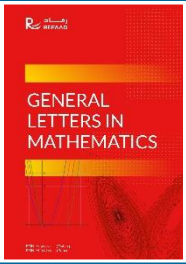
Available online at www.refaad.com
Gen. Lett. Math., 12(1) (2022), 40- 48
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: https://www.refaad.com/Journal/Index/1

ISSN: 2519-9277 (Online) 2519-9269 (Print)



Conjugated Gradient with Four Terms for Nonlinear Unconstrained Optimization

Ahmed Anwer Mustafa [a,*]

[a] Department of Mathematics, Faculty of Education, University of Zakho, Kurdistan Region, Iraq
Email: ahmed.mustafa@uoz.edu.krd

Abstract

The nonlinear conjugate gradient (GJG) technique is an effective tool for addressing minimization on a huge scale. It can be used in a variety of applications., We presented a novel conjugate gradient approach based on two hypotheses, and we equalized the two hypotheses and retrieved the good parameter in this article. To get a new conjugated gradient, we multiplied the new parameter by a control parameter and substituted it in the second equation. a fresh equation for β_k is proposed. It has global convergence qualities. When compared to the two most common conjugate gradient techniques, our algorithm outperforms them in terms of both the number of iterations (NOIS) and the number of functions (NOFS). The new technique is efficient in real computing and superior to previous comparable approaches in many instances, according to numerical results.

Keywords: Unconstrained; Four-Terms; Conjugate Gradient; Sufficient Descent; Convergent Property.

2010 MSC: 42A50, 90C52.

Abbreviations: CJG conjugation gradient, NOIS number of iterations, NOFS number of functions, Min minimum, $g_k=\nabla f(x_k)$ gradient, β_k A parameter has different formulas, ELS exact line search, F.T first term, S.T second term, T.T third term.

1. Introduction

In resolving the optimization with no constraints questions the conjugate gradient methodology (CJG) is quite useful. The technique can be described this way

$$\text{Min.} f\,(x),\ x \in R^n \tag{1}$$

where $f:\ R^n \rightarrow R$ it's continually differentiable There are several steps to the CJG technique, including iteration.

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

Or $v_k = \alpha_k d_k$ where $v_k = x_{k+1} - x_k$

where x_k is the iteration point at the moment, $\alpha_k > 0$ is the length of a step, and d_k is the search's general direction. d_k is characterized by the following formula:

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

where g_k is the gradient of $f(x)$ at the point x_k. $\beta_k \in R$ determines which CJG technique will be used. There are numerous formulae for β_k.

*Corresponding author
Email address: ahmed.mustafa@uoz.edu.krd (Ahmed Anwer Mustafa)
doi: 10.31559/glm2022.12.1.5
Received: 22 Jan 2022 Revised: 5 Mar 2022 Accepted: 16 Mar 2022

$\beta_k^{HAS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}$, $\beta_k^{FAR} = \frac{g_{k+1}^T g_k}{g_k^T g_k}$, $\beta_k^{PAR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}$, $\beta_k^{COD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k}$, $\beta_k^{BA1} = \frac{y_k^T y_k}{-d_k^T g_k}$ $\beta_k^{LAS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}$,

$\beta_k^{DAY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k}$, $\beta_k^{RMIAL} = \frac{g_k^T y_k}{d_k^T (d_k - g_{k+1})}$, $\beta_k^{MIMS} = \frac{\frac{g_k^T y_k}{d_{k-1}^T (d_{k-1} - g_k)} + \frac{g_k^T y_k}{\|d_{k-1}\|^2}}{2}$

