STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, December 2025, pp 0–8.
Published online in International Academic Press (www.IAPress.org)

Extra Dai-Liao Method in Conjugate Gradient Method for Solving Minimization Problems

Waleed Abdulazeez Majeed, Basim A. Hassan*

Department of Mathematics,College of Computers Sciences and Mathematics University of Mosul, IRAQ

Abstract This study explores different strategies for setting parameters in optimization algorithms, focusing on refining the Dai–Liao (DL) conjugate gradient method by using a modified quasi-Newton framework. The DL version of the conjugate gradient method is known for its effectiveness in addressing large-scale unconstrained optimization challenges. Nonetheless, conventional implementations often depend on differences between successive iterates and gradient vectors, which can limit adaptability and convergence capabilities in certain circumstances. To overcome these limitations, the proposed method introduces an innovative parameter formula that utilizes the curvature condition differently by incorporating objective function values, instead of just relying on point and gradient differences. This use of function values offers more detailed insights into the optimization landscape, thereby enhancing both the stability and accuracy of the search direction. The main benefit of this modification is its augmented computational efficiency and its capacity to ensure global convergence under relatively mild and realistic conditions. Theoretical analysis, including a proof of global convergence for the new method, supports these assertions. To verify the practical effectiveness of this approach, extensive numerical experiments were carried out on various standard test problems. The results consistently show that the modified method surpasses the traditional DL conjugate gradient algorithm in terms of convergence speed and robustness, confirming the theoretical enhancements and underscoring its potential for wider use in nonlinear optimization.

Keywords Extra Dai-Liao, Conjugate Gradient, Convergence, Optimization

AMS 2010 subject classifications 90C30, 90C26, 65F10

DOI: 10.19139/soic-2310-5070-2896

1. Introduction

Our solution to the unconstrained optimization issue is formulated using the conjugate gradient methodology:

$$Minf(x), \quad \mathrm{x} \in R^n \tag{1}$$

where f is a smooth function, [2]. First, we study the conjugate gradient method (CG), which repeatedly applies the following to create a series of iterates:

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

In this case, α_k represents the step size and d_k the search direction, which is as follows:

$$d_{k+1} = -g_{k+1} + \beta_k s_k \tag{3}$$

where the conjugate gradient methodology is delineated by the parameter β_k, refer to [3, 5]. While a subset of these methodologies possesses theoretical utility, others exhibit efficacy in numerical applications. The nonlinear

*Correspondence to: Dr. Basim A. Hassan (Email: basimah@uomosul.edu.iq). Department of Mathematics,College of Computers Sciences and Mathematics University of Mosul, IRAQ.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

conjugate gradient method was originally formulated by Hestenes-Stiefel [4], wherein β_k is represented in the subsequent manner:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \tag{4}$$

Step length α_k is conventionally chosen in iterative methodologies to fulfill specific criteria. The Wolfe conditions are frequently employed in the analysis of convergence and the execution of conjugate gradient techniques among the conditions presented, and they must adhere to the following stipulations:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \tag{5}$$

$$d_k^T g(x_k + \alpha_k d_k) \geq \sigma \, d_k^T g_k \tag{6}$$

where $0 < \delta < \sigma < 1$, see [10]. It is unfortunate that in instances where the objective function is nonconvex and the conventional Wolfe line search is applied, none of the proposed methodologies can achieve simultaneous convergence.

One of the well-known conjugate gradient strategies intended to improve the effectiveness of unconstrained nonlinear optimization algorithms is the Dai-Liao approach. In [6], Dai and Liao developed this technique in 2001 to overcome some of the drawbacks of more conventional approaches, which don't always ensure a significant enough decrease in the objective function value throughout the optimization process. The Dai-Liao approach uses a new formula to determine the beta coefficient β_k in order to increase numerical stability and guarantee trustworthy global convergence. This approach uses the following connection to calculate β_k as follows:

$$\beta_k^{DL} = \frac{g_{k+1}^T y_k - t g_{k+1}^T s_k}{d_k^T y_k} \tag{7}$$

This formula eliminates the requirement for restart approaches while allowing for the maintenance of an efficient fall direction. This method's ability to provide enough descent at each iteration and achieve good convergence, even in complicated circumstances where other approaches fail, is one of its primary characteristics. The Dai-Liao technique is the best option for constants because, according to numerical data, it is more stable and effective than other conjugate gradient methods, [7].

