Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM
e-ISSN: 2788-5879 p-ISSN: 2788-5887

A New Conjugate Gradient for Efficient Unconstrained Optimization with Robust Descent Guarantees

Hussein Saleem Ahmed[1*], **Salah Gazi Shareef** [2]

[1,2]Department of Mathemetics, University of Zakho, Kurdistan Region, 42001, Iraq

*Corresponding Author: Hussein Saleem Ahmed

DOI: https://doi.org/10.31185/wjcms.358
Received 21 March 2025; Accepted 24 June 2025; Available online 30 June 2025

ABSTRACT: The Conjugate Gradient method is a powerful iterative algorithm designed to find the minimum of a function by iteratively searching along conjugate directions. This work presents a nonlinear conjugate gradient approach for unconstrained optimization, resulting from the resolution of a new optimization problem. The theoretical foundations of the proposed method are analyzed, with particular attention given to satisfying the descent condition. To evaluate its performance, numerical experiments were conducted, comparing the proposed method with well-known algorithms, including the Dai and Yuan,and Lia and Story methods. The algorithms were implemented in Fortran 95 to record the number of iterations and functions. Matplotlib was used for visualizing performance. The results demonstrate that the new method not only exhibits enhanced efficiency but also significantly outperforms the (Dai and Yuan) and (Lia and Story) methods in terms of optimization effectiveness. These findings suggest that the proposed approach offers a competitive and promising alternative for solving unconstrained optimization problems.

Keywords: Unconstrained optimization, Descent and Sufficient descent condition, Global convergence.

1. INTRODUCTION

Unconstrained optimization involves minimizing or maximizing an objective function without restrictions on the variables. Typically, the objective function is smooth and differentiable. The Conjugate Gradient (CG) method is especially effective for quadratic objective functions, commonly found in linear regression and machine learning, though it can also be adapted for non-quadratic functions.

A general unconstrained optimization problem is represented as:

$$\{\min f(x) | x \epsilon R^n\}, \text{where } f(x): R^n \to R \, , \tag{1}$$

is continuously differentiable and its gradient $g(x) = \nabla f(x)$ where $g(x): R^n \to R^n$ Is available. Conjugate Gradient (CG) methods are important for solving (1), especially for large-scale problems. The CG method has the following form:

$$x_{i+1} = x_i + v_i, i = 0,1,\ldots \tag{2}$$

Where x_0 *is a given initial point*, $v_i = \alpha_i d_i$, α_i Is the step size, d_i The search direction is defined as $d_0 = -g_0$, if $i = 0$

$$d_{i+1} = -g_{i+1} + B_i d_i, \tag{3}$$

where B_i is scaler and $i \geq 0$,

A scaler B_i Given by different formulas and different results in distinct CG methods. There is a cardinal number of CG given by: ($B^{HS} = \frac{g_{i+1}^T y_i}{d_i^T y_i}$ Hestenes and Stiefel [1], $B^{FR} = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}$ R. Fletcher, C. M. Reeves [2], $B^{PRP} = \frac{g_{i+1}^T y_i}{g_i^T g_i}$ Polak and Ribière [3] ,$B^{DY} = \frac{g_{i+1}^T g_{i+1}}{d_i^T y_i}$ Dai and Yuan [4]In addition to the global CG, researchers are looking for new CGs, for example, (see [5],[6],[7],[8], and [9]).

When researching convergence, the researcher frequently needs to conduct an exact or approximate line of research . Applying the CG method. Like the strong Wolf conditions .The strong Wolf conditions are to find α_i,

$$f(x_{i+1} + \alpha_i x_i) \leq f(x) + \delta_i \alpha_i g_i^T d_i, \tag{4}$$

$$|g_{i+1}^T d_i| \leq -\delta_k d_i^T g_i, \tag{5}$$

*Corresponding author: hosensalimramrom@gmail.com

Where $0< \delta_i < \delta_K < 1$ are constants according to Li and Weijun [10].

The study likely explores how these different CG methods perform in terms of convergence speed, computational efficiency, and robustness across various optimization problems. By comparing new formulations with traditional methods in ([4], [11]), the research aims to highlight advancements in optimization techniques that can lead to improved performance in practical applications.

