Hindawi
Journal of Applied Mathematics
Volume 2019, Article ID 8728196, 6 pages
https://doi.org/10.1155/2019/8728196

Research Article

A New Hybrid Algorithm for Convex Nonlinear Unconstrained Optimization

Eman T. Hamed,[1] Huda I. Ahmed ,[1] and Abbas Y. Al-Bayati[2]

[1] *Department of Operation Research and Intelligent Techniques, College of Computer Sciences and Mathematics, University of Mosul, Iraq*
[2] *University of Telafer, Iraq*

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

Received 24 December 2018; Revised 16 February 2019; Accepted 13 March 2019; Published 1 April 2019

Academic Editor: Urmila Diwekar

In this study, we tend to propose a replacement hybrid algorithmic rule which mixes the search directions like Steepest Descent (SD) and Quasi-Newton (QN). First, we tend to develop a replacement search direction for combined conjugate gradient (CG) and QN strategies. Second, we tend to depict a replacement positive CG methodology that possesses the adequate descent property with sturdy Wolfe line search. We tend to conjointly prove a replacement theorem to make sure global convergence property is underneath some given conditions. Our numerical results show that the new algorithmic rule is powerful as compared to different standard high scale CG strategies.

1. Introduction

The nonlinear CG technique could be a helpful procedure to search out the minimum value of any nonlinear function through exploitation unconstrained nonlinear optimization strategies.

Let us contemplate the subsequent unconstrained minimization problem:

$$\underset{x\in R^n}{\text{minimize}} \quad f(x) \tag{1}$$

where $f : R^n \longrightarrow R$ is a real-valued smooth function. The repetitious formula is given as

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

where α_k is associate optimum step-size computed by any line search procedure [1]. The search direction d_k is defined as

$$d_{k+1} = \begin{cases} -g_{k+1} & \text{for k} = 0 \\ -g_{k+1} + \beta_k d_k & \text{for k} \geq 1 \end{cases} \tag{3}$$

and $g_k = g(x_k)$ denotes $\nabla f(x_k)$, while β_k is a positive scalar.

Well-known established instances of β_k square measure are from Hestenes-Stiefel, Fletcher-Reeves, Polak-Ribière, Liu-Storey, Dai-Yuan, and Dai-Liao (see [2, 3, 3, 4, 4–7], respectively), within the already-existing convergence analysis and implementation of the CG methodology, the weak Wolfe conditions square measure [8]:

$$\text{"} f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k \nabla f(x_k)^T d \text{"} \tag{4}$$

$$\text{"} \nabla f(x_k + \alpha_k d_k)^T d_k \geq \sigma \nabla f(x_k)^T d_k \text{"} \tag{5}$$

If we choose $0 < \delta < \sigma < 1$, also the strong Wolfe conditions [8] consist of (4) and

$$\text{"} \left| g(x_k + \alpha_k d_k)^T d_k \right| \leq -\sigma g_k^T d_k \text{"} \tag{6}$$

Now, this allows us to review Ibrahim et al. work [9] that could be a work that considers unconstrained minimization problems. Ibrahim et al. recommend a search direction that is outlined as

$$d_k = -B_k^{-1} g_k + \lambda_{k+1} \mathrm{d}_{k-1} \tag{7}$$

and B_k is the BFGS updating matrix if step-size is

$$\lambda_k = \frac{\eta \mathrm{g}_{\mathrm{k}}^{\mathrm{T}} g_k}{\mathrm{g}_{\mathrm{k}}^{\mathrm{T}} d_{k-1}} \tag{8}$$

The scalar $\eta \in [0,1]$ is chosen to ensure conjugacy.

Also, another addition is regarded as an equivalent unconstrained minimization problem. Ibrahim et al. [10] recommend another search direction outlined as

$$\begin{aligned} d_k &= -B_k^{-1} g_k + \rho g_k, \\ \left(\mathrm{B_k} - \rho \mathrm{I}\right) \mathrm{d_k} &= -\mathrm{g_k} \end{aligned} \tag{9}$$

Matrix I is the identity and ρ <0.

In addition, Ibrahim et al. [11] propose another search direction that is outlined as

$$d_k = \begin{cases} -B_k^{-1} g_k & \mathrm{k} = 0 \\ -B_k^{-1} g_k + \eta \left(-g_k + \beta_k \mathrm{d}_{k-1}\right) & \mathrm{k} \geq 1 \end{cases} \tag{10}$$

The positive scalar η and β_k are the Hestenes-Stiefel parameters.