In a multitude of other scholarly investigations, extended formulations of the traditional CG parameters have been proposed, primarily focusing on the transition from the Hestenes–Stiefel framework to the DL methodology, [17]. As a consequence, several one-parameter extended classes of conventional CG algorithms have been formulated to enhance their adaptability for specific objectives. Notably, modified secant (quasi–Newton) equations have been instrumental in the advancement of generalized versions of the CG methods, thereby ensuring the sufficient descent property through a straightforward yet significant methodology, [13, 14, 15]. Subsequently, many conjugate gradient algorithms based on various modified secant equations were scrutinized, as indicated in [16, 17].

In the quest to accomplish both global convergence utilizing the conventional Wolfe line search and the essential descent property for any line search, we propose modifications to the conjugate gradient methodology. Moreover, a novel framework is introduced along with the foundational inspiration.

2. Extra Dai-Liao Method:

The Taylor series constitutes a fundamental concept essential for the determination of the novel conjugate gradient parameter. As we progress:

$$f(x) = f(x_{k+1}) - g_{k+1}^T s_k + \frac{1}{2} s_k^T Q(u_{k+1}) s_k \tag{8}$$

The computation of the derivative is articulated as:

$$g_{k+1} = g_k + Q(u_k) s_k \tag{9}$$

The calculation of second-order curvature can be achieved through the application of equations (8) and (9):

$$s_k^T Q(u_k) s_k = 2(f_k - f_{k+1}) + 2y_k^T s_k + 2g_k^T s_k \tag{10}$$

By substituting $B_{k+1} \cong Q(u_{k+1})$ into (10), we derive:

$$s_k^T B_{k+1} s_k \cong 2(f_k - f_{k+1}) + 2y_k^T s_k + 2g_k^T s_k \tag{11}$$

An optimal choice of the parameter β_k leads to improved accuracy and stability, by using:

$$d_{k+1}^T y_k = -g_{k+1}^T s_k \tag{12}$$

This criterion is called as Perry's conjugacy condition, as noted in [9]. By employing (12) in (11), we have derived:

$$d_{k+1}^T B_{k+1} s_k \cong -\left[2 + \frac{2(f_k - f_{k+1}) + 2g_k^T s_k}{s_k^T y_k}\right] g_{k+1}^T s_k \tag{13}$$

Currently, by utilizing the direction (3) in (13), we determine that:

$$(-g_{k+1} + \beta_k s_k)^T y_k \cong -\left[2 + \frac{2(f_k - f_{k+1}) + 2g_k^T s_k}{s_k^T y_k}\right] g_{k+1}^T s_k \tag{14}$$

This suggests:

$$\beta_k = \frac{g_{k+1}^T y_k - t g_{k+1}^T s_k}{s_k^T y_k}, t_k^{BJ} = 2 + \frac{2(f_k - f_{k+1}) + 2g_k^T s_k}{s_k^T y_k} \tag{15}$$

It is called Dai-Liao (BJ) method.

Algorithm: Dai-Liao (BJ).

Input: Initial guess $x_0 \in R^n$ set a small tolerance ε.

Output: Minimizer x such that gradient is approximately zero.

1. Compute $g_0 = \nabla f_0$, set $d_0 = -g_0$, and initialize $k = 0$.
2. While the norm of the gradient $gk + 1 > 0$, then:
   (a) Perform a line search to compute step size α_k satisfying the Wolfe-conditions.
   (b) Update $x_{k+1} = x_k + \alpha_k d_k$.
   (c) Compute $t_k^{BJ} = 2 + \frac{2(f_k - f_{k+1}) + 2g_k^T s_k}{s_k^T y_k}$. Then calculate $\beta_k^{\text{Dai-Liao (BJ)}} = \frac{g_{k+1}^T y_k - t^{BJ} g_{k+1}^T s_k}{s_k^T y_k}$.
   (d) Update the search direction: $d_{k+1} = -g_{k+1} + \beta_k s_k$.
   (e) Increment $k \leftarrow k + 1$.
3. End while.
4. Return the final iterate x_{k+1}.

