

A Dai-Liao Hybrid Hestenes-Stiefel and Fletcher-Reeves Methods for Unconstrained Optimization

Nasiru Salihu ^{a*}, Mathew Remilekun Odekunle ^a, Also Mohammed Saleh ^a, Suraj Salihu ^b

^aDepartment of Mathematics, School of Physical Science, Moddibo Adama University of Technology, Yola, Nigeria

^bDepartment of Computer Science, Faculty of Science, Gombe State University, Gombe, Nigeria

nasirussalihu@gmail.com; odekunle@mautech.edu.ng; moalso81@gmail.com; surajsalihu@gmail.com

*Corresponding Author: nasirussalihu@gmail.com

ARTICLE INFO

Keywords:

Unconstrained optimization;
Descent property;
Conjugacy condition;
Global convergence;
Conjugate gradient algorithm.

Article history

Received:
November 5, 2020
Revised:
November 25, 2020
Accepted:
January 22, 2021
Available online:
February 24, 2021

ABSTRACT

Some problems have no analytical solution or too difficult to solve by scientists, engineers, and mathematicians, so the development of numerical methods to obtain approximate solutions became necessary. Gradient methods are more efficient when the function to be minimized continuously in its first derivative. Therefore, this article presents a new hybrid Conjugate Gradient (CG) method to solve unconstrained optimization problems. The method requires the first-order derivatives but overcomes the steepest descent method's shortcoming of slow convergence and needs not to save or compute the second-order derivatives needed by the Newton method. The CG update parameter is suggested from the Dai-Liao conjugacy condition as a convex combination of Hestenes-Stiefel and Fletcher-Reeves algorithms by employing an optimal modulating choice parameter t to avoid matrix storage. Numerical computation adopts an inexact line search to obtain the step-size that generates a decent property, showing that the algorithm is robust and efficient. The scheme converges globally under Wolfe line search, and it's like is suitable in compressive sensing problems and M-tensor systems.

1. Introduction

Conjugate Gradient (CG) method was initially proposed for solving linear systems and unconstrained minimization. The method is an excellent choice for solving optimization problems by scientists, engineers, and mathematicians. The method has the following form

$$\min f(x), \tag{1}$$

$$x \in \mathbb{R}^n$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth nonlinear function.

The iterative scheme of the method is computed by

$$x_0 \in \mathbb{R}^n \tag{2}$$

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

in which $\alpha_k > 0$ is a step length obtained by a suitable exact or inexact line search. However, α_k is usually generated by an inexact line search, such as the standard Wolfe line search

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (3)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k. \quad (4)$$

Or using strong Wolfe condition, which consists of (3) and

$$|g(x_k + \alpha_k d_k)^T d_k| \leq -\sigma g_k^T d_k \quad (5)$$

where $0 < \delta < \sigma < 1$ and d_k is the search direction given by

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad (6)$$

where β_k is a scalar called CG (update) parameter (Babaie -Kafaki, 2011).

Liu & Du (2019) proposed a CG method by transforming M - tensor system to general unconstrained minimization problem and solving a kind of nonsmooth optimization problems with l_1 - norm, the given numerical experiments show the efficiency of the suggested method due to simplicity, low storage and nice convergence properties of the CG methods. Esmaeili, Rostami & Kimiaei (2018) employed a new CG method to solve compressive sensing problems that play an important role in medical and astronomical imaging, file restoration, image, video coding, etc applications. The method is characterized by l_0 - minimization that is NP-hard in general. Hence, replaced the l_0 - norm by the closest convex norm, which is the l_1 - norm leads to the minimization problem. Similarly, Guo & Wan (2019) developed a CG algorithm to solve an engineering problem originated from compressed sensing of sparse signals. Numerical tests and preliminary application in recovering sparse signals indicate that the established algorithm outperforms similar algorithms in the literature, especially for solving large-scale problems and singular ones. It was shown that the compressed sensing of sparse signals does not involve computing the Jacobian matrix or its approximation, both information storage and computational cost of the algorithm are lower. Recently, Liu, Du & Chen (2020) suggested a kind of important tensor optimization problem with higher-order nonlinear equations, widely used in engineering and economics. The algorithm is concerned with solving M -tensor equations by transforming the equations to nonlinear unconstrained optimization problems. The effectiveness of the proposed nonlinear conjugate gradient method was compared with the three-term conjugate gradient method and Newton method. Numerical results show that the proposed nonlinear conjugate gradient method is potentially efficient.

Different values of the scalar parameters β_k correspond to several CG schemes. Some excellent CG algorithms employed in practice to obtain new formulas include; Hestenes & Stiefel (HS) (1952), Polak, Ribie're & Polyak (PRP) (1967), Liu & Storey (LS) (1991), Fletcher & Revees (FR) (1964), Fletcher (Conjugate Descent (CD)) (1987), and Dai & Yuan (1991) schemes. Let $\|\cdot\|$ denotes Euclidean norm and define $s_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$ (Dai & Yuan, 2001). Numerical experiments show that the FR, DY and CD conjugate schemes are characterized by strong global convergence properties and have poor practical performances due to jamming. On the contrast, LS, HS and PRP have better

practical performances, but may not always be convergent (Babaie-Kafaki & Ghanbari, 2014c). To improve these schemes' behavior and avoid numerical uncertainty, researchers were interested in combining CG schemes of the two groups (Babaie-Kafaki & Mahdavi-Amiri, 2013).

