International Journal of Mathematical Analysis and Modelling

(Formerly Journal of the Nigerian Society for Mathematical Biology)

Volume 5, Issue 4 (Nov.), 2022

ISSN (Print): 2682 - 5694
ISSN (Online): 2682 - 5708

Developments in computational optimization techniques of conjugate gradient coefficient, search direction, Broyden-Fletcher-Goldfarb-Shanno, symmetric-rank one and Davidon-Fletcher-Powell quasi-Newton methods.

A. Adams*†, U.Y. Abubakar*‡, U. Abubakar* and Y. Yakubu*

Abstract
In this paper, we propose a modified Conjugate Gradient Coefficient (β) for solving unconstrained minimization problems as well as the Broyden-Fletcher-Goldfarb-Shanno, Davidon-Fletcher-Powell (DFP) and Symmetric-Rank-One (SR1) updates. The modified. It is proved that the resulting Conjugate Gradient Coefficient have global convergence under some mild conditions as well as the search direction(d). It is also proved that the search direction plays a key role in the line search method and the step size approaches mainly guarantee global convergence in general cases. The convergence rate of this method is also investigated. Some numerical results show that the modified Conjugate Gradient Coefficient algorithm is effective in practical computation.

Keywords: unconstrained optimization; quasi-newton method; hybridization; global convergence; symmetric-rank-one (sr1); Davidon-Fletcher-Powell (DFP).

1 Introduction

Quasi-Newton methods are distinguished by their use of approximate Hessian matrices. These approximate matrices are evaluated with respect to some iterative update formula as the algorithm progresses. The update procedure only requires the gradient of the objective function at each iteration. Thus, these methods provide a way of obtaining some curvature information without evaluating the exact Hessian. This is particularly useful when Hessian is very demanding to compute or cannot be computed at all for some reason. Because they are known to be generally more applicable and quite efficient, quasi-Newton methods are still widely used tools of nonlinear programming even after the development of automatic differentiation packages, Nocedal and Wright (2006). There are numerous works on the use of quasi-Newton methods either in line-search or trust-region applications.

* Department of Mathematics, Federal University of Technology, Minna, Niger State, Nigeria.
† Corresponding Author. Email: adamsayuba90@gmail.com
‡ Email: abbkruy@yahoo.com

Consider the unconstrained minimization problem

$$\min_{x \in \mathcal{R}^n} f(x), \tag{1}$$

where $f: \mathcal{R}^n \to \mathcal{R}$ is continuously differentiable function, the Conjugate Gradient (CG) methods are the best methods for solving (1), especially when the dimension n is large. The iterates of Conjugate Gradient (CG) methods for solving (1) are obtained as

$$x_{k+1} = x_k + \alpha_k d_k ; \qquad k = 0, 1, 2, \dots \dots \dots \tag{2}$$

where x_k is the current iterate point and $\alpha_k > 0$ is the step size. The step size is computed by carrying out some line search, especially the exact line search as given below

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha_k d_k) \tag{3}$$

The search direction d_k is defined as

$$d_k = \begin{cases} -g_k & if\ k = 0 \\ -g_k + \beta_k d_{k-1} & if\ k \geq 1 \end{cases} \tag{4}$$

where the Conjugate Gradient Coefficient, $\beta_k \in \mathcal{R}$ is a scalar. Some of the classical formulas for Conjugate Gradient Coefficient (β_k) are the Hestenes Stiefel (HS) in 1952, the Fletcher-Reeves (FR) in 1964, the Polak-Ribiere-Polyak (PRP) in 1969, the Conjugate Descent (CD) method in 1987, the Liu-Storey (LS) in 1992, and the Dai-Yuan (DY) in 2000. The parameters of the β_k are given as follows;

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, \qquad \beta_k^{FR} = \frac{g_k^T g_k}{\|g_{k-1}\|^2}, \qquad \beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}$$

$$\beta_k^{CD} = \frac{-g_k^T g_k}{d_{k-1}^T g_{k-1}}, \qquad \beta_k^{LS} = \frac{g_k^T(g_k - g_{k-1})}{-d_{k-1} g_{k-1}}, \qquad \beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}$$

where $\|.\|$ is the Euclidean norm of vectors.