3. Essential Assumptions for Establishing Convergence

1. Convex Level Set: The level set:
$$L_0 = \{x/_{f(x) \leftarrow x_0}\} \tag{16}$$
is convex.
2. Lipschitz gradient: The gradient ∇f is Lipschitz continuous on L_0; that is, there is $L > 0$ such that:
$$\left(\nabla f\left(o^-\right) - \nabla f\left(v^+\right)\right) \leq L\left\|o^- - v^+\right\|, \quad \forall o^-, v^+ \in L_0 \tag{17}$$
3. Bounded Gradient Norm: Under these assumptions, a constant $\Pi > 0$ exists with:
$$\|g_{k+1}\| \leq \Pi \tag{18}$$

See [11].

Theorem 1
If d_{k+1} is derived utilizing the Dai-Liao (BJ) methodology and $0 < L < 1$, then: $d_{k+1}^T g_{k+1} \leq -c\|g_{k+1}\|^2$ holds.

Proof
Starting from the definition of d_{k+1} by Multiply g_{k+1}^T, we can write:

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{\left[2 + \frac{2(f_k - f_{k+1}) + 2g_k^T s_k}{s_k^T y_k}\right] s_k^T g_{k+1}}{s_k^T y_k}\right] s_k^T g_{k+1} \tag{19}$$

Combining (13) and (19), we get:

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{s_k^T g_{k+1}}{s_k^T y_k}\right] s_k^T g_{k+1} \tag{20}$$

Based on the Lipschitz continuity property (17), we obtain:

$$y_k^T g_{k+1} \leq L s_k^T g_{k+1} \tag{21}$$

Inserting (20) into (21) yields:

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \left[L\frac{s_k^T g_{k+1}}{s_k^T y_k} - \frac{s_k^T g_{k+1}}{s_k^T y_k}\right] s_k^T g_{k+1} \tag{22}$$

The result above leads to the conclusion that:

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + [L-1]\frac{(s_k^T g_{k+1})^2}{s_k^T y_k} \tag{23}$$

Additionally,

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 < 0 \tag{24}$$

The proof is finished. □

To prove the comprehensive convergence of the new conjugated gradient method, we present the important lemma, presented by Dai et. al. according to the source [8].

Lemma 1
Suppose $(\alpha_k, x_{k+1}, g_{k+1}, d_{k+1})$ generated by the conjugate gradient algorithm, since the search direction satisfies the regression condition and satisfies the Wolfe condition, if:

$$\sum_{k>1} \frac{1}{\|d_{k+1}\|^2} = \infty \tag{25}$$

Then it will be:

$$\liminf_{k\to\infty} \|g_{k+1}\| = 0 \tag{26}$$

Theorem 2
We assume $(\alpha_k, x_{k+1}, g_{k+1}, d_{k+1})$ generated from the Dai-Liao (BJ) algorithm. Suppose that for every existence there is a constant $\Phi > 0$, such that:

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq \Phi\|x - y\|^2 \tag{27}$$

Then it will be:

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{28}$$

Proof
Using the direction of the search line and the new correlation coefficient known as equation (3), we obtain the following equation:

$$\|d_{k+1}\| = \| - g_{k+1} + \beta_k^{\text{Dai-Liao (BJ)}}\ s_k\| = \| - g_{k+1} + \frac{g_{k+1}^T y_k}{d_k^T y_k} s_k - \frac{s_k^T g_{k+1}}{d_k^T y_k} s_k\| \tag{29}$$