There are some strengths and weaknesses in the theory of CG schemes. The first global convergent property of FR method with exact line was proved by Zoutendijk (1970); where Al-Baali (1985) extended this result to an inexact line search and show that the method generates sufficient descent direction under the strong Wolfe conditions using the constraint $\sigma < \frac{1}{2}$. However, the HS and PRP schemes possess an automatic approximate restart feature that addresses a jamming problem that makes them numerically efficient (Babaie-Kafaki, 2013). Here this research considers a new convex combination of HS and FR conjugate gradient methods. The corresponding conjugate gradient parameters are

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \tag{7}$$

and

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \tag{8}$$

The structures of CG update parameters were obtained from conjugacy condition and secant equation which depends on the exact line search. These procedures require computation and storage of the Hessian matrix, respectively. However, the practical numerical analysis adopts inexact line searches instead of exact line searches to obtain the step-size. To address these drawbacks, this article, therefore, presents a hybrid method from Dai-Liao conjugacy condition so that, if the modulating parameter $t = 1$, then it reduces to a method that uses pure conjugacy condition. The numerical performance of the Dai-Liao CG method depends on the parameter t . The best choice of t remains subject of consideration (Babaie-Kafaki & Ghanbari, 2017).

The article aims to modify the CG methods using classical HS and FR method by employing optimal choice of the parameter t for solving large scale unconstrained optimization problems. This paper is organized as follows: Section 2 presents the proposed method. Convergence results are presented in Section 3. Some numerical results are reported in Section 4. Finally, conclusions are made in Section 5.

2. Literature Review

A large number of hybrid conjugate gradient techniques were proposed. The idea is to combine different conjugate algorithms to use the projection to form a new hybrid convex-combination algorithm to avoid jamming and improve the convergence analysis (Mohammed, et al., 2020). Djordjevic (2016; 2017; 2018) proposed hybrid conjugate gradient algorithms. The conjugate gradient parameters β_k are computed as a convex combination of the hybrid parameter θ_k , where they are computed in such a way that the conjugacy condition(s) are satisfied using strong Wolfe line search conditions, which has the following formulas for β_k respectively

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{PRP} + \theta_k \beta_k^{FR}. \tag{9}$$

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{CD}. \quad (10)$$

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{HS} + \theta_k \beta_k^{FR}. \quad (11)$$

On the other hand; Djordjevic (2019), Al-Namat & Al-Naemi (2020) and Salihu et al. (2020) recently derived new hybrid schemes for solving large scale unconstrained optimization algorithms. The hybrid schemes satisfy the sufficient descent condition in such a way that Newton directions are employed, global convergence analysis were proved under the same conditions above, and the algorithms are characterized by following β_k .

$$\beta_k^{hyb} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{FR}. \quad (12)$$

$$\beta_k^{CLCS} = (1 - \theta_k) \beta_k^{LS} + \theta_k \beta_k^{CD}. \quad (13)$$

$$\beta_k^{FG} = (1 - \theta_k) \beta_k^{MMWU} + \theta_k \beta_k^{RMAR}. \quad (14)$$

Numerical comparisons show that these algorithms behave better than some known methods. Based on the modified BFGS method proposed by Li & Fukushima (2001), Lotfi & Hosseini (2019) presented a new value of the parameter t in Dai-Liao CG scheme. The proposed method's global convergence property was established, and numerical results illustrated the computational efficiency of the new method. Considerable efforts have recently been made to extend CG methods to solve monotone nonlinear equations, Abubakar et-al. (2019) presented a modification of the FR CG method for constrained monotone nonlinear equations. The method possesses sufficient descent property, and its global convergence was proved. Numerical experiments show efficiency of the proposed method using some benchmark test problems while applying the method in signal and image recovery problems arising from compressive sensing. Also, Xue et-al. (2018) suggested DY CG method for solving large-scale unconstrained optimization problems, which possesses a spectral CG parameter in which the search direction generated at each iteration is independent of any line search. Global convergence of the method is also established using strong Wolfe conditions. Finally, comparison experiments on impulse noise removal are reported and demonstrated the effectiveness of the method.

These new methods are based on secant equations or conjugacy condition, for nonlinear conjugate gradient methods, the conjugacy condition is given by

$$d_{k+1}^T y_k = 0. \quad (15)$$

Perry (1978) extended the result in (15) by exploiting the following secant condition of quasi-Newton scheme $B_{k+1} s_k = y_k$ and quasi-Newton search direction given by $B_{k+1} d_{k+1} = -g_{k+1}$, where B_{k+1} is a square matrix of the Hessian approximation; as

$$d_{k+1}^T y_k = -g_{k+1}^T s_k, \quad (16)$$

which implies that (16) holds for exact line search, where $-g_{k+1}^T s_k = 0$, but practical numerical computations normally adopt inexact line search; that is, $-g_{k+1}^T s_k \neq 0$. For this reason, Dai and Liao (2001) replaces (16) with a condition called extended conjugacy condition:

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k, \text{ where } t \geq 0. \quad (17)$$

Due to the simpler structure and low memory requirements of Dai-Liao conjugate gradient methods, Yao et-al. (2019) combined Dai-Liao conjugacy condition with a modified symmetric Perry matrix to propose a class of three-term Dai-Liao conjugate gradient algorithms. The method possesses both Dai-Liao conjugacy conditions and sufficient descent conditions. Meanwhile, the global convergence is established under Wolfe line search for general objective functions. Numerical experiments show that the proposed method is promising. The best choice of t remains subject of consideration, with several optimal choices for the parameter proposed in (Babaie-Kafaki, 2015; Babaie-Kafaki & Ghanbari, 2014a; Babaie-Kafaki & Ghanbari, 2014b; Babaie-Kafaki & Ghanbari, 2015; Babaie-Kafaki & Ghanbari, 2017; Waziri, Ahmed & Sabi'u, 2019). Motivated by the above, this research proposes a hybrid parameter by employing the choice of the parameter t in Andrei (2017) using (17) to access and combine the CG update parameters' strength.

