS
https://e-journal.uum.edu.my/index.php/jcia

How to cite this article:
Musa, A. M., Mandara, A. V., Markus, S., Lawal, A. A., & Yakubu, U. A. (2025). A modification of the conjugate gradient parameter and its global convergence for solving unconstrained optimization problems. *Journal of Computational Innovation and Analytics*, *4*(2), 38-54. https://doi.org/10.32890/jcia2025.4.2.3

A MODIFICATION OF CONJUGATE GRADIENT PARAMETER AND ITS GLOBAL CONVERGENCE FOR SOLVING UNCONSTRAINED OPTIMIZATION PROBLEMS

[1]Ahmadu M. Musa, [2]Abba V. Mandara, [3]Samaila Markus, [4]Ahmed A. Lawal & [5]Usman A. Yakubu

[1,2,3,4]Faculty of Physical Science, Department of Mathematics,
University of Maiduguri, Borno State, Nigeria.
[5]Faculty of Science, Department of Mathematical Sciences,
Maitama Sule University Kano, Kano State, Nigeria.

[1]*Corresponding author: ahmadumohammedmusa2024@gmail.com*

Received: 25/09/2024 Revised: 29/01/2025 Accepted: 10/03/2025 Published: 31/07/2025

ABSTRACT

The methods of nonlinear conjugate gradient coefficients are significant and helpful in solving large-scale unconstrained optimization problems, due to their simplicity and lower storage requirement. Research activities on its applications to handle higher-dimensional systems of nonlinear equations are of paramount importance. Many authors studied and developed different kinds of conjugate gradient coefficients. Recently, some conjugate gradients were proposed with large dimensions, but they have small sample sizes. So, they cannot solve problems that have higher dimensions. Therefore, this research proposed a modified algorithm with a huge dimension and a large sample size. The strategy of strong Wolfe line search was applied in the convergence analysis, which makes it possible to converge globally with a descent property. Finally, the numerical results show that the proposed algorithm performs more efficiently and is superior to the existing CG coefficients.

Keywords: Conjugate Gradient, global convergence, Modified Conjugate Gradient Coefficient, unconstrained optimization, Wolfe line search.

INTRODUCTION

Conjugate gradient (CG) methods play a significant and useful role in solving unconstrained optimization. The CG methods are substantial and useful for finding the minimum value of a function for unconstrained optimization. In general, the method has the form:

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

Where $f: R^n \to R$ is continuously differentiable and R^n denotes n-dimensional Euclidean space, x is the variable or unknown, and f is the objective function. The CG methods are given by an iterative of the form:

$$x_{k+1} = x_k + \alpha_k d_k. \qquad k= 0,1, 2, 3, \ldots \tag{2}$$

Where x_k is the current iteration point, $\alpha_k > 0$ is a step size, and d_k is the search direction. The step size is a one-dimensional minimization problem known as the “line search” in the scalar form that will search along the d_k for the best x_{k+1}. The most common line search is the exact line search, which is given by Equation (3).

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k). \tag{3}$$

The CG search direction, d_k is defined by

$$d_k = \begin{cases} -g_k \; for \; k = 0, \\ -g_k + \beta_k d_{k-1} \; for \; k \geq 1, \end{cases} \tag{4}$$

Where g_k denotes the gradient of $f(x)$ at the point x_k. $\beta_k \epsilon R$ is a scalar known as the CG coefficient. Some well-known formulas for β_k include:

$$\beta_k^{FR} = \frac{g_k^T g_k}{\|g_{K-1}\|^2} \tag{5}$$

$$\beta_k^{PR} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \tag{6}$$

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \tag{7}$$

$$\beta_k^{LS} = \frac{g_k^T (g_k - g_{k-1})}{- d_{k-1} g_{k-1}} \tag{8}$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}} \tag{9}$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{(d_{k-1}^T g_{k-1})} \tag{10}$$

$$\beta_k^{NMRMIL+} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T d_{k-1}|}{\|d_{k-1}\|^2} \tag{11}$$

The corresponding methods are known as Fletcher and Reeves (1964), Polak and Ribiere (1969), Hestenes and Stiefel (1952), Liu and Storey (1992), Dai and Yuan (2000), and finally, Conjugate Descent (Fletcher, 1987), respectively. The norm of vectors is represented as $\|\ .\ \|$. For (x), which is strictly a convex quadratic function, all these methods exhibit finite convergence properties under the exact line search. However, they behave quite differently for general non-quadratic functions (Dai & Yuan, 2000; Yuan & Sun, 1999; Yuan & Wei, 2010; Yakubu et al., 2023; Yakubu, Morch, & Saputra, 2022).

The most extensively studied properties of CG methods are their global convergence properties. The most notable work in this regard is by Zoutendijk (1970), who proved the convergence of the Fletcher–Reeves (FR) method under the exact line search. However, Powell (1984) identified a significant drawback in the FR method. Powell (1986) also demonstrated that the Polak–Ribiere (PR) method can cycle infinitely without reaching the minimizer, thereby lacking global convergence. Gilbert and Nocedal (1992), further analyzed the global convergence of algorithms related to the FR method using strong Wolfe conditions. Notably, Powell (1986) reaffirmed that FR is superior to other methods. Since Powell's proof, the global convergence of PR, LS, and HS methods has not yet been fully established, primarily because these methods cannot generate descent objective function values at each iteration (Hager & Zhang, 2005).

Andrei (2008) classified CG methods into three categories: classical CG methods, scaled CG methods, and hybrid/parameterized CG methods. The pioneering CG methods fall under the classical category. While producing CG methods in this category is challenging, their applications are relatively straightforward. Consequently, researchers have focused on developing scaled and hybrid CG methods. Significant contributions to these types of CG methods include the works of Fletcher (1987), Powell (1986), Sun and Zhang (2001), Yuan and Wei (2009), and Shi and Guo (2009).