Based on the condition of Lipschitz and equation (29), we get:

$$\|d_{k+1}\| \leq \|g_{k+1}\| + \frac{\|g_{k+1}\| L \|s_k\|^2}{\Phi \|s_k\|^2} \leq \left(1 + \frac{L}{\Phi} + \frac{1}{\Phi}\right) \|g_{k+1}\| \leq \left[\frac{\Phi + L + 1}{\Phi}\right] \|g_{k+1}\| \tag{30}$$

Using the hypothesis 3 on f, there is a constant number with the direction of the search that produces:

$$\sum_{k\geq 1} \frac{1}{\|d_{k+1}\|^2} \geq \left(\frac{\Phi}{\Phi + L + 1}\right) \frac{1}{\Pi} \sum_{k\geq 1} 1 = \infty \tag{31}$$

Thus, the proof is completed. □

4. Numerical Results

In this study, we examine the impact of a newly proposed algorithm on a set of numerical optimization problems, [1]. The unconstrained optimization test problems considered are listed in Table 1.

Table 1. Table of problems, starting points and dimensions

Problem	**Dim**	**Function No.**	**Starting Points**
Freudenstein & Roth	100, 1000	1, 2	$x_0 = [0.5, -2, 0.5, -2, ..., 0.5, -2]$
Trigonometric	100, 1000	3, 4	$x_0 = [0.2, 0.2, 0.2, 0.2, ..., 0.2, 0.2]$
Extended Rosenbrock	100, 1000	5, 6	$x_0 = [-1.2, 1, -1.2, 1, ..., -1.2, 1]$
Penalty	100, 1000	7, 8	$x_0 = [1, 2, 1, 2, 1, 2, ...1, 2, 1, 2]$
Perturbed Quadratic	100, 1000	9, 10	$x_0 = [0.5, 0.5, 0.5, 0.5, ..., 0.5, 0.5]$
Extended Tridiagonal 1	100, 1000	11, 12	$x_0 = [2.0, 2.0, 2.0, ...2.0, 2.0]$
Extended Three Expo	100, 1000	13, 14	$x_0 = [0.1, 0.1, 0.1, 0.1, ..., 0.1, 0.1]$
Generalized Tridiagonal 2	100, 1000	15, 16	$x_0 = [-1, -1, -1, ... - 1, -1]$
Extended PSC1	100, 1000	17, 18	$x_0 = [3, 0.1, 3, 0.1, ...3, 0.1, 3, 0.1]$
Extended Cliff	100, 1000	19, 20	$x_0 = [-1, 0, -1, 0, ..., -1, 0, -1, 0]$
Extended Hiebert	100, 1000	21, 22	$x_0 = [0, 0, 0, 0, ..., 0, 0, 0, 0]$
Extended Tridiagonal 2	100, 1000	23, 24	$x_0 = [1, 1, 1, 1, ..., 1, 1, 1]$
STAIRCASE S1	100, 1000	25, 26	$x_0 = [1, 1, 1, 1, ..., 1, 1, 1]$
DIXON3DQ (CUTE)	100, 1000	27, 28	$x_0 = [-1, -1, -1, ... - 1, -1]$
SINCOS	100, 1000	29, 30	$x_0 = [-1, -1, -1, ... - 1, -1]$
Generalized Quartet GQ2	100, 1000	31, 32	$x_0 = [1, 1, 1, 1, ..., 1, 1, 1]$

We evaluate the performance of our proposed method, referred to as the Dai-Liao (BJ) algorithm, in combination with a modified version of the HS algorithm, for problem dimensions n=100 and n=1000. The performance is compared against the standard HS algorithm on the same set of problems.

This section outlines the stopping criteria, problem dimensions, and key parameter settings used in the numerical experiments $\delta = 0.001$ and $\sigma = 0.9$. The termination condition is defined by $\|g_{k+1}\| \leq 10^{-6}$.

The comparison metrics include:

- **NI**: Number of Iterations
- **NR**: Number of Restarts
- **NF**: Number of function evolution

The numerical results demonstrate that the modified DL (Dai-Liao) method is effective and robust in practical computations.

