

Robust Dai-Liao method for conjugate gradient method to solving iteration problems

Waleed Abdulazeez Majeed and Basim Abbas Hassan*

Department of Mathematics, College of Computers Sciences and Mathematics University of Mosul, Iraq.

Abstract

The examination of alternative methodologies for the formulation of parameter t constitutes a captivating area of inquiry; we propose that certain modified Dai-Liao methods employing distinct parameter t under the modified quasi–Newton equation merit exploration.

This paper presents an adaptive selection for the parameter of the Dai–Liao conjugate gradient method, derived through the modified quasi–Newton equation. Consequently, we introduce a modified Dai-Liao conjugate gradient method. A noteworthy characteristic of the proposed methodology is the incorporation of both gradient and function value information into the parameter t of the modified Dai-Liao conjugate gradient method. We establish the global convergence properties of the modified Dai-Liao conjugate gradient method under certain appropriate assumptions. Numerical results demonstrate that the modified Dai-Liao method is efficacious in practical computations.

Keywords. Robust Dai-Liao Method, Conjugate Gradient Method, Global convergence, Iteration Problems. 2010 Mathematics Subject Classification. 65K05, 65F10, 65J05.

1. Introduction

We employ the conjugate gradient methodology to formulate our solution to the unconstrained optimization problem:

$$Minf(x)$$
 , $x \in \mathbb{R}^n$, (1.1)

where f is a smooth function, [5]. We commence by reviewing the conjugate gradient method (CG), which generates a sequence of iterates through repetitive application of:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where α_k denotes the step size and d_k signifies the search direction, which is defined as follows:

$$d_{k+1} = -g_{k+1} + \beta_k s_k, \tag{1.3}$$

where the conjugate gradient methodology is characterized by the parameter β_k , refer to [6, 10]. Although a portion of these methodologies possesses theoretical significance, others demonstrate effectiveness in numerical applications. The nonlinear conjugate gradient method was initially articulated by Hestenes-Stiefel [13], in which β_k is expressed in the subsequent manner:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}. (1.4)$$

Received: 31 August 2025; Accepted: 27 October 2025.

Step length α_k is traditionally selected in iterative methodologies to satisfy specific criteria. The Wolfe conditions are routinely utilized in the analysis of convergence and the implementation of conjugate gradient techniques among the conditions delineated, and they must conform to the following stipulations:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k, \tag{1.5}$$

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma \ d_k^T g_k,\tag{1.6}$$

where $0 < \delta < \sigma < 1$, see [16]. It is regrettable that in scenarios where the objective function is nonconvex and the conventional Wolfe line search is utilized, none of the proposed methodologies can converge simultaneously.

One of the well-known conjugate gradient techniques created to enhance the efficiency of unconstrained nonlinear optimization algorithms is the Dai-Liao approach, [2]. The drawbacks of conventional techniques, which don't always provide enough descent in the objective function value, prompted researchers Yuan-Hai Dai and Li-Qun Liao to suggest it in 2001. In order to improve numerical stability and guarantee global convergence, this approach uses a novel formula to calculate the beta coefficient, β_k . The following ratio is used to calculate β_k in the Dai-Liao method:

$$\beta_k^{DL} = \frac{g_{k+1}^T y_k - t g_{k+1}^T s_k}{d_k^T y_k}.$$
(1.7)

Using this method eliminates the requirement for restart approaches while keeping an efficient fall direction. The fact that this approach ensures a substantial drop at each iteration and achieves notable convergence even when other approaches encounter challenges is one of its main advantages, [9]. According to numerical research, the Dai–Liao approach is more stable and efficient than other conjugate gradient techniques, which makes it appropriate for a variety of optimisation applications.

To augment the precision of the Conjugate Gradient (CG) method, scholars have integrated the modified secant (quasi-Newton) equations into their analyses, [7, 11]. Consequently, numerous revised iterations of the Dai-Liao (DL) method have been proposed within the academic literature, which leverage both the values of the cost function and the gradient information, or demonstrate global convergence independent of convexity prerequisites, [8, 12]. Later, numerous conjugate gradient algorithms grounded on various modified secant equations were examined, as noted in [17].

In pursuit of achieving both global convergence utilizing the conventional Wolfe line search and the requisite descent property for any line search, we introduce modifications to the conjugate gradient approach. Furthermore, a novel framework is presented alongside the underlying inspiration.

