A dai-liao hybrid conjugate gradient method for unconstrained optimization

Nasiru Salihu[a]*, Mathew Remilekun Odekunle[a], Mohammed Yusuf Waziri[b], Abubakar Sani Halilu[c], Suraj Salihu[d]

[a] Department of Mathematics, School of Physical Science, Moddibo Adama University of Technology, Yola, Nigeria

[b] Department of Mathematical Sciences, Bayero University, Kano, Nigeria

[c] Department of Mathematics and Computer Science, Sule Lamido University, Kafin Hausa, Nigeria

[d] Department of Computer Science, Faculty of Science, Gombe State University, Gombe, Nigeria

nasirussalihu@gmail.com; remi_odekunle@yahoo.com; mywaziri.mth@buk.edu.ng; abubakars.halilu@slu.edu.ng; surajsalihu@gmail.com

*Corresponding Author: nasirussalihu@gmail.com

ARTICLE INFO

ABSTRACT

One of todays' best performing CG methods is Dai-Liao (DL) method which depends on non-negative parameter t and conjugacy conditions for its computation. Although numerous optimal selections for the parameter were suggested, the best choice of t remains a subject of consideration. The pure conjugacy condition adopts an exact line search for numerical experiments and convergence analysis. Though, a practical mathematical experiment implies using an inexact line search to find the step size. To avoid such drawbacks, Dai and Liao substituted the earlier conjugacy condition with an extended conjugacy condition. Therefore, this paper suggests a new hybrid CG that combines the strength of Liu and Storey and Conjugate Descent CG methods by retaining a choice of Dai-Liao parameter t that is optimal. The theoretical analysis indicated that the search direction of the new CG scheme is descent and satisfies sufficient descent condition when the iterates jam under strong Wolfe line search. The algorithm shown to converge globally using standard assumptions, where the numerical experimentation of the scheme demonstrated that the proposed method is robust and promising than some known methods applying the performance profile presented by Dolan and Mor´e on 250 unrestricted problems. Numerical assessment of the tested CG algorithms with sparse signal reconstruction and image restoration in compressive sensing problems, file restoration, image video coding and other applications show that these CG schemes are comparable and can be apply in different fields such as temperature, fire, seismic sensors and humidity detectors in forest and so on using the wireless sensor network techniques.

1. Introduction

Conjugate Gradient (CG) method was initially suggested for solving linear system of equation. Subsequently, the solution of a linear system is comparable to minimizing a positive definite quadratic function, for this reason (Babaie-Kafaki & Ghanbari, 2014a), the method was later modified to solve unconstrained minimization problems (Rao, 2009). Therefore, the method constitutes an excellent choice for solving optimization problems by scientists, engineers and mathematicians (Babaie-Kafaki, 2011). The method is categorized by absent of matrix storage with powerful theoretical properties (Djordjevic, 2017). The problem as the form:

$$min\, f(x), \quad (1)$$
$$x \in R^n$$

where $f: R^n \rightarrow R$ is a function that is twice continuously differentiable, and the CG scheme that iteratively solves the problem is given

$$x_0 \in R^n \quad (2)$$
$$x_{k+1} = x_k + \alpha_k d_k,$$

where $\alpha_k > 0$ is a step-size obtained by a suitable line search and d_k here is a search direction (Gilbert & Nocedal, 1992).

Generally, the distance to move along the search direction d_k can be attained by solving one-dimensional minimization called an exact line search such that the objective function is minimized to find α_k, that is,

$$f(x_k + \alpha_k d_k) = min\, \alpha > 0 \ (x_k + \alpha\, d_k) \quad (3)$$

However, for large scale problems usually, an exact line search is not possible so any value of α_k that satisfies certain properties called Wolfe conditions is accepted (Nocedal & Wright, 2006):

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (4)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma\, g_k^T d_k. \quad (5)$$

where $0 < \delta < \sigma < 1$, and d_k that is a path towards minimum needs to be descent (Babaie-Kafaki & Ghanbari, 2014c). Whereas other value of α_k constitute up of (4) and

$$|g(x_k + \alpha_k d_k)^T d_k| \leq -\sigma\, g_k^T d_k \quad (6)$$

known strong Wolfe condition is also accepted (Nocedal & Wright, 2006).

Therefore, the direction towards minimum can be obtained by the formula

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad (7)$$

where β_k is a scalar CG (update) parameter which is determined by some inner products (Ding, et al., 2010). Mainly, the CG schemes vary by the selection of β_k coefficient. Some well-known CG schemes can be divided into two (Babaie-Kafaki, et al,2010). The schemes in the first category may perform poorly theoretically but numerically behave well due to an important feature known as restart that helps them circumvent jamming automatically (Babaie-Kafaki & Ghanbari, 2014c). These CG parameters were initially suggested by Hestenes and Stiefel (1952), Polak, Ribie're and Polyak (1967), Liu and Storey (1991) with the following coefficients, respectively:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \qquad \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \qquad \beta_k^{LS} = -\frac{g_{k+1}^T y_k}{d_k^T g_k}, \tag{8}$$

where $\|.\|$ symbolizes Euclidean norm and define $s_k = x_{k+1} - x_k$ and $g_k = g_{k+1} - g_k$ (Andrei, 2008a).