A key factor in achieving global convergence is the selection of αk. The most common line search technique is the exact line search, which involves determining the precise value of α_k (Rivaie et al., 2012). In recent years, research efforts have increasingly focused on modifying and developing new CG formulas that exhibit strong numerical performance and global convergence properties.

THE NEW CG COEFFICIENT AND ITS ALGORITHM

In this section, we introduce a modified conjugate gradient (CG) algorithm, denoted as $\beta_k^{NMRMIL+}$, which is a modification of the algorithms proposed by Olowo and Sulaiman (2021) and Rivaie, Mamat, and Abdelrhaman (2015). The new approach introduces a novel formula for the numerator while retaining the denominator structure as in Rivaie et al. (2012) and Rivaie et al. (2015).

$$\beta_k^{NMRMIL+} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T d_{k-1}|}{\|d_{k-1}\|^2} \qquad (12)$$

The algorithm of the above formula is given below.

Step 1: Initialization, given an arbitrary value $x_{0,}$ set k = 0

Step 2: Compute $\beta_k^{NMRMIL+}$ based on (11)

Step 3: Compute $d_k = -g_k + \beta_k d_{k-1.}$ If $\|g_k\|$ =0, then stop.

Step 4: Solve $\alpha_k = \min_{\alpha>0} f(x_k + \alpha d_k)$.
Step 5: Update the new iterative point using $x_{k+1} = x_k + \alpha_k d_k$
Step 6: Convergent test and stopping criteria If $f(x_{k+1} < f(x_k)$ and $\|g_k\| \le \varepsilon$. Then stop, otherwise go to Step 1 with $k =: k+1$

METHODOLOGY

In this section, the research shows that the CG given by $\beta_k^{NMRMIL+}$ converges globally, by strong Wolfe line-search, i.e.,

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

$$|g_{k+1}^T d_k| \le \sigma |g_k^T d_k| \tag{13}$$

$$0 < \delta < \sigma < 1$$

Obtained using the iterative of the form

$$x_{k+1} = x_k + \alpha_k d_k. \qquad \text{k= 0,1, 2, 3, …} \tag{14}$$

Where the CG search direction, d_k is defined by

$$d_k = \begin{cases} -g_k \; for \; k = 0, \\ -g_k + \beta_k^{NMRMIL+} d_{k-1} \; for \; k \ge 1, \end{cases} \tag{15}$$

And

$$\beta_k^{NMRMIL+} = \begin{cases} \dfrac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k^T d_{k-1}|}{\|d_{k-1}\|^2} & if \; 0 \le |g_k^T g_{k-1}| \le \|g_k\|^2 \\ 0 & otherwise \end{cases} \tag{*}$$

Global Convergence

The proof of the global convergence of our newly proposed conjugate gradient (CG) algorithm, with the above choice of parameter (*), relies on the strong Wolfe line search strategy (Fletcher, 1987; Jorge & Stephen, 1999). The following inequalities are essential for the convergence analysis.

$$0 \le \beta_k^{NMRMIL+} \le \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \tag{16}$$

$$\forall\, k \ge 1$$

and

$$\frac{\|g_k\|}{[\![d_k]\!]} < 1 \tag{17}$$

$$\forall\, k \ge 0$$

These inequalities are shown as follows:

The first inequality can be established from $\beta_k^{NMRMIL+}$. We should note that for a positive constant. σ, $0 < \sigma < \frac{1}{2} \Rightarrow 2\sigma < 1$

$$\Rightarrow 2\sigma - 2 < -1$$

$$\Rightarrow 2(\sigma - 1) < -1$$

$$\Rightarrow \sigma - 1 < -\frac{1}{2} < 0$$

$$\Rightarrow \sigma - 1 < 0$$

Similarly,

$$0 < \delta < \frac{1}{2} \quad \Rightarrow \delta - 1 < -\frac{1}{2}$$

$$\Rightarrow 1 - \delta > \frac{1}{2}$$

$$\Rightarrow \frac{1}{1-\delta} < 2$$

Theorem 1

Suppose that the proposed CG method generates sequences $\{d_k\}$ and $\{g_k\}$ and let a strong wolf line-search incorporated in the $0 < \sigma < 1$, then the inequality in Equation (17) holds good.

Proof:

The proof is by induction, we can see that from Equation (13), it is true when $k = 0$.

Also let $\frac{\|g_k\|}{\|d_k\|} < 1$ holds true whenever $k \geq 0$. Thus,

$$g_{k+1} = -d_{k+1} + \beta_k^{NMRMIL+} \; d_k \tag{18}$$

Multiplying Equation (18) on both sides with g_{k+1}^T

$$\|g_{k+1}\|^2 = -g_{k+1}^T d_{k+1} + \beta_k^{NMRMIL+} \; g_{k+1}^T d_k$$

Now, using the triangle inequality, we will have

$$\|g_{k+1}\|^2 \leq \left| g_{k+1}^T d_{k+1} \right| + \left| \beta_k^{NMRMIL+} \right| \left| g_{k+1}^T d_k \right|$$

$$\leq \left| g_{k+1}^T d_{k+1} \right| + \left| \beta_k^{NMRMIL+} \right| \left| g_k^T d_k \right| \text{ (Using strong Wolfe line-search)}$$

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

$$= \|g_{k+1}\| \|d_{k+1}\| + \sigma \|g_{k+1}\|^2$$

i.e

$$\|g_{k+1}\| \leq \|d_{k+1}\| + \sigma\|g_{k+1}\|.\frac{\|g_k\|}{\|d_k\|}$$

Now, applying the inequality in Equation (17) gives

$$\|g_{k+1}\| < \|d_{k+1}\| + \sigma\|g_{k+1}\|$$

Therefore, we have

$$\|g_{k+1}\|(1-\sigma) < \|d_{k+1}\|$$

But, since $1-\sigma > 0$ and from Equation (17), it implies

$$\frac{\|g_{k+1}\|}{\|d_{k+1}\|} < \frac{1}{1-\sigma} < 1$$

Thus, Equation (17) holds true for $k+1$

Next, we can easily rectify that from the inequality in Equation (17), and squaring it on both sides gives:

$$\frac{1}{\|d_k\|^2} < \frac{1}{\|g_k\|^2}, \forall\, k \geq 0$$

However, we will show the sufficient decency property, which is crucial in obtaining the global convergence.