$\beta_{k+1}^{hybrid} = \frac{g_{k+1}^T (y_k - t s_k)}{\max\{y_k^T d_k, \|g_k\|^2\}}$, $\beta_k^{MMR} = \frac{m_k \|g_k\|^2 - (g_k^T g_{k-1})}{m_k \|g_{k-1}\|^2}, where\ m_k = \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|}$

$$\beta_k^{LS+} = \begin{cases} \frac{\|g_k\|^2 - \mu_k |g_k^T g_{k-1}|}{\|g_k\|^2}, if\ \|g_k\|^2 > \mu_k |g_k^T g_{k-1}|\ where\ \mu_k = \frac{\|x_k - x_{k-1}\|}{\|y_k\|} \\ \beta_k^{DL-HS} = -\mu_k \frac{g_k^T s_{k-1}}{d_{k-1}^T y_{k-1}} \qquad otherwise \end{cases}$$

the gradients of $f\ (x)$ at x_k and k_{k+1}, respectively, are g_k and g_{k+1} The related techniques are described above., Hestenes and Steifel are two names for HAS. [9], Fletcher and Reeves are known as FAR. [11], PAR is Polak and Ribiere [4], COD stands for Conjugate Descent. [12], BAA1 is AL - Bayati, A.Y. and AL-Assady [1], LAS is Liu and Storey [13], DAY is Dai and Yuan [14], RMIAL is Rivaie, Mustafa, Ismail, and Leong [8]. MIMS is Mamat, Ibrahim and Mohammed Sulaiman [5], hybrid by Zhang, L. [17] MMR is Mouiyad, Mustafa, and Rivaie [7], and lastly, LS+ is the modification of Liu and Storey [16]. The standard formulae (HAS) and (PAR) are compared to our unique β_k^{AA4} the formula in this study. The remaining sections of the text are included below. It's given as a modern conjugate gradient formula with a novel algorithm approach in part 2, and as a descent condition, appropriate descent condition, and global convergence proof in section 3. Section 4 contains graphs, percentages, and visualizations. Finally, we arrive at the conclusion end in section 5.

2. Technique Update Proposed and Algorithm

2.1. The Modern Concept

As you see in that part, we will suggest based on hybridization for the new conjugate gradient method, there are many proposed hybrid gradient conjugate formulas for example

$$\beta_k^{AN} = (1 - \theta_k)\beta_k^{PAR} + \theta_k \beta_k^{DAY}, 0 < \theta_k < 1 \text{ see[2]} \tag{4}$$

$$\beta_k^{hDY} = \max\{c\beta_k^{DAY}, \min\{\beta_k^{HAS}, \beta_k^{DAY}\}\}, c = \frac{1-\sigma}{1+\sigma} \text{ see [3]} \tag{5}$$

$$\beta_k = \begin{cases} \frac{g_{k+1}^T y_k}{\max\{d_k^T y_k, -d_k^T g_k\}}, if\ \ 0 < g_{k+1}^T g_k < g_{k+1}^T g_{k+1} \\ \beta_k^{CAD}, \qquad otherwise \end{cases}, \text{ see [15]} \tag{6}$$

The main idea of proposing the new method for the new conjugated gradient is to rely on two hypotheses, the first is a conjugated gradient for Hestenes and Steifel multiplying a parameter, and the second depends on the hybridization between Polak-Ribiere's method and Hestenes-Steifel's method, where we used to hybridize the two methods the idea of the concavity function and we equalized the two hypotheses and extracted the good parameter. We multiply the new parameter by a control parameter and substituted it in the second equation to get a new conjugated gradient and the hypotheses that We suggested it as follows

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

$$\beta_k = \tau \frac{g_{k+1}^T y_k}{g_k^T g_k} + (1 - \tau) \frac{g_{k+1}^T y_k}{d_k^T y_k} = \tau \frac{g_{k+1}^T y_k}{g_k^T g_k} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \tau \frac{g_{k+1}^T y_k}{d_k^T y_k} \tag{8}$$

From the equations above, we obtain

$$\tau \frac{g_{k+1}^T y_k}{d_k^T y_k} = \tau \frac{g_{k+1}^T y_k}{g_k^T g_k} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \tau \frac{g_{k+1}^T y_k}{d_k^T y_k} \Rightarrow \tau \left(\frac{g_{k+1}^T y_k}{d_k^T y_k} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T y_k}{g_k^T g_k} \right) = \frac{g_{k+1}^T y_k}{d_k^T y_k}$$

$$\Rightarrow \tau \left(2 \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T y_k}{g_k^T g_k} \right) = \frac{g_{k+1}^T y_k}{d_k^T y_k} \Rightarrow \tau \left(\frac{2 g_{k+1}^T y_k g_k^T g_k - g_{k+1}^T y_k d_k^T y_k}{d_k^T y_k g_k^T g_k} \right) = \frac{g_{k+1}^T y_k}{d_k^T y_k}$$

$$\tau = \frac{g_k^T g_k}{2 g_k^T g_k - d_k^T y_k}$$

Now we add a controlling parameter η to the formula we have

$$\tau = \frac{\eta g_k^T g_k}{2g_k^T g_k - d_k^T y_k} = \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k}, \text{ where } \eta \in (0,1) \tag{9}$$

