Hindawi
Journal of Applied Mathematics
Volume 2019, Article ID 6378368, 6 pages
http://dx.doi.org/10.1155/2019/6378368

Research Article

Spectral Three-Term Constrained Conjugate Gradient Algorithm for Function Minimizations

Huda I. Ahmed[1] **Rana Z. Al-Kawaz,**[2] **and Abbas Y. Al-Bayati**[3]

[1] *Department of Operations Researches and Intelligent Techniques, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq*
[2] *Department of Mathematics, College of Basic Education, University of Telafer, Mosul, Iraq*
[3] *University of Telafer, Mosul, Iraq*

Correspondence should be addressed to Huda I. Ahmed; hudaea72@gmail.com

Received 20 May 2019; Revised 8 August 2019; Accepted 4 September 2019; Published 25 December 2019

Academic Editor: Xiaohui Yuan

In this work, we tend to deal within the field of the constrained optimization methods of three-term Conjugate Gradient (CG) technique which is primarily based on Dai–Liao (DL) formula. The new proposed technique satisfies the conjugacy property and the descent conditions of Karush–Kuhn–Tucker (K.K.T.). Our planned constrained technique uses the robust Wolfe line search condition with some assumptions. We tend to prove the global convergence property of the new planned technique. Numeral comparisons for (30-thirty) constrained optimization issues make sure the effectiveness of the new planned formula.

1. Introduction

All strategies for constrained problems will be classified into (2) basic categories; specifically, direct and indirect ways. Generating uncontained sub-problem ways for the later kind square measure vital even for a few special optimization interior and exterior penalty function techniques transform the constrained problem into unconstrained optimization type problems. The technique in the main easy and quite sturdy for a previous technique known as Sequential Unconstrained Minimization Technique (SUMT). The essential optimization problem with inequality constrained of this way outlined as

$$\min f(x) \text{ s.t. } \quad c_i(x) \le 0 \quad i = 1, \cdots m. \tag{1}$$

This problem is regenerate into unconstrained minimization technique by constructing a function of the shape

$$\varphi(x, \mu) = f(x) + \mu B(x), \tag{2}$$

where $\mu \to 0$ and $B(x)$ is defined by [1]:

$$B(x) = \sum_{j=1}^{m} \frac{1}{c_j(x)}. \tag{3}$$

Therefore, we can rewrite the Equation (2) as follows

$$\varphi(x, \mu) = f(x) + \mu \sum_{j=1}^{m} \frac{1}{c_j(x)}. \tag{4}$$

The derivatives of this functions are $\nabla f(x)$ and $\nabla c_i(x)$, for $i = 1, \cdots, n$ are linear independent, so that

$$\nabla\varphi(x, \mu) = \nabla f(x) + \mu \sum_{j=1}^{m} \left(\frac{-1}{c_j^2(x)}\right) \nabla c_j(x). \tag{5}$$

Now we turn to the second part parallel to the importance of the previous part, which is unconstrained optimization technique and let us know the problem (2), where $\varphi: R^n \to R$ is a real-valued continuous and scalable derivation function. The iterative is

$$x_{k+1} = x_k + \alpha_k d_k, \tag{6}$$

whereas α_k is step-length. The new search direction d_{k+1} is:

$$d_{k+1} = \begin{cases} -\nabla\varphi(x_{k+1}, \mu_{k+1}) & \text{for} \quad \mathrm{k} = 0, \\ -\nabla\varphi(x_{k+1}, \mu_{k+1}) + \beta_k d_k & \text{for} \quad k \ge 1. \end{cases} \tag{7}$$

The value of the derivative function at the current point is $g(x_{k+1}) = \nabla\varphi(x_{k+1}, \mu_{k+1})$ and β_k is a positive scalar called the conjugate gradient parameter.

There are some known formulas of β_k are from Hestenes–Stiefel (HS) [2], Fletcher–Reeves (FR) [3], Polak–Ribière (PR) [4], Liu–Storey (LS) [5] and Dai–Liao (DL) [6].

In the existing convergence analysis and implementation of the CG technique, the weak Wolfe condition [7] are defined as:

$$\varphi(x_{k+1}, \mu_{k+1}) - \varphi(x_k, \mu_k) \leq \delta\alpha_k \nabla\varphi(x_k, \mu_k)^T d_k, \quad (8)$$

$$\nabla\varphi(x_{k+1}, \mu_{k+1})^T d_k \geq \sigma\nabla\varphi(x_k, \mu_k)^T d_k, \quad (9)$$

and $0 < \delta < \sigma < 1$.