Moreover, this paper introduces an adaptive strategy for selecting the parameter in the Dai-Liao conjugate gradient method, derived using a Taylor series. As a result, we develop an improved extra Dai-Liao method conjugate gradient algorithm.

Table 2. Test results for *method Dai-Liao (BJ) with HS method*

P.No.	**N**	**HS**			**Dai-Liao (BJ)**		
		NI	NR	NF	NI	NR	NF
1	100	102	95	2709	12	7	27
	1000	14	8	32	14	8	31
2	100	19	10	35	19	10	36
	1000	39	22	67	38	22	67
3	100	34	18	72	34	18	72
	1000	35	19	77	35	19	78
4	100	9	6	25	9	6	25
	1000	92	84	2256	39	30	576
5	100	102	33	155	94	26	146
	1000	352	100	543	372	98	577
6	100	10	5	21	7	4	15
	1000	14	7	27	13	7	26
7	100	13	8	23	12	7	18
	1000	52	45	1225	10	7	73
8	100	42	17	62	42	18	61
	1000	67	26	102	72	31	114
9	100	8	6	17	8	6	17
	1000	26	25	505	7	5	15
10	100	9	7	23	9	7	22
	1000	37	35	844	15	11	34
11	100	83	52	182	79	50	174
	1000	79	50	171	79	50	171
12	100	36	13	59	35	10	58
	1000	37	14	59	36	16	56
13	100	477	130	746	373	103	596
	1000	F	F	F	F	F	F
14	1000	523	152	813	508	152	794
	1000	F	F	F	F	F	F
15	100	8	6	17	8	6	17
	1000	26	25	505	7	5	15
16	100	34	10	55	36	11	57
	1000	38	12	93	32	7	58

The aforementioned results underwent further scrutiny employing a performance profile tool as delineated by Dolan and More [12]. This instrument serves as a cumulative distribution function, delineating the probability that a specified methodology will successfully resolve a problem within a multiple of the optimal observed performance. The x-axis of the plotted curve specifies the performance ratio, whereas the y-axis signifies the proportion of test problems resolved within that ratio. Any algorithm whose curve is positioned above other curves is deemed superior, signifying that the algorithm resolves a greater proportion of functions with enhanced efficiency.

Figure 1 depicts the performance profile curve of the proposed algorithm in juxtaposition with the classical HS algorithm based on the NI metric, thereby elucidating the rate at which each algorithm converges. The curve illustrates that the proposed Dai-Liao (BJ) method surpassed the classical HS algorithm, as it achieved a higher cumulative performance more expediently. This observation denotes that the proposed algorithm necessitates fewer iterations to reach convergence across the majority of test functions. In contrast, the classical HS curve trails behind, indicating that it generally necessitates a greater number of iterations, which could adversely affect computational efficiency.

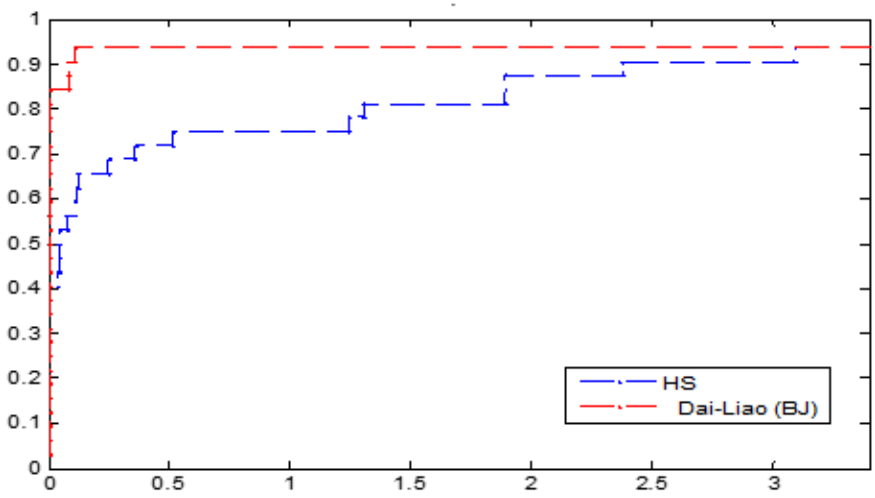

Figure 1. Performance profiles for Iter.