Replacement (9) in (8), we acquire

$$\beta_k^{AA4} = \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{\|g_k\|^2} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{d_k^T y_k} \tag{10}$$

Put (10) in (3) We've been given a new search direction.

$$d_{k+1} = -g_{k+1} + \beta_k^{AA4} d_k \tag{11}$$

$$\Rightarrow d_{k+1} = -g_{k+1} + \left(\frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{\|g_k\|^2} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{d_k^T y_k} \right) d_k \tag{12}$$

2.2. Outline of The AA4 Method

Step(1): Select $x_0 \in R^n$, $\varepsilon = 10^{-5}$, $\eta \in (0,1)$

Step(2): Set $k = 0$, Find $f(x_0)$, g_0, $d_k = -g_k$

Step(3): Compute $\alpha_k > 0$ meetings the stringent Wolfe requirement

$$f(x_k + \alpha_k d_k) \le f(x_k) + c_1 \alpha_k {g_k}^T d_k$$
$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \le c_2 |{g_k}^T d_k|$$

Let $0 < c_1 < c_2 < 1$

Step(4): Evaluate $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = \nabla f(g_{k+1})$, If $\|g_{k+1}\| < \varepsilon$ break.

Step(5): Calculate $d_{k+1} = -g_{k+1} + \beta_k^{AA4} d_k$

$$\beta_k^{AA4} = \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{\|g_k\|^2} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{d_k^T y_k}$$

Step(6): If $|g_{k+1}^T g_k| > 0.2\|g_{k+1}\|^2$ go to step(2) else $k = k + 1$ and then head to step(3)

3. The New Method's Global Convergent Analysis

The properties of β_k^{AA4}'s convergence should be analyzed. An algorithm's convergence must be proven. the condition of descent, a descent of sufficient condition, and Properties of global convergence.

Theorem 3.1 Consider this sequence $\{x_k\}$ is created by (2), then the direction of search in (3) with the fresh technique of conjugate gradient (10) fulfills the requirement of descent, that is $d_{k+1}^T g_{k+1} \le 0$ using line search (exact and inexact).

Proof:

From (3) and (10) we have

$$d_{k+1} = -g_{k+1} + \left(\frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{\|g_k\|^2} + \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{d_k^T y_k} \right) d_k \tag{13}$$

Multiply all ends of the preceding equation by g_{k+1}^T to obtain

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{\|g_k\|^2} g_{k+1}^T d_k + \frac{g_{k+1}^T y_k g_{k+1}^T d_k}{d_k^T y_k} - \frac{\eta \|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k} \frac{g_{k+1}^T y_k}{d_k^T y_k} g_{k+1}^T d_k \tag{14}$$

An exact line search, which needs a step size α_k which requires $g_{k+1}^T d_k = 0$. The evidence is then complete. An inexact line search that requires a step size of α_k is $g_{k+1}^T d_k \neq 0$. Because (HS) meets the descent assumption, the third term in equation (14) will be less than or equal to zero.

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

Since by strong Wolfe condition $-g_{k+1}^T d_k \le -c_2 g_k^T d_k = c_2 \|g_k\|^2$ *and* $g_{k+1}^T d_k \le d_k^T y_k \Rightarrow c_2 \|g_k\|^2 \ge -d_k^T y_k$ so,