3. Research Methodology

In this section, this research combines the CG update parameters proposed by Hestenes & Stiefel (1952) with Fletcher & Reeves (1964) conjugate descent based on Dai-Liao conjugacy condition as a convex combination as follows:

$$\beta_k^{DHF} = (1 - \theta_k) \beta_k^{HS} + \theta_k \beta_k^{FR}. \quad (18)$$

From relations (7) and (8), it obtains

$$\beta_k^{DHF} = (1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{\|g_k\|^2} \right), \quad (19)$$

where θ_k is the hybridization scalar parameter satisfying $\theta_k \in [0,1]$. It is obvious that if $\theta_k \leq 0$, set $\theta_k = 0$, then $\beta_k^{DHF} = \beta_k^{HS}$ and if $\theta_k \geq 1$, set $\theta_k = 1$, then $\beta_k^{DHF} = \beta_k^{FR}$. On the other hand, if $0 < \theta_k < 1$, then β_k^{DHF} is a proper convex combination of β_k^{HS} and β_k^{FR} . Therefore, from relation (6) and by taking the inner product with the vector y_k^T it obtains

$$d_{k+1} = -g_{k+1} + \left((1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{\|g_k\|^2} \right) \right) s_k, \quad (20)$$

$$d_{k+1}^T y_k = -g_{k+1}^T y_k + \left((1 - \theta_k) \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} \right) + \theta_k \left(\frac{\|g_{k+1}\|^2}{\|g_k\|^2} \right) \right) s_k^T y_k. \quad (21)$$

Applying y_k^T on $d_{k+1} = -\nabla^2 f(x_{k+1})^{-1} g_{k+1}$ and equating with (21), lead to the following hybridization parameter of Djordjevic (2018), which imply that d_{k+1} satisfies Newton direction:

$$\theta_k = \frac{(-s_k^T g_{k+1}) \|g_k\|^2}{(y_k^T s_k) \|g_{k+1}\|^2 - (g_{k+1}^T y_k) \|g_k\|^2} \quad (22)$$

Here, this research uses Dai-Liao conjugacy condition (17) on (21), and after some algebra, this research proposes another hybridization parameter as follows:

$$\theta_k = \frac{(-t s_k^T g_{k+1}) \|g_k\|^2}{(y_k^T s_k) \|g_{k+1}\|^2 - (g_{k+1}^T y_k) \|g_k\|^2} \quad (23)$$

However, for large-scale problems, the update parameter choices that do not require evaluation of the Hessian matrix are often required. Therefore, to have an algorithm for solving large-scale problems, this research computes the modulating parameter t from optimal choice obtain in Babaie-Kafaki and Ghanbari (2015) and Andrei (2017).

$$t^* = \frac{s_k^T y_k}{s_k^T s_k} . \quad (24)$$

3.1 Dai-Liao Hybrid Hestenes-Stiefel and Fletcher-Reeves (DHF) Algorithm

Step 1. *Initialization.* Select $x_0 \in \mathbb{R}^n$, $\varepsilon > 0$ and parameter $0 < \delta < \sigma < 1$. Compute $f(x_0)$ and g_0

Step 2. *Test for Continuation of Iterations.* If $\|g_k\| \leq \varepsilon$, then stop.

Step 3. *Line Search.* Compute $\alpha_k > 0$ satisfying Wolfe conditions (3) and (5).

Step 4. *Computation of θ_k .* If $(y_k^T s_k) \|g_{k+1}\|^2 - (g_{k+1}^T y_k) \|g_k\|^2 = 0$, then set $\theta_k = 0$; otherwise, Compute θ_k by (23) and (24).

Step 5. *Computation of β_k^{DHF} .* If $0 < \theta_k < 1$, then compute β_k^{DHF} by (19).

Step 6. *Computation of Search Direction.* Compute $d = -g_{k+1} + \beta_k^{DHF} s_k$. If restart criterion of Powell

$$|g_{k+1}^T g_k| > c \|g_{k+1}\|^2, \quad (25)$$

It is satisfied, then set $d_{k+1} = -g_{k+1}$; otherwise, define $d_{k+1} = d$. Compute α_k , set $k = k + 1$ and go to step 2.

3.2 Convergence Analysis

In this section, the convergence result of the hybrid CG method is analyzed base on strong Wolfe condition, an algorithm has to possess both sufficient descent condition and global convergence properties to be convergent.

3.2.1. Sufficient Descent Condition

Definition: Search direction satisfies descent directions (or equivalently, satisfy the decent condition) if an only if

$$d_k^T g_k < 0, \tag{26}$$

and also satisfies sufficient descent condition if and only if

$$d_k^T g_k \leq -c \|g_k\|^2, \forall k \geq 0, \tag{27}$$

where c is positive constant.

Theorem 3.1. Consider a CG method with search direction (6) and β_k^{DHF} generated by (19), then condition (27) holds.

Proof: From DHF algorithm, suppose the restart criterion of Powell (1984) condition (25) holds, then $d_k = -g_k$ and (27) holds. So, this research assumes that (25) does not hold. Then this research has

$$|g_{k+1}^T g_k| \leq c \|g_{k+1}\|^2, \text{ where } c = 0.2. \tag{28}$$

If $k = 0$, it to see that it holds; $d_0 = -g_0$, so $g_0^T d_0 = -\|g_0\|^2$, then it can be concluded that (25) holds for $k = 0$. Next is to show that it holds for $k > 0$.