2. Robust Dai-Liao Method

The efficient convergence of the Conjugate Gradient method is contingent upon the maintenance of conjugacy among the search directions. A fundamental mathematical principle that facilitates this preservation is known as the curvature condition. Basim et al. [1] have proposed an enhanced iteration of this condition, which explicitly ensures satisfactory performance, articulated as follows:

$$s_k^T B_{k+1} s_k \cong 1/2 s_k^T y_k + (f_{k+1} - f_k) - g_k^T s_k.$$
 (2.1)

This condition is instrumental in the derivation of the Conjugate Gradient update formula, which subsequently enhances numerical stability and accelerates convergence. Initiating from Equation (2.1), we reformulate it algebraically as follows:

$$d_{k+1}^T B_{k+1} s_k \cong [1/2 + ((f_{k+1} - f_k) - g_k^T s_k)/s_k^T y_k] dk + 1Tyk. \tag{2.2}$$

In [15], by applying Perry's conjugacy condition $(d_{k+1}^T y_k = -g_{k+1}^T s_k)$ in conjunction with Equation (2.1), we derive:

$$d_{k+1}^T B_{k+1} s_k \cong -[1/2 + ((f_{k+1} - f_k) - g_k^T s_k)/s_k^T y_k] g_k + 1T s_k.$$
(2.3)

Perry's condition underpins stable and efficient updates within the Conjugate Gradient method. By utilizing the search direction articulated in 1.1, we attain:

$$(-g_{k+1} + \beta_k s_k)^T y_k \cong -\left[\frac{1}{2} + \frac{(f_{k+1} - f_k) + g_k^T s_k}{s_k^T y_k}\right] g_{k+1}^T s_k.$$
(2.4)



This results in:

$$\beta_k = \frac{g_{k+1}^T y_k - \left[\frac{1}{2} + \frac{(f_{k+1} - f_k) + g_k^T s_k}{s_k^T y_k}\right] g_{k+1}^T s_k}{s_k^T y_k}.$$
(2.5)

Within the framework established by Dai and Liao, it can be articulated as follows:

$$\beta_k = \frac{g_{k+1}^T y_k - t g_{k+1}^T s_k}{s_k^T y_k},\tag{2.6}$$

with

$$t_k^{BK} = \frac{1}{2} + \frac{(f_{k+1} - f_k) + g_k^T s_k}{s_k^T y_k}, \tag{2.7}$$

The new method is called Dai-Liao (BK).

The method is recognized as this derivation emphasizes the significance of the refined curvature condition in maintaining conjugacy and refining the update formula. Contextualizing the result within the Dai-Liao framework bolsters both theoretical comprehension and practical convergence in the conjugate gradient method.

Algorithm: Dai-Liao (BK).

Input: Initial approximation establish a minimal $x_0 \in \mathbb{R}^n$ Tolerance ε .

Output: Minimizer x such that the gradient is approximately nullified.

- (1) Execute the computation $g_0 = \nabla f_0$, establish parameters $d_0 = -g_0$, and initialize values k=0.
- (2) While the gradient's norm $||g_{k+1}|| > 0$, then:
 - (a) Conduct a line search to ascertain the step size α_k satisfying the Wolfe conditions.

 - (a) Conduct a line scarch to describe the European She scarch and She s

 - (e) Increment the iteration index k + 1.
- (3) Conclude the while loop.
- (4) Return the ultimate iterate x_{k+1} .

Theorem 2.1. The new research d_{k+1} generated using the new formula $\beta_k^{\text{Dai-Liao (BA)}}$ achieves the regression property.

Proof. Multiplying both sides of the search direction by g_{k+1}^T and substituting formula $\beta_k^{\text{Dai-Liao (BK)}}$, this gives us:

$$d_{k+1}^{T}g_{k+1} = -\|g_{k+1}\|^{2} + \frac{g_{k+1}^{T}y_{k} - \left[\frac{1}{2} + \frac{(f_{k+1} - f_{k}) + g_{k}^{T}s_{k}}{s_{k}^{T}y_{k}}\right]g_{k+1}^{T}s_{k}}{s_{k}^{T}y_{k}}s_{k}^{T}g_{k+1}.$$

$$(2.8)$$

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{s_k^T g_{k+1}}{s_k^T y_k}\right] s_k^T g_{k+1}. \tag{2.9}$$

Equation (2.9) produces:

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k s_k^T g_{k+1}}{s_k^T y_k} - \frac{(s_k^T g_{k+1})^2}{s_k^T y_k}.$$
(2.10)

Taking advantage of Cauchy's inequality $w^T v \leq \frac{1}{2}(\|w\|^2 + \|v\|^2)$ and for simplicity, we define $w = (y_k^T s_k)g_{k+1}$ and $v = (s_k^T g_{k+1}) y_k$. This gives us:

$$\frac{g_{k+1}^T y_k s_k^T g_{k+1}}{s_k^T y_k} \le \frac{\frac{1}{2} [\|g_{k+1}\|^2 (y_k^T s_k)^2 + (s_k^T g_{k+1})^2 \|y_k\|^2]}{(s_k^T y_k)^2}.$$
(2.11)