Section 2 (Derivation of New Parameter) is to derive a new parameter by reformulation of existing conjugate gradient equations analytically. The described parameter is aimed at preserving the descent properties and enhancing robustness. The full outline of the new algorithm is filed in section 3 (Roadmap of New Algorithm). It incorporates the new parameter into a modified conjugate gradient scheme, whereby search directions satisfy descent conditions. The part describes the process of initialization to convergence checks. Section 4 contains a rigorous demonstration that the algorithm fulfils the criteria of descent and sufficient descent, required to ensure convergence. They also deliver the convergence of the algorithm globally to standard mathematical assumptions, e.g., Lipschitz continuity of the gradient. At last, Section 5 (Results and Conclusion) provides numerical results of comparing the new algorithm to the existing ones.

2. DERIVATION OF NEW PARAMETER

The main idea of the new CG algorithm is to replace y_i with y_i^* where $y_i^* = y_i + \frac{0.2-\rho_i}{1-\rho_i}(B_i v_i - y_i)$[12], In search direction. This adjustment improves the algorithm's efficiency and convergence properties when solving optimization problems, particularly for large-scale systems.

$$y_i^* = y_i + \frac{0.2-\rho_i}{1-\rho_i}(B_i v_i - y_i),$$

where $\rho_i < 0.2$, $\delta = \frac{2\sqrt{\epsilon}(1+\|x_{i+1}\|)}{\|v_i\|}$, and $B_i v_i = \frac{y_i}{\delta} = \frac{y_i}{\frac{2\sqrt{\epsilon}(1+\|x_{i+1}\|)}{\|v_i\|}} = \frac{\|v_i\|\, y_i}{2\sqrt{\epsilon}(1+\|x_{i+1}\|)}$, then

$$y_i^* = [1 + (\frac{0.2-\rho_i}{(1-\rho_i)})(\frac{1}{\delta} - 1)]y_i$$

Multiply both sides by g_{i+1}^T, We obtain

$$g_{i+1}^T y_i^* = [1 + (\frac{0.2-\rho_i}{(1-\rho_i)})(\frac{1}{\delta} - 1)]g_{i+1}^T y_i, \quad (6)$$

We can write (3) as:

$$d_{i+1}^T = -g_{i+1}^T + B_i^{new} d_i^T$$

Multiplying both sides by y_i and by equation (6), along with the modification of the conjugacy condition ($d_{i+1}^T y_i = -t g_{i+1}^T v_i$), We obtain:

$$-t g_{i+1}^T v_i + (1 + (\frac{0.2-\rho_i}{(1-\rho_i)})(\frac{1}{\delta} - 1))g_{i+1}^T y_i = B_i^{new} d_i^T y_i$$

Let

$$\mu = 1 + (\frac{0.2-\rho_i}{(1-\rho_i)})(\frac{1}{\delta} - 1), \quad (7)$$

Finally, we get

$$B_i^{new} = \frac{-t g_{i+1}^T v_i}{d_i^T y_i} + \mu \frac{g_{i+1}^T y_i}{d_i^T y_i}, \quad (8)$$

And the new search direction update gives us:

$$d_{i+1} = -g_{i+1} + [\frac{-t g_{i+1}^T v_i}{d_i^T y_i} + \mu \frac{g_{i+1}^T y_i}{d_i^T y_i}]d_i, \quad (9)$$

3. ROADMAP OF NEW ALGORITHM

Data: $0< \delta_i < \delta_K <1$, $\mu = 0.98$, $\rho_i < 0.2$, $\delta = 2\sqrt{\epsilon}$, $\sqrt{\epsilon}=10^{-3}$ *and* $\epsilon = 10^{-6}$

- **Initialization:** start with an initial point x_0 *and initial value* $\rho_i = 1.0001$,$d_0 = -g_0$, $\epsilon = 10^{-6}$ k=0 and $\mu = 0.98$.
- **Convergence check:** if the norm of the gradient $g_i \le \epsilon$ Stop; otherwise, proceed to the next step.
- **Step length calculation:** calculate the step size by (4) and (5).
- **Update position:** update the point $x_{i+1} = x_i + v_i, i = 0,1, \ldots$ If the norm of the gradient $g_{i+1} \le \epsilon$, stop.
- **Search direction:** Compute the new search direction by using (9) where B_i^{new} calculate by (8)
- **Iteration:** increment i and repeat it.

3.1 DESCENT AND SUFFICIENT DESCENT CONDITIONS

The CG method's descent condition ensures each iteration reduces the objective function or error, and a descent direction guarantees movement towards a minimum (optimal solution). But the second condition, this stronger condition, guarantees that the reduction in the function value from one iteration to the next is significant enough.