Theorem 2

Suppose that the proposed CG method generates sequences $\{d_k\}$ and $\{g_k\}$ using a strong Wolfe line search where $\sigma \in (0,\frac{1}{2})$ then

$$-1-\sigma < \frac{g_k^T d_k}{\|g_k\|^2} < -1+\sigma \tag{19}$$

$$\forall\, k \geq 0$$

Thus, the sufficient descent property holds.

Proof:

We can easily see that the result holds true for any $k = 0$.

Now for $k > 0$, we have as follows:

$$g_k^T d_k = -\|g_k\|^2 + \beta_k^{NMRMIL+} g_k^T d_{k-1} \tag{20}$$

Thus, using the strong Wolfe condition and the inequality (16), we have

$$-\sigma\, \beta_k^{NMRMIL+} \left|g_{k-1}^T d_{k-1}\right| \leq \beta_k^{NMRMIL+} d_{k-1} \leq \sigma\, \beta_k^{NMRMIL+} \left|g_{k-1}^T d_{k-1}\right| \tag{21}$$

Now, subtracting $\|g_k\|^2$ from both sides of Equation (21) and then substituting Equation (20) gives

$$-\|g_k\|^2 - \sigma\, \beta_k^{NMRMIL+} \left|g_{k-1}^T d_{k-1}\right| \leq -\|g_k\|^2 + \beta_k^{NMRMIL+} g_k^T d_{k-1}$$

$$\leq -\|g_k\|^2 + \sigma\, \beta_k^{NMRMIL+} \left|g_{k-1}^T d_{k-1}\right|$$

$$-\|g_k\|^2 - \sigma \beta_k^{**} \left|g_{k-1}^T d_{k-1}\right| \leq g_{kd_k}^T \leq -\|g_k\|^2 + \sigma \beta_k^{**} \left|g_{k-1}^T d_{k-1}\right|$$

(For simplicity, we represent $\beta_k^{NMRMIL+}$ with β_k^{**})

Now, by applying Cauchy Schwartz inequality, we get:

$$-\|g_k\|^2 - \sigma \beta_k^{**} \|g_{k-1}\| \|d_{k-1}\| \leq g_{kd_k}^T \leq -\|g_k\|^2 + \sigma \beta_k^{**} \|g_{k-1}\| \|d_{k-1}\| \tag{22}$$

We can now put inequality (16) into Equation (22), or we can establish a relationship as follows:

$$(-1-\sigma)\frac{\|g_{k-1}\|}{\|d_{k-1}\|} \leq \frac{g_{kd_k}^T}{\|g_k\|^2} \leq (-1+\sigma)\frac{\|g_{k-1}\|}{\|d_{k-1}\|} \tag{23}$$

Thus, the result follows by using the result of Theorem 1 into Equation (23), which yields.

$$(-1-\sigma) \leq \frac{g_{kd_k}^T}{\|g_k\|^2} \leq (-1+\sigma) \blacksquare$$

Next, to prove the global convergence, we have to assume the following basic conditions:

Assumption 1: The level set $\mathbb{N} = \{ x \in \mathbb{R}; f(x) \leq f(x_o)\}$ is bound with the x_o being the starting point.

Assumption 2: The objective function f is continuously differentiable and its gradient is Lipschitz continuously in the neighbourhood, say $\mathbb{N} of \chi$ i.e there exists $\mu > 0$ such that

$$\|g(x) - g(y)\| \leq \mu \|x - y\|, \forall\, x, y \in \mathbb{N}$$

More so, we need the following lemma and the subsequent theorem.

Lemma 1

Let the stated assumptions in Theorem 1 and 2 hold with the formulations defined in Equations (14) and (15), whose step-size is obtained through strong Wolfe line search, the following condition strictly follows:

$$\sum_{k=0}^{\infty} \|g_k\|^2 cos^2 x\theta_k < \infty$$

The above condition is fulfilled with

$$\theta_k = \frac{g_{kd_k}^T}{\|g_k\|\|d_k\|}$$

Theorem 3

Let the assumptions hold and suppose the sequences $\{d_k\}$ and $\{g_k\}$ are generated by the proposed algorithm with the step size obtained via strong Wolfe line-search with $0 < \sigma < 1$ then;

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

Proof:

We begin this proof first by multiplying Equation (15) by $-\frac{\|g_k\|}{\|d_k\|}$ and using Lemma (14) to have:

$$\delta_1 \frac{\|g_k\|}{\|d_k\|} < cos\theta_k < \delta_2 \frac{\|g_k\|}{\|d_k\|}, \text{ in which } \delta_1 = 1 - \sigma \text{ and } \delta_2 = 1 + \sigma$$

Since $\delta_1 > 0$, $0 < \sigma < 1$, then it follows that:

$$cos\theta_k > 0$$

Hence, we will have as follows after squaring.

$$\delta_1^2 \frac{\|g_k\|^2}{\|d_k\|^2} < cos^2\theta_k \tag{24}$$

Thus, multiplying $\|g_k\|^2$ on both sides of Equation (24) and summing over gives:

$$\delta_1^2 \sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \sum_{k=0}^{\infty} \|g_k\|^2 \, cos^2\theta_k < \infty$$