By combining inequality 2.11 with Equation (2.10) to become:

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 \frac{\frac{1}{2} [\|g_{k+1}\|^2 (y_k^T s_k)^2 + (s_k^T g_{k+1})^2 \|y_k\|^2]}{(s_k^T y_k)^2} - \frac{(s_k^T g_{k+1})^2}{s_k^T y_k}. \tag{2.12}$$

According to Bechitz's condition that it be $y_k^T y_k \leq L s_k^T y_k$ and by substituting it with the inequality 2.12, we have:

$$d_{k+1}^T g_{k+1} \le -\frac{1}{2} \|g_{k+1}\|^2 + \left[\frac{1}{2}L - 1\right] \frac{(s_k^T g_{k+1})^2}{s_k^T y_k} \le 0.$$
(2.13)

This completes the proof.

3. Prerequisites required for establishing convergence

Convexity of the level set: The collection of points delineated by must be identified as a convex set:

$$L_0 = \{x/f(x) \le f(x_0)\}. \tag{3.1}$$

Lipschitz continuity of the gradient: The gradient is required to exhibit Lipschitz continuity, which implies the existence of a constant such that:

$$(\nabla f(o^{-}) - \nabla f(v^{+})) \le L \|o^{-} - v^{+}\| , \forall o^{-}, v^{+} \in L_{0}.$$
(3.2)

In light of these two prerequisites, it can be demonstrated that there exists a positive constant $\Pi > 0$ such that:

$$||g_{k+1}||^2 \le \Pi. \tag{3.3}$$

See [14].

The theorem proposed by Dai et al. [3], plays a key role in analyzing the convergence behavior of optimization algorithms like gradient descent and the conjugate gradient method. In essence, the theorem states that:

$$\sum_{k>0} \frac{1}{\|d_{k+1}\|^2} = \infty,\tag{3.4}$$

This means that:

$$\liminf_{k \to \infty} ||g_{k+1}|| = 0.$$
(3.5)

Theorem 3.1. Let the sequence $\{x_k\}$ be generated by Dai-Liao (BK), and suppose the Assumptions holds, If:

$$(\nabla f(x) - \nabla f(y))^{T} (x - y) \ge \Phi \|y - y\|^{2}. \tag{3.6}$$

Then:

$$\liminf_{k \to \infty} \|g_k\| = 0.$$
(3.7)

Proof. The parameter delineated by Dai–Liao (BK) satisfies:

$$|\beta_k| \le \frac{\left| g_{k+1}^T y_k \right| + \left| s_k^T g_{k+1} \right|}{s_k^T y_k}. \tag{3.8}$$

Utilizing inequality 3.6 and the Lipschitz continuity property of ∇f , it follows that:

$$|\beta_k| \le \frac{L\|g_{k+1}\| \|s_k\| + \|g_{k+1}\| \|s_k\|}{\Phi \|s_k\|^2}.$$
(3.9)

The norm of the search direction update is given by:

$$||d_{k+1}|| = || - g_{k+1} + \beta_k d_k||. (3.10)$$

Plugging 3.9 and 3.3 into 3.10 yields:

$$||d_{k+1}|| \le \Pi + \frac{\Pi L ||s_k|| + \Pi ||s_k||}{\Phi ||s_k||^2} ||s_k|| \le (1 + C)\Pi, \tag{3.11}$$



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

Table 1. Table of problems, starting points and dimensions.

where $C = (L+1)/\Phi$. As a result, we have:

$$\sum_{k>1} \frac{1}{\|d_{k+1}\|^2} \ge \left(\frac{\Phi}{\Phi + L + 1}\right) \frac{1}{\Pi} \sum_{k>1} 1 = \infty.$$
(3.12)

Based on the equations 3.4 and 3.5, we obtained:

$$\liminf_{k \to \infty} \|g_k\| = 0.$$
(3.13)

4. Numerical Results

We test the effect of new algorithm on numerical problems, and the unconstrained optimization problems, [1] are listed in Table 1. We will report on various numerical implementations variable counts of n=100 and 1000 of our presented method (BK algorithm) with the modified HS on these problems and compare its performance with the HS algorithm for the problems. In this part of the numerical experiment, we introduce the stopping rules, dimensions, and some parameters as follows, $\delta_1 = 0.001$ and $\delta_2 = 0.9$, with the termination criterion as defined by $||g_{k+1}|| \le 10^{-6}$. The comparison data can be shown as follows: NI: number of iterations. IS: number of restarts. NF: Number of function evolution.