In the Conjugate Gradient (CG) methods, many researchers focus on its global convergence properties and descent condition. For the Fletcher-Reeves (FR) method, Zoutendijk (1970) has proved the global convergence properties under the exact line search in (3). Al-Baali (1985) also shows the Fletcher-Reeves (FR) method fulfills global convergence properties under inexact line search. The Conjugate Descent (CD) method generated a descent search direction in each iteration for the parameter $\sigma < 1$ under the strong Wolfe line search, but its global convergence properties are not excellent. The Dai-Yuan (DY) method is a modification of the Fletcher-Reeves (FR) method, under the strong Wolfe line search, the Dai-Yuan (DY) method fulfills the descent condition, but this method has bad numerical results. Under strong Wolfe line search, Yuhong (2002) proved that Polak-Ribiere-Polyak (PRP) method has the global convergence properties and fulfills the descent condition. The (WYL) method is the modification of the Polak-Ribiere-Polyak (PRP) method, this method satisfied both the descent condition and global convergence properties under an exact line search and strong Wolfe line search. The Ibrahim and Rohanin (2020) method is a modification of the Polak-Ribiere-Polyak (PRP) method using an exact line search.

In this paper, the performance of the propose coefficient β_k with the hybridization of Symmetric-Rank One (SR1) and Davidon-Fletcher-Powell (DFP) updates is compared with some classical conjugate gradient methods. The organization of our paper takes the following format. In section 2, a new CG coefficient, hybridization of SR1 and DFP with the general algorithm are presented.

In section 3, the global convergence of β_k is presented. Section 4 covers the numerical experiments and discussions of Tables. Finally, Section 5 deals with conclusion.

2 New CG coefficient, hybridization of SR1-DFP and algorithm

In this section, we propose new β_k defined by:

$$\beta_k^{AMCGC} = \begin{cases} \beta_k^{PRP} & if \ \left|1 - \frac{g_{k+1}^T g_k}{\|g_k\|^2}\right| \geq \mu, \\ \beta_k^{New} & otherwise. \end{cases} \quad (5)$$

Where

$$\beta_k^{New} = \frac{\mu g_{k+1}^T (g_{k+1} - \lambda g_k)}{\|g_k\|^2}$$

$$\beta_K^{AMCGC} = \frac{\mu\|g_{k+1}\|^2 - \mu\lambda\frac{\|g_{k+1}\|}{\|g_k\|}|g_{k+1}^T g_k|}{\|d_k\|^2} \quad (6)$$

$\mu \geq \frac{\lambda-\sigma}{1-\sigma} > 0, \sigma < \lambda \leq 1.$ Obviously, if $\mu = 1$ and $\lambda = 1$, then $\beta_k^{New} = \beta_k^{PRP}$. If $\mu = 0$, then the method reduces to the steepest descent method. From β_k^{AMCGC}, AMCGC represent Adams Modified Conjugate Gradient Coefficient as defined in (5).

The hybrid conjugate gradient algorithm proposed in Touati-Ahmed and Storey (1990) [16] was a break through which since then, some research work focus on employing hybridization principles to have better algorithms that can handle large scale unconstrained optimization problems such as in [3, 14].