The other category is prone to poor numerical performance as a result of jamming, but they have powerful theoretical properties (Andrei, 2008b). These schemes were earlier proposed by Fletcher and Revees (1964), Dai and Yuan (1991) and Fletcher (1987) with the following CG parameters respectively:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \qquad \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}, \qquad \beta_k^{CD} = -\frac{\|g_{k+1}\|^2}{d_k^T g_k}. \tag{9}$$

The schemes in (9) vary with other β_k selections in theory because their theoretical properties require Lipchitz assumption only but not including boundedness assumption (Hager, & Zhang, 2006). The FR method's poor practical performance is associated to taking tiny steps without meaningful progress to reach the minimum (Powell, 1984). Specifically, if a bad path is taken, then tiny steps from x_{k-1} and x_k will be generated, the next path along d_k step α_k are likely to be poor except a restart along the gradient direction is made (Babaie-Kafaki, 2013). Babaie-Kafaki, et al. (2011) Pointed out that despite such deficiency, the FR method was proved to be theoretically powerful with exact line search on general functions; later this result was extended to an in-exact line search to improve the efficiency of the scheme (Hager & Zhang, 2006). In general, the performance of the methods in the first category is efficient but their convergence is uncertain (Hager & Zhang, 2006). The behavior of these schemes needs to be improved to avoid jamming (Djordjevic, 2017). Therefore, Researchers were interested in combining CG schemes of the two schemes (Babaie-Kafaki & Mahdavi-Amiri, 2013). Although, CD scheme is closely related to FR scheme with exact line search but, the restriction $c < \frac{1}{2}$ in FR is not required for CD to attain sufficient descent using strong Wolfe line condition. Moreover, CD scheme is theoretically powerful for the generalized Wolfe conditions with $\delta < 1$ and $\sigma = 0$ restrictions (Hager & Zhang, 2006). Meanwhile, Djordjevic (2017) pointed out that no much research has been done on the choice of β_k^{LS} except for the work of Liu & Storey (1991) initially, but the analysis of PRP techniques should be applied to the LS method (Hager & Zhang, 2006). Since LS and PRP schemes are identical when an exact line search is used (Dai, 2001).

Babaie-Kafaki & Ghanbari (2014c) suggested two Hybrid Conjugate Gradient (HCG) methods where the CG coefficients were obtained from standard and modified secant equations respectively. The parameters are calculated as an affine combination of β_k^{HS} and β_k^{DY}. While Djordjevic (2017) suggested a hybrid parameter using Liu & Storey (1991) with Conjugate

Descent CG parameters Convex Combination (LSCDCC) from conjugacy condition as an affine combination of β_k^{LS} and β_k^{CD}. To achieve global convergence for general function, HCG adopted $\beta_k \geq 0$ restriction, while the other β_k coefficient is hypothetically superior for uniformly convex function. Subsequently, the selection of β_k^{LS} CG parameter received little attention by researchers except for the work of (Djordjevic, 2017; Salihu, et al, 2020) recently motivated this work. Given the above, a large number of hybrid conjugate gradients techniques were proposed (Andrei, 2008a) that modified different β_k coefficients to maximize their strengths and minimize their weaknesses (Babaie-Kafaki, et al. 2010). Among them: (e.g. see (Yuan, 1991; Gilbert & Nocedal, 1992; Andrei,2008c; Dai & Yuan, 2001; Sabiu, & Waziri, 2017; Sabiu, et al, 2017). The excellent contributions of Andrei and Babaie-Kafaki on hybridization using convex combination and that of Djordjevic motivated us to extend their approaches to access and combine the strength of LS and CD CG update parameters.