By updating one of the conditions also strong Wolfe conditions [7] consist of (8) and

$$\left|\nabla\varphi(x_{k+1}, \mu_{k+1})^T d_k\right| \leq -\sigma \nabla\varphi(x_k, \mu_k)^T d_k. \quad (10)$$

Furthermore, the sufficient descent property, namely

$$d_{k+1}^T \nabla\varphi(x_{k+1}, \mu_{k+1}) \leq -c\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2. \quad (11)$$

The constant c is a positive number that satisfies the descent condition.

2. A Modified Dai–Liao Three-Term CG Technique

Many researchers have provided different updates which are suitable for the parameter of Dai–Liao (DL) CG-method consisting of:

$$\beta_k^{DL} = \frac{g_{k+1}^T(y_k - ts_k)}{s_k^T y_k}. \quad (12)$$

Recall the work of Liveries and Pintelas [8] which they forward a new update to the parameter β_k^{DL} which was based on the modified secant equation and they replaced y_k with this new one. Other researchers, e.g. Babaie-Kafaki and Ghanbari [9] present in their work a derivation of two modified CG-methods which are based on Perry's work; they got better numerical results than the original one given by DL. The researchers continued various updates of the DL-parameter in order to obtain some suitable formulas. See for example [10–12]. Moreover, the researcher's Zhang et al. [13] prompt a three-term CG-technique based mostly of the DL-technique as follows:

$$d_{k+1} = -g_{k+1} + \frac{g_{k+1}^T(y_k - ts_k)}{s_k^T y_k}s_k - \frac{g_{k+1}^T d_k}{s_k^T y_k}(y_k - ts_k), t > 0. \quad (13)$$

This direction satisfying the condition ($d_{k+1}^T g_{k+1} \leq -c_1\|g_{k+1}\|^2$) for all k. Now, exploitation (13) within the constrained CG-technique outlined in (1)–(4) yields

$$\begin{aligned} d_{k+1} = &- \nabla\varphi(x_{k+1}, \mu_{k+1}) + \frac{\nabla\varphi(x_{k+1}, \mu_{k+1})^T(y_k - ts_k)}{s_k^T y_k}s_k \\ &- \frac{s_k^T\nabla\varphi(x_{k+1}, \mu_{k+1})}{s_k^T y_k}(y_k - ts_k). \end{aligned} \quad (14)$$

By updating this formula using the modified techniques of Dai–Liao CG in (14), we obtain:

$$\begin{aligned} d_{k+1} = &-\frac{s_k^T(y_k - ts_k)}{s_k^T y_k}\nabla\varphi(x_{k+1}, \mu_{k+1}) \\ &+ \frac{\nabla\varphi(x_{k+1}, \mu_{k+1})^T(y_k - ts_k)}{s_k^T y_k}s_k \\ &- \frac{s_k^T\nabla\varphi(x_{k+1}, \mu_{k+1})}{s_k^T y_k}(y_k - ts_k). \end{aligned} \quad (15)$$

When rewriting the new search direction this is as follows:

$$d_{k+1} = -Q_{k+1}\nabla\varphi(x_{k+1}, \mu_{k+1}), \quad (16)$$

where

$$Q_{k+1} = \frac{1}{s_k^T y_k}\left[s_k^T(y_k - ts_k).I + \left(y_k s_k^T - s_k y_k^T\right)\right]. \quad (17)$$

Since $y_k^T s_k>0$ (by the strong Wolfe condition), through these inequality and Quasi Newton condition we get:

$$Q_{k+1}s_k = y_k \Rightarrow s_k^T Q_{k+1}s_k > 0, \quad (18)$$

this means that Q_{k+1} is a positive definite matrix.

3. New Theorem

The new direction d_{k+1} in (15) satisfying the sufficiently descent condition (11).