Before proving the descent condition, we first verify that the scalar. μ Is positive.
According to algebraic properties, we start with the inequality:

$$\|x_{i+1} - x_i\| \leq \|x_{i+1}\|$$

We can write by

$$\|x_{i+1} - x_i\| \leq 1 + \|x_{i+1}\|$$

Now, divide both sides by $(1 + \|x_{i+1}\|)$ and multiply by $(\frac{1}{2\sqrt{\epsilon}})$ gives:

$$\frac{\|x_{i+1} - x_i\|}{2\sqrt{\epsilon}(1 + \|x_{i+1}\|)} \leq \frac{1}{2\sqrt{\epsilon}}$$

We know $\|v_i\| = \|x_{i+1} - x_i\|$, this gives

$$\frac{\|v_i\|}{2\sqrt{\epsilon}(1 + \|x_{i+1}\|)} - 1 \leq \frac{1}{2\sqrt{\epsilon}} - 1$$

Multiplying both sides by $(\frac{0.2-\rho_i}{1-\rho_i})$ and adding 1 to both sides yields:

$$\left[\frac{1}{\delta} - 1\right]\left(\frac{0.2 - \rho_i}{1 - \rho_i}\right) + 1 \leq \left(\frac{1}{2\sqrt{\epsilon}} - 1\right)\left(\frac{0.2 - \rho_i}{1 - \rho_i}\right) + 1$$

From (7) we get

$$\mu \leq \left(\frac{1}{2\sqrt{\epsilon}} - 1\right)\left(\frac{0.2 - \rho_i}{1 - \rho_i}\right) + 1$$

since both $(\frac{1}{2\sqrt{\epsilon}} - 1)$, and $\frac{0.2-\rho_i}{1-\rho_i}$ They are positive. We conclude that. $\mu > 0$. For this study, the value of μ Chosen to be 0.98.

Theorem 1: The sequence $\{x_i\}$It is generated by (2) and (9), where the step size is determined using both Exact Line Search (ELS) and Inexact Line Search (ILS) methods. Then the search direction (3) with a new parameter (8) of the conjugate The gradient method is given as (9) satisfies the descent direction, i.e. $g_{i+1}^T d_{i+1} \leq 0$.

Proof: Multiply both sides of (9) by g_{i+1}^Ton the left , we obtain

$$g_{i+1}^T d_{i+1} = -g_{i+1}^T g_{i+1} + \left[\frac{-t g_{i+1}^T v_i}{d_i^T y_i} + \mu \frac{g_{i+1}^T y_i}{d_i^T y_i}\right] g_{i+1}^T d_i, \quad (10)$$

I)If the step length is determined by (ELS), which requires $g_{i+1}^T d_i = 0$
Thus,

$$g_{i+1}^T d_{i+1} = -\|g_{i+1}\|^2 \leq 0.$$

Then the proof is completed.

II)On the other hand, if the step length is determined by (ILS), which requires $g_{i+1}^T d_i \neq 0$. We know that the first term is less than or equal to zero; we need to prove that the second term is also less than or equal to zero.

$$g_{i+1}^T d_{i+1} = -g_{i+1}^T g_{i+1} - \frac{\alpha t (g_{i+1}^T d_i)^2}{d_i^T y_i} + \mu \frac{g_{i+1}^T g_{i+1}}{d_i^T y_i} g_{i+1}^T d_i - \mu \frac{g_{i+1}^T g_i}{d_i^T y_i} g_{i+1}^T d_i$$

since $d_i^T y_i \geq g_{i+1}^T d_i$ then

$$g_{i+1}^T d_{i+1} \leq -g_{i+1}^T g_{i+1} - \alpha t d_i^T y_i + \mu g_{i+1}^T g_{i+1} - \mu g_{i+1}^T g_i, \quad (11)$$

Multiply wolf conditions $g_{i+1}^T d_i \geq c\, d_i^T g_i$ by (-1) we have $g_{i+1}^T g_i \geq c\, g_i^T g_i$
Or $-g_{i+1}^T g_i \leq -c\, g_i^T g_i$, then

$$g_{i+1}^T d_{i+1} \leq -(1 - \mu)\|g_{i+1}\|^2 - \alpha t d_i^T y_i - c\, \mu \|g_i\|^2, \quad (12)$$

It is clearly $1 - \mu$, $\alpha, t, d_i^T y_i$ And c, are less than or equal to zero, s.t

$$g_{i+1}^T d_{i+1} \leq 0. \blacksquare$$

Theorem 2: Assume that the step length satisfies strong Wolfe conditions, and then the following result: $g_{i+1}^T d_{i+1} \leq -c\|g_{i+1}\|^2$ holds when $0 \leq i$.