2. Extended Conjugacy Condition of Dai and Liao CG Method (ECCDL)

In the earlier CG methods; conjugacy condition $d_{k+1}^T y_k = 0$ that rests on the exact line search plays a significant part in the mathematical experiment and convergence analysis (Sun & Yuan, 2006). Though, a practical mathematical experiment implies using an inexact line search to find the step-size α_k. Especially in a situation where $d_{k+1}^T y_k$ is not equal to zero, then it may take other form, to avoid such defect Dai & Liao (2001) substituted the pure conjugacy condition with extended conjugacy condition. Due to simpler structure and low memory requirements of Dai-Liao conjugate gradient methods; Yao, et al. (2019) proposed some three-term Dai-Liao CG algorithms that possess efficient conjugate gradient structures. Esmaeili, et al. (2018) suggested a new CG scheme to solve problems emerged from astronomical imaging, file restoration, image video coding and other applications. Similarly, Guo & Wan (2019) developed CG algorithm for sparse engineering signal problem. Numerical tests indicated that the algorithm is an alternative for recovering sparse signal problems and beats earlier methods. Recently, Liu, et al. (2020) transformed nonlinear unconstrained optimization problems as M-tensor equation to solve real-life issues originating from engineering and economics. Numerical results revealed that the proposed CG scheme is efficient than some known methods. One of todays' best performing CG method is Dai & Liao (DL) method which depends on non-negative parameter *t* for its computation. Although numerous optimal selections for the parameter were suggested as in (Babaie-Kafaki, & Ghanbari 2014b; Babaie-Kafaki,2015; Babaie-Kafaki & Ghanbari, 2017; Waziri, et al, 2019; Salihu, et al, 2021), the best choice of *t* still remains subject of consideration (Babaie-Kafaki, & Ghanbari, 2015). Motivated by the above, in this section, using similar approach in (Andrei, 2008b; Andrei, 2009; Babaie-Kafaki, et al, 2010; Djordjevic, 2017), this research will combine the attractive features of CG update parameters proposed by Liu and Storey (1991) with CG parameter proposed by Fletcher (1987) conjugate descent using Extended Conjugacy Condition of Dai and Liao (2001) CG method called (ECCDL) as follows:

$$\beta_k^{ECCDL} = (1 - \theta_k)\, \beta_k^{LS} + \theta_k\, \beta_k^{CD}. \tag{10}$$

From relations (8) and (9), we can write (10) as

$$\beta_k^{ECCDL} = (1 - \theta_k)\left(-\frac{g_{k+1}^T y_k}{d_k^T g_k}\right) + \theta_k\left(-\frac{g_{k+1}^T g_{k+1}}{d_k^T g_k}\right), \tag{11}$$

Therefore, using vector y_k^T on relations (7) and (10) we obtain

$$d_{k+1} = -g_{k+1} + \left((1 - \theta_k)\left(-\frac{g_{k+1}^T y_k}{d_k^T g_k}\right) + \theta_k\left(-\frac{g_{k+1}^T g_{k+1}}{d_k^T g_k}\right)\right)d_k, \tag{12}$$

$$d_{k+1}^T y_k = -g_{k+1}^T y_k - \frac{g_{k+1}^T y_k}{d_k^T g_k}d_k^T y_k + \theta_k\left(\frac{g_{k+1}^T y_k}{d_k^T g_k} - \frac{g_{k+1}^T g_{k+1}}{d_k^T g_k}\right)d_k^T y_k. \tag{13}$$

Applying $d_{k+1}^T y_k = 0$ on (13) will lead to the following hybridization parameter in (Djordjevic ,2017):

$$\theta_k = -\frac{(g_{k+1}^T y_k)(d_k^T g_{k+1})}{(g_{k+1}^T g_k)(d_k^T y_k)} . \tag{14}$$

Similarly, if we apply Dai-Liao extended conjugacy condition:

$$d_{k+1}^T y_k = -t d_k^T g_{k+1}, \; where \; t \geq 0 \tag{15}$$

on (13) and after some algebra another new hybridization parameter is propose as:

$$\theta_k = \frac{t(d_k^T g_{k+1})(d_k^T g_k) - (g_{k+1}^T y_k)(d_k^T g_{k+1})}{(g_{k+1}^T g_k)(d_k^T y_k)}. \tag{16}$$

The justification of the choice of the method in this work, an algorithm with update parameter that do not require the calculation of the Hessian matrix for solving large scale problems is preferred. For this reason, we assume that θ_k does not satisfies $d_{k+1}^T y_k = 0$. Therefore, the beautiful structures of CG update parameters proposed by Liu and Storey (1991) with CG parameter suggested by Fletcher conjugate descent using Extended Conjugacy Condition of Dai and Liao (ECCDL) CG method is proposed in such a way that, if the modulating parameter $t = 0$, then (16) reduces to the method in (Djordjevic, 2017). So, as for the optimal choice of the method we assume $t \neq 0$ in the term.

Next the algorithm of the proposed method is presented as follows:

Algorithm 1 (ECCDL).

Step 1. Guess $x_0 \in R^n$ and parameter $0 < \delta < \sigma < 1$. Compute $f(x_0)$ and g_0.

Step 2. If $\|g_k\|_\infty \leq 10^{-5}$, then stop.

Step 3. Compute $\alpha_k > 0$ satisfying (4) and (6).

Step 4. If $(g_{k+1}^T g_k)(d_k^T y_k) = 0$, then set $\theta_k = 0$; otherwise, compute θ_k by (16).

Step 5. If $0 < \theta_k < 1$, then compute β_k^{ECCDL} by (10).

Step 6. Compute d_k from (7) using (10). Set $d_{k+1} = -g_{k+1}$ if restart principle of Powell (1984) is fulfilled, then

$$|g_{k+1}^T g_k| \geq 0.2 \|g_{k+1}\|^2, \tag{17}$$

otherwise, define $d_{k+1} = d$. Compute α_k, go to Step 2 and set $k = k + 1$.