In a similar vein, by evaluating the second performance curve, which assessed the number of restarts required for convergence, the proposed algorithm further demonstrates its superior efficacy, resolving a larger proportion of problems with fewer restarts. The findings suggest that the classical HS algorithm consumes a more considerable duration to achieve the same cumulative fraction of resolved functions, implying that it may encounter greater challenges in sustaining stable convergence and necessitates more frequent restarts.

Finally, the results pertaining to function evaluations, as illustrated in Figure 3, evidenced a direct correlation with the computational cost of an optimization algorithm. Nonetheless, the curve further substantiates the efficiency of the proposed algorithm, as its trajectory (the red line) predominates over the green line, indicating that the proposed algorithm consistently requires fewer function evaluations to attain an optimal solution. Given that function evaluations frequently constitute the most resource-intensive component of an optimization process, this outcome underscores the computational superiority of the proposed Dai-Liao (BJ) method in comparison to the classical HS algorithm.

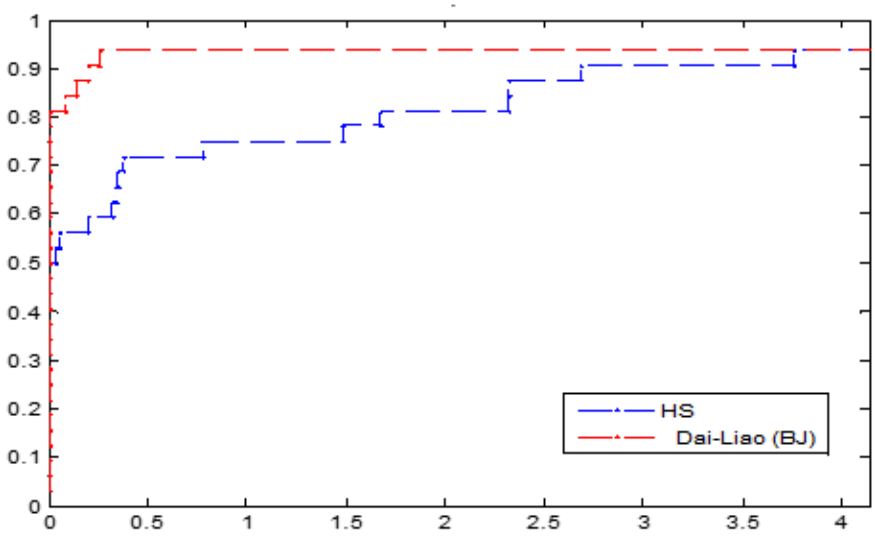

Figure 2. Performance profiles for Res.

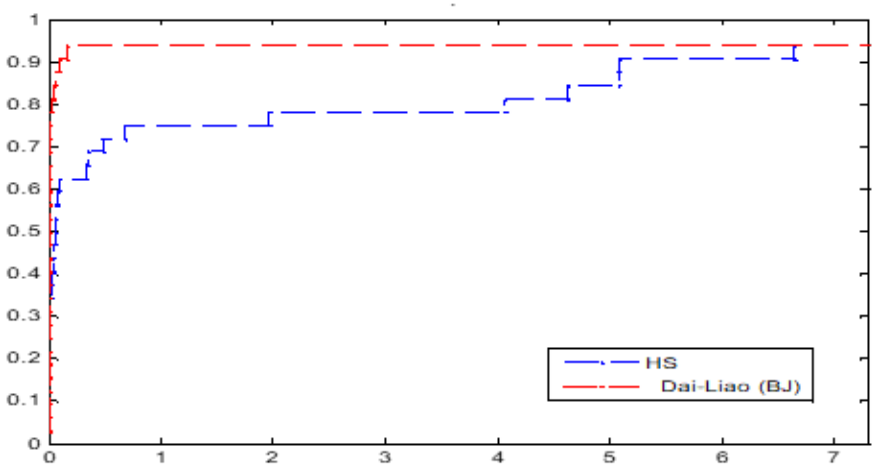

Figure 3. Performance profiles for Fval.