Proof: From equation (11) we have

$$g_{i+1}^T d_{i+1} \leq -(1 - \mu)\|g_{i+1}\|^2 - \alpha t d_i^T y_i - c\, \mu \|g_i\|^2 \leq 0, \quad (13)$$

According to algebraic rules, we can be written can be write (13) by

$$g_{i+1}^T d_{i+1} \leq -(1 - \mu)\|g_{i+1}\|^2,$$

Let c= $1-\mu$ And c is positive
Finally, we have
$g_{i+1}^T d_{i+1} \leq -c\|g_{i+1}\|^2$

3.2 THE GLOBAL CONVERGENCE OF THE NEW ALGORITHM

Global convergence is a key property of optimization algorithms. Ensuring they reach a stationary point (or minimizer) of the objective function from any initial point. F conjugate gradient method, this means that the sequence of iterations. Subscript $\{x_i\}$ will satisfy the condition $\lim_{i\to\infty}\|g_i\| = 0$.

Lema 1. If $\sum_{k\geq 1}\frac{1}{\|d_k\|^2} = \infty \quad then \ \lim_{k\to\infty} inf\|g_k\| = 0$

In this section, the global convergence of the new CG algorithm is analyzed under assumptions.

Assumptions (E)

(E1) Lipschitz Continuity of the Gradient:

Assuming η is a neighborhood of a point or set Ω, and f is continuously differentiable in η and its gradient is Lipschitz continuous, i.e., there exists a constant L > 0 such that $\|g(x_{i+1}) - g(x_i)\| \leq L\|x_{i+1} - x_i\| \ for\ all\ x$ ∈n

We can be written by

$$\|y_i\| \leq L\|v_i\|, \tag{14}$$

(E2) **Convex function:** There exists a constant $\vartheta \geq 0$ such that for all x, y∈ φ

$$(g(x) - g(y))^T(x - y) \geq \vartheta\|x - y\|^2, \tag{15}$$

If f Is a uniformly convex function, we can write Eq. (15) by

$$y_i^T v_i \geq \vartheta\|v_i\|^2 \text{ or } y_i^T d_i \geq \frac{\vartheta}{\alpha}\|v_i\|^2, \tag{16}$$

(E3) **Bounded level sets:** The level set Ω = $\{x \in R^I / f(x) \leq f(x_0) + \varepsilon\}$ The function f is bounded below.

By algebraic role, we can write.

$$\|f(x)\| \ \| g(x) \| \geq f(x)^T g(x), \tag{17}$$

On the objective function f(x), there exists a constant. $\gamma \geq 0$ such that

$$\|g_{i+1}\| \leq \gamma, \tag{18}$$

Theorem 3: Suppose that assumption (E) holds and that the objective function f(x) is uniformly convex. The new algorithm of the form (14), (15), (16), (17) and (18), where satisfies the descent condition and is obtained by the strong Wolfe conditions (4) and (5), satisfies the global convergence, i.e.

$$\lim_{k\to\infty} inf\|g_i\| = 0$$

Proof:

We can rewrite equation (9) as

$$\|d_{i+1}\| = \|-g_{i+1} + B_i^{new} d_i\|$$

Taking norms on both sides gives

$\|d_{i+1}\| \leq \|g_{i+1}\| + |B_i^{new}| * \|d_i\|$

Substituting the expression for B_i^{new}, we obtain

$\|d_{i+1}\| \leq \|g_{i+1}\| + \left|\frac{t g_{i+1}^T v_i}{d_i^T y_i}\right| \|d_i\| + \left|\mu \frac{g_{i+1}^T y_i}{d_i^T y_i}\right| * \|d_i\|$

We know $d_i^T y_i > g_{i+1}^T d_i$, $g_{i+1}^T y_i < \|g_{i+1}\|\|y_i\|$ and $y_i^T d_i \geq \frac{\vartheta}{\alpha}\|v_i\|^2$

$\|d_{i+1}\| \leq \|g_{i+1}\| + \alpha t\|d_i\| + \left|\mu\alpha \frac{\|g_{i+1}\|\|y_i\|}{\vartheta\|v_i\|^2}\right| * \|d_i\|$

Under Assumptions (E1), (E2) and (E3) gives as:

$\|d_{i+1}\| \leq \gamma + \alpha t\|d_i\| + \frac{\mu\alpha}{\vartheta}\gamma$

Let $\|v_i\| = \|x_{i+1} - x_i\|$, D= $max\{\|x_{i+1} - x_i\|, for\ all\ x\epsilon R\}$ $x_{i+1}$and x_i They are consecutive iterates of the optimization algorithm.