Remark: The update parameter θ_k computed by (16) may be outside interval [0, 1]. However, to have a proper convex combination in (10)-(11). The following rules are applied: if $\theta_k \leq 0$, set $\theta_k = 0$, i.e, $\beta_k^{ECCDL} = \beta_k^{LS}$; if $\theta_k \geq 1$ set $\theta_k = 1$, i.e, $\beta_k^{ECCDL} = \beta_k^{CD}$. Therefore, under this selection for θ_k, the direction d_{k+1} in (12)-(13) is a proper convex combination of β_k^{CD} and β_k^{LS}.

3. Theoretical Analysis

To demonstrate the sufficient descent condition of the ECCDL method, we apply the following theorem.

Theorem 3.1 Suppose ECCDL algorithm generates sequences $\{g_k\},\{x_k\}$ and $\{\alpha_k\},\{d_k\}$. Then the search direction d_k satisfies:

$$d_k^T g_k \leq -c\|g_k\|^2, \ for \ all \ k \geq 0, \tag{18}$$

where $c = (1 - 2.2\sigma)$.

Proof: From ECCDL algorithm, if the criterion in (17) holds, then it is easy to see that $d_k = -g_k$ and (18) also holds. Assume that the condition in (17) does not hold. Then

$$|g_{k+1}^T g_k| < 0.2 \|g_{k+1}\|^2. \tag{19}$$

The proof is using induction. For $k = 0$; $d_0^T g_0 = -\|g_0\|^2$. Since $c < 1$, it is to see that (18) is satisfied. Subsequently, for $k \geq 1$ it not difficult from (6) to show that (18) holds and we have

$$d_k^T y_k = d_k^T g_{k+1} - d_k^T g_k \geq -(1-\sigma)d_k^T g_k \geq 0. \tag{20}$$

Applying vector g_{k+1}^T on (7), we get

$$d_{k+1}^T g_{k+1} = \|g_{k+1}\|^2 + \beta_k d_k^T g_{k+1}. \tag{21}$$

Firstly, when $\theta_k \geq 1$, from Step 5 we have $\beta_k = \beta_k^{CD}$, it follows easily from (6), (9) and (21) that

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

$$\leq -(1 + \sigma)\|g_{k+1}\|^2. \tag{22}$$

Secondly, we obtain $\beta_k = \beta_k^{LS}$ when $\theta_k \leq 0$ according to Step 5 so that from (6), (8) and (21), it holds that

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

$$\leq -(1 - 1.2\sigma)\|g_{k+1}\|^2. \tag{23}$$

Finally, if $\theta_k \in (0,1)$, then θ_k is computed by (10), from (6) and (19) we have

$$d_{k+1}^T g_{k+1} \leq -\|g_{k+1}\|^2 + |\beta_k^{LS}| \cdot |d_k^T g_{k+1}| + |\beta_k^{CD}| \cdot |d_k^T g_{k+1}|$$
$$\leq -\|g_{k+1}\|^2 + \sigma|\beta_k^{LS}| \cdot |d_k^T g_k| + \sigma|\beta_k^{CD}| \cdot |d_k^T g_k| = -\|g_{k+1}\|^2 + \sigma|g_{k+1}^T y_k| + \sigma\|g_{k+1}\|^2$$
$$\leq -\|g_{k+1}\|^2 + 2\sigma\|g_{k+1}\|^2 + 0.2\sigma\|g_{k+1}\|^2 \leq -(1 - 2.2\sigma)\|g_{k+1}\|^2.$$

Then we have

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

This inequality implies (18) is satisfied for $k + 1$. ∎

3.1. Convergence Analysis

The following assumptions are required to establish the global convergence of ECCDL method:

Boundedness Assumptions: Assumption 3.1. *The set $S = \{x \in R : f(x) \leq f(x_0)\}$ is bounded from below where x_0 is the starting point of CG method in (2) and (7). That is, there exist a positive constant B such that*

$$\|x\| \leq B, \forall\, x \in S. \tag{25}$$

Lipschitz Assumptions: Assumption 3.2. *In a neighborhood N of S, the objective function f is continuously differentiable and its gradient $g(x)$ is Lipchitz continuous on N, that is, there exist a constant $L > 0$ such that*

$$\|g(x) - g(y)\| \leq L\|x - y\|, \tag{26}$$

for all $x, y \in N$.

Under Assumptions 3.1 and 3.2 on f, there exist a constant $\Gamma > 0$ such that

$$\|g(x)\| \leq \Gamma, \text{ for all }\, x \in S. \tag{27}$$

Lemma 3.1 (Djordjevic,2017). Let $f \in (R^n)$ and d_k be a descent direction in the point x_k. Suppose that the function f is bounded from below along direction $\{x_k + \alpha d_k \mid \alpha > 0\}$ and if $0 < \delta < \sigma < 1$,

then there exist an interval inside which the step size satisfies Wolfe conditions and strong Wolfe condition

Lemma 3.2 Suppose Assumption 3.1 holds, consider ECCDL Algorithm, where d_k is a descent direction, and step length α_k is obtained using the Wolfe conditions. Then.

$$\sum_{k \geq 1} \frac{(d_k^T g_k)^2}{\|d_k\|^2} < +\infty. \tag{28}$$

Remark: Since α_k satisfies (4) and (6), it must also hold for (4) and (5). Consequently, Lemma 3.2 also holds for (4) and (6) equally.