5. Conclusions

The Conjugate Gradient update formula is meticulously derived using an enhanced curvature condition. To ensure stability and efficacy, it is crucial to maintain conjugacy between search directions, which is essential for the process. By embedding the derivation within the Dai-Liao framework, a better theoretical understanding and advantageous convergence behavior are achieved. Due to its enhanced numerical stability and quicker convergence, the resulting update formula is well-suited for addressing large-scale problems.

REFERENCES

1. Andrie N. (2008). An Unconstrained Optimization Test Functions Collection. *Advanced Modeling and Optimization*, **10**, 147–161.
2. Basim A. Hassan and Abdulameer A. Saad. (2023). Elastic conjugate gradient methods to solve iteration problems. *Journal of Interdisciplinary Mathematics*, **26**(6), 1207–1217.
3. Basim A. Hassan, Hawraz N. Jabbar and Yoksal A. Laylani. (2023). Upscaling parameters for conjugate gradient method in unconstrained optimization. *Journal of Interdisciplinary Mathematics*, **26**(6), 1171–1180.
4. Hestenes, M. R. and Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of National Bureau of Standards*, **49**, 409–436.
5. Basim A. Hassan and Abdulameer A. Saad. (2023). Explaining new parameters conjugate analysis based on the quadratic model. *Journal of Interdisciplinary Mathematics*, **26**(6), 1219–1229.
6. Dai Y.H. and Liao L.Z. (2001). New conjugacy conditions and related conjugate gradient methods. *Applied Mathematics and Optimization*, **43**, 87–101.
7. Zhang L., Zhou W.J., and Li D.H. (2006). A descent modified Polak–Ribière–Polyak conjugate gradient method and its global convergence. *IMA Journal of Numerical Analysis*, **26**, 629–640.
8. Dai Y., Han J., Liu G., Sun D., Yin H., and Yan Y. (1999). Convergence properties of nonlinear conjugate methods. *SIAM Journal on Optimization*, **2**, 345–358.
9. Perry A. (1978). Technical Note—A modified conjugate gradient algorithm. *Applied Mathematics Operations Research*, **26**(6), 1073–1078.
10. Wolfe P. (1969). Convergence conditions for ascent methods. *SIAM Review*, **11**, 226–235.
11. Li G., Tang C., and Wei Z. (2007). New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. *Journal of Computational and Applied Mathematics*, **202**(2), 523–539.
12. Dolan E. and Moré J.J. (2002). Benchmarking optimization software with performance profiles. *Mathematical Programming*, **91**, 201–213.
13. Basim A. Hassan and Hameed M. Sadiq. (2022). Efficient New Conjugate Gradient Methods for Removing Impulse Noise Images. *European Journal of Pure and Applied Mathematics*, **15**(4), 2011–2021.
14. Basim A. Hassan and Ali Ahmed A. Abdullah. (2023). Improvement of conjugate gradient methods for removing impulse noise images. *Indonesian Journal of Electrical Engineering and Computer Science*, **29**(1), 245–251.
15. Basim A. Hassan and Haneen A. Alashoor. (2022). Pediment New Parameters for a Conjugate Gradient Method And Using it in Restoring Distorted Images. In: *8th International Conference on Contemporary Information Technology and Mathematics (ICCITM2022)*, Mosul University, Mosul-Iraq, pp. 385–390.
16. Basim A. Hassan and Haneen A. Alashoor. (2022). Involving New Coefficients Conjugate Gradient Method for Restoring Distorted Images. In: *8th International Conference on Contemporary Information Technology and Mathematics (ICCITM2022)*, Mosul University, Mosul-Iraq, pp. 380–384.
17. Basim A. Hassan, Issam A. R. Moghrabi and Ibrahim M. Sulaiman. (2023). New conjugate gradient image processing methods. *Asian-European Journal of Mathematics*, pp. 1–14. DOI: 10.1142/S1793557123500997.