$\|d_{i+1}\| \leq \gamma + tD + \mu\gamma\frac{L}{\vartheta} = \varphi$

Now, by the lemma.1 If $\sum_{i\geq 1}\frac{1}{\|d_{i+1}\|^2} = \sum_{i\geq 1}\frac{1}{\varphi^2} = \infty$

then

$\lim_{i\to\infty} \inf\|g_i\| = 0$. ■

4. NUMERICAL RESULT

The section is devoted to comparative numerical analysis of the proposed method with two classical methods of computing conjugate gradients:(DY-CG) and (LS-CG). The benchmark test functions employed in the evaluation are seven in number, comprising different dimensions of the problem: 5, 100, 3000, and 5000. In Tables 1, 2, and Figure 1,2, the results are outlined. The algorithms were implemented in Fortran 95 to record the number of iterations and functions. Matplotlib was used for visualizing performance.

They are then evaluated based on the number of iterations (NOI) and number of function evaluations (NOF) a method may need to complete. An overall lower number in either of the measures indicates more efficiency and strength of the optimization strategy. In all experiments, the following parameter settings were fixed:

$$\mu = 0.98, \rho_i = 1.9998, t = 1, \varepsilon = 10^{-6} \; and \; \delta = 2\sqrt{\epsilon} .$$

As evidenced in the results, the proposed method also used a lesser number of iteration steps and evaluations of the functions in most of the test cases, and in particular, the high-dimensional problems (see Tables 1,2, and Figure 1, 2)

Table 1. - Comparison between the new technique with two well-known CG methods across the number of iterations (NOI) and the number of functions (NOF)

Function Test	Dimensions N	NEW Method NOI-NOF	Dai-Youan NOI-NOF	Lia- Story NOI-NOF
	5	5-24	6-39	6-39
SUM	100	14-79	14-85	14-80
$\mathbf{x_0} = \mathbf{2}$	3000	42-217	31-166	32-167
	5000	29-132	33-149	37-202
SHALLOW	5	8-21	8-21	8-21
$\mathbf{x_0} = (\mathbf{-2}, \mathbf{-2})$	100	8-21	8-21	8-21
	3000	9-24	9-24	9-24
	5000	9-24	9-24	9-24
WOLF	5	15-31	13-27	14-29
$\mathbf{x_0} = \mathbf{-1}$	100	49-99	45-91	49-99
	3000	174-359	125-263	176-364
	5000	166-347	159-327	166-347
BEAL	5	11-28	11-28	11-28
$\mathbf{x_0} = (\mathbf{0}, \mathbf{0})$	100	11-28	12-30	12-30
	3000	12-30	12-30	12-30
	5000	12-30	12-30	12-30
ROSEN	5	30-84	30-82	30-85
$\mathbf{x_0} = (\mathbf{-1.2}, \mathbf{1})$	100	30-84	30-82	30-85
	3000	31-86	30-82	30-85
	5000	31-86	30-82	30-85
CUBIC	5	12-35	14-39	15-45
$\boldsymbol{x_0} = (\mathbf{-1.2}, \mathbf{1})$	100	13-37	15-43	16-47
	3000	13-37	15-43	16-47
	5000	13-37	15-43	16-47
OSP	5	8-41	9-45	12-71
$\mathbf{x_0} = \mathbf{1}$	100	47-163	52-180	72-227
	3000	218-668	421-1388	400-1438
	5000	254-785	555-1857	729-2706
Total	--------	1296-3698	1724-5321	1915-6483

Key Finding

• Hybrid parameter formulation: Since the technique is a combination of the advantages of various beta update strategies, it has the advantage that it will exhibit the global convergence benefits as well as local acceleration effects.

• Adaptive scaling μ: Use of the μ Scaler in the beta formula is a balancing agent to the method. It was observed that after setting dozens of values, the best option available was 0.98 because it produced the best performance concerning the efficiency of optimization. It was a value that considerably enhanced the outcome as opposed to the common method, which does not incorporate the said parameter.