Lemma 3.3 Let Assumption 3.1 holds and d_k is a descent direction and the step-size α_k satisfies

$$g_{k+1}^T d_k \geq \sigma d_k^T g_k, \ \sigma < 1, \tag{29}$$

then

$$\alpha_k \geq \frac{(1-\sigma)}{L} \cdot \frac{|d_k^T g_k|}{\|d_k\|^2}. \tag{30}$$

Proof: Using (26) and (29), it holds that
$$-(1 - \sigma)d_k^T g_k \geq d_k^T(g_{k+1} - g_k) \leq L\alpha_k \|d_k\|^2.$$
Subsequently, when $\sigma > 1$ and the search direction is $d_k^T g_k < 0$, it is not difficult to claim that (30) holds. Clearly, from relations (6) and (18), the step-size α_k in the ECCDL algorithm fulfils (30). Consequently, according to (18) and since $\alpha_k = 0$ does not fulfil (6), we can easily conclude that $d_k^T g_k \neq 0 \ \forall \ k \geq 0$, which means $\alpha_k \neq 0$ and as such there exist $\mu > 0$ so that

$$\alpha_k \geq \mu, \forall \ k \geq 0. \tag{31}$$

Generally, any CG method with strong Wolfe line search converges. However, only weak form of the (Zoutendijk, 1970) condition is required for general function (Dai & Liao, 2001). The theorem below establishes the useful theoretical property of ECCDL through the strong Wolfe line search.

Theorem 3.2 Let Assumptions 3.1 and 3.2 hold. Let the ECCDL algorithm generates sequences $\{d_k\}$ and $\{g_k\}$. Then, Zoutendijk condition below holds

$$inf \ \|g_k\| = 0. \tag{32}$$

The proof is using contradiction, that theorem (3.2) is not true.

Proof: Let $g_k \neq 0$ and assume that (32) does not hold. Then, we have a constant $\epsilon > 0,$ such that

$$\|g_k\| \geq \epsilon. \tag{33}$$

Let $\|s_k\| \leq D$. Then, from (10), (18) and (33) we have

$$\left| \beta_k^{ECCDL} \right| \leq \left| \beta_k^{LS} \right| + \left| \beta_k^{CD} \right| = \frac{\left| g_{k+1}^T y_k \right|}{\left| d_k^T g_k \right|} + \frac{\| g_{k+1} \|^2}{\left| d_k^T g_k \right|}$$

$$\leq \frac{\| g_{k+1} \| \cdot \| y_k \| + \| g_{k+1} \|^2}{c \| g_k \|^2} \leq \frac{\Gamma L D + \Gamma^2}{c \epsilon^2} = P. \tag{34}$$

It follows from the first inequality that $\theta_k \in (0,1)$ and subsequently the second inequality holds from the Cauchy Schwartz inequality. But when $\theta_k \notin (0,1)$, it is easy to get the above inequality according to β_k selection in Step 5 of algorithm 1. Thus, from (7) and (34) we get

$$\| d_{k+1} \| \leq \| g_{k+1} \| + |\beta_k| \cdot \| d_k \| = \| g_k \| + \frac{|\beta_k| \cdot \| s_k \|}{\alpha_k} \leq \Gamma + \frac{PD}{\mu} = M, \tag{35}$$

and this indicates that

$$\sum_{k \geq 0} \frac{1}{\| d_k \|^2} + \infty. \tag{36}$$

Contrarily, from (18), (28) and (33), we can achieve

$$c^2 \epsilon^4 \sum_{k \geq 0} \frac{1}{\| d_k \|^2} \leq \sum_{k \geq 0} \frac{c^2 \| g_k \|^4}{\| d_k \|^2} \leq \sum_{k \geq 0} \frac{(d_k^T g_k)^2}{\| d_k \|^2} < +\infty. \tag{37}$$

Also, applying Lemma (3.1), we conclude that

$$inf \ \| g_k \| = 0. \tag{38}$$

Obviously, this is a contradiction of (36) and hence (33) is not satisfied which implies that (32) is proved. ∎

4. Results and Discussion

In engineering, medical sciences, biological and other areas of science; digital image processing plays an important role. Therefore; Ibrahim, et al. (2020) utilized Hybrid Liu and Storey and Fletcher and Revees (HLSFR) algorithm of Djordjevic (2019) in restoring one dimensional signal sparse problem using mean squared error (MSE) with the LS and FR CG algorithm to suggest a hybrid algorithm for unconstrained minimization problems and extend the result to convex monotone equations. Numerical assessments with some image restoration in compressive sensing CG algorithms show that the proposed scheme is efficient and promising than other schemes with smaller number of iterations, computing time and MSE on different noise sample problems. The HLSFR algorithm is the foundation of Ibrahim et al. (2020) work and is similar to Hybrid Hestenes and Stiefel and Fletcher and Revees (HHSFR) algorithm of Djordjevic (2018), LSCDCC of Djordjevic (2017) and ECCDL algorithms which can be applied to compressive sensing problems that has wide variety of applications as shown in Ibrahim et al.