We formulate a new hybridization of SR1 and DFP for solving problem (1) by considering the summation of SR1and DFP quasi-Newton methods as follows:

$$B_{k+1} = B_k + \alpha uu^T + B^* + \theta vv^T + \varphi\omega\omega^T \quad (7)$$

where $\alpha, \ \theta, \ \varphi \neq 0; \ u, \ v, \omega \neq 0$
choose $\alpha, \theta, \varphi, u, v$ and ω such that B_{k+1} satisfies Quasi-Newton condition

$$B_{k+1}\gamma_k = \delta_k \quad (8)$$

$$(B_k + uu^T)\gamma_k = \delta_k \quad (9)$$

also

$$(B_k^* + \theta vv^T + \varphi\omega\omega^T)\gamma_k = \delta_k \quad (10)$$

from equation (9), we have

$$B_k\gamma_k + \alpha uu^T\gamma_k = \delta_k$$

$$\alpha u^T\gamma_k u = \delta_k - B_k\gamma_k \quad (11)$$

setting $u = \delta_k - B_k\gamma_k$ (12)

$\therefore \alpha u^T\gamma_k = 1 \Longrightarrow \ \alpha^{-1}(\delta - B_k\gamma_k)^T\gamma_k$ (13)

from equation (10)

$$B_k^*\gamma_k + \theta vv^T\gamma_k + \varphi\omega\omega^T\gamma_k = \delta_k \quad (14)$$

$$\theta vv^T\gamma_k + \varphi\omega\omega^T\gamma_k = \delta_k - B_k^*\gamma_k \quad (15)$$

where

$\theta v^T\gamma_k = \varphi\omega^T\gamma_k = 1$, we have

$$v = \omega = \delta_k - B_k^* \gamma_k \tag{16}$$

also,

$$\theta v^T \gamma_k = 1$$

$$\theta^{-1} = v^T \gamma_k = (\delta_k - B_k^* \gamma_k)^T \gamma_k \tag{17}$$

$$\varphi^{-1} = \omega^T \gamma_k = (\delta_k - B_k^* \gamma_k)^T \gamma_k \tag{18}$$

substitute equations (12), (16), (17) and (18) into (7), we have

$$B_{k+1} = B_k + \frac{1}{(\delta_k - B_k\gamma_k)^T} \times (\delta_k - B_k\gamma_k)(\delta_k - B_k\gamma_k)^T + B_k^T + \frac{1}{\delta_k^T \gamma_k} \times \delta_k \delta_k^T - \frac{1}{\gamma_k^T B_k^* \gamma_k} \times -B_k^* \gamma_k (-B_k \gamma_k^T) \tag{19}$$

where $B_k = B_k^*$

$$B_{k+1}^{AHM} = [2B_k(\delta_k - B_k\gamma_k)^T \gamma_k \delta_k^T \gamma_k^T B_k^* + (\delta_k - B_k\gamma_k)(\delta_k - B_k\gamma_k)^T \delta_k^T \gamma_k^T B_k^* + \delta_k \delta_k^T (\delta_k - B_k\gamma_k)^T \gamma_k^T B_k^* - B_k^* \gamma_k B_k^* \gamma_k^T (\delta_k - B_k\gamma_k)^T \delta_k^T] \div [(\delta_k - B_k\gamma_k)^T \gamma_k \delta_k^T \gamma_k^T B_k^*] \tag{20}$$

Algorithm 2.1

Step 1: $x_o \epsilon \mathcal{R}^n$ and a positive definite matrix $B_o = I_n$.
Choose $\mu \geq 0, \varepsilon > 0$ and set $k \coloneqq 0$.
Step 2: Terminate if $\|g(x_{k+1})\| < \varepsilon$.
If $\|g_k\| \leq \varepsilon$, the algorithm stops. Otherwise, go to step 3.
Step 3: Calculate the search direction by

$$d_k = \begin{cases} -H_k g_k\,; & k = 0 \\ -H_k g_k + \beta_k^{AMCGC} d_{k-1}\,; & k \geq 1. \end{cases}$$

where H_k is the Hybridization Method (AHM) updating matrix
with $\sigma \epsilon (0, 1]$ is chosen to ensure conjugacy.
Step 4: Calculate the step size α_k by