2. A New Proposed Search Direction

In this section, we advise a replacement search direction as deduced from Ibrahim et al. [9–11]. The new search direction is outlined as

$$\mathrm{d}_{k+1} = -\lambda_k \mathrm{g}_{k+1} + \mathrm{H}_{k+1} \mathrm{g}_{k+1} \tag{11}$$

whereas H_{k+1} (the approximation matrix of BFGS updating matrix) denotes approximations of Hessian matrix G, and λ_{k} could be a positive constant. In order to drive the value of λ_{k}, we have a tendency to multiply either side of (11) by $\mathrm{y}_{\mathrm{k}}^{\mathrm{T}}$ to induce

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

Since $y_k^T H_{k+1} = \mathrm{s}_{\mathrm{k}}^{\mathrm{T}}$ and $y_k^T d_{k+1} = -\mathrm{t} s_k^T g_{k+1}$ (Perry condition [12]), then

$$\begin{aligned} -\mathrm{t} s_k^T g_{k+1} &= -\lambda_k y_k^T g_{k+1} + s_k^T g_{k+1} \\ \lambda_k &= \frac{(1+\mathrm{t})\, s_k^T g_{k+1}}{y_k^T g_{k+1}} \end{aligned} \tag{13}$$

In order to see value of ρ_k, we tend to additionally multiply either side of (11) to induce

$$d_{k+1} = -\lambda_k g_{k+1} + \rho_k H_{k+1} g_{k+1} \tag{14}$$

$$y_k^T d_{k+1} = -\lambda_k y_k^T g_{k+1} + y_k^T H_{k+1} g_{k+1} \tag{15}$$

$$\rho_k = \frac{\lambda_k y_k^T g_{k+1} - t s_k^T g_{k+1}}{s_k^T g_{k+1}} \tag{16}$$

As a result of processes of multiplication shown in (13) and (16), we tend to reach the following search directions in (17a), (17b), and (17c). The subsequent new search directions are our new projected algorithmic program:

$$d_{k+1} = -\lambda_k g_{k+1} + \rho_k H_{k+1} g_{k+1} \tag{17a}$$

$$\lambda_k = \frac{(1+\mathrm{t})\, s_k^T g_{k+1}}{y_k^T g_{k+1}} \tag{17b}$$

$$\rho_k = \frac{\lambda_k y_k^T g_{k+1} - t s_k^T g_{k+1}}{s_k^T g_{k+1}} \tag{17c}$$

In the following step, we have a tendency to assume that each search direction got to satisfy the subsequent descent condition ($g_k^T d_k < 0$, for all k). Also, there should exist a constant c>0 in order to get

$$"d_k^T g_k \leq -c_1 \left\| g_k \right\|^2" \tag{18}$$

For all $k \geq 0$, the new direction that is outlined in (18) ought to satisfy the sufficient descent condition. The enough descent conditions are going to be used later to prove our new theorem (see Section 2.2). So to prove our new theorem, we have a tendency to necessarily use the subsequent given assumptions (see Section 2.1).

2.1. Assumptions in [9, 11, 13]

(A1) $\mathrm{f} : \mathrm{R}^{\mathrm{n}} \longrightarrow \mathrm{R}$ is twice continuously differentiable.

(A2) f is uniformly convex; that is, m and M are positive constants, such that

$$"m \left\| z \right\|^2 \leq z^T G(x) z \leq M \left\| z \right\|^2" \tag{19a}$$

for all $\mathrm{x}, \mathrm{z} \in \mathrm{R}^{\mathrm{n}}$, and G is the Hessian matrix of f.

(A3) The matrix G is Lipschitz continuous at the point x^*; that is, there exists the positive constant L satisfying

$$"\left\| g(x) - g(x^*) \right\| \leq L \left\| x - x^* \right\|" \tag{19b}$$

for all x in a neighborhood of x^*.

2.2. A New Theorem for Proving Sufficient Descent Property. To prove that our new projected algorithm defined in (17a), (17b), and (17c) satisfies sufficiently descent condition, we tend to suppose that assumptions in (Section 2.1) square measure are true. Additionally, the sequence H_k is bounded. Then, sufficient descent condition (18) is true for all $\mathrm{k} \geq 0$.