(2020) for example; the wireless sensor networks that are usually placed in field can be used in temperature, fire, seismic sensors and humidity detectors in forest, etc.

Therefore, in this section, we present the performance of ECCDL and compare with that of LSCDCC of Djordjevic (2017) and HCG method of Babaie-Kafaki & Ghanbari (2014c). To implement the hybridize CG parameters, the codes were run on a computer with a processor and memory of $2.20\ GHz\ CPU$ and $3.0\ GB\ RAM$, respectively, using Matlab $8.3\ (R2014a)$ on 250 unconstrained optimization problems. The test problems are the unconstrained problems obtained from (Andrei, 2008c; and Gould, et al, 2003). Since CG schemes are used to solve largescale unconstrained optimization, we choice 25 problems that are tested 10 times for: 100; 200; 500; 1,000; 2,000; 5,000; 10,000; 20,000; 50,000 and 100,000 with summary of the numerical results and list of test functions shown in table 1-2 respectively. All the algorithms were implemented using (4) and (6) with $\delta = 0.0001$ and $\sigma = 0.001$, and the step length is computed with initial trail value $\alpha_k = 1$ and the modulating parameter $t = 0.5$ .The same stopping criterion $\|g_k\|_\infty \leq 10^{-5}$ is used. All the test functions were minimizing from standard starting points.

Test function results are obtained by running a solver on set of problems and recording the number of iterations and the computing time. Interpretation of figures 1-2 show the performance of these methods using (Dolan & Mor´e, 2002) profile. The $P(\tau)$ is the portion of problems with performance ration τ, thus, a solver with high values $P(\tau)$ or at the top right of the figures are preferable. That is, for each method, we plot the percentage $P(\tau)$ of the problems for the best time for each algorithm within a factor of $P(\tau)$ versus time τ. The left side gives the percentage of the test problems of the method that is fastest. The right side gives the percentage of the test functions successfully solved by each method. The interpretation of figure 1 shows the probability of ECCDL method is the winner on a given problem is 61%. While LSCDCC and HCG methods win 39% and 15% percentages respectively, when the factor τ is chosen within the interval $0 < \tau < 0.5$. Clearly, ECCDL method has the most wins, because it has the highest probability of being closer to the optimal solution. However, if we extend our τ of interest to $\tau \geq 0.5$, ECCDL and HCG algorithms solved the test functions in a given time and reach 88% and 87% respectively, while LSCDCC method is 85% to. It is easy to see that the performance of ECCDL and HCG algorithms are computationally efficient than LSCDCC scheme.

Since the computing time is also affected by the computer atmosphere like operating system and busy status, we additionally compare the number of iterations of the algorithms. Figure 2 shows that the fraction of ECCDL method is the winner on a given problem is 82%. While LSCDCC and HCG methods win 80% and 73% percentages respectively, when the factor τ is chosen within the interval $0 < \tau < 0.5$ Clearly, ECCDL method wins, since it has the highest possibility of being closer to the best solution. However, if we extend our τ of interest to $\tau \geq 0.5$, ECCDL and HCG algorithms solved the test functions in a given number of iterations and reach 88% and 87% respectively, while LSCDCC method is 85%, it is easy to see that the performance of ECCDL and HCG algorithms are computationally efficient than LSCDCC scheme.

Table 1: Summary of Numerical Results of ECCDL, LSCDCC and HCG Methods

	ECCDL	LSCDCC	HCG	
Successful	105	20	33	CPU Time
	199	177	178	No. of Iteration
Not Successful	116	200	180	CPU Time
	21	36	35	No. of Iteration
Failure	29	30	37	CPU Time
	30	37	36	No. of Iteration
Total	250	250	250	CPU Time
	250	250	250	No. of Iteration