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \mu \alpha_k (g_k^T d_k - \tfrac{1}{2} \alpha_k \eta L_k \|d_k\|^2)$$

where $\mu \epsilon (0, 1), \eta [0, +\infty), \rho \epsilon (0, 1)$ are given constant scalars.
where $L_k = \frac{S_{k-1}^T y_{k-1}}{\|S_{k-1}\|^2}$; for $k > 1$.
Step 5: Compute the difference:
$S_k = x_{k+1} - x_k$ and $y_k = g_{k+1} - g_k$.
Step 6: Update B_k by $B_{k+1}^{AHM} = B_{k+1}^{DFP} + B_{k+1}^{SR1}$ to obtain B_{k+1}^{AHM}.
Step 7: Set $k \coloneqq k + 1$ and go to step 1.

3 Global convergence analysis

In this section, we study the global convergence of β_k^{AMCGC} and start with sufficient descent condition. The sufficient descent condition is defined by

$$g_k^T d_k \leq -c\|g_k\|^2 \quad \forall \quad k \geq 0, \; c > 0. \tag{21}$$

The following theorem shows that AMCGC with the modified exact line search possess the sufficient descent condition.

Theorem 1: Suppose the initial value of x_k and the search direction $d_k$as contain in (2) and (4), while the modified exact line search is used to determine the step size (α_k) then, the condition below holds for all $k \geq 0.$ The theorem helps us to determine the number of iterations.

$g_k^T d_k \leq -c\|g_k\|^2 \quad \forall \quad k \geq 0,\ c > 0.$

Proof:
The proof is by mathematical induction, if $k = 0$ then we have
$$g_0^T d_0 = -c\|g_0\|^2 \tag{22}$$
Hence the condition (21) holds true.

Next, we need to show that it holds for $k = 1$, condition (21), we have
$$g_1^T d_1 = -c\|g_1\|^2 \tag{23}$$
Also, to show that it holds for $k \geq 1$, we multiply equation (4) by g_{k+1}^T then,
$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1} + \beta_{k+1}^{AMCGC} d_k) \tag{24}$$
$$g_{k+1}^T d_{k+1} = -g_{k+1}^T g_{k+1} + \frac{\mu\|g_{k+1}\|^2 - \lambda\frac{\|g_{k+1}\|}{\|g_k\|}|g_{k+1}^T g_k|}{\|d_k\|^2} g_{k+1}^T d_k$$
$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \frac{\mu\|g_{k+1}\|^2 - \lambda\frac{\|g_{k+1}\|}{\|g_k\|}|g_{k+1}^T g_k|}{\|d_k\|^2} g_{k+1}^T d_k$$

For exact line search, we know that $g_{k+1}^T d_k = 0$, thus
$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 \tag{25}$$
Hence the sufficient condition holds true for $k + 1$.
First, we must prove that our method (β_k^{AMCGC}) are always not less than zero in order to establish the global convergence.
$$\beta_{k+1}^{AMCGC} = \frac{\mu\|g_{k+1}\|^2 - \lambda\frac{\|g_{k+1}\|}{\|g_k\|}|g_{k+1}^T g_k|}{\|d_k\|^2} \geq \frac{\mu\|g_{k+1}\|^2 - \lambda\frac{\|g_{k+1}\|}{\|g_k\|}\|g_{k+1}^T\|\|g_k\|}{\|d_k\|^2} \tag{26}$$
$$\beta_{k+1}^{AMCGC} = \frac{\mu\|g_{k+1}\|^2 - \lambda\frac{\|g_{k+1}\|}{\|g_k\|}|g_{k+1}^T g_k|}{\|d_k\|^2} \leq \frac{\mu\|g_{k+1}\|^2}{\|d_k\|^2} \tag{27}$$