By taking the inner product of (6) with vector g_{k+1}^T this research has

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \beta_k d_k^T g_{k+1}. \tag{29}$$

Firstly, suppose that $\theta_k \leq 0$, then $\beta_k = \beta_k^{HS}$, it follows from (7) and (29) with triangular inequality that

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \left| \frac{d_k^T g_{k+1}}{d_k^T y_k} \right| |g_{k+1}^T y_k|.$$

From (5), this research has

$$|d_k^T g_{k+1}| \leq -\sigma d_k^T g_k \tag{30}$$

$$d_k^T y_k = d_k^T g_{k+1} - d_k^T g_k \leq -\sigma d_k^T g_k - d_k^T g_k$$

$$\geq -(1 - \sigma) d_k^T g_k \geq 0 \tag{31}$$

Using relations (30) and (31) with the above inequality, it obtains

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{(-\sigma) d_k^T g_k}{-(1 - \sigma) d_k^T g_k} |g_{k+1}^T y_k|,$$

But from (28) and $y_k = g_{k+1} - g_k$ it holds $|g_{k+1}^T y_k| \leq \|g_{k+1}\|^2 + |g_{k+1}^T g_k| \leq 1.2 \|g_{k+1}\|^2$, so that we have

$$\begin{aligned} &\leq -\|g_{k+1}\|^2 + \frac{1.2\sigma}{(1-\sigma)}\|g_{k+1}\|^2 \\ &\leq -\left(1 - \frac{1.2\sigma}{1-\sigma}\right)\|g_{k+1}\|^2 \leq -\left(\frac{1-2.2\sigma}{1-\sigma}\right)\|g_{k+1}\|^2. \end{aligned}$$

Denote $c_1 = \left(\frac{1-2.2\sigma}{1-\sigma}\right)$, so that this research has

$$d_{k+1}^T g_{k+1} \leq -c_1 \|g_{k+1}\|^2. \quad (32)$$

Secondly, suppose that $\theta_k \geq 1$, then $\beta_k = \beta_k^{FR}$, it follows from (8) and (29) with triangular inequality that

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{\|g_k\|^2} |g_{k+1}^T d_k|.$$

Let remind the fact that (27) holds for FR method in the presence of (3) and (5) was initially mentioned in (Hager & Zhang, 2006) and later in (Djordjevic; 2018; 2019).

So, there exists a constant $c_2 > 0$, such that

$$d_{k+1}^T g_{k+1} \leq -c_2 \|g_{k+1}\|^2 \quad (33)$$

Finally, if $\theta_k \in (0,1)$, then $0 < a_1 < \theta_k < a_2 < 1$. Therefore, from (14) and (20) it can write

$$d_{k+1}^T g_{k+1} = \theta_k g_{k+1}^T d_{k+1}^{FR} + (1 - \theta_k) g_{k+1}^T d_{k+1}^{HS},$$

which implies that

$$g_{k+1}^T d_{k+1}^{DHF} = a_1 g_{k+1}^T d_{k+1}^{FR} + (1 - a_1) g_{k+1}^T d_{k+1}^{HS}.$$

Denote $c = a_1 c_2 + (1 - a_2) a_1$. Then finally, we get

$$d_{k+1}^T g_{k+1} \leq -c \|g_{k+1}\|^2. \quad (34)$$

3.3. Convergence Analysis

In this section, this research applies the following theorems to illustrate the global convergence of DHF method. It is necessary to show that θ_k and t^* are bounded. Therefore, the following basic assumptions are:

Assumption (i). The level set $S = \{x \in \mathbb{R} : f(x) \leq f(x_0)\}$, is bounded from below. That is, there exists a positive constant B such that $\|x\| \leq B, \forall x \in S$.

Assumption (ii). In a neighborhood N of S , the objective function f is continuously differentiable and its gradient $g(x)$ is Lipschitz continuous on N that is, there exist a constant $L > 0$ such that

$$\|f(x) - f(y)\| \leq L\|x - y\|, \text{ for all } x, y \in N. \quad (35)$$

Under Assumptions (i) and (ii) on f , there exist a constant Γ such that

$$\|g(x)\| \leq \Gamma, \text{ for all } x \in S. \quad (36)$$

For any conjugate gradient method with a strong Wolfe line search, the convergence holds. But, for general function, only a weak form of the Zoutendijk condition is needed (Dai & Liao, 2001).

Lemma 3.1. Let Assumptions (i) and (ii) hold. Consider the methods (2) and (6), where d_k is a descent direction and α_k satisfies (3) and (5). If

$$\sum_{k \geq 1} \frac{1}{\|d_k\|^2} = \infty, \quad (37)$$

then

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (38)$$

A CG method converges globally if $g_k = 0$ for some k or (38) holds.

Theorem 3.2. Consider the iterative method, defined by DHF method. Let d_{k+1} be a descent direction, then either $g_k = 0$, for some k , or

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (39)$$

The proof is using contradiction, that theorem (3.1) is not true.