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 - \frac{\eta\|g_k\|^2 g_{k+1}^T y_k d_k^T y_k}{\left(\frac{2d_k^T y_k}{c_2} + d_k^T y_k\right)\|g_k\|^2} - \frac{\eta\|g_k\|^2 g_{k+1}^T y_k}{\left(\frac{2d_k^T y_k}{c_2} + d_k^T y_k\right) d_k^T y_k} c_2\|g_k\|^2 \Rightarrow d_{k+1}^T g_{k+1}$$

$$\le -\|g_{k+1}\|^2 - \frac{\eta\|g_k\|^2 g_{k+1}^T y_k}{\left(\frac{2d_k^T y_k}{c_2} + d_k^T y_k\right)}\left(\frac{d_k^T y_k}{\|g_k\|^2} + \frac{c_2\|g_k\|^2}{d_k^T y_k}\right)$$

$$\Rightarrow d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 - K g_{k+1}^T y_k$$

Where $K = \frac{\eta\|g_k\|^2}{\left(\frac{2d_k^T y_k}{c_2} + d_k^T y_k\right)}\left(\frac{d_k^T y_k}{\|g_k\|^2} + \frac{c_2\|g_k\|^2}{d_k^T y_k}\right)$, which is positive, since $c_2, \eta, \|g_k\|^2, d_k^T y_k \ge 0$

Multiplying and dividing the right-hand side by $g_{k+1}^T y_k$ we get

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

$$\Rightarrow d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + \frac{K(g_{k+1}^T y_k)^2}{(-\|g_{k+1}\|^2 - d_k^T y_k)}$$

$$\Rightarrow d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 - K\frac{\left(g_{k+1}^T y_k\right)^2}{(\|g_{k+1}\|^2 + d_k^T y_k)} \tag{15}$$

Since it is $K, (g_{k+1}^T y_k)^2, \|g_{k+1}\|^2$ and $d_k^T y_k$ are greater than zero then $d_{k+1}^T g_{k+1} \le 0$ ■

Theorem 3.2 Suppose that (2) generates the sequence $\{x_k\}$, then the search direction in (3) with a conjugate gradient approach is proposed (10) fulfill the condition for a sufficient descent, that is $d_{k+1}^T g_{k+1} \le -C\|g_{k+1}\|^2$ using line search, both exact and inexact.

Proof:

Multiply the two sides of the equation (*12*) by g_{k+1}^T we have

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

Now, since a parameter (HAS) provides the requirements for the descent, and the first third term in the above equation will be less than or equal to zero, then the equation above becomes

$$d_{k+1}^T g_{k+1} \le \frac{\eta\|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k}\frac{g_{k+1}^T y_k}{\|g_k\|^2} g_{k+1}^T d_k - \frac{\eta\|g_k\|^2}{2\|g_k\|^2 - d_k^T y_k}\frac{g_{k+1}^T y_k}{d_k^T y_k} g_{k+1}^T d_k \tag{16}$$

Some after algebraic operations and the application of some conditions, as in the theorem (13) above, we get

$$d_{k+1}^T g_{k+1} \le -K\frac{\left(g_{k+1}^T y_k\right)^2}{(\|g_{k+1}\|^2 + d_k^T y_k)} \quad \Rightarrow d_{k+1}^T g_{k+1} \le -\frac{K\left(g_{k+1}^T y_k\right)^2}{(\|g_{k+1}\|^2 + d_k^T y_k)}\frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \tag{17}$$

Let $C = \frac{K\left(g_{k+1}^T y_k\right)^2}{(\|g_{k+1}\|^2 + d_k^T y_k)}\frac{1}{\|g_{k+1}\|^2}$ which is positive, then $d_{k+1}^T g_{k+1} \le -C\|g_{k+1}\|^2$ ■

Global Convergent:

Suppose that the assumptions below are usually required to establish the convergence of the Methodology for nonlinear conjugated gradients, see [6].

Assumptions:

a. f is confined below on the level set R^n continuous and differentiable in a neighborhood N of the level set $S = \{x \in R^n : f(x) \le f(x_0)\}$ at the start point x_0.

b. Because the gradient $g(x)$ is Lipschitz continuous in N, there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \le L\|x - y\|. x, y \in N$.

We have the following theorem when it was shown using these assumptions [18].

Theorem 3.3 Assume that your assumption is right. Take into account any gradient that is conjugated from (2) where d_k is a descent search direction and we use α_k in both situations, precise and inexact line searches are used. Then comes the condition called as Zoutendijk condition holds

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

For proof see [10,18]. From the previous information, we can obtain the following convergence theorem of the conjugate gradient methods.