The following are needed to establish the global convergence of our formula;
Assumption 1
The level set $M = \{x \in \mathcal{R}^n; f(x) \leq f(x_0)\}$ is bounded, where x_0 is the starting point.
In some neighborhood N of M, the function is continuously differentiable and its gradient is Lipchitz continuous. That is, there exist a constant $L > 0$ such that
$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|, \quad \text{for all } x, y \in N. \tag{28}$$
Under this assumption, we have the following Lemma, which was proved by Zoutendijk (1970).
Lemma 1
Suppose Assumption 1 holds and $\{x_k\}$ is generated by (2) and the search direction $\{d_k\}$ satisfies (21) and the step size $\{\alpha_k\}$ satisfies (3), then
$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty. \tag{29}$$

From Lemma 1, we have the following theorem.

Theorem 2
Suppose that Assumption 1 holds true. Let $\{x_k\}$ and $\{d_k\}$ be generated by equations (2) and (21) with β_k^{AMCGC} and the sufficient descent condition holds true. Then either
$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \quad \text{or} \quad \lim_{k\to\infty}\|g_k\| = 0 \tag{30}$$

Proof: We proceed by contradiction to prove Theorem 2. Suppose Theorem 2 is not true, then $\exists$ a constant $c > 0$ such that

$$\|g_k\| \geq c \quad \forall \; k \geq 0 \tag{31}$$

Rewriting (4) as

$$d_{k+1} + g_{k+1} = \beta_{k+1} d_k \tag{32}$$

And squaring both sides of the equation above then

$$\|d_{k+1}\|^2 = (\beta_{k+1})^2 \|d_k\|^2 - 2g_{k+1}^T d_{k+1} - \|g_{k+1}\|^2. \tag{33}$$

Dividing both sides of equation (33) by $(g_{k+1}^T d_{k+1})^2$ then,

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} = \frac{(\beta_{k+1})^2 \|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} - \frac{2}{g_{k+1}^T d_{k+1}} - \frac{\|g_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}$$

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} = \frac{(\beta_{k+1})^2 \|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \left[\frac{1}{\|g_{k+1}\|} + \frac{\|g_{k+1}\|}{g_{k+1}^T d_{k+1}}\right]^2 + \frac{1}{\|g_{k+1}\|^2}$$

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} \leq \frac{(\beta_{k+1})^2 \|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} + \frac{1}{\|g_{k+1}\|^2}. \tag{34}$$

Applying equation (26) to (34) above we get

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} \leq \frac{1}{\|g_{k+1}\|^2}. \tag{35}$$

Hence

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2}$$

$$\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \geq \frac{c^2}{k}. \tag{36}$$

From equations (36) and (31), it follows that

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} = \infty.$$

This contradicts the Zoutendijk condition in Lemma 1.

Theorem 3.3

Suppose that assumption 1 holds, consider any Conjugate Gradient (CG) methods of the form (3) and (4), step size (α_k) is obtained by the exact line search and β_k is obtained using (6), then either $\lim_{k\to\infty} \|g_k\| = 0$ or $\sum_{k=0}^{\infty} \frac{\left(g_k^T\right)^4}{\|d_k\|^2} < \infty$.

Proof:

From (31) and (25)

$$\|d_{k+1}\|^2 = \left(\frac{\|g_{k+1}\|^2}{\|d_k\|^2}\right)^2 \|d_k\|^2 - 2g_{k+1}^T d_{k+1} - \|g_{k+1}\|^2$$

$$\|d_{k+1}\|^2 = \frac{\|g_{k+1}\|^4}{\|d_k\|^2} - 2g_{k+1}^T d_{k+1} - \|g_{k+1}\|^2. \tag{37}$$

We have already proved that sufficient descent condition holds using mathematical induction.