Proof. When taking (17a), (17b), and (17c) and achieving the descent condition, we can see the following:

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\lambda_k g_{k+1}^T g_{k+1} + \rho_k g_{k+1}^T H_{k+1} g_{k+1} \\ &\leq -\lambda_k \left\| g_{k+1} \right\|^2 + \delta_k \left\| g_{k+1} \right\|^2 \leq c \left\| g_{k+1} \right\|^2 \end{aligned} \tag{20}$$

We get value $c = -(\lambda_k - \delta_k)$, which is bounded away from zero. Therefore (18) is true. □

2.3. Lemma in [10, 14]. Suppose that assumptions in Section 2.1 are true. Then, the step-size α_k which is determined by (2) satisfies

$$"f_{k+1} - f_k \leq -c_3 \frac{\left(g_k^T d_k\right)^2}{\left\|d_k\right\|^2}" \tag{21}$$

when c_3 is a positive constant.

2.4. New Theorem for Proving Global Convergence Property. Having demonstrated the important and necessary properties of regression algorithms, we now come to the proof of the necessary property to be present in all numerical optimization algorithms. Let us demonstrate the new algorithm defined in (17a), (17b), and (17c). To achieve a global convergence property, assume that the theory in Section 2.2 and the assumptions in Section 2.1 are correct. Then

$$"\lim_{k \longrightarrow \infty} \left\|g_k\right\|^2 = 0" \tag{22}$$

Proof. By linking the theory in Section 2.2 and lemma in Section 2.3 we give the following result:

$$\sum_{k=0}^{\infty} \frac{\left\|g_k\right\|^4}{\left\|d_k\right\|^2} < \infty \tag{23}$$

Hence, from our new theorem in Section 2.2, we can define that $\|d_k\| \leq -c\|g_k\|$, and we can therefore simplify (23) as $\sum_{k=0}^{\infty} \|g_k\|^2 < \infty$. Thus, the proof is established. □

3. A New Form for the Parameter β_k in Conjugate Gradient Method

To obtain an updated version of the conjugate gradient method associated with a new parameter β_k, we compare the standard CG-method specified in (3) with the proposed new algorithm specified in (17a), (17b), and (17c):

$$-\lambda_k g_{k+1} + \rho_k H_{k+1} g_{k+1} = -g_{k+1} + \beta_k s_k \tag{24}$$

Multiplying both sides of (24) by y_k^T we get

$$-\lambda_k y_k^T g_{k+1} + \rho_k y_k^T H_{k+1} g_{k+1} = -y_k^T g_{k+1} + \beta_k y_k^T s_k \tag{25}$$

Or

$$H_{k+1} = \frac{-y_k g_{k+1}^T + \beta_k y_k s_k^T + \lambda_k y_k g_{k+1}^T}{\rho_k y_k^T g_{k+1}} \cdot I \tag{26}$$

knowing that I is the identity matrix. Moreover,

$$s_k^T H_{k+1} s_k = \frac{-s_k^T y_k g_{k+1}^T s_k + \beta_k s_k^T y_k s_k^T s_k + \lambda_k s_k^T y_k g_{k+1}^T s_k}{\rho_k y_k^T g_{k+1}} \tag{27}$$

Then, the new β_k is

$$\beta_k^{hh} = \frac{\left(s_k^T H_{k+1} s_k\right) \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k}{\left\|s_k\right\|^2 s_k^T y_k} \tag{28}$$

and $y_k = g_{k+1} - g_k$, $s_k = x_{k+1} - x_k$. It should be noted that when using exact line searches assuming that $\rho_k = 1$ and the matrix H_{k+1} is the identity matrix, the new standard will be reduced to HS. This condition must also be met, with $s_k^T H_{k+1} s_k > 0$ (positive constant), and since

$$s_k^T H_{k+1} s_k = \psi_k > 0 \tag{29a}$$

after these conditions we know the new parameter β_k^{hh} is

$$\beta_k^{hh} = \frac{\psi_k \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k}{\left\|s_k\right\|^2 s_k^T y_k} \tag{29b}$$

In conjunction with both parameters

$$\lambda_k = \frac{(1+t)\, s_k^T g_{k+1}}{y_k^T g_{k+1}} \tag{29c}$$

$$\rho_k = \frac{\lambda_k y_k^T g_{k+1} - t s_k^T g_{k+1}}{s_k^T g_{k+1}} \tag{29d}$$

and the search direction is defined as

$$d_{k+1} = -g_{k+1} + \frac{\psi_k \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k}{\left\|s_k\right\|^2 s_k^T y_k} s_k \tag{29e}$$

3.1. Outlines of the New Proposed CG-Algorithm in (29a), (29b), (29c), (29d), and (29e). By assuming $x \in R^n$, $\epsilon \geq 0$, and by setting the iteration k=1, we get the following steps.