Proof. Now multiply each side of (15) by $\nabla\varphi(x_{k+1}, \mu_{k+1})$ that capable for unconstrained optimization then we have a tendency to get

$$\begin{aligned} &\nabla\varphi(x_{k+1}, \mu_{k+1})^T d_{k+1} \\ &\quad = -\frac{(y_k - ts_k)^T s_k}{s_k^T y_k}\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2 \\ &\qquad + \frac{(y_k - ts_k)^T\nabla\varphi(x_{k+1}, \mu_{k+1})}{s_k^T y_k}\nabla\varphi(x_{k+1}, \mu_{k+1})s_k \\ &\qquad - \frac{s_k^T\nabla\varphi(x_{k+1}, \mu_{k+1})}{s_k^T y_k}\nabla\varphi(x_{k+1}, \mu_{k+1})(y_k - ts_k) \\ &\quad = -\left[\frac{y_k^T s_k}{y_k^T s_k}\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2 - t\frac{\|s_k\|^2}{y_k^T s_k}\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2\right] \\ &\quad = \left[-1 + t\frac{\|s_k\|^2}{y_k^T s_k}\right]\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2. \end{aligned} \quad (19)$$

Let $s_k = \alpha_k d_k$

$$\begin{aligned} \nabla\varphi(x_{k+1}, \mu_{k+1})^T d_{k+1} &= \left[-1 + t\frac{\alpha_k^2\|d_k\|^2}{\alpha_k y_k^T d_k}\right]\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2 \\ &= \left[-1 + t\alpha_k\frac{\|d_k\|^2}{y_k^T d_k}\right]\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2. \end{aligned} \quad (20)$$

The scalar β_k^{DL} is known, this means ($d_k = -g_k$). Moreover, when multiply the other end of the direction by y_k we get:

$$\begin{aligned} y_k^T d_k &= \nabla\varphi(x_{k+1}, \mu_{k+1})^T d_k - \nabla\varphi(x_k, \mu_k)^T d_k \\ &= \nabla\varphi(x_{k+1}, \mu_{k+1})^T d_k + \|d_k\|^2 \geq \|d_k\|^2 \end{aligned} \quad (21)$$

$$\begin{aligned} \nabla\phi(x_{k+1}, \mu_{k+1})^T d_{k+1} &\leq \left[-1 + t\alpha_k \frac{\|d_k\|^2}{\|d_k\|^2}\right] \|\nabla\phi(x_{k+1}, \mu_{k+1})\|^2 \\ &\leq [-1 + t\,\alpha_k] \|\nabla\phi(x_{k+1}, \mu_{k+1})\|^2. \end{aligned} \quad (22)$$

Where $c = -(1 - t\alpha_k)$ (is a positive constant). Now, we have

$$\begin{aligned} &\nabla\varphi(x_{k+1}, \mu_{k+1})^T d_{k+1} \\ &\quad \leq -c\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|^2 \\ &\quad \left(\nabla f(x_{k+1}) - \mu_{k+1} \sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1})\right)^T d_{k+1} \\ &\quad \leq -c\left[\nabla f(x_{k+1}) + \mu_{k+1} \sum_{j=1}^{m} \left(\frac{-1}{c_j^2(x_{k+1})}\right) \nabla c_j(x_{k+1})\right]^2 \\ &\quad \left(\nabla f(x_{k+1}) - \mu_{k+1} \sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1})\right)^T d_{k+1} \\ &\quad \leq -c\Bigg(\nabla f(x_{k+1})^T \nabla f(x_{k+1}) - 2\mu_{k+1} \nabla f(x_{k+1})^T \sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1}) \\ &\quad + \mu_{k+1}^2 \left(\sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1})^T \sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1})\right)\Bigg), \end{aligned} \quad (23)$$

which can be written differently

$$\begin{aligned} &\left(\nabla f(x_{k+1}) - \mu_{k+1} \sum_{j=1}^{m} \frac{1}{c_j^2(x_{k+1})} \nabla c_j(x_{k+1})\right)^T d_{k+1} \\ &\quad \leq -c[\nabla f(x_{k+1}) + \mu_{k+1} \nabla B(x_{k+1})]^2, \end{aligned} \quad (24)$$

and $B(x_{k+1})$ is the Barrier function at point $k+1$.