Therefore, we know that

$$g_{k+1}^T d_{k+1} \leq -c\|g_{k+1}\|^2 \tag{38}$$

Hence from (37),

$$\|d_{k+1}\|^2 = \frac{\|g_{k+1}\|^4}{\|d_k\|^2} + 2c\|g_{k+1}\|^2 - \|g_{k+1}\|^2$$

$$\|d_{k+1}\|^2 = \frac{\|g_{k+1}\|^4}{\|d_k\|^2} - \|g_{k+1}\|^2(1-2c). \tag{39}$$

Multiplying both sides of (39) with $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$

then,

$$\|d_{k+1}\|^2 \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}\left[\frac{\|g_{k+1}\|^4}{\|d_k\|^2} - \|g_{k+1}\|^2(1-2c)\right]$$

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left[(2c-1) + \frac{\|g_{k+1}\|^2}{\|d_k\|^2}\right]$$

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} \leq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \tag{40}$$

Based on theorem 2, we know that $\lim_{k\to\infty} \frac{\left(g_{k+1}^T d_{k+1}\right)^2}{\|d_{k+1}\|^2} < 0.$

This will imply that if theorem (3) is not true, then we have $\lim_{k\to\infty} \frac{\left(g_{k+1}^T d_{k+1}\right)^2}{\|d_{k+1}\|^2} = \infty.$

From (40) we obtain

$$\infty \leq \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}. \tag{41}$$

Hence, theorem 3 holds true for sufficiently large k.

The proof of Theorems 1, 2 and 3 showed that the formulas converge analytically. Next is to validate numerically that the method also converge using Rosenbrock test function. This, we will have in the next section.

4 Numerical results

In this section, experimentation of our proposed method was carried out against some classical methods in the literature, to weigh the significance of our algorithm with $\beta_k = \beta_K^{AMCGC}$. To effect this, Rosenbrock standard test function was considered. The dimension of the function is varied to compare the computational strength against Fletcher-Reeves (FR), Dai-Yuan (DY), (PRP), (FRPRP) which is the hybridization of (FR) and (PRP) and Ibrahim-Rohanin (IR) methods. The modified exact line search condition was used in the computation.

In carrying out the simulation, the number of iterations (NI), the number of function evaluations (NF), CPU time (CT) and the dimension of the standard test problems is 500, 1000, 5000, 10000 and 100000 variables were put into consideration to determine the numerical strength of the proposed algorithms AMCGC with BFGS, DFP and SR1. The choice of values of $\{\mu\}$ and $\{\lambda\}$ must satisfy the conditions $\mu \geq \frac{\lambda-\sigma}{1-\sigma} > 0$ and $\sigma < \lambda \leq 1$. After numerous experimentation with the randomly selected values for the parameters, the values $\mu = 0.7$ and $\lambda = 1.0$ were taken into consideration to be the best values for the parameters which make our algorithms robust to obtain the results presented.

$\|g_k\| \leq \varepsilon$, where $\varepsilon = 10^{-5}$ is considered as the stopping criterion. The implementation of the algorithm was done using MATLAB R2021 on a window 10 machines with the following specification; processor: Intel Core i7-17050Hcpu@2.6GHz2.59GHz, with RAM of 16GB DDR4 and 1TB SSD storage. Also, comes with a GPU card: 4GB of GDDR5, GDDR6 memory clocked at 8GHz are supplied, altogether it has 128-Bit memory interface that creates a bandwidth of 112.1GB/s. In Tables 1, 2, 3, 4 and 5, we reported the numerical results with different dimension (Dim).

Table 1: Numerical Results of $\boldsymbol{AMCGC^{SR1}, IR, FR, DY, PRP\ and\ FRPRP}$ Methods for Rosenbrock Test Function

Problem	$Dim.$	$AMCGC^{SR1}$			IR			FR			DY			PRP		
		NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT
Rosenbrock	500	55	57	1.22	56	57	1.73	78	79	1.61	59	60	1.38	56	57	1.21
	1000	53	55	1.37	55	56	1.79	77	78	1.72	58	59	1.50	55	56	1.32
	5000	49	52	2.45	52	53	3.11	74	75	3.53	55	56	2.98	52	53	2.33
	10000	40	41	3.35	40	41	3.48	63	64	5.09	44	45	4.16	40	41	3.03
	100000	12	13	9.15	12	13	10.92	13	14	12.11	13	14	12.81	12	13	11.43