Therefore,

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

Theorem 4

The proposed algorithm is globally convergent in the sense that,

$$\lim_{k\to\infty} \|g_k\| = 0 \tag{25}$$

Provided that assumption in Theorem 1 holds and the step-size is obtained via a strong Wolfe line-search with

$$0 < \sigma < 1.$$

Proof:

The proof follows through contradiction. Suppose Equation (25) fails; in that case, there exist Q and K ∈ ℤ such that

$$\|g_k\|^2 \geq Q^2, \text{ for all } k \gg K,$$

$$\Rightarrow \frac{1}{\|g_k\|^2} \leq \frac{1}{Q^2},$$

Now, adopting Equation (15) and then rearranging into:

$$d_k + g_k = \beta_k^{**} d_{k-1} \tag{26}$$

Consequently, squaring Equation (26) on both sides yields

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

$$\Rightarrow \|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + (\beta_k^{**})^2 \|d_{k-1}\|^2 \tag{27}$$

We can see that Equation (26) can be expressed as:

$$\|g_k\|^2(-1-\sigma) < g_k^T d_k < (-1+\sigma)\|g_k\|^2 \tag{28}$$

This can be written as:

$$2(1-\sigma)\|g_k\|^2 < -2g_k^T d_k < 2(1+\sigma)\|g_k\|^2 \tag{29}$$

Now, since $-2g_k^T d_k < 2(1+\sigma)\|g_k\|^2$ holds for all $k \geq 0$.

Thus, Equation (27) can also be expressed as:

$$\|d_k\|^2 < -\|g_k\|^2 + 2(1+\sigma)\|g_k\|^2 + (\beta_k^{**})^2 \|d_{k-1}\|^2$$

$$= (1+\sigma)\|g_k\|^2 + (\beta_k^{**})^2 \|d_{k-1}\|^2 < (1+\sigma)\|g_k\|^2 + \frac{\|g_k\|^4}{\|d_{k-1}\|^4} \|d_{k-1}\|^2 \tag{30}$$

We can divide Equation (30) by $\|g_k\|^4$ to have:

$$\frac{\|d_k\|^2}{\|g_k\|^4} < (1+\sigma)\frac{1}{\|g_k\|^2} + \frac{1}{\|d_{k-1}\|^2} < (1+\sigma)\frac{1}{\|g_k\|^2} + \frac{1}{\|g_{k-1}\|^2}$$

Hence, we have

$$\frac{\|d_k\|^2}{\|g_k\|^4} < \frac{(1+\sigma)}{Q^2} + \frac{1}{Q^2}, \forall\, k \geq k+1$$

Therefore, we have

$$\frac{\|g_k\|^4}{\|d_k\|^2} > \frac{Q^2}{2+\sigma},\ \forall\, k \geq k+1$$

Thus, we have $\forall k > k_1 + 1$, as follows:

$$\sum_{k=k_1+1}^{n} \frac{\|g_k\|^4}{\|d_k\|^2} > \sum_{k=k_1+1}^{n} \frac{Q^2}{2+\sigma}$$

$$= (n - k_1)\,\frac{Q^2}{2+\sigma}$$

Finally, if we take the limit as n→ ∞, we have

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} > \sum_{k=k_1+1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2}$$

$$= \lim_{n\to\infty} \sum_{k=k_1+1}^{n} \frac{\|g_k\|^4}{\|d_k\|^2}$$

$$> \lim_{n\to\infty} (n - k_1)\frac{Q^2}{2+\sigma} = \infty.$$

However, this is a contradiction. Therefore

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

NUMERICAL RESULTS AND DISCUSSION

In this section, we compared the performance of the modified method with other existing methods of conjugate gradient coefficient for solving unconstrained optimization problems. These methods include: "Performance Analysis of a Modified Conjugate Gradient Algorithm for Optimization Models" (Olowo & Sulaiman, 2021) known as SMO, "A New Class of Conjugate Gradient Coefficient with Global Convergence Properties" (Rivaie et al., 2012) known as RMIL1, "A New Class of Nonlinear Conjugate Gradient Coefficients with Exact and Inexact Line Searches" (Rivaie et al., 2015) known as RMIL+, "A New Conjugate Gradient Methods with Descent Properties and its Application to Regression Analysis"

(Sulaiman et al., 2020), known as RMIL2. In order to check the efficiency of our proposed method on some standard optimization test functions, numerical computations have been performed in MATLAB R2014a on a PC with an Intel Core i5 processor with 4064MB and a 1.80GHz CPU. We use 123 test functions with dimensions ranging from 2 to 500,000 to test the performance of the proposed method in terms of the number of iterations (NI) and the CPU time (in seconds). We consider $\varepsilon = 10^{-6}$ for termination criterion as suggested by Hilstromm (1977) Dolan and More (2002), Lawal (2016), Abba et al. (2017), Abba et al. (2018), and Yakubu et al. (2018).

Table 1 lists the numerical results, where the NI, FE, and CPU stand for the total number of iterations, number of function evaluations, and the CPU time in seconds, respectively. We also used “fail” to represent failure during the iteration process due to one of the following factors:

(1) Failure on code;
(2) If “FE” is not a real number;
(3) If the number of iterations and/ or CPU time in seconds reaches 1000.

However, our modified CG method always converges within the neighborhood of x^*.

Table 1
Numerical results for the proposed and existing methods with their given NI, FE, and CPU time.