Theorem 3.4 Suppose the hypotheses are accurate. Consider any conjugate gradient strategy of the sort (2) and (3) where α_k is acquired through both exact and inexact line searches, and d_k is the descent search direction than either

$$\lim_{k\to\infty}\|g_k\|=0 \quad \text{Or} \quad \sum_{k\geq 1}\frac{(g_k^T d_k)^2}{\|d_k\|^2}<\infty$$

Proof:

Contradiction is used to prove theorem 15. It is false if theorem 15., then there exists a constant $\mu>0$, such that

$$\|g_i\|\geq\mu,\ \forall i\geq 0. \tag{18}$$

Rewrite (12), we get

$$d_{k+1}+g_{k+1}=\beta_k^{AA4}d_k \tag{19}$$

Squaring the above equation, we get

$$\|d_{k+1}\|^2=(\beta_k^{AA4})^2\|d_k\|^2-2g_{k+1}^T d_{k+1}-\|g_{k+1}\|^2 \tag{20}$$

Divide the equation's two sides. (20) by $(g_{k+1}^T d_{k+1})^2$, therefore we end up with

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}=\frac{(\beta_k^{AA4})^2\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}-\frac{2}{g_{k+1}^T d_{k+1}}-\frac{\|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}$$

$$=\frac{(\beta_k^{AA4})^2\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}-\left(\frac{1}{\|g_{k+1}\|}+\frac{\|g_{k+1}\|}{g_{k+1}^T d_{k+1}}\right)^2+\frac{1}{\|g_{k+1}\|^2}$$

$$\leq\frac{(\beta_k^{AA4})^2\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}+\frac{1}{\|g_{k+1}\|^2}$$

Substitute β_k^{AA4}, we have

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

$$\Rightarrow\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}\leq\frac{\left(\frac{\eta g_{k+1}^T y_k}{2\|g_k\|^2-d_k^T y_k}+\frac{g_{k+1}^T y_k}{d_k^T y_k}\right)^2\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}$$
$$-2\left(\frac{\eta g_{k+1}^T y_k}{2\|g_k\|^2-d_k^T y_k}+\frac{g_{k+1}^T y_k}{d_k^T y_k}\right)\left(\frac{\eta\|g_k\|^2}{2\|g_k\|^2-d_k^T y_k}\frac{g_{k+1}^T y_k}{d_k^T y_k}\right)\frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}$$
$$+\left(\frac{\eta\|g_k\|^2}{2\|g_k\|^2-d_k^T y_k}\frac{g_{k+1}^T y_k}{d_k^T y_k}\right)^2\frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2}+\frac{1}{\|g_{k+1}\|^2}$$

$$\text{F.T}=\left(\frac{\eta^2\left(g_{k+1}^T y_k\right)^2}{4\|g_k\|^4-4\|g_k\|^2 d_k^T y_k+\left(d_k^T y_k\right)^2}+\frac{2\eta\left(g_{k+1}^T y_k\right)^2}{2\|g_k\|^2 d_k^T y_k-\left(d_k^T y_k\right)^2}+\frac{\left(g_{k+1}^T y_k\right)^2}{\left(d_k^T y_k\right)^2}\right)\frac{\|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}$$

We know that $g_{k+1}^T d_k\leq d_k^T y_k$ and by Wolfe's condition $c_2 {g_k}^T d_k \leq d_k^T y_k\Rightarrow -c_2 {g_k}^T d_k \geq -d_k^T y_k$ This implies that $\|g_k\|^2\geq\frac{-1}{c_2}d_k^T y_k$, so

$$\text{F.T}\leq\left(\frac{\eta^2\left(g_{k+1}^T y_k\right)^2}{\frac{-4\|g_k\|^2 d_k^T y_k}{c_2}-4\|g_k\|^2 d_k^T y_k-c_2\|g_k\|^2 d_k^T y_k}+\frac{2\eta\left(g_{k+1}^T y_k\right)^2}{-2c_2\|g_k\|^4-\left(d_k^T y_k\right)^2}+\frac{\left(g_{k+1}^T y_k\right)^2}{-c_2\|g_k\|^2 d_k^T y_k}\right)\frac{\|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}$$