Step 1. Set $d_k = -g_k$, if $\|g_k\| < \epsilon$, then stop.

Step 2. Compute α_k by strong Wolfe line search conditions in (4) and in (6).

Step 3. Generate d_{k+1} by implementing (29a), (29b), (29c), (29d), and (29e).

Step 4. If $|g_{k+1}^T g_k| \geq 0.2\|g_{k+1}\|^2$ [15] is satisfied, then go to Step 1; if not, then continue.

3.2. Assumptions for Proving the Convergence Analysis Property of the New Algorithm in (29a), (29b), (29c), (29d), and (29e). Let us suppose the following.

(i) The level set

$$"S = \left\{x \in R^n : f(x) \leq f\left(x_0\right)\right\}" \text{ is bonded.} \tag{30}$$

(ii) The condition $\|\nabla f(x)\| \leq \gamma$ is satisfied where $\gamma > 0$; also since f is a uniformly convex function on S, then there exists a constant $\mu > 0$, such that

$$"(\nabla f(x) - \nabla f(y))^T (x-y) \geq \mu \|x-y\|^2, \quad \forall x, y \in S" \tag{31}$$

From both (19a), (19b), and (31), we can get the following:

$$"\mu \|s_k\|^2 \leq y_k^T s_k \leq L \|s_k\|^2" \tag{32}$$

3.3. New Theorem for Proving Sufficiently Descent Directions of the New Algorithm in (29a), (29b), (29c), (29d), and (29e). If we have a tendency to assume that (A3) in Section 2.1 is true and if we assume that conditions (i) and (ii), outlined in (30) and in (31), respectively, are true, then the new proposed search direction d_{k+1} defined in (29e) satisfies the sufficient descent condition.

Proof. By using the mathematical induction we demonstrate this new theorem; for initial direction (k=1) we have

$$d_1 = -g_1 \longrightarrow d_1^T g_1 = -\|g_1\|^2 \leq 0 \tag{33}$$

which satisfies (18).

We suppose that $d_i^T g_i \leq -c\|g_i\|^2 \leq 0$, $\forall i = 1, 2, \ldots, k$.

And if we multiply both sides of (29a) by $(g_{k+1}/\|g_{k+1}\|^2)$, then we can get

$$\begin{aligned} &\frac{d_{k+1}^T g_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \\ &= \frac{\psi_k \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k}{\|s_k\|^2 s_k^T y_k} \frac{s_k^T g_{k+1}}{\|g_{k+1}\|^2} \end{aligned} \tag{34}$$

since $s_k^T g_{k+1} \leq s_k^T y_{k+1}$, since $y_k^T g_{k+1} \leq \|y_k\| \cdot \|g_{k+1}\|$:

$$\begin{aligned} &\frac{d_{k+1}^T g_{k+1} + \|g_{k+1}\|^2}{\|g_{k+1}\|^2} \\ &\leq \frac{\psi_k \rho_k \|y_k\| \cdot \|g_{k+1}\| + s_k^T y_k s_k^T y_k - \lambda_k s_k^T y_k s_k^T y_k}{\|s_k\|^2 s_k^T y_k} \left(\frac{s_k^T y_k}{\|g_{k+1}\|^2} \right) \\ &\leq \frac{\psi_k \rho_k \|y_k\| \cdot \|g_{k+1}\| + (s_k^T y_k)^2 - \lambda_k (s_k^T y_k)^2}{\|s_k\|^2 s_k^T y_k} \left(\frac{s_k^T y_k}{\|g_{k+1}\|^2} \right) \\ &\leq \frac{\psi_k \rho_k \|y_k\| \cdot \|g_{k+1}\| + (1-\lambda_k)(s_k^T y_k)^2}{\|s_k\|^2 \|g_{k+1}\|^2} \end{aligned} \tag{35}$$

and since $0 < \lambda < 1 \longrightarrow (1-\lambda) > 0$

$$\begin{aligned} &let \ c_1 = \left(\frac{\psi_k \rho_k \|y_k\| \cdot \|g_{k+1}\| + (1-\lambda_k)(s_k^T y_k)^2}{\|s_k\|^2 \|g_{k+1}\|^2} \right) \\ &\quad > 0 \\ &d_{k+1}^T g_{k+1} \leq -(1-c_1)\|g_{k+1}\|^2. \end{aligned} \tag{36}$$