Then depending on one of Karush–Kuhn–Tucker, KKT's [14] optimal conditions and some regularity conditions of [15] such that if $\nabla c(x)^T d \leq 0$ as well as formula

$$\mu_{k+1} = \frac{\mu_k}{10}, \mu_0 > 0. \quad (25)$$

So, when $\mu_{k+1} \to 0$ and in order to get a min of the function $f(x)$ we take the limit for the function $\varphi(x, \mu)$ when $\mu \to 0$, in form:

$$\nabla f(x_{k+1})^T d_{k+1} \leq -c\|\nabla f(x_{k+1})\|^2, \quad (26)$$

we get the required, a sufficient descent direction of our new algorithm. □

Lemma 1 [16]. *The new direction d_{k+1} defined in (15) is satisfying the conjugacy condition.*

Proof. Let $\tilde{y} = (y_k - ts_k)$

$$\begin{aligned} \tilde{y}_k^T d_{k+1} = &-\frac{\tilde{y}_k^T s_k}{y_k^T s_k} \tilde{y}_k^T \nabla\varphi(x_{k+1}, \mu_{k+1}) + \frac{\tilde{y}_k^T \nabla\varphi(x_{k+1}, \mu_{k+1})}{y_k^T s_k} \tilde{y}_k^T s_k \\ &- \frac{s_k^T \nabla\varphi(x_{k+1}, \mu_{k+1})}{y_k^T s_k} \tilde{y}_k^T y_k. \end{aligned} \quad (27)$$

$$\tilde{y}_k^T d_{k+1} = -\frac{\|\tilde{y}\|^2}{y_k^T s_k} s_k^T \nabla\varphi(x_{k+1}, \mu_{k+1}) = -\phi s_k^T \nabla\varphi(x_{k+1}, \mu_{k+1}), \quad (28)$$

where $\phi > 0$ and this condition is equal to $\left(y_k^T d_{k+1} = -ts_k^T g_{k+1}\right)$ where

$$g_{k+1} = \nabla\varphi(x_{k+1}, \mu_{k+1}). \quad (29)$$

□

4. Global Convergence Property

In this part of the article, we will address the convergence analysis of the new algorithm where the following assumptions are often used in CG techniques.

Assumptions [17] and [18].

(i) Let the level set $S = \{x : \phi(x, \mu) \leq \phi(x_0, \mu)\}$ bounded i.e., there exists a constant $q > 0$ such that

$$\|x\| \leq q, \ \forall x \in S. \quad (30)$$

(ii) Clearly there is some neighborhood N of S, the function f is continuously differentiable, and its gradient is Lipschitz continuous, i.e. there exists a constant $L > 0$ such that

$$\|\nabla\phi(x, \mu) - \nabla\phi(y, \mu)\| \leq L\|x - y\|, \ \forall x, y \in N. \quad (31)$$

Assuming that conditions (i) and (ii) are satisfy, we can deduce that there exists a constant $\gamma > 0$ such that

$$\|\nabla\phi(x, \mu)\| \leq \gamma. \quad (32)$$

6. Global Convergence for New Theorem

Consider the new three-term CG-technique (15) which is satisfying (13) and assume that the step-size α_k satisfies (8) and (10) then

$$\lim_{k\to\infty} \|\nabla\varphi(x_{k+1}, \mu_{k+1})\| = 0. \quad (33)$$

Proof. The new search direction is:

$$\begin{aligned} \|d_{k+1}\| \leq &\frac{|y_k - ts_k|\|s_k\|}{\|y_k\|\|s_k\|} \|\nabla\varphi(x_{k+1}, \mu_{k+1})\| \\ &+ \frac{|y_k - ts_k|\|s_k\|}{\|y_k\|\|s_k\|} \|\nabla\varphi(x_{k+1}, \mu_{k+1})\| \\ &+ \frac{\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|\|s_k\||y_k - ts_k|}{\|y_k\|\|s_k\|} \\ \leq &(\|y_k\| + t\|s_k\|)\left(\frac{3\|\nabla\varphi(x_{k+1}, \mu_{k+1})\|}{\|y_k\|}\right). \end{aligned} \quad (34)$$