One adaptive choice for the parameter of the Dai–Liao conjugate gradient method is suggested in this paper, which is obtained with modified quasi-Newton equation. So we get a modified Dai-Liao conjugate gradient method.

The aforementioned results were subjected to further scrutiny utilizing a performance profile instrument as delineated by Dolan and More [4]. This analytical tool operates as a cumulative distribution function, depicting the likelihood that a specified methodology will proficiently resolve a problem within multiple iterations of the optimal performance recorded. The x-axis of the resultant curve illustrates the performance ratio, while the y-axis signifies the proportion of test problems resolved within that ratio. Any algorithm whose curve is positioned above those of its competitors is deemed superior, indicating that the algorithm tackles a larger percentage of functions with enhanced efficiency.

Figure 1 illustrates the performance profile curve of the proposed algorithm in juxtaposition with the classical HS algorithm based on the NI metric, thereby clarifying the convergence rate of each algorithm. The curve reveals that the proposed Dai-Liao (BK) method surpassed the classical HS algorithm, as it achieved a higher cumulative



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

P.No.	N		HS			Dai-Liao (BK)		
	11	NI	NR	NF	NI	NR	NF	
1	100	102	95	2709	F	F	F	
	1000	14	8	32	12	7	27	
2	100	19	10	35	18	10	34	
	1000	39	22	67	35	20	62	
3	100	34	18	72	33	17	69	
	1000	35	19	77	35	19	78	
6	100	9	6	25	9	6	25	
	1000	92	84	2256	55	47	993	
7	100	102	33	155	97	26	151	
	1000	352	100	543	335	86	526	
15	100	10	5	21	10	5	21	
	1000	14	7	27	13	7	26	
16	100	13	8	23	12	7	18	
	1000	52	45	1225	39	36	974	
17	100	42	17	62	40	15	61	
	1000	67	26	102	60	23	94	
22	100	8	6	17		6	17	
	1000	26	25	505	8	5	15	
26	100	9	7	23	15	11	53	
	1000	37	35	844	46	42	1070	
29	100	83	52	182	79	50	174	
	1000	79	50	171	79	50	171	
35	100	36	13	59	37	13	60	
	1000	37	14	59	44	24	73	
53	100	477	130	746	435	133	681	
	1000	\mathbf{F}	F	F	\mathbf{F}	\mathbf{F}	\mathbf{F}	
56	100	523	152	813	468	131	744	
	1000	\mathbf{F}	F	\mathbf{F}	\mathbf{F}	\mathbf{F}	\mathbf{F}	
69	100	8	6	17	8	6	17	
	1000	26	25	505	7	5	15	
70	100	34	10	55	31	8	53	
	1000	38	12	93	32	8	57	

performance in a more expeditious manner. This finding suggests that the proposed algorithm necessitates fewer iterations to attain convergence across the majority of test functions. In contrast, the classical HS curve is observed to lag, implying that it typically requires a larger number of iterations, which may adversely affect computational efficiency. In a similar context, the examination of the secondary performance curve, which assessed the number of restarts requisite for convergence, further substantiates the proposed algorithm's enhanced efficacy, as it successfully resolves a greater proportion of problems with fewer restarts. The findings suggest that the classical HS algorithm consumes a substantially longer duration to achieve the same cumulative fraction of resolved functions, indicating that it may encounter increased challenges in sustaining stable convergence and necessitates more frequent restarts. Lastly, the findings related to function evaluations, as illustrated in Figure 3, exhibited a direct relationship with the computational cost associated with an optimization algorithm. Nevertheless, the curve further reinforces the efficiency of the proposed algorithm, as its trajectory (the red line) consistently exceeds the green line, signifying



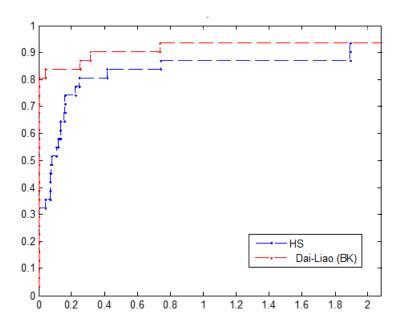


FIGURE 1. Performance profiles for iter.

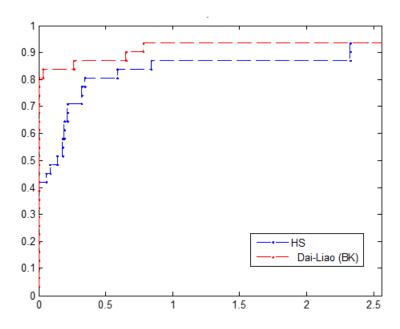


FIGURE 2. Performance profiles for Res.