			RMIL1			RMIL+			NMRMIL+			SMO			RMIL2		
S/n	**Problem**	**Dim**	**NI**	**FE**	**CPU**	**NI**	**FE**	**CPU**	**NI**	**FE**	**CPU**	**NI**	**FE**	**CPU**	**NI**	**FE**	**CPU**
1	Extended Penalty	100	15	411	0.000432	14	41	0.00099	6	33	0.012571	6	33	0.000721	14	162	0.000736
2	Extended Maratos	2	29	228	0.00026	Fail	Fail	Fail	23	162	0.001089	78	384	0.00089	Fail	Fail	Fail
3	Extended Maratos	30	29	228	0.000308	Fail	Fail	Fail	23	162	0.001805	84	408	0.002622	Fail	Fail	Fail
4	Diagonal 5	10	3	62	0.002174	4	13	0.001903	2	13	0.027347	2	13	0.003372	4	25	0.002911
5	Diagonal 5	50000	3	66	0.5484	4	13	0.000865	2	13	0.10013	2	13	0.11229	4	25	0.074463
6	Diagonal 5	100000	3	66	1.0004	4	13	0.000897	2	13	0.33619	2	13	0.15418	4	25	0.20299
7	Trecanni	2	10	42	0.000251	22	46	0.000259	5	23	0.000493	18	75	0.000364	14	56	0.001071
8	Trecanni	2	8	35	0.000236	30	64	0.000193	7	32	0.000405	15	64	0.00028	30	110	0.000317
9	Quadratic Penalty 1	4	11	46	0.002495	26	54	0.008255	7	31	0.021461	22	91	0.008836	24	87	0.00296
10	Quadratic Penalty 1	100	7	30	0.001861	26	54	0.004806	7	30	0.003935	23	94	0.00739	10	38	0.00196
11	Quadratic Penalty 1	1000	Fail	Fail	Fail	30	62	0.005296	7	30	0.046716	Fail	Fail	Fail	Fail	Fail	Fail
12	Quadratic Penalty 2	500	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	10	58	0.007267
13	Quadratic Penalty 2	100	5	86	0.003714	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	10	58	0.00208
14	Quadratic Penalty 2	10	5	86	0.003774	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	10	58	0.001949
15	Quadratic Function 1	100	23	469	0.017194	78	313	0.0077	19	448	0.038606	27	785	0.01644	54	828	0.011393
16	Quadratic Function 1	10	23	469	0.011684	78	313	0.007735	19	448	0.026441	27	785	0.016506	54	828	0.010401
17	Quadratic Function 1	10	35	803	0.020286	42	108	0.003722	32	907	0.051629	26	815	0.016527	48	699	0.009589
18	Quadratic Function 2	50	5	119	0.008559	6	13	0.001535	2	10	0.008875	2	10	0.001126	6	22	0.00124
19	Quadratic Function 2	1000	5	119	0.006843	6	13	0.000902	1	6	0.002812	1	6	0.001584	6	22	0.001801
20	Quadratic Function 2	5000	5	119	0.027872	6	13	0.000718	1	6	0.004272	1	6	0.001964	6	22	0.00336
21	Power	2	25	689	0.000413	26	53	0.000193	6	31	0.000447	6	31	0.00025	12	43	0.00021
22	Power	2	29	803	0.000258	30	61	0.000224	6	31	0.000456	6	31	0.000187	12	43	0.000236
23	Zettl	2	10	207	0.000238	20	59	0.000269	6	38	0.000566	9	50	0.000406	20	201	0.000224
24	Diagonal 2	1000	667	4259	0.84888	26	70	0.002919	428	2702	0.62376	Fail	Fail	Fail	Fail	Fail	Fail
25	Diagonal 2	10000	Fail	Fail	Fail	34	94	0.003164	1249	8119	14.556	Fail	Fail	Fail	Fail	Fail	Fail