$$\Rightarrow \text{F.T} \le -\left(\frac{\eta^2\left(g_{k+1}^T y_k\right)^2}{\frac{4\|g_k\|^2 d_k^T y_k}{c_2}+4\|g_k\|^2 d_k^T y_k+c_2\|g_k\|^2 d_k^T y_k}+\frac{2\eta\left(g_{k+1}^T y_k\right)^2}{2c_2\|g_k\|^4+\left(d_k^T y_k\right)^2}+\frac{\left(g_{k+1}^T y_k\right)^2}{c_2\|g_k\|^2 d_k^T y_k}\right)\frac{\|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}$$

$$\text{S.T} = -2\left(\frac{\eta g_{k+1}^T y_k d_k^T y_k + g_{k+1}^T y_k(2\|g_k\|^2 - d_k^T y_k)}{(2\|g_k\|^2 - d_k^T y_k)d_k^T y_k}\right)\left(\frac{\eta\|g_k\|^2 g_{k+1}^T y_k}{(2\|g_k\|^2 - d_k^T y_k)d_k^T y_k}\right)\frac{\|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2}$$

$$\Rightarrow \text{S.T} = -2\left(\frac{\left(\eta^2(g_{k+1}^T y_k)^2 d_k^T y_k\|g_k\|^2 + \eta\|g_k\|^2(g_{k+1}^T y_k)^2(2\|g_k\|^2 - d_k^T y_k)\right)\|d_k\|^2}{\left((2\|g_k\|^2 - d_k^T y_k)d_k^T y_k\right)^2(g_{k+1}^T d_{k+1})^2}\right)$$

$$\Rightarrow \text{S.T} = 2\left(\frac{\left(-\eta^2(g_{k+1}^T y_k)^2 d_k^T y_k\|g_k\|^2 + \eta\|g_k\|^2(g_{k+1}^T y_k)^2(-2\|g_k\|^2 + d_k^T y_k)\right)\|d_k\|^2}{\left((2\|g_k\|^2 - d_k^T y_k)d_k^T y_k\right)^2(g_{k+1}^T d_{k+1})^2}\right)$$

$$\Rightarrow \text{S.T} = 2\left(\frac{\left(-\eta^2(g_{k+1}^T y_k)^2 d_k^T y_k\|g_k\|^2 + \eta\|g_k\|^2(g_{k+1}^T y_k)^2(-2\|g_k\|^2 + d_k^T g_{k+1} + \|g_k\|^2)\right)\|d_k\|^2}{\left((2\|g_k\|^2 - d_k^T y_k)d_k^T y_k\right)^2(g_{k+1}^T d_{k+1})^2}\right)$$

Since $g_{k+1}^T d_k \le -c_2 g_k^T d_k$ by strong Wolfe condition, then $g_{k+1}^T d_k \le c_2\|g_k\|^2$

$$\Rightarrow \text{S.T} = 2\left(\frac{\left(-\eta^2(g_{k+1}^T y_k)^2 d_k^T y_k\|g_k\|^2 + \eta\|g_k\|^2(g_{k+1}^T y_k)^2(c_2\|g_k\|^2 - \|g_k\|^2)\right)\|d_k\|^2}{\left((2\|g_k\|^2 - d_k^T y_k)d_k^T y_k\right)^2(g_{k+1}^T d_{k+1})^2}\right)$$

$$\Rightarrow \text{S.T} = -2\left(\frac{\left(\eta^2(g_{k+1}^T y_k)^2 d_k^T y_k\|g_k\|^2 + \eta(1-c_2)\|g_k\|^4(g_{k+1}^T y_k)^2\right)\|d_k\|^2}{\left((2\|g_k\|^2 - d_k^T y_k)d_k^T y_k\right)^2(g_{k+1}^T d_{k+1})^2}\right)$$

$$\text{T.T} = \frac{\eta^2\|g_k\|^4\left(g_{k+1}^T y_k\right)^2}{\left(2\|g_k\|^2 d_k^T y_k - \left(d_k^T y_k\right)^2\right)^2}\frac{\|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} = \frac{\eta^2\|g_k\|^4\left(g_{k+1}^T y_k\right)^2\|d_k\|^2}{\left(4\|g_k\|^4\left(d_k^T y_k\right)^2 - 4\|g_k\|^2\left(d_k^T y_k\right)^3 + \left(d_k^T y_k\right)^4\right)\left(g_{k+1}^T d_{k+1}\right)^2}$$