Table 2: Numerical Results of $\boldsymbol{AMCGC^{DFP}, IR, FR, DY, PRP\ and\ FRPRP}$ Methods for Rosenbrock Test Function

$Problem$	$Dim.$	$AMCGC^{DFP}$			IR			FR			DY			PRP			$FRPRP$		
		NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT
Rosenbrock	500	55	57	1.67	56	57	1.73	78	79	1.61	59	60	1.38	56	57	1.21	55	56	1.26
	1000	54	55	1.75	55	56	1.79	77	78	1.72	58	59	1.50	55	56	1.32	54	55	1.44
	5000	51	52	3.43	52	53	3.11	74	75	3.53	55	56	2.98	52	53	2.33	52	53	2.07
	10000	39	40	3.45	40	41	3.48	63	64	5.09	44	45	4.16	40	41	3.03	40	41	3.39
	100000	12	13	10.32	12	13	10.92	13	14	12.11	13	14	12.81	12	13	11.43	12	13	11.66

Table 3: Numerical Results of $\boldsymbol{AMCGC^{BFGS}, IR, FR, DY, PRP\ and\ FRPRP}$ Methods for Rosenbrock Test Function

$Problem$	$Dim.$	$AMCGC^{BFGS}$			IR			FR			DY			PRP			$FRPRP$		
		NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT
Rosenbrc	500	54	56	1.24	56	57	1.73	78	79	1.61	59	60	1.38	56	57	1.21	55	56	1.26
	1000	53	54	1.35	55	56	1.79	77	78	1.72	58	59	1.50	55	56	1.32	54	55	1.44
	5000	50	51	2.43	52	53	3.11	74	75	3.53	55	56	2.98	52	53	2.33	52	53	2.07
	10000	40	41	3.37	40	41	3.48	63	64	5.09	44	45	4.16	40	41	3.03	40	41	3.39
	10000	12	13	9.02	12	13	10.9	13	14	12.1	13	14	12.8	12	13	11.4	12	13	11.6

Table 4: Numerical Results of $\boldsymbol{AMCGC^{BFGS}}$,$\boldsymbol{AMCGC^{SR1}}$, $\boldsymbol{AMCGC^{DFP}}$ $\boldsymbol{FRPRP}$ and $\boldsymbol{IR}$ Methods for Rosenbrock Test Function.

$Problem$	Dim.	$AMCGC^{SR1}$			$AMCGC^{DFP}$			$AMCGC^{BFGS}$			IR			$FRPRP$		
$Rosenbrock$		NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT	NI	NF	CT
	500	55	57	1.22	55	57	1.67	54	56	1.24	56	57	1.73	55	56	1.26
	1000	53	55	1.37	54	55	1.75	53	54	1.35	55	56	1.79	54	55	1.44
	5000	49	52	2.45	51	52	3.43	50	51	2.43	52	53	3.11	52	53	2.07
	10000	40	41	3.35	39	40	3.45	40	41	3.37	40	41	3.48	40	41	3.39
	100000	12	13	9.15	12	13	10.32	12	13	9.02	12	13	10.92	12	13	11.66

4.1 Discussion

In Table 1, it is shown that the developed algorithm of AMCGC and the modified exact line search with the SR1 method is promising, because it requires less number of iterations (NI), less number function evaluation (NF) as well as less CPU time (CT) as compare to the existing methods. It is observed that, the (NI) and (NF) increasing down the table with increment in dimension so as the CPU time as well. In Table 2, we also use the CPU time, NI and NF to compare the performance of the AMCGC and the modified exact line search combined with the DFP method. The Table 2 above shown that $AMCGC^{DFP}$ is fastest, then FRPRP, then IR, then PRP, then DY, and FR. These only differ in their choice of the search direction and Conjugate Gradient Coefficient (CGC), then the numerical results show that the proposed method is promising.