Proof: Let $g_k \neq 0$, for all k . Then it has to prove (39). Suppose, on the contrary, that (39) does not hold, which means the gradient is bounded away from zero. Then there exists a constant $r > 0$, such that

$$\|g_k\| \geq r. \quad (40)$$

Let D be the diameter of the level set S , then

$$\|s_k\| \leq D. \quad (41)$$

Because the descent condition holds for DHF method. Since it has $d_{k+1} \neq 0$, it is sufficient to prove that d_{k+1} is bounded above, so from relation (20), it has

$$\|d_{k+1}\| = \|g_{k+1} + |1 - \theta_k| \cdot |\beta_k^{HS}| + |\theta_k| \cdot |\beta_k^{FR}| \cdot d_k\|. \quad (42)$$

But it holds from (27), (28), (31), (36), (40), and (41) that

$$|\beta_k^{HS}| = \left| \frac{g_{k+1}^T y_k}{d_k^T y_k} \right| \leq \frac{\|g_{k+1}\| \|y_k\|}{|d_k^T y_k|} \leq \frac{\Gamma L \|s_k\|}{c(1-\sigma) \|g_k\|^2} \leq \frac{\Gamma LD}{c(1-\sigma)r^2}. \quad (43)$$

$$|\beta_k^{FR}| = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \leq \frac{\Gamma^2}{r^2}. \quad (44)$$

Applying Lipschitz condition $\|y_k\| \leq L \|s_k\|$ and (24) implies that

$$|t| = \left| \frac{s_k^T y_k}{s_k^T s_k} \right| \leq \frac{\|s_k\| \|y_k\|}{\|s_k\|^2} \leq \frac{L \|s_k\|^2}{\|s_k\|^2} \leq L. \quad (45)$$

It shows that θ_k is bounded using (23), so that

$$|\theta_k| = \left| \frac{(-t s_k^T g_{k+1}) \|g_k\|^2}{(y_k^T s_k) \|g_{k+1}\|^2 - (g_{k+1}^T y_k) \|g_k\|^2} \right|$$

Since $|s_k^T g_{k+1}| \leq |g_k^T s_k| + L \|s\|^2$, $|s_k^T g_k| \leq \|g_k\| \cdot \|s_k\|$, $y_k^T s_k = y_k^T d_k$ and $|g_{k+1}^T y_k| \leq (1 - 0.2) \|g_{k+1}\|^2$, clearly, it gets

$$\begin{aligned} |\theta_k| &\leq \frac{t [\|g_k\| \cdot \|s_k\| + L \|s\|^2] \cdot r^2}{[c(1-\sigma) \|g_{k+1}\|^2 \cdot r^2 + (1-0.2) \|g_{k+1}\|^2 \cdot r^2]} \\ &\leq \frac{t [rD + LD^2] \cdot r^2}{[c(1-\sigma)\Gamma^2 + (1-0.2)\Gamma^2] \cdot r^2}, \\ &\leq \frac{t [rD + LD^2]}{\Gamma^2 [c(1-\sigma) + (0.8)]}, \end{aligned}$$

using (45) implies

$$\leq \frac{LD[r + LD]}{\Gamma^2 [c(1-\sigma) + (0.8)]}$$

$$|\theta_k| \leq A. \quad (46)$$

Finally, from (42), using (43), (44), and (45), it gets as follows

$$\begin{aligned} \|d_{k+1}\| &\leq \|g_{k+1}\| + [1 - \theta_k] \cdot |\beta_k^{HS}| + |\theta_k| \cdot |\beta_k^{HS}| \cdot \|d_k\| \\ &\leq \Gamma + \left[|1 - A| \cdot \frac{\Gamma LD}{c(1-\sigma)r^2} + A \cdot \frac{\Gamma^2}{r^2} \right] \cdot E \end{aligned}$$

$$\leq \Gamma + EF$$

$$\|d_{k+1}\| \leq \Gamma + EF. \tag{47}$$

Therefore, from

$$\sum_{k \geq 1} \frac{1}{\|d_k\|^2} = \infty.$$

Applying Lemma 3.1, this research concludes that

$$\lim_{k \rightarrow \infty} \inf \|g_k\| = 0.$$

This is a contradiction of (38), so it has been proved (39).

4. Results and Discussion

In this section, we present computational performance of DHF method and compare with Hybrid Hestenes-Stiefel and Fletcher-Reeves (HHSFR) of Djordjevic (2018) method. To implement the hybridize CG parameters, the codes were written in Matlab 9.2 (R2018a) and run on a personal computer 2.20 GHz CPU processor and 3.0 GB RAM memory. The test problems are the unconstrained problems from (Andrei, 2008) and (Gould, Orban & Toint, 2003). Since CG schemes are mainly designed to solve large-scale unconstrained optimization, we select 24 unconstrained optimization problems and tested them on a gradually increasing number of dimension(s) from 100 to 1000000 as shown in Table 1. The stopping criterion is set to $\|g_k\|_\infty \leq 10^{-4}$. Numerical results were compared based on the performance profile of Dolan and Mor'e (2002) and shown graphically in figures 1-2.

Benchmark results are generated by running a solver on a set of problems and recording information of interest such as the number of iterations and the computing time. A solver has higher efficiency when its value of $P_s(t)$ is higher. The $P_s(t)$ from the performance, the profile is the fraction of the problem with a high ratio performance t . In a set of problem P and a set of optimization solver S , a performance comparison of problem $p \in P$ by a particular algorithm $s \in S$ is measured. Let, $t_{p,s}$ be the number of iterations or CPU time required when solving a problem $p \in P$ with solver $s \in S$. The performance ratio is defined by $r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s}:s \in S\}}$. From this expression, it is assumed that $r_{p,s} \in [1, r_M]$, where $r_M \geq r_{p,s}$ and $r_{p,s} = r_M$ only when problem P is not solved by the solver. Then, graphically, a graph of $P_s(t)$ versus $t \in [1, r_M]$ is plotted. In a graph of performance profile, the smallest performance ratio is 1, and it will be located at the most left of t -axis hence, the top curve represents the most efficient method. In particular, if the set of problems P is suitably large and representative of problems that are likely to occur in applications, then solver with large probability $P_s(t)$ are to be preferred.