Table 2. List of Test Functions

| NO	Faction
1	Extended White & Holst
2	Extended Rosenbrock
3	Extended Freudenstein & Roth
4	Extended Tridiagonal 1
5	Extended Himmelblau
6	Extended Powel 1
7	Fletcher Function (Cute)
8	Extended Powel
9	Nonscomp Function (Cute)
10	Extended Denschnb Function (Cute)
11	Extended Quadratic Penalty Qp1
12	Hager
13	Extended Maratos
14	Shallo
15	Quardratic Qf2
16	Generalized Tridiagonal 1
17	Generalized Tridiagonal 2
18	Power
19	Quadratic Qf1
20	Extended Quadratic Penalty
21	Extended Penalty
22	Extended Beale
23	Raydan
24	Diagonal 1
25	Quadratic Qf2
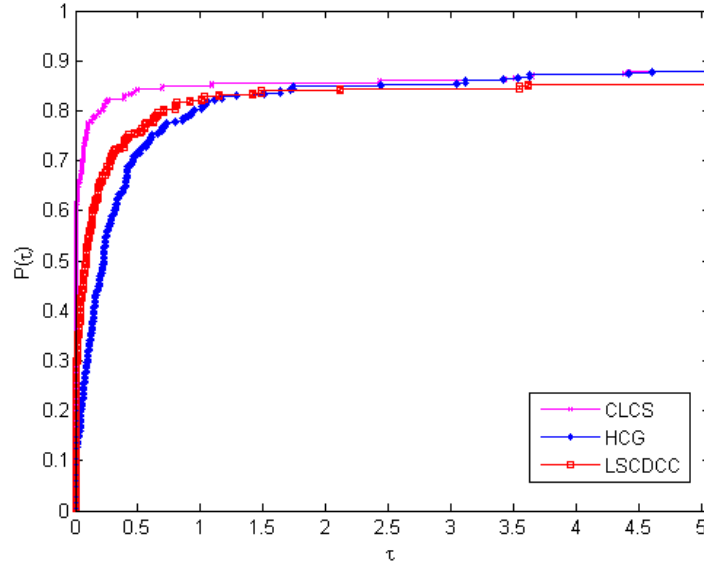
Figure 1: CPU time performance profile for ECCDL, HCG and LSCDCC schemes.
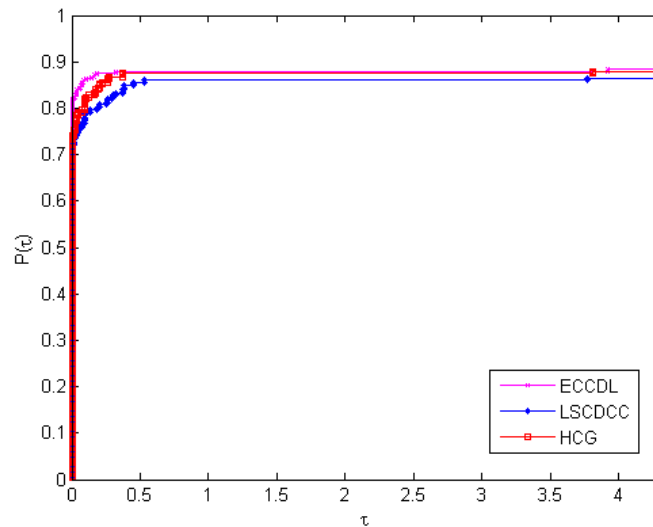

Figure 2: Number of iterations performance profile for ECCDL, HCG and LSCDCC schemes.

5. Conclusion

One of todays' best performing CG method is DL method which depends on non-negative parameter t for its computation. Although numerous optimal selections for the parameter were suggested, the best choice of t remains a subject of consideration. In this paper, we have presented a new hybrid Dai-Liao conjugate algorithm in which the parameter β_k is computed from β_k^{LS} and β_k^{CD} in such a way that if the modulating parameter $t = 0$ then it reduces to the method

that uses the pure conjugacy condition. Theoretical and numerical computations adopt inexact line search when compared with some known CG coefficients using strong Wolfe condition show the algorithm is robust, efficient and converge globally compared to LSCDCC and HCG methods on 250 unconstrained optimization problems. Numerical assessments of these CG algorithms show that the schemes are comparable with smaller iterations and computing time and can be applied to compressive sensing problems with a wide variety of applications, as shown in Ibrahim et al.(2020) method.

References

Andrei, N. (2008a). Another hybrid conjugate gradient algorithm for unconstrained optimization, *Numerical Algorithm,* 47, 143-156.

Andrei, N. (2008b). A hybrid conjugate gradient algorithm for unconstrained optimization as convex combination of Hestenes- Stiefel and Dai-Yuan. *Studies in Informatics and Control,* 17, 55-70.

Andrei, N. (2008c). An unconstrained optimization test function collection. *Advanced Modeling and Optimization*, 10, 147-161.

Andrei, N. (2009). Hybrid conjugate gradient algorithm for unconstrained optimization. *Journal of Optimization Theory Application,* 141, 249-264.

Babaie-Kafaki, S., Ghanbari, R. & Mahdavi-Amiri, N. (2010). Two new conjugate gradient methods based on modified secant equations. *Journal of Computational and Applied Mathematics ,* 234, 1374-1386.

Babaie-Kafaki, S. (2011). A modified BFGS algorithm based on a hybrid secant equation. *Science China Mathematics,* 54, (9), 2019-2036.

Babaie-Kafaki, S., Fatemi, M., & Mahdavi-Amiri, N. (2011). Two effective hybrid conjugate gradient algorithms based on modified BFGS updates. *Science China Math.*, *58,* 315-331.