The above equations ensure that condition (18) is satisfied. Hence, the proof is complete. ☐

3.4. Lemma in [1, 16]. By assuming that assumptions in Section 2.1 are true, by supposing that any CG-method with search direction d_{k+1} could be a descent direction provided that the step-size α_k is obtained by the strong Wolfe line search conditions, and if

$$"\sum_{k \geq 1} \frac{1}{\|d_{k+1}\|^2} = \infty" \tag{37}$$

then

$$"\lim_{k \longrightarrow \infty} (\inf \|g_k\|) = 0" \tag{38}$$

3.5. A New Theorem for Proving the Global Convergence Property of the New Algorithm in (29a), (29b), (29c), (29d), and (29e). If we suppose that assumptions (i) and (ii) in (30) and in (31), respectively, are true and if we have a tendency to assume that (A3) in Section 2.1 is additionally true, then the search directions d_{k+1} outlined in (29e) are descent provided that the step-size α_k is computed using (4) and (6), and then

$$\lim_{k \longrightarrow \infty} (\|g_k\|) = 0. \tag{39}$$

Proof. since

$$\begin{aligned} \beta_k^{hh} &= \frac{\psi_k \rho_k y_k^T g_{k+1} + s_k^T y_k g_{k+1}^T s_k - \lambda_k s_k^T y_k g_{k+1}^T s_k}{\|s_k\|^2 s_k^T y_k} \\ \left|\beta_k^{hh}\right| &\leq \left| \frac{\psi_k \rho_k y_k^T g_{k+1}}{\|s_k\|^2 s_k^T y_k} \right| + \left| \frac{s_k^T y_k g_{k+1}^T s_k}{\|s_k\|^2 s_k^T y_k} \right| \\ &\quad + \left| \frac{\lambda_k s_k^T y_k g_{k+1}^T s_k}{\|s_k\|^2 s_k^T y_k} \right| \end{aligned} \tag{40}$$

and since $s_k^T g_{k+1} \leq s_k^T y_{k+1}$ and $y_k^T g_{k+1} \leq \|y_k\| \cdot \|g_{k+1}\|$,

$$\begin{aligned} \left|\beta_k^{hh}\right| &\leq \frac{\psi_k \rho_k \|y_k\| \|g_{k+1}\|}{\|s_k\|^2 s_k^T y_k} + \left| \frac{(s_k^T y_k)^2}{\|s_k\|^2 s_k^T y_k} \right| \\ &\quad + \left| \frac{\lambda_k (s_k^T y_k)^2}{\|s_k\|^2 s_k^T y_k} \right| \\ &\leq \left| \frac{\psi_k \rho_k \|y_k\| \|g_{k+1}\|}{\|s_k\|^2 s_k^T y_k} \right| + \left| \frac{(s_k^T y_k)}{\|s_k\|^2} \right| + \left| \frac{\lambda_k (s_k^T y_k)}{\|s_k\|^2} \right| \end{aligned} \tag{41}$$

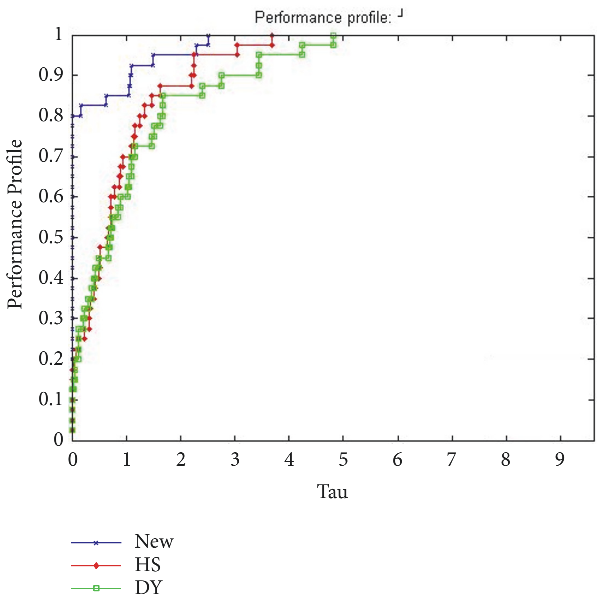

FIGURE 1: Performance profiles based on number of iterations.