We know that $g_{k+1}^T d_k \le d_k^T y_k$ and by Wolfe's condition $c_2 {g_k}^T d_k \le d_k^T y_k \Rightarrow -c_2 {g_k}^T d_k \ge -d_k^T y_k$ This implies that $\|g_k\|^2 \ge \frac{-1}{c_2} d_k^T y_k \Rightarrow -c_2\|g_k\|^2 \le d_k^T y_k$

$$\text{T.T} \le \frac{\eta^2\|g_k\|^4\left(g_{k+1}^T y_k\right)^2\|d_k\|^2}{\left(-c_2 4\|g_k\|^6 d_k^T y_k - 4\|g_k\|^2\left(d_k^T y_k\right)^3 - c_2\|g_k\|^2\left(d_k^T y_k\right)^3\right)\left(g_{k+1}^T d_{k+1}\right)^2}$$

$$\Rightarrow \text{T.T} \le \frac{-\eta^2\|g_k\|^4\left(g_{k+1}^T y_k\right)^2\|d_k\|^2}{\left(c_2 4\|g_k\|^6 d_k^T y_k + 4\|g_k\|^2\left(d_k^T y_k\right)^3 + c_2\|g_k\|^2\left(d_k^T y_k\right)^3\right)\left(g_{k+1}^T d_{k+1}\right)^2}$$

Therefore $\eta, c_2, \|g_k\|^2, (g_{k+1}^T y_k)^2, d_k^T y_k, \|d_k\|^2$ and $(g_{k+1}^T d_{k+1})^2$ are greater than zero, then

$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} \le \frac{1}{\|g_{k+1}\|^2}$ Hence $k = 0$ the above inequality yields $\frac{\|d_1\|^2}{\left(g_1^T d_1\right)^2} \le \frac{1}{\|g_1\|^2}$

Hence for all k, we conclude that $\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \le \frac{1}{\|g_k\|^2}$. Therefore $\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \le \sum_{i=0}^{k}\frac{1}{\|g_i\|^2}$ So, by (18)

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \sum_{i=0}^{k}\frac{1}{\mu^2} \Rightarrow \frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \frac{1}{\mu^2}\sum_{i=0}^{k}1 \Rightarrow \frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \frac{k}{\mu^2} \Rightarrow \frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \frac{\mu^2}{k}$$

When we add both sides together, we receive $\sum_{k\ge1}\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \ge \mu^2\sum_{k\ge1}\frac{1}{k} = \infty$

∎

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \infty$$

Which contradicts Theorem 15's Zountendijk condition. After then, the proof is finished.

4. Outcomes of Numerical Examination

A novel approach is tested for this section. The comparison tests comprise well-known nonlinear functions (classical function to be tested) with distinct bounds $5 \leq \mathrm{N} \leq 5000$, and we compare our technique with Conjugate Gradient methods ((HAS) and (PAR)), all cases the stopping condition is $|g_k{}^T g_{k+1}| > 0.2\|g_{k+1}\|^2$, the outcomes are given in table (1) specifically quote the number of functions NOFS and the number of iteration NOIS. More findings from the experiments and table (2) certify that the fresh method CJG is good than the usual CJG (HAS) and standard CG (PAR) concerning the NOIS and NOFS.