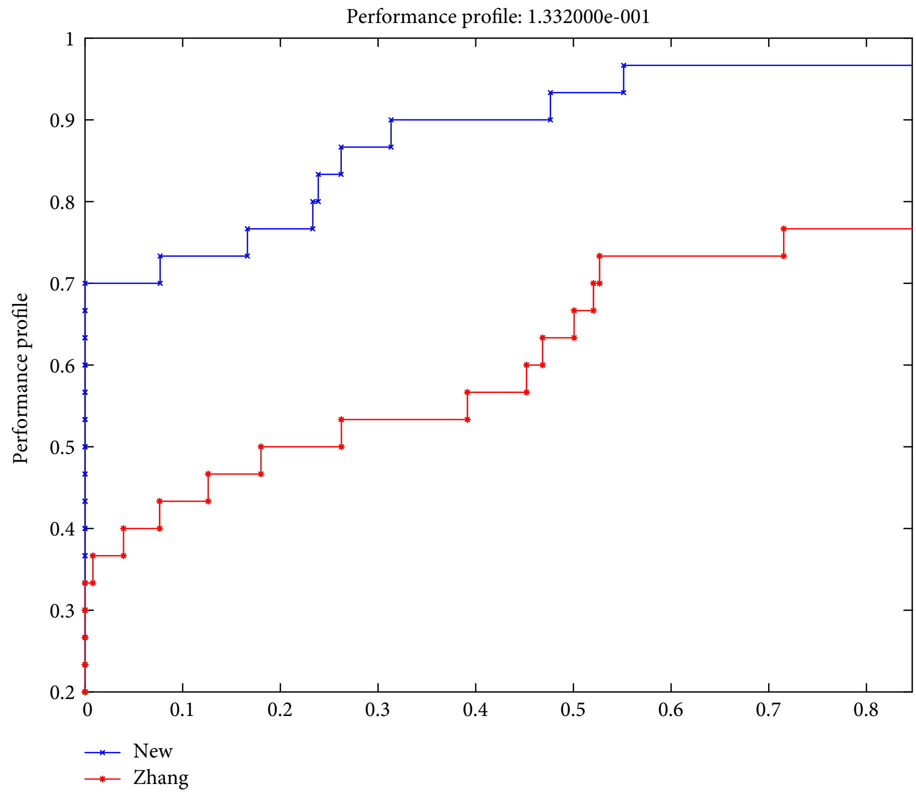

FIGURE 1: Performance profiles based on function evaluation.

From Lipschitz condition and

$$\begin{aligned}\mu\|s_k\|^2 &\le y_k^T s_k \le L\|s_k\| \le (L\|s_k\| + t\|s_k\|)\left(\frac{3\gamma}{\mu\|s_k\|}\right)\\ &\le (L+t)\left(\frac{3\gamma}{\mu}\right) = r.\end{aligned} \tag{35}$$

Hence, by taking the summation of the search direction we get:

$$\sum_{k\ge1} \frac{1}{\|d_{k+1}\|^2} \ge \frac{1}{r}\sum_{k\ge1} 1 = \infty. \tag{36}$$

This means that (33) is true. □

7. Numerical Experiments

In order to assess the performance of the planned new algorithm outlined in (15). The new constrained CG technique is checked over thirty nonlinear-selected test functions (see the Appendix of [1, 19, 20] for the details of those test problems). For all cases the stopping criterion is taken to be

$$\begin{aligned}\|\nabla\varphi(x_{k+1}, \mu_{k+1})\| \le 0.000001 \text{ or } &\left|\frac{\varphi(x_{k+1}, \mu_{k+1}) - \varphi(x_k, \mu_k)}{\varphi(x_{k+1}, \mu_{k+1})}\right|\\ &\le 0.000001.\end{aligned} \tag{37}$$

The comparative performance of all thought of algorithms is evaluated by considering NOF, NOI, NOC where NOF denotes the number of perform function evaluations, NOI denotes the number of iterations required to minimize the test problem and NOC denotes the number of constrained evaluations. We adopt the performance profiles given by Dolan and More [21].

The following three forms have the task of clarifying the performance of the algorithm more clearly as follows:

(i) Figure 1 illustrates the activity of the new algorithm in calculating the number of function values.

(ii) Figure 2 shows the level of improvement of the number of iterations.

(iii) Figure 3 illustrates the efficiency of the new algorithm in the calculation of constraints.

TABLE 1: Percentage performance of algorithm (15) against algorithm (13).

Tools	Algorithm (13) Zhang algorithm	New algorithm (15)
NOF	100%	40.35
NOI	100%	52.84
NOC	100%	92.43

To measure the percentages of optimization for better accuracy we give the following Table 1 showing the percentage of effectiveness of the new algorithm and the efficiency of the number of updates to reach the optimal solution.