26	Test	3	35	222	0.009144	24	83	0.002116	7	69	0.018656	Fail	Fail	Fail	Fail	Fail	Fail
27	Test	3	17	114	0.003106	30	101	0.002288	8	48	0.003326	Fail	Fail	Fail	Fail	Fail	Fail
28	Sum Of Squares	100	31	1052	0.000401	22	78	0.000224	9	163	0.001096	9	163	0.000195	42	992	0.000243
29	Shallow	1000	27	118	0.000307	390	796	0.000279	14	77	0.000368	147	605	0.000335	166	608	0.000276
30	Shallow	10000	14	119	0.001275	274	567	0.000422	14	80	0.000873	161	661	0.000731	176	645	0.000698
31	Quartic	100	4	97	0.022604	6	29	0.001271	3	29	0.012542	3	29	0.003127	6	62	0.003375
32	Quartic	1000	4	97	0.03498	6	29	0.001167	3	29	0.011961	3	29	0.007116	6	62	0.007228
33	Quartic	5000	5	141	0.18836	4	14	0.00088	3	35	0.052053	3	35	0.034948	8	99	0.049111
34	Quartic	10000	5	141	0.31176	4	14	0.000981	3	35	0.09488	3	35	0.17629	8	99	0.091847
35	Matyas	2	1	10	0.000702	2	10	0.000776	1	10	0.001408	1	10	0.000918	2	18	0.000567
36	Matyas	2	1	10	0.000277	2	10	0.000759	1	10	0.000414	1	10	0.000344	2	18	0.000782
37	Diagonal 1	10	50	227	0.015342	290	612	0.037525	60	276	0.028358	117	505	0.020309	134	500	0.012556
38	Diagonal 1	100	50	227	0.01013	290	612	0.037893	60	276	0.020808	117	505	0.020262	134	500	0.013848
39	Diagonal 1	10	43	193	0.009929	228	478	0.028622	55	248	0.021644	89	380	0.017147	110	416	0.012394
40	Diagonal 1	100	43	193	0.007475	228	478	0.026183	55	248	0.019782	89	380	0.015313	110	416	0.01093
41	Hager	500	Fail	Fail	Fail	26	53	0.00024	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
42	Hager	300	Fail	Fail	Fail	20	41	0.000297	36	158	0.000635	Fail	Fail	Fail	Fail	Fail	Fail
43	Zirilli	2	6	26	0.00043	12	26	0.000626	7	30	0.000454	6	27	0.000431	12	47	0.00025
44	Raydan 1	100	101	406	0.022549	40	108	0.003523	97	402	0.043409	307	1231	0.045256	316	1109	0.027388
45	Raydan 1	10	19	107	0.003201	24	79	0.003837	28	164	0.010223	36	211	0.006813	50	268	0.005844
46	Raydan 1	10	19	110	0.003916	8	31	0.001752	28	185	0.011687	36	214	0.009863	50	272	0.00666
47	Raydan 1	50	59	237	0.006405	22	74	0.003284	50	206	0.011375	159	637	0.017851	174	610	0.009077
48	Raydan 2	10000	2	19	0.018881	4	9	0.001197	2	9	0.018937	2	9	0.007793	4	17	0.006971
49	Raydan 2	50000	2	19	0.048303	4	9	0.001115	2	9	0.044482	2	9	0.026071	4	17	0.029638
50	Raydan 2	100000	2	19	0.10007	4	9	0.000842	2	9	0.085262	2	9	0.048139	4	17	0.047398
51	Fletchcer	10	3	64	0.000272	4	11	0.000231	2	11	0.000629	2	11	0.000265	4	19	0.000178
52	Fletchcer	1000	3	64	0.00049	4	11	0.000193	2	11	0.000559	2	11	0.000274	4	19	0.000239
53	Fletchcer	50000	5	121	0.003947	4	11	0.000188	3	15	0.005728	3	15	0.004125	4	19	0.005288
54	Diagonal 3	10	28	119	0.012473	76	157	0.006622	34	143	0.016775	43	179	0.008364	60	221	0.013047
55	Diagonal 3	10	28	119	0.005556	76	157	0.009211	34	143	0.010692	43	179	0.009202	60	221	0.006221
56	Diagonal 3	2	8	38	0.002895	22	48	0.001596	6	29	0.003307	9	42	0.002497	14	104	0.002933
57	Diagonal 3	10	26	108	0.008004	74	154	0.005951	29	126	0.009358	43	181	0.011208	66	242	0.00687
58	Extended Denschn B	100	7	29	0.000272	18	37	0.000994	5	21	0.000628	10	41	0.000394	18	64	0.000242
59	Extended Denschn B	5000	7	29	0.00082	10	21	0.000324	5	21	0.000544	5	21	0.00042	18	64	0.000449
60	Extended Denschn B	10000	7	29	0.001452	10	21	0.000238	5	21	0.001644	5	21	0.000893	18	64	0.000809

DISCUSSION

We have tested our proposed method, NMRMIL+ within the neighborhood x^* by taking 60 results from 123 benchmark nonlinear system of equations solved. The results demonstrate that the use of NMRMIL+ reduced the number of iterations, number of function evaluations, and the CPU time for solving the tested problems compared to RMIL1, RMIL+, RMIL2, and SMO. This happened because of the low computational cost of the proposed algorithm, which was achieved by incorporating a strong Wolfe line search into the convergence analysis. Also, from Figure 1 to 3, it can be seen clearly that the curve of the proposed CG performs better than the remaining CGs in this research. Based on the above tables, the performance of SMO is 83.3%, RMIL2 is 88.3%, and RMIL1 and RMIL+ have 91.7% each. While the proposed method performed better at 93.3%. So, we regard the proposed CG as superior in solving problems.

Figure 1

Performance profile based on the number of iterations (NI).

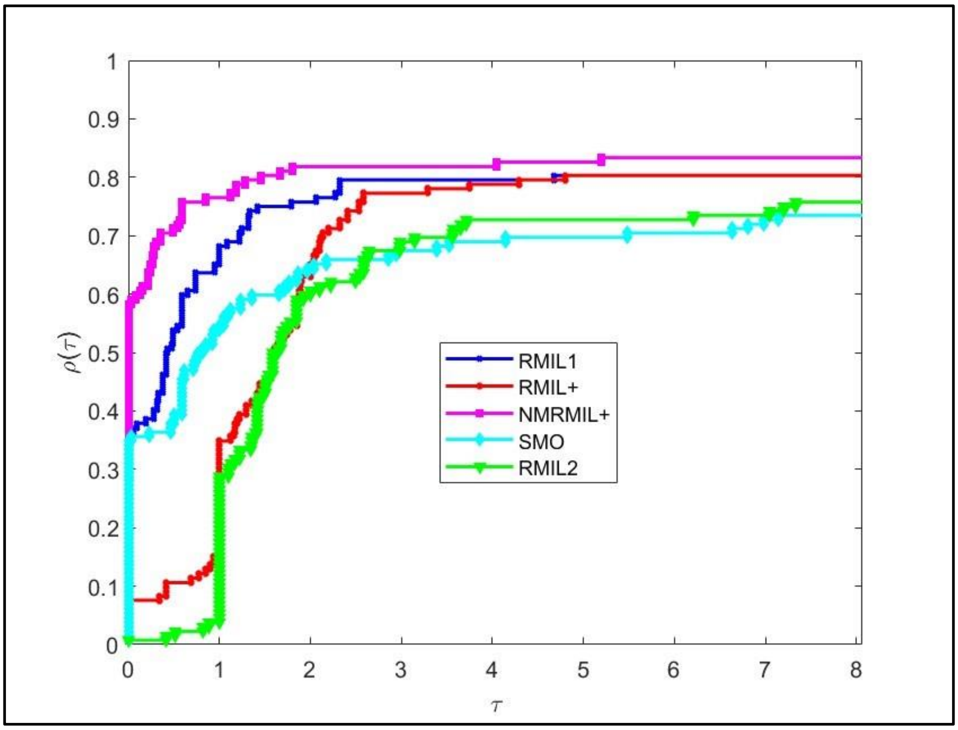


Figure 2

Performance profile based on the number of function evaluations (FE)