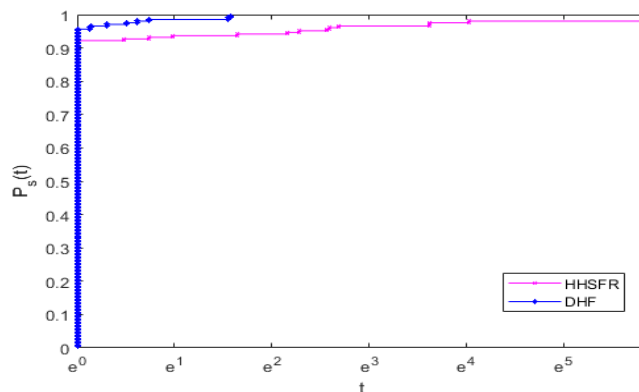


Figure 1. Number of iterations Performance Profile of HHSFR and DFR schemes

Figures 1 and Figure 2 show the hybrid coefficients' performance based on the number of iteration and central processing time per unit. The top left curved indicated fraction or percentage of how fast the coefficient converges, while the top right determines the fraction or percentage how many test functions can be tested on a given coefficient. Both figures clearly indicate that the DHF hybrid comparable and outperformed the HHSFR CG coefficient.

Digital image processing plays an important role in medical sciences, biological engineering, and other science and engineering areas. Ibrahim et al. (2020) combined Solodov and Svaiter method with the Liu-Storey and Fletcher-Reeves conjugate gradient algorithm of Djordjevic unconstrained minimization problems to propose a hybrid conjugate gradient algorithm and extend the result to solve convex constrained nonlinear monotone equations. The global convergence established and applied to solve the l_1 -norm regularized problems to restore sparse signal and image in compressive sensing. Numerical comparisons of the algorithm with some sparse signal reconstruction and image restoration in compressive sensing CG algorithms show that the proposed scheme is computationally more efficient and robust than the compared schemes. Ibrahim et al. (2020) utilized HLSFR Algorithm of Djordjevic (2019) in the restoration of one-dimensional sparse signal and image restoration using mean squared error (MSE). The performance of HLSFR won and proves to be more efficient in decoding sparse signals in compressive sensing by a lesser number of iterations, computing time, and lesser MSE by repeated experiment on 10 different noise samples. The HLSFR algorithm is similar to the HHSFR of Djordjevic (2018) and DHS algorithms here.

Table 1. List of Test Functions

NO	Function	Dimension/s	Initial Points
1	Extended White & Holst	100,200,600,1000,2000,6000,10000,20000,60000,100000,200000	(5,5,...,5)
2	Extended Rosenbrock	100,200,600,1000,2000,6000,10000,20000,60000,100000,200000	(-1.2,1, -1.2,1,..., -1.2,1)
3	Extended Freudenstein & Roth	100,200,600,1000,2000,6000,10000,20000,60000,100000,200000	(5,5,...,5)

4	Extended Beale	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000	(1.8,1.8,...,1.8)
5	Extended Tridiagonal 1	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(2,2,...,2)
6	Extended Himmelblau	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(1,1,...,1)
7	Extended Powel 1	100,200,600,1000,2000,6000,10000, 20000,60000,100000	(0,0,...,0)
8	Fletcher Function (Cute)	100,200,600,1000,2000,6000,10000, 20000	(0,0,...,0)
9	Extended Powel	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(-1,-1,...,-1)
10	Nonscomp Function (Cute)	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(3,3,...,3)
11	Extended Denschnb Function (Cute)	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(-6,-6,...,-6)
12	Extended Quadratic Penalty Qp1	100,200,600,1000,2000,6000,10000, 20000	(1,1,...,1)
13	Hager	100	(1,1,...,1)
14	Extended Maratos	100,200,600,1000,6000	(1,1,...,1)
15	Shallo	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000 500000,1000000	(-2,-2,...,-2)
16	Quadratic Qf2	100,200	
17	Generalized Tridiagonal 1	100,200,600,1000,2000	(0.5,0.5,...,0.5)
18	Generalized Tridiagonal 2	100,200,600,1000,2000,6000	(2,2,...,2)
19	Power	100,200,600,1000,2000,6000,10000	(1,1,...,1)
20	Quadratic Qf1	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000	(1,1,...,1)
21	Extended Quadratic Penalty	100,200,600,1000,2000,6000,10000	(1,1,...,1)
22	Extended Penalty	100,200,600,1000,2000	(1,1,...,1)
23	Dixon Price Function	100	(1,2,3,...)
24	Sum of Squares	100,200,600,1000,2000,6000,10000, 20000,60000,100000,200000	(-3,-3,...,-3)