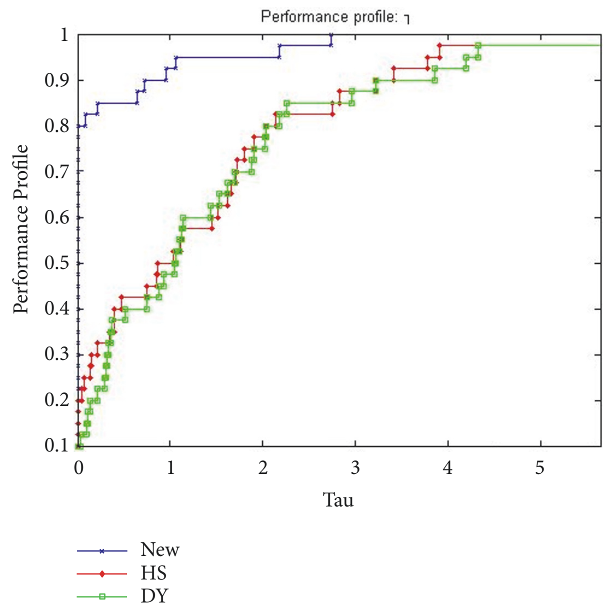

FIGURE 2: Performance profiles based on CPU time.

then from (A3) in part (Section 2.1), from condition (ii) in (31), and from (28) we get

$$\begin{aligned}
\left|\beta_k^{hh}\right| &\leq \frac{\psi_k \rho_k L \left\|s_k\right\| \gamma}{\left\|s_k\right\|^2 \mu \left\|s_k\right\|^2} + \frac{L \left\|s_k\right\|^2}{\left\|s_k\right\|^2} + \lambda_k \frac{L \left\|s_k\right\|^2}{\left\|s_k\right\|^2} \\
&\leq \frac{\psi_k \rho_k L \gamma}{\left\|s_k\right\| \mu \left\|s_k\right\|^2} + L + L\lambda_k = c_2 > 0 \\
\left\|d_{k+1}\right\| &\leq \left\|g_{k+1}\right\| + c_2 \left\|s_k\right\| = D \\
\sum_{k\geq 1} \frac{1}{\left\|d_{k+1}\right\|^2} &\geq \frac{1}{D} \sum_{k\geq 1} 1 = \infty. \\
\text{i.e. } \lim_{k\longrightarrow\infty} \left\|g_k\right\| &= 0
\end{aligned} \tag{42}$$

□

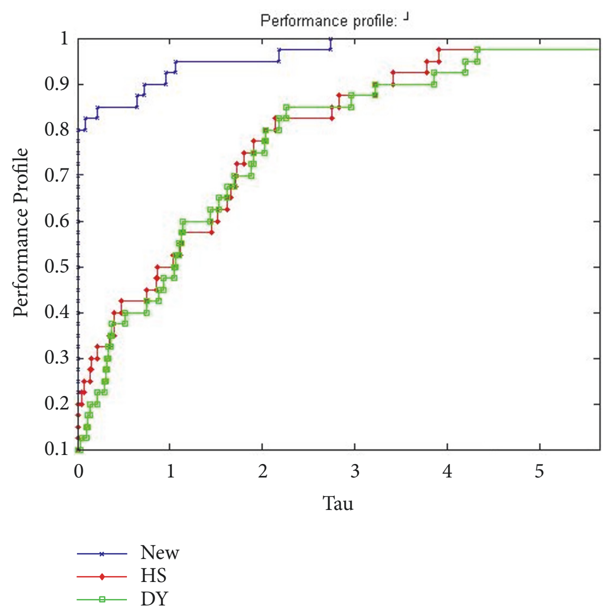

FIGURE 3: Performance profiles based on function evaluation.