From the last Table 1, it is evident that the new proposed constrained CG-technique formulated in (15) outdo the

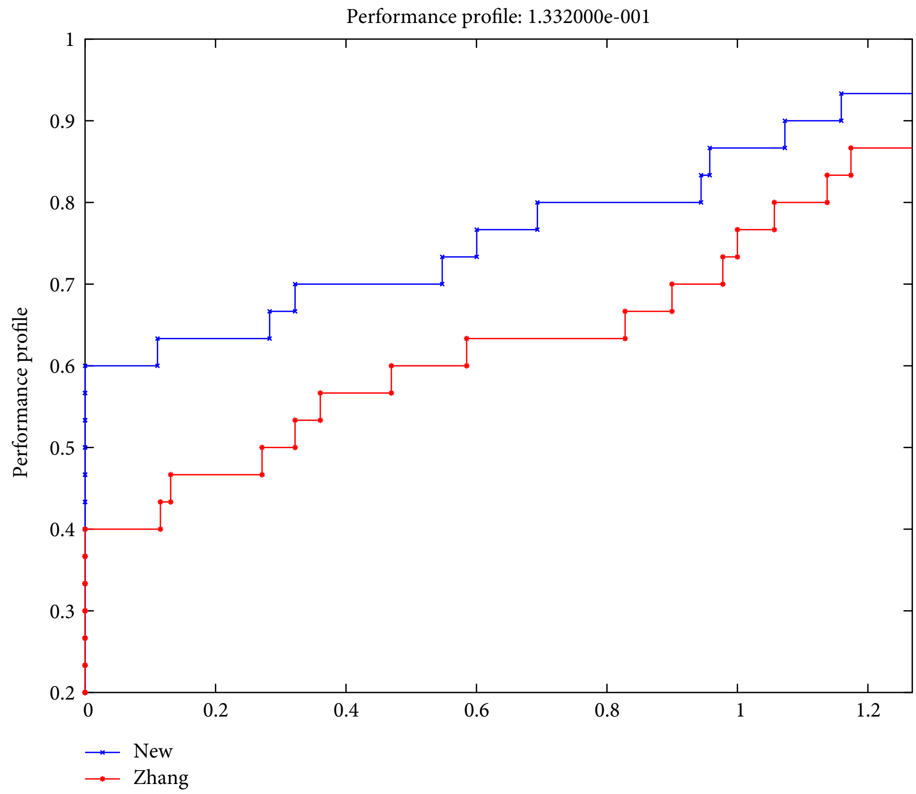

FIGURE 2: Performance profiles based on number of iterations.

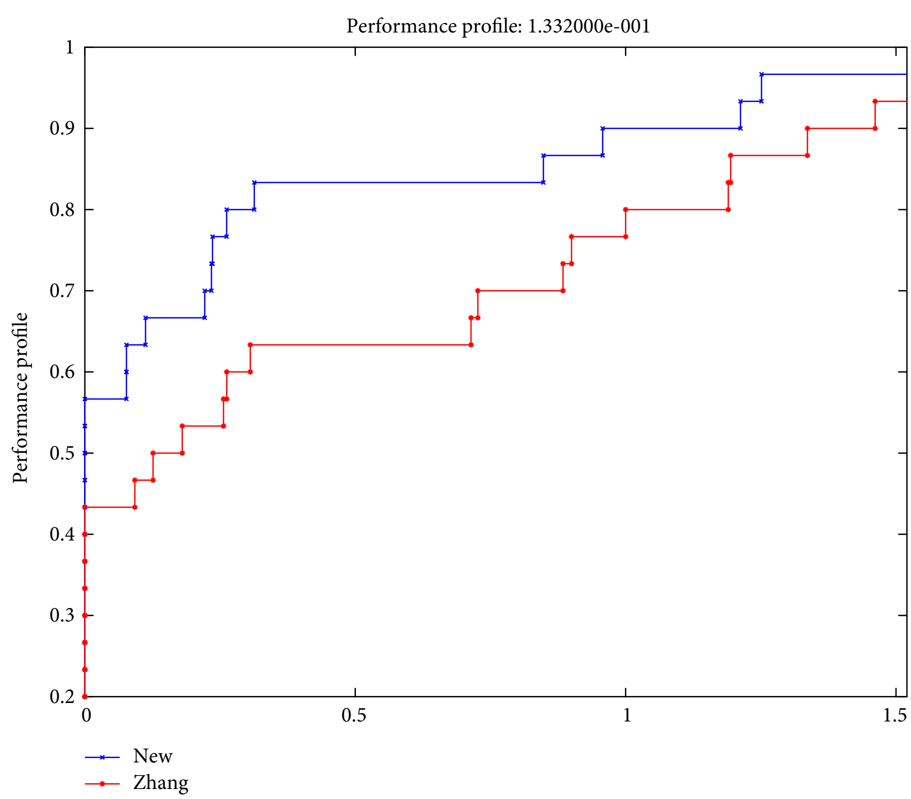

FIGURE 3: Performance profiles based on number of constrained.

standard three-term CG-technique formulated in (12) in about (59)% NOF; (47)% in NOI and (7)% in NOC, respectively.