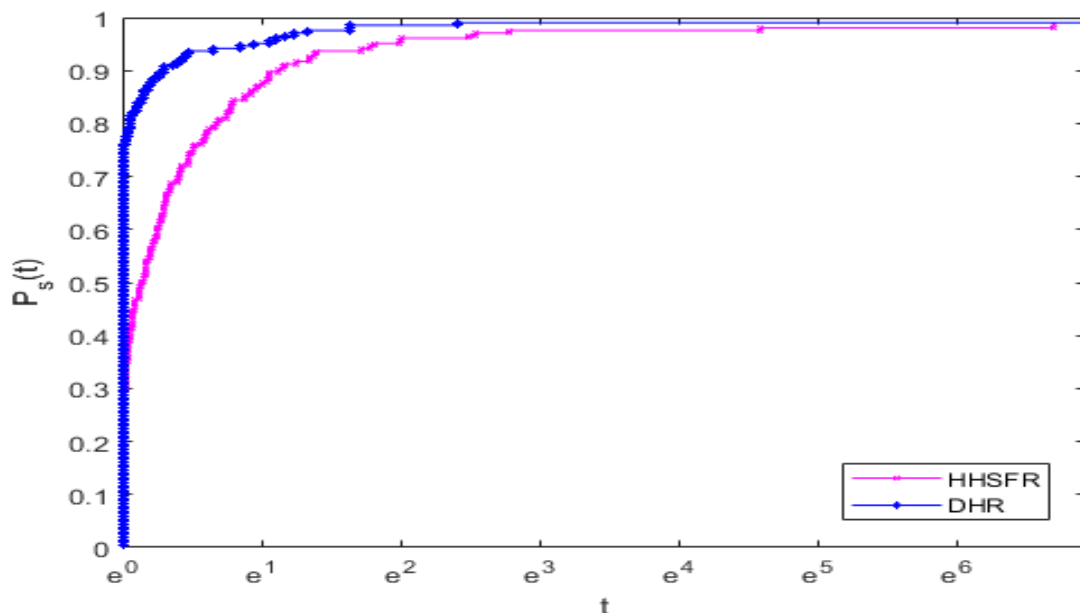


Figure 2. CPU time Performance Profile of HHSFR and DFR schemes.

5. Conclusion

This paper has presented a new hybrid conjugate hybrid algorithm in which the CG parameter is computed as a convex combination of β_k^{HS} and β_k^{FR} from Dai-Liao conjugacy condition by employing an optimal choice of the modulating parameter t . Numerical computation adopts inexact line search, which is compared with HHSFR conjugate gradient coefficient proposed by Djordjevic. The method requires the first-order derivatives but overcomes the steepest descent method's shortcoming of slow convergence and needs not to save or compute the second-order derivatives needed by the Newton method. Numerical results show that DHR coefficient outperforms the HHSFR scheme and suitable in compressed sensing. The algorithm converges globally using strong Wolfe conditions.

References

- Abubakar, A. B., Kumam, P., Mohammad, H., Awwal., A. M. & Sitthithakerngkiet, K.(2019). A modified fletcher–reeves conjugate gradient method for monotone nonlinear equations with some applications, *MPDI, mathematics*, 7, 745 doi:10.3390/math7080745.
- Al-Baali, M. (1985). Descent property and global convergence of the Fletcher Reeves method with inexact line search. *IMA Journal of Numerical Analysis*, 5, 121-124.
- Al-Namat, F.N. and Al-Naemi, G.M. (2020). Global convergence property with inexact line search for a new hybrid conjugate gradient method. *Open Access Library Journal*, 7: e6048. <https://doi.org/10.4236/oalib.1106048>

-
- Andrei, N. (2008). An unconstrained optimization test function. *Advanced Modeling and Optimization. An Electronic International Journal*, 10, 172-182.
- Andrei, N. (2017). A Dai-Liao nonlinear conjugate gradient algorithm. *Numer. Algor.*, DOI 10.1007/s11075-017-0362-5.
- Babaie-Kafaki, S. (2011). A modified BFGS algorithm based on a hybrid secant equation. *Science China Mathematics*, 58, 315-331.
- Babaie-Kafaki, S. (2013). A hybrid conjugate gradient method based on quadratic relaxation of Dai-Yuan hybrid conjugate gradient parameter. *Journal of Mathematical Programming and Operation Research*, 62(7), 929-941.
- Babaie-Kafaki, S. & Mahdavi-Amiri, N. (2013). Two hybrid conjugate gradient methods based on hybrid secant equation. *Mathematical Modeling and Analysis*, 18(1):32-52.
- Babaie-Kafaki, S. & Ghanbari, R. (2014c). Two hybrid nonlinear conjugate gradient methods based on a modified secant equation. *Journal of Mathematical Programming and Operation Research*, 63(7), 1027-1042.
- Babaie-Kafaki, S., & Ghanbari, R. (2014a). The Dai-Liao nonlinear conjugate gradient method with optimal parameter choices. *Eur. J. Oper. Res.* 234(3), 625-630.
- Babaie-Kafaki, S., & Ghanbari, R. (2014b). A descent family of Dai-Liao conjugate gradient methods. *Optim. Methods Softw.*, 29(3): 583-591.
- Babaie-Kafaki, S., & Ghanbari, R. (2015). Two optimal Dai-Liao conjugate gradient methods. *Optimization*, 64(1): 2277-2287.
- Babaie-Kafaki, S. (2015). On optimality of two adaptive choices for the parameter of Dai-Liao method. *Optim. Lett.*, DOI 10.1007/s11590-015-0965-5.
- Babaie-Kafaki, S., & Ghanbari, R. (2017). Two adaptive Dai-Liao nonlinear conjugate gradient methods. *Iran J. Sci. Technol. Trans. Sci.*, DOI 10.1007/s40995-017- 027 1-4.
- Dai, Y.H. & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.*, 10, 177-182.
- Dai, Y.H & Liao, L.Z. (2001). New conjugacy conditions and related nonlinear conjugate gradient methods, *Appl. Math. Optim.*, 43, 87-101.
- Dai, Y.H. & Yuan, Y. (2001). An efficient hybrid conjugates gradient method for unconstrained optimization. *Annals of Operations Research*, 103, 33-47.
- Djordjevic, S.S. (2016). New hybrid conjugate gradient method as a convex combination of FR and PRP methods. *Published by faculty of science and mathematics. University of Nis, Serbia, Filomat*, 31, 3083-3100. <https://doi.org/10.2298/FIL16083D>