Babaie-Kafaki, S. (2013). A hybrid conjugate gradient method based on quadratic relaxation of Dai-Yuan hybrid conjugate gradient parameter. A *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. (2014a). The Dai-Liao nonlinear conjugate gradient method with optimal parameter choices. *European Journal of Operation Re*search, 234(3), 625-630.

Babaie-Kafaki, S., & Ghanbari, R. (2014b). A descent family of Dai-Liao conjugate gradient methods. *Optimization and Methods Software*, 29(3): 583-591.

Babaie-Kafaki, S., & Ghanbari, R. (2014c). Two hybrid nonlinear conjugate gradient methods based on modified hybrid secant equation. *Journal of Mathematical Programming and Operation Research, 63*(7), 1027-1042.

Babaie-Kafaki, S. (2016). On optimality of two adaptive choices for the parameter of Dai- Liao method. *Optimization Letters,* 10, 1789-1797.

Babaie-Kafaki, S., & Ghanbari, R. (2015). Two optimal Dai-Liao conjugate gradient methods. *Optimization*, 64(11): 1-11.

Babaie-Kafaki, S., & Ghanbari, R. (2017). Two adaptive Dai-Liao nonlinear conjugate gradient methods. *Iranian Journal of Science and Technology Transactions A: Science,* 42, 1505-1509.

Dai, Y.H., & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on Optimization*,*10(1)*, 177-182.

Dai, Y. H. (2001). New properties of nonlinear conjugate gradient method, *Numerical Mathematics,* 89(1), 83-98.

Dai, Y.H., & Liao, L.Z. (2001). New conjugacy conditions and related nonlinear conjugate gradient methods. *Applied Mathematics and Optimization,* 43(1), 87-101.

Dai, Y.H., & Yuan, Y. (2001). An efficient hybrid conjugates gradient method for unconstrained optimization. *Annals of Operations Research,* 103, 33-47.

Ding, Y., Lushi, E., & Li, Q. (2010). Investigation of quasi-newton methods for unconstrained optimization. *International Journal of Computer Application,* 29, 48-58.

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(9), 366-378.

Djordjevic, S.S. (2019). New hybrid conjugate gradient method as a convex combination of LS and FR conjugate gradient methods. *Acta Mathematica Scientia,* 39B(1), 214-228.

Dolan, E., & Mor´ e, J. (2002). Benchmarking optimization software with performance profiles. *Journal of Mathematical 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.

Fletcher, R., & Reeves, C. (1964). Function minimization by conjugate gradients. *Computational Journal,* 7, 149-154.

Fletcher, R. (1987). *Practical Methods of Optimization*, New York: John Wiley & Sons.

Gilbert, J.C., & Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on Optimization,* 2(1), 21-42.

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*, 2019, 1-17.

Hager, W.W., & Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific Journal of Optimization*, 2, 35-58.

Hestenes, M.R & Stiefel, E.L. (1952). Methods of conjugate gradients for solving linear systems. J*ournal of Research of National Bureau of Standarts,* 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, 6, 1-17.*

Liu, Y., & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: Theory. *Journal of optimization theory and applications. 69*, 129-137.

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,* 371, 112-129.

Nocedal, J., & Wright, S. (2006). *Numerical Optimization*. New York: Springer.

Polyak, B.T. (1967). A general method of solving extremal problems. *Soviet Math. Doklady*, 8, 593-597.

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*, 1066, 122-141.

Rao, S.S. (2009). *Engineering Optimization Theory and Practice,* 4th ed, New Jersey: John Wiley & Sons, Inc.

Sabiu, J., & Waziri, M.Y. (2017). Effective modified hybrid conjugate gradient method for largescale symmetric nonlinear equation. *Application and Applied Mathematics,* 12, 1036-1056.

Sabiu, J., Waziri, M.Y., & Idris, A. (2017). A new hybrid Dai-Yuan and Hestenes Stiefel conjugate gradient parameter for solving system of equation. *Mayfeb Journal of Mathematics,* 1, 44-55.

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.

Salihu, N., Odekunle, M.R., Saleh, A. M. & Salihu, S. (2021). A Dai-Liao Hybrid Hestenes-Stiefel and Fletcher-Revees Methods for Unconstrained Optimization. *International Journal of Industrial Optimization*, 2(1), 33-50.

Sun, W., & Yuan, Y.X. (2006). *Optimization Theory and Methods: Nonlinear Programming,* New York: Springer.

Waziri, M. Y., Ahmed, K., & Sabi'u, J. (2019). A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations. *Arabian Journal of Mathematics,* 9, 443-457.

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.

Yuan, Y. X. (1991). A modified BFGS algorithm for unconstrained optimization. *IMA Journal of Numerical Analysis,* 11, 325-332.

Zoutendijk, G. (1970). Nonlinear programming computational methods, in integer and nonlinear programming. *J. Abadie ed.*, North-Holland, Amsterdam, 3, 37-86.