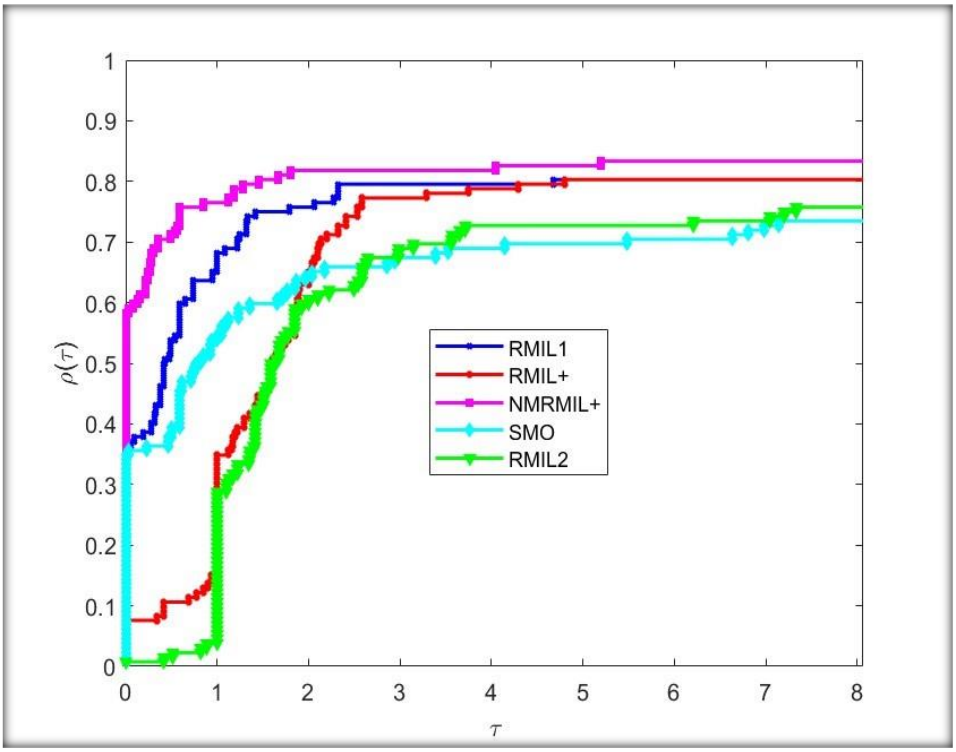
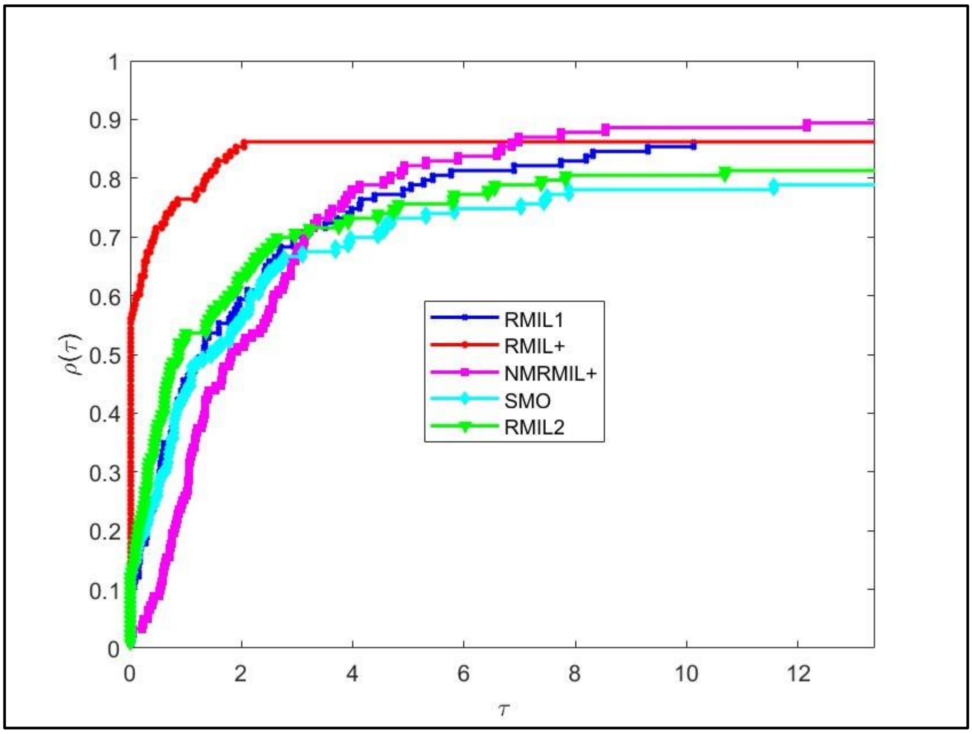


Figure 3

Performance profile based on the CPU time in seconds (CPU)



CONCLUSION

This research presented a modified conjugate gradient coefficient for solving unconstrained optimization problems using inexact line search. The algorithm ($\beta_k^{NMRMIL+}$) is superior to other existing conjugate gradient coefficients. The global convergence and the descent condition were derived with the help of a strong Wolfe line search. The comprehensive numerical results on some benchmark nonlinear system of equations of different characteristics within the neighborhood of x^* showed that the proposed method ($\beta_k^{NMRMIL+}$) is faster and more efficient compared to RMIL1, RMIL+, RMIL2, and SMO respectively.

ACKNOWLEDGEMENT

The authors thank the reviewers and editors for their valuable contribution to improving the manuscript. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abba, V. M., Mustafa, M., Waziri, M. Y., & Afendee, M. M. (2018). A new conjugate gradient with exact line search for unconstrained optimization. Far East Journal of Mathematical Sciences (FJMS), 105(2), 193–206.