- Djordjevic, S.S. (2017). New hybrid conjugate gradient methods as a convex combination of LS and CD methods. *Published by faculty of science and mathematics. University of Nis, Serbia*, 31(6), 1813-1825.
- Djordjevic, S.S. (2018). New hybrid conjugate gradient method as a convex combination of HS and FR conjugate gradient methods. *Journal of Applied Mathematics and Computation*, 2, 366-378. <https://doi.org/10.26855/jamc.2018.09.002>
- Djordjevic, S.S. (2019). New hybrid conjugate gradient method as a convex combination of LS and FR conjugate gradient methods. *Acta Mathematica Scientia*, 39, 214-228.
- Dolan, E.D. & Moré, J.J. (2002). Benchmarking optimization software with performance profiles. *Journal of Math. Program*, 91(2), 201-213.
- Esmaeili H., Rostami M., & Kimiaei M. (2018). Extended Dai–Yuan conjugate gradient strategy for large-scale unconstrained optimization with applications to compressive sensing. *Published by faculty of sciences and mathematics, University of Nis, Serbia, Filomat*, 32(6), 2173–2191. Available at: <http://www.pmf.ni.ac.rs/filomat>
- Fletcher, R., & Reeves, C. (1964). Function minimization by conjugate gradients. *Computational Journal*, 7, 149-154.
- Fletcher, R. (1987). *Practical methods of optimization*, (vol. 1), *Unconstrained optimization*. New York: John Wiley & Sons.
- Gould, N.I.M., Orban, D. & Toint, P.L. (2003). CUTEr: a constrained and unconstrained testing environment, revisited. *ACM Transactions on Mathematical Software*, 29(4), 373-394.
- Guo, J., & Wan, Z. (2019). A modified spectral PRP conjugate gradient projection method for solving large-scale monotone equations and its application in compressed sensing. *Hindawi Mathematical Problems in Engineering*. Volume 2019, Article ID 5261830, 17 pages. <https://doi.org/10.1155/2019/5261830>
- Hager, W.W. & Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific J. Optim.*, 2, 35-58.
- Hestenes, M.R., & Stiefel, E.L. (1952). Methods of conjugate gradients for solving linear systems. *Journal Res. Nat. Bur. Stand.*, 49, 409-436
- Ibrahim, A.H., Kumam, P., Abubakar, A. B., Jirakitpuwapat, W., & Abubakar, J. (2020). A hybrid conjugate gradient algorithm for constrained monotone equations with application in compressive sensing. *Heliyon*.e03466. <https://doi.org/10.1016/j.heliyon.2020.e03466>
- Li, D.H., & Fukushima, A. M. (2001). modified BFGS method and its global convergence in nonconvex minimization, *J. Comput. Appl. Math.* 129(1) 15-35.

- Liu, K., & Du, S. (2019). Modified three-term conjugate gradient method and its applications. *Hindawi Mathematical Problems in Engineering*. Volume 2019, Article ID 5976595, 9 pages. <https://doi.org/10.1155/2019/5976595>
- Liu, J., Du, S., & Chen, Y. (2020). A sufficient descent nonlinear conjugate gradient method for solving M-tensor equations. *Journal of Computational and Applied Mathematics*, 371112709
- Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of optimization theory and applications*. 69, 129-137.
- Lotfi, M., & Hosseini, S.M. (2019). An efficient Dai–Liao type conjugate gradient method by reformulating the CG parameter in the search direction equation. *Journal of Computational and Applied Mathematics*, 371. 112708. <https://doi.org/10.1016/j.cam.2019.112708>
- Mohammed, N. S., Mustafa M., Mohd R. & Shazlyn M. S. (2020). A new hybrid coefficient of conjugate gradient method. *Indonesian Journal of Electrical Engineering and Computer Science*, Vol. 18, No. 3, June 2020, pp. 1454-1463.
- Perry, A. (1978). A modified conjugate gradient algorithm. *Oper. Res. Tech. Notes*,26(6): 1073-1078.
- Polak, B.T. (1969). The conjugate gradient method in extreme problems. *USSR Comput. Math. Math. Phys.* 4, 94-112.
- Polyak, B.T. (1967). A general method of solving extremal problems. *Soviet Math. Doklady*, 8, 14-29.
- Powell, M.J.D. (1984). Nonconvex minimization calculations and the conjugate gradient method, *Numerical Analysis (Dundee, 1983)*, D.F. Griffiths, ed., *Lecture Notes in Mathematics*, Vol. 1066, Springer, Berlin. 122-141.
- Salihu, N., Odekunle, M., Waziri, M. & Halilu, A. (2020). A new hybrid conjugate gradient method based on secant equation for solving large scale unconstrained optimization problems. *Iranian Journal of Optimization*, 12(1): 33-44.
- Waziri, M. Y., Ahmed, K., & Sabi'u, J. (2019). A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations. *Arab. J. Math.* <https://doi.org/10.1007/s40065-019-0264-6>.
- Xue, W., Ren, J., Zheng, X., Liu, Z., & Liang, Y. (2018). A New DY conjugate gradient method and applications to image denoising. *IEICE TRANS. INF. & SYST.*, VOL.E101–D, NO.1 2, December 2018. DOI: 10.1587/transinf.2018EDP7210
- Yao, S., Feng, Q., Li, L., & Xu, J. (2019). A class of globally convergent three-term Dai-Liao conjugate gradient methods. *Applied Numerical Mathematics*, 151 354-366. <https://doi.org/10.1016/j.apnum.2019.12.026>.
- Zoutendijk, G. (1970). Nonlinear programming computational methods, in integer and nonlinear programming. *J. Abadie ed.*, North-Holland, Amsterdam 37-86.

This page is intentionally left blank.