Data Availability

The data used the support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The research is supported by College of Computer Sciences and Mathematics, University of Mosul, Republic of Iraq, under Project No. 6378368.

References

[1] S. S. Rao, *Engineering Optimization Theory and Practice*, John Wiley & Sons Inc., Hoboken, New Jersey, Canada, 4th edition, 2009.

[2] M. R. Hestenes and E. Stiefel, "Techniques of conjugate gradients for solving linear system," *Journal of Research of the National Bureau of Standards*, vol. 49, pp. 409–436, 1952.

[3] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, pp. 149–154, 1964.

[4] E. Polak and G. Ribière, "Note sur la convergence de méthodes de directions conjuguées," *ESAIM, Mathematical Modeling and Numerical Analysis*, vol. 3, pp. 35–43, 1969.

[5] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part1: theory," *Journal of Optimization Theory and Applications*, vol. 69, pp. 129–137, 1991.

[6] Y. Dai and L. Z. Liao, "New conjugacy conditions and related nonlinear conjugate gradient methods," *Applied Mathematical Optimization*, vol. 43, no. 1, pp. 87–101, 2001.

[7] P. Wolfe, "Convergence condition for ascent methods. II: some corrections," *SIAM Review*, vol. 13, pp. 185–188, 1971.

[8] I. E. Liveries and P. Pintelas, "A descent Dai–Liao conjugate gradient method based on a modified secant equation and its global convergence," *ISRN Computational Mathematics*, vol. 2012, Article ID 435495, 8 pages, 2012.

[9] S. Babaie-Kafaki and R. Ghanbari, "The Dai–Liao nonlinear conjugate gradient method with optimal parameter choices," *European Journal of Operational Research*, vol. 234, no. 3, pp. 625–630, 2014.

[10] M. R. Arazm, S. Babaie-Kafaki, and R. Ghanbari, "An extended Dai–Liao conjugate gradient method with global convergence for nonconvex functions," *Glasnik Matematicki*, vol. 52, no. 2, pp. 361–375, 2017.

[11] S. Babaie-Kafaki and R. Ghanbari, "A descent family of Dai–Liao conjugate gradient methods," *Optimization Methods and Software*, vol. 29, no. 3, pp. 583–591, 2014.

[12] M. Y. Waziri, K. Ahmed, and J. Sabi'u, "A Dai–Liao conjugate gradient method via modified secant equation for system of nonlinear equations," *Arabian Journal of Mathematics*, 2019.

[13] J. Zhang, Y. Xiao, and Z. Wei, "Nonlinear conjugate gradient methods with sufficient descent condition for large-scale unconstrained optimization," *Mathematical Problems in Engineering*, vol. 2009, Article ID 243290, 16 pages, 2009.

[14] M. Freund Robert, *Optimality Condition for Constrained Optimization Problems*, Massachusetts Institute of Technology (K.K.T.), 2004.

[15] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, 2nd edition, 1999.

[16] G. Li, C. Tang, and Z. Wei, "New conjugacy condition and related new conjugate gradient techniques for unconstrained optimization," *Journal of Computation and Applied Mathematics*, vol. 202, pp. 523–539, 2007.

[17] N. Andrei, "A modified Ribière–Polyak conjugate gradient algorithm for unconstrained optimization," *Optimization*, vol. 60, no. 12, pp. 1457–1471, 2011.

[18] S. Babaie-Kafaki and R. Ghanbari, "An adaptive Hager-Zhang conjugate gradient technique," *Filomat*, vol. 40, pp. 3715–3723, 2016.

[19] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, John Wiley and Sons, 2006.

[20] B. S. Gottefred and J. Weisman, *Introduction to Optimization Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[21] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming*, vol. 91, no. 2, pp. 201–213, 2002.