In Table 3, it is shown that the results obtained from the modified exact line search and the Conjugate gradient Coefficient (CGC) with the cautious BFGS method indicates that $AMCGC^{BFGS}$ is fastest relative to the less number of iterations (NI) and number of function evaluation (NF), while the CPU time (CT) increased down the table, but relatively low as compare to IR, FR, DY, PRP and FRPRP methods. In Table 4, we also use CPU time (CT), number of iterations (NI) and number of function evaluation (NF) to compare the performance of the $AMCGC^{BFGS}$,$AMCGC^{SR1}$ $AMCGC^{DFP}$ $FRPRP$ and IR algorithms, from the numerical results shown above, $AMCGC^{BFGS}$ is fastest as compare to others for the Rosenbrock standard test function on average with less CPU time (CT), less number of iterations (NI) and less number of function evaluation (NF) as compare to others.

Finally, the effectiveness of our proposed Algorithm AMCGC was shown in Table 4 to be more effective against its counterparts.

5 Conclusion

In this paper, we proposed a new CG coefficient β_k^{AMCGC} with the $HSR1DFP$ methods for the solution of unconstrained optimization problem. The choice of values of μ and λ which must satisfy the conditions $\mu \geq \frac{\lambda-\sigma}{1-\sigma} > 0$ and $\sigma < \lambda \leq 1$. After numerous experimentation with the randomly selected values for the parameters, the values $\mu = 0.7$ and $\lambda = 1.0$ were taken into consideration to be the best values for the parameters which make our algorithms robust to obtain the results presented. This new CG coefficient β_k^{AMCGC} possesses the descent property with exact line search condition. We established the global convergence of the method using Zoutendijk condition given in [10]. The experimentation of the formulas on standard test function showed that our proposed Algorithm AMCGC is efficient and robust.

References

[1] Al-Baali M. (1985). Descent property and global convergence of the fletcher-reeves method with inexact line search. IMA Journal of Numerical Analysis. 5(1):121 – 124.

[2] Al-Baali, M., Narushima, Y. and Yabe, H. (2015). A family of three-term conjugate gradient methods with sufficient descent property for unconstrained optimization. Computational Optimization and Applications. 8(5):1 – 22.

[3] Babaie-Kafaki, S. and Ghanbari, R. (2014a). Two hybrid nonlinear conjugate gradient methods based on a modified secant equation. Optimization. 63(7):1027 – 1042.

[4] Ibrahim A. and Rohanin A. (2020). convergence analysis of a new conjugate gradient method for unconstrained optimization. Applied Mathematical Sciences. 9(140):6969 – 6984.

[5] Liu J and Du S. (2019). Modified three-term conjugate gradient method and its applications. Mathematical Problems in Engineering, https,//doi.org/10.1155/2019/5976595.

[6] Lu J., Li Y. and Pham H. (2020). A modified Dai-Liao conjugate gradient method with a new parameter for solving image restoration problems. Mathematical Problems in Engineering, https,//doi.org/10.1155/2020/6279543.

[7] Lu, Y., Li, W., Zhang, C. and Yang, Y. (2015). A class of new hybrid conjugate gradient method for unconstrained optimization. Journal of Information and Computational Science, 12(5):1941 – 1949.

[8] Touati-Ahmed, D. and Storey, C. (1990). Efficient hybrid conjugate gradient techniques. Journal of Optimization Theory and Applications, 64(2):379 – 397.

[9] Yuhong, D. (2002). A Non-monotone conjugate gradient algorithm for unconstrained optimization. Journal of Systems Science and Complexity, 15(2):139 – 145.

[10] Zoutendijk, G. (1970). Nonlinear programming computational methods. Integer and Nonlinear Programming, 143(1):37 – 86.