• Numerical behavior on large-scale problems: The approach is robust with dimension, as in the case of 3000 and the 5000 variable benchmarks.

• Smooth updates to the search direction: The three-term form keeps away useless zigzagging and aids in staying nearer to a straight shot towards the minimizer.

Table 2. - Total number of iterations and functions of the new method with the DY and LS methods

Method	Total NOI	Total NOF
New CG method	1296	3698
Dai-Yoan method	1724	5321
Lia-Story method	1915	6483

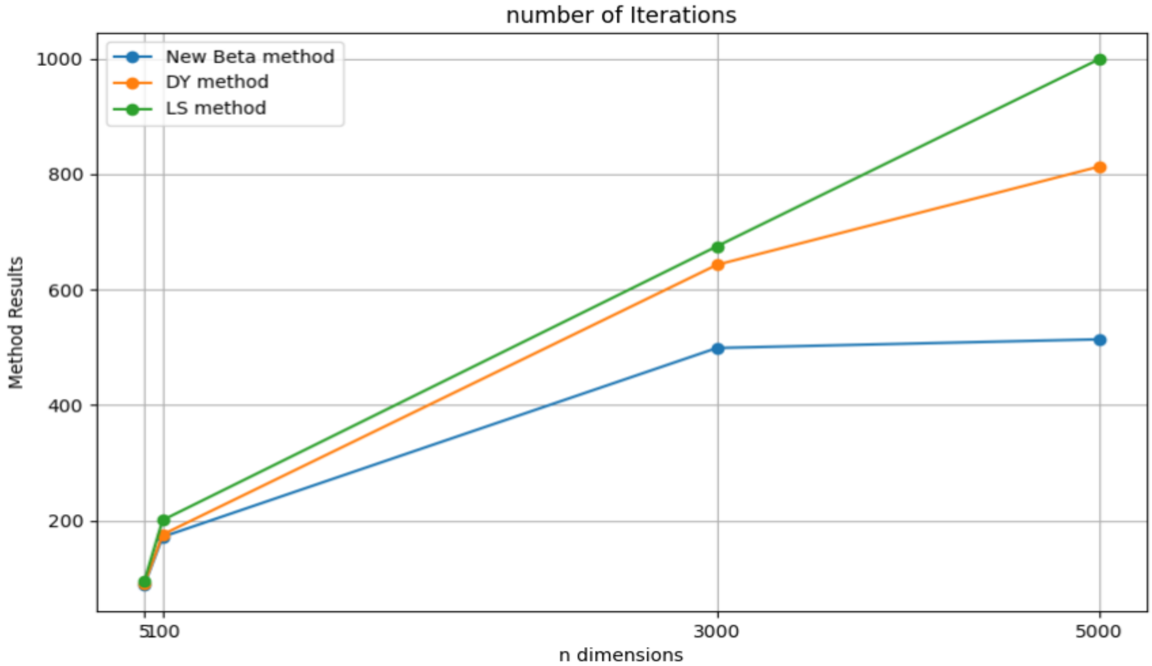

FIGURE 1. - Comparison of the new method with the DY and LS methods in terms of the number of iterations across different problem dimensions, 5 to 5000