4. Numerical Results and Comparisons

In this work, we have a tendency to compare our new proposed CG-method with some normal classical CG strategies like Hestenes-Stiefel [HS] and Dai-Yuan [DY] by exploitation of fifty unconstrained nonlinear cases; take a look at functions obtained from Andrei [17, 18]. As for the computer program, it was stopped when $\|g_k\| \leq 10^{-6}$. In addition, the term $s_k^T H_{k+1} s_k$, which is defined in (29a), can be computed as $s_k^T H_{k+1} s_k = 2(f_k - f_{k+1}) + 2 s_k^T g_{k+1}$. Numerical results for new algorithm in (29a), (29b), (29c), (29d), and (29e) with $\lambda = 1$ and $\rho = 0.1$ are for the total of 50 test problems from the CUTE library. The Sigma plotting software was used to graph the data. We adopt the performance profiles given by Dolan and Moré [19]. Thus, new, HS, and DY strategies are compared in terms of NOI, CPU, and NOF in Figures 1–3. For each method, we plotted the fraction of problems that were solved correctly within a factor of the best time. In the figures, the uppermost curve is the method that solves the most problems within a factor t of the best time. In Figures 1–3, the new method outperforms the HS algorithm and DY method in terms of NOI, CPU, and NOF. If the solution had not converged after 800 seconds, the program was terminated. Generally, convergence was achieved within this time limit; functions for which the time limit was exceeded are denoted by "F" for fail-in.

5. Conclusions

At the end of this work we were able to obtain a new direction of research defined in (29a), (29b), (29c), (29d), and (29e). This new trend is a hybrid trend that combines pedigree techniques with Quasi-Newton ones. Through the theories presented in the research, the new trend in (29a), (29b), (29c), (29d), and (29e) proved that it satisfies the requirement of sufficient proportions and ensures the property of global convergence. In addition, we have presented a new scalar (β_k) which ensures the sufficient descent directions. Moreover, under some conditions, we have established that the new proposed algorithm is a globally convergent algorithm for uniformly convex functions under the strong Wolfe line search conditions. The numerical results show that when we choose the value of parameters λ in (13) and ρ in (16), we obtain best numerical results as compared to other similar numerical results.

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 there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The research is supported by College of Computer Sciences and Mathematics, University of Mosul, Republic of Iraq, under Project no. 8728196.

References

[1] W. Sun and Y.-X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, NY, USA, 2006.

[2] Y. H. Dai and L. Z. Liao, "New conjugacy conditions and related nonlinear conjugate gradients methods," *Applied Mathematical Optimization*, vol. 43, pp. 87–101, 2001.

[3] Y. H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999.

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

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

[6] E. Polak and G. Ribière, "Note Sur la convergence de methods de directions Conjugue'e, ESAIM," *Mathematical Modeling and Numerical Analysis*, vol. 3, no. 16, pp. 35–43, 1969.

[7] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, Part 1," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129–137, 1991.

[8] P. Wolfe, "Convergence conditions for ascent methods(II): some corrections," *SIAM Review*, vol. 13, no. 2, pp. 185–188, 1971.

[9] M. A. Ibrahim, M. Mamat, and W. J. Leong, "BFGS method: a new search direction," *Saints Malaysian*, vol. 43, pp. 1593–1599, 2014.

[10] M. A. H. Ibrahim, M. Mamat, P. L. Ghazali, and Z. Salleh, "The scaling of hybrid method in solving unconstrained optimization method," *Far East Journal of Mathematical Sciences*, vol. 99, no. 7, pp. 983–991, 2016.

[11] M. A. H. Ibrahim, M. Mamat, and A. Z. M. Sofi, "A modified search direction of Broyden family method and its global convergence," *Journal of Engineering and Applied Sciences*, vol. 12, no. 17, pp. 4504–4507, 2017.

[12] A. Perry, "A modified conjugate gradient algorithm," *Operations Research*, vol. 26, pp. 1073–1078, 1987.

[13] R. Jaafar, M. Mamat, and I. Mohd, "A new scaled hybrid modified BFGS algorithms for unconstrained optimization," *Applied Mathematical Sciences*, vol. 7, pp. 263–270, 2013.

[14] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 2006.

[15] M. J. Powell, "Restart procedure for the conjugate gradient methods," *Mathematics Programming*, vol. 12, pp. 241–254, 1977.

[16] H. Yabe and M. Takano, "Global convergence properties of new nonlinear conjugate gradient methods for unconstrained optimization," *Computational Optimization and Application*, vol. 28, pp. 203–225, 2004.

[17] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization. The Electronic International Journal*, vol. 10, no. 1, pp. 147–161, 2008.

[18] N. Andrei, "Open problems in nonlinear conjugate gradient algorithms for unconstrained optimization," *Bulletin of the Malaysian Mathematical Sciences Society. Second Series*, vol. 34, no. 2, pp. 319–330, 2011.

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