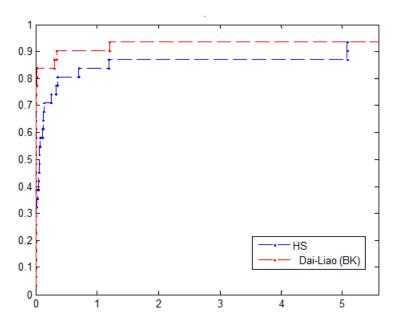


FIGURE 3. Performance profiles for Fval.

that the proposed algorithm reliably demands fewer function evaluations to reach an optimal solution. Considering that function evaluations frequently constitute the most resource-intensive component of an optimization process, this outcome accentuates the computational superiority of the proposed Dai-Liao (BK) method in comparison to the classical HS algorithm.

5. Conclusions

In this work, we derive a new conjugate gradient technique of the Dai–Liao type for unconstrained optimization problems. The suggested techniques add a new secant equation to the conjugacy condition of Dai and Liao that was previously used in [2]. When using the strong Wolfe line search, we demonstrate the techniques' global convergence under specific assumptions. The suggested techniques can exceed a number of current approaches in terms of resilience and efficiency, as shown by numerical trials.

References

- [1] N. Andrie, An Unconstrained Optimization Test functions collection, Advanced Modeling and Optimization, 10 (2008), 147–161.
- [2] Y. H. Dai and L. Z. Liao, New conjugacy conditions and related conjugate gradient methods, Appl. Math. Optim., 43 (2001), 87–101.
- [3] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin and Y. Yan, Convergence properties of nonlinear conjugate methods, SIAM J. Optim., 2 (1999), 345–358.
- [4] E. Dolan and J. J. Moré, Benchmarking optimization software with performance profiles, Math. Program., 91 (2002), 201–213.
- [5] B. A. Hassan and A. A. Saad, Elastic conjugate gradient methods to solve iteration problems, J. Interdiscip. Math., 26(6) (2023), 1207–1217.
- [6] B. A. Hassan and A. A. Saad, Explaining new parameters conjugate analysis based on the quadratic model, J. Interdiscip. Math., 26(6) (2023), 1219–1229.



REFERENCES

[7] B. A. Hassan and A. A. A. Abdullah, Improvement of conjugate gradient methods for removing impulse noise images, Indones. J. Electr. Eng. Comput. Sci., 29(1) (2023), 245–251.

- [8] B. A. Hassan and H. A. Alashoor, *Pediment new parameters for a conjugate gradient method and using it in restoring distorted images*, Proc. 8th Int. Conf. Contemporary Information Technology and Mathematics (ICCITM 2022), Univ. Mosul, Mosul, Iraq, (2022), 385–390.
- [9] B. A. Hassan and H. A. Alashoor, Involving new coefficients conjugate gradient method for restoring distorted images, Proc. 8th Int. Conf. Contemporary Information Technology and Mathematics (ICCITM 2022), Univ. Mosul, Mosul, Iraq, (2022), 380–384.
- [10] B. A. Hassan, H. N. Jabbar, and Y. A. Laylani, Upscaling parameters for conjugate gradient method in unconstrained optimization, J. Interdiscip. Math., 26(6) (2023), 1171–1180.
- [11] B. A. Hassan and H. M. Sadiq, Efficient new conjugate gradient methods for removing impulse noise images, Eur. J. Pure Appl. Math., 15(4) (2022), 2011–2021.
- [12] B. A. Hassan, I. A. R. Moghrabi, and I. M. Sulaiman, New conjugate gradient image processing methods, Asian-Eur. J. Math., (2023), 1–14.
- [13] M. R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Nat. Bur. Standards, 49 (1952), 409–436.
- [14] G. Li, C. Tang, and Z. Wei, New conjugacy condition and related new conjugate gradient methods for unconstrained optimization, J. Comput. Appl. Math., 202(2) (2007), 523–539.
- [15] A. Perry, Technical Note A modified conjugate gradient algorithm, Appl. Math. Oper. Res., 26(6) (1978), 1073–1078.
- [16] P. Wolfe, Convergence conditions for ascent methods, SIAM Rev., 11 (1969), 226–235.
- [17] L. Zhang, W. J. Zhou, and D. H. Li, A descent modified Polak–Ribière–Polyak conjugate gradient method and its global convergence, IMA J. Numer. Anal., 26 (2006), 629–640.