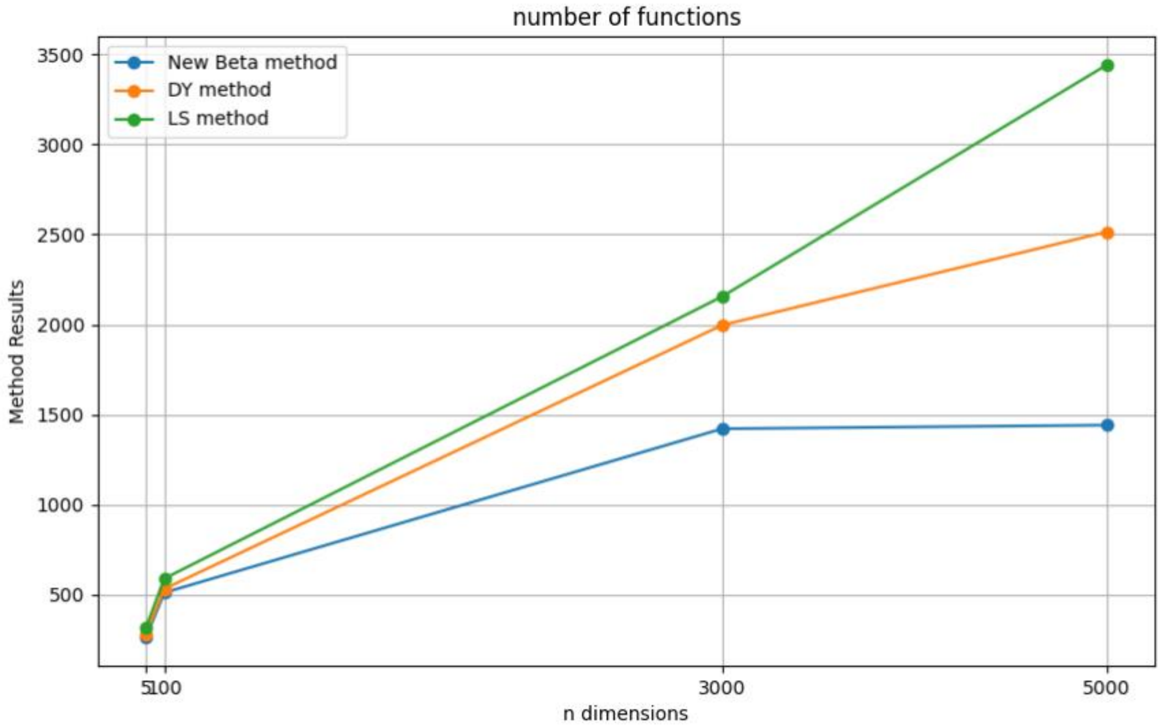

FIGURE 2. - Comparison of the new method with the DY and LS methods in terms of the number of function evaluations across different problem dimensions, 5 to 5000

5. CONCLUSION

In the work, a new beta formula was presented in the scope of conjugate gradient methods of unconstrained nonlinear optimization problems. The suggested formula uses a mechanism that is underpinned by a tunable (miu =0.98), which lets the search direction be adaptively mined. Based on an extensive numerical testing of a collection of standard test functions with dimensions up to 5000-dimensional, the new method was compared in terms of its performance against two of the most famous conjugate gradient methods: DY-CG and LS-CG.

The findings demonstrated the fact that the new beta-based procedure revealed consistent success on both measures of iterations and the functional call in high-dimensional problems. This is improved by the effects of the adaptive term in the formula, in the descent properties, and improved numerical stability of the formula. The procedure also preserved competitive convergence behavior as well as lowering computational cost, suggesting its robustness and scalability.

Funding

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems (Vol. 49, No. 1). Washington, DC: NBS.

[2] Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. The computer journal, 7(2), 149-154.

[3] Polak, E., & Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. Revue française d'informatique et de recherche opérationnelle. Série rouge, 3(16), 35-43.

[4] Dai, Y. H., & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence

property. SIAM Journal on Optimization, 10(1), 177-182.
[5] Jahwar, B. H., Ajeel, S. M., & Shareef, S. G. (2024). Two new classes of the conjugate gradient method based on logistic mapping. TELKOMNIKA (Telecommunication Computing Electronics and Control), 22(1), 86-94.
[6] Khatab, H. A., & Shareef, S. G. (2024). Two new limited-memory preconditioned conjugate gradient algorithms for nonlinear optimization problems. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-14.
[7] SHAREEF, S. G. (2022). A NEW CONJUGATE GRADIENT WITH GLOBAL CONVERGENCE FOR NONLINEAR PROBLEMS. Journal of Duhok University, 25(2), 573-578.
[8] Dwail, H. H., Mahdi, M. M., & Shiker, M. A. (2022). CG method with modifying βk for solving unconstrained optimization problems. Journal of Interdisciplinary Mathematics, 25(5), 1347-1355.
[9] Ibrahim, A. L., Sadiq, M. A., & Shareef, S. G. (2019). A New Conjugate Gradient Coefficient for Unconstrained Optimization Based On Dai-Liao. Science Journal of University of Zakho, 7(1), 34-36.
[10] Nocedal, J., & Wright, S. J. (Eds.). (1999). Numerical optimization. New York, NY: Springer New York.
[11] Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: theory. Journal of optimization theory and applications, 69, 129-137.
[12] Powell, M. J. (1978). Algorithms for nonlinear constraints that use Lagrangian functions. Mathematical programming, 14, 224-248.