Abba, V. M., Mustafa, M., Waziri, M. Y., Afendee, M. M., & Usman, A. Y. (2018). A class of conjugate gradient parameters and its global convergence for solving unconstrained optimization. Far East Journal of Mathematical Sciences (FJMS), 106(1), 43–58.

Andrei, N. (2008). ICI Technical Report 13/08.

Dai, Y. H., & Yuan, Y. (2000). Nonlinear conjugate gradient methods. Beijing: Shang Hai Science and Technology Publisher.

Dolan, E., & More, J. J. (2002). Benchmarking optimization software with performance profile. Mathematical Programming, 91(2), 201–213.

Fletcher, R. (1987). Practical methods of optimization: Unconstrained optimization (Vol. 1). New York: Wiley.

Fletcher, R., & Reeves, C. (1964). Function minimization by conjugate gradient. Computational Journal of Science, 7(1), 149–154.

Gilbert, J. C., & Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. SIAM Journal on Optimization, 2(1), 21–42.

Hager, W. W., & Zhang, H. C. (2005). A new conjugate gradient method with guaranteed descent and efficient line search. SIAM Journal on Optimization, 16(1), 170–192.

Hestenes, M. E., & Stiefel, E. (1952). Method of conjugate gradient for solving linear equation. Journal of Research of the National Bureau of Standards, 49(6), 409–436.

Hilstromm, K. E. (1977). A simulation test approach to the evaluation of nonlinear optimization algorithm. ACM Transactions on Mathematical Software, 3(4), 305–315.

Jorge, N., & Stephen, J. W. (1999). Numerical optimization (5th ed.). New York: Springer.

Lawal, M. (2016). On three conjugate gradient methods for solving nonlinear system of equations. Department of Mathematical Sciences, Bayero University, Kano State, Nigeria.

Liu, Y., & Storey, C. (1992). Efficient generalized conjugate gradient algorithms: Part 1 theory. Journal of Computational and Applied Mathematics, 69(1), 129–137.

Mandara, A. V., Mamat, M., Waziri, M. Y., & Afendee, M. M. (2018). A modification of conjugate gradient method using strong Wolfe line search. International Journal of Engineering & Technology, 7(3.28), 163–167.

Olowo, S. E., & Sulaiman, I. M. (2021). Performance analysis of a modified conjugate gradient algorithm for optimization models. IOP Conference Series: Materials Science and Engineering, 1115(1), 012004.

Polak, E., & Ribiere, G. (1969). Note sur la convergence de directions conjuguées. Revue Française d'Informatique et de Recherche Opérationnelle, 16, 35–43.

Powell, M. J. D. (1984). Nonconvex minimization calculations and the conjugate gradient method. In Lecture Notes in Mathematics (Vol. 1066). Berlin: Springer–Verlag.

Powell, M. J. D. (1986). Convergence properties of algorithms for nonlinear optimization. SIAM Review, 28(4), 487–500.

Rivaie, M., Fauzi, M., Mamat, M., & Ismail, M. (2012). A new class of conjugate gradient coefficient with global convergence properties. AIP Conference Proceedings, 1482(1), 486–490.

Rivaie, M., Mamat, M., & Abdelrhaman, A. (2015). A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches. Journal of Applied Mathematics and Computation, 268, 1152–1163.

Shi, Z. J., & Guo, J. (2009). Conjugate gradient method for nonlinear equations. Journal of Computational and Applied Mathematics, 224(2), 444–457.

Sulaiman, I. M., Mamat, M., Abashar, A., Rivaie, M., Mohamad, F. S., & Saleh, Z. M. (2020). A new conjugate gradient method with descent properties and its application to regression analysis. Journal of Numerical Analysis, Industrial and Applied Mathematics, 14(3–4), 1–15.

Sun, J., & Zhang, J. (2001). Annals. Operational Research, 103, 161–173.

Yakubu, U. A., Abdullahi, I., Murtala, S., & Mandara, A. V. (2020). Application of spectral PRP conjugate gradient parameter for unconstrained optimization problems. Malaysia Journal of Computing and Applied Mathematics, 3(2), 45–53.

Yakubu, U. A., Mamat, M., Afendee, M. M., Ghazali, P. L., & Rivaie, M. (2018). Secant condition for a spectral Hestenes conjugate gradient method and its sufficient descent properties. International Journal of Engineering & Technology, 7(3.28), 312–315.

Yakubu, U. A., Mamat, M., Mandara, A. V., Iguda, A., Murtala, S., & Afendee, M. M. (2019). Decent condition for a scalar parameter on spectral HS (SPHS) conjugate gradient. International Journal of Recent Technology and Engineering (IJRTE), 8, 1–5.

Yakubu, U. A., Mandara, A. V., Abdullahi, I., & Kankarofi, R. H. (2023). Production maximization: A simplex method approach (A case study of Samfarera Rice Mills Kano-Nigeria). Journal of Scientific Research, 2(2), 107–114.

Yakubu, U. A., Morch, M., & Saputra, P. A. (2022). Time series model analysis using autocorrelation function (ACF) and partial autocorrelation function (PACF) for e-wallet transaction during a pandemic. International Journal of Global Operations Research.

Yuan, G., & Wei, Z. (2009). New line search methods for unconstrained optimization. Journal of the Korean Statistical Society, 38(1), 29–39.

Yuan, G., & Wei, Z. (2010). Conjugate gradient of nonlinear problems. Journal of Software Engineering and Applications, 3(5), 503–509.

Yuan, Y., & Sun, W. (1999). Theory and methods of optimization. Beijing: Science Press of China.

Zoutendijk, G. (1970). Nonlinear programming computational methods. In J. Abadie (Ed.), Integer and nonlinear programming (pp. 37–86).