Table (1): Speed Equivalences for Three Methods traditional HAS, PAR, and AA4

Test No.	Test Functions	N	Classical Algorithm (HAS)		Classical Algorithm (PAR)		New Algorithm (AA4)	
			NOIS	NOFS	NOIS	NOFS	NOIS	NOFS
1	Powell	5	38	108	35	87	31	86
		50	38	108	35	87	31	86
		500	41	124	35	87	31	86
		1000	41	124	35	87	31	86
		5000	41	124	35	87	31	86
2	OSP	5	10	56	10	56	10	56
		50	36	133	37	147	35	121
		500	112	353	110	341	105	332
		1000	156	473	161	493	143	435
		5000	256	744	276	844	256	776
3	Wood	5	30	68	29	67	27	62
		50	30	68	29	67	29	66
		500	30	68	30	69	29	66
		1000	30	68	30	69	29	66
		5000	30	68	30	69	30	69
4	Sum	5	6	39	6	39	6	39
		50	11	60	11	60	11	60
		500	21	124	21	123	20	109
		1000	23	128	23	127	23	126
		5000	31	159	31	145	28	129
5	Miele	5	28	85	37	116	34	108
		50	31	102	44	148	37	126
		500	40	146	44	148	38	128
		1000	46	176	50	180	38	128
		5000	54	211	50	180	44	162
6	Rosen	5	30	83	30	85	27	78
		50	30	83	30	85	27	78
		500	30	83	30	85	27	78
		1000	30	83	30	85	27	78
		5000	30	83	30	85	29	82
7	G-Biggs	5	20	64	20	63	17	52
		50	FF	FF	FF	FF	FF	FF
		500	FF	FF	FF	FF	FF	FF
		1000	FF	FF	FF	FF	FF	FF
		5000	FF	FF	FF	FF	FF	FF
8	Cubic	5	12	35	15	45	12	35
		50	13	37	16	47	12	35
		500	13	37	16	47	12	35
		1000	13	37	16	47	12	35
		5000	13	37	16	47	12	35
9	Dixon	5	13	30	13	30	13	30
		50	4044	8090	532	1188	458	1035
		500	499	1122	5395	10793	3842	7687
		1000	2366	4737	511	1157	500	1118
		5000	476	1068	516	1159	497	1126
10	G-Central	5	22	159	22	159	22	159
		50	22	159	22	159	22	159
		500	23	171	23	171	23	171
		1000	23	171	23	171	23	171
		5000	28	248	30	270	27	235
11	Fred	5	8	23	6	19	6	19
		50	8	23	6	19	6	19
		500	8	23	6	19	6	19
		1000	8	23	6	19	6	19
		5000	8	23	6	19	6	19
12	Powell3	5	16	36	FF	FF	16	36
		50	16	36	FF	FF	16	36
		500	16	36	FF	FF	16	36
		1000	16	36	FF	FF	16	36
		5000	16	36	16	58	16	36
13	Non-Diagonal	5	25	64	24	64	24	62
		50	29	79	27	74	27	71
		500	FF	FF	27	73	27	71
		1000	29	79	27	73	27	71
		5000	30	81	27	73	27	71
14	Shallow	5	8	21	8	21	8	21
		50	8	21	8	21	8	21
		500	8	21	8	21	8	21
		1000	9	24	9	24	9	24
		5000	9	24	9	24	9	24
	Totals							

Note, FF stands for failure. When both scenarios fail, we overlook the outcomes. When there is success in one instance and failure in the other, we double the values for failure.

Table (2): compares the pace of change progress between the novel (AA4) and conventional (HAS) and (PAR)

Tools	Standard (HAS)	New (AA4)	Standard (PAR)	New (AA4)
NOIS	100%	75.9178%	100%	79.0760%
NOFS	100%	78.7374%	100%	80.9130%

Table (2) displays the rate at which the progress has occurred. in the novel method (AA4) with the classical algorithms (HAS) and (PAR), The new method produces better numerical results than the traditional method., As we observe, (NOIS), (NOFS) of the tradition algorithm (HAS) is about 100%, That is to say, the new algorithm has improved on a normal algorithm (HAS) prorate (24.0822%) in (NOIS) and prorate (21.2626%) in (NOFS) and normal algorithm (PAR) is about 100%, So the new method outperforms the traditional algorithm. (PAR) prorate (20.924%) in (NOIS) and prorate (19.087%) in (NOFS), Overall, the new way (AA4) has been prorated enhanced (21.3390%) compared with standard algorithms (HAS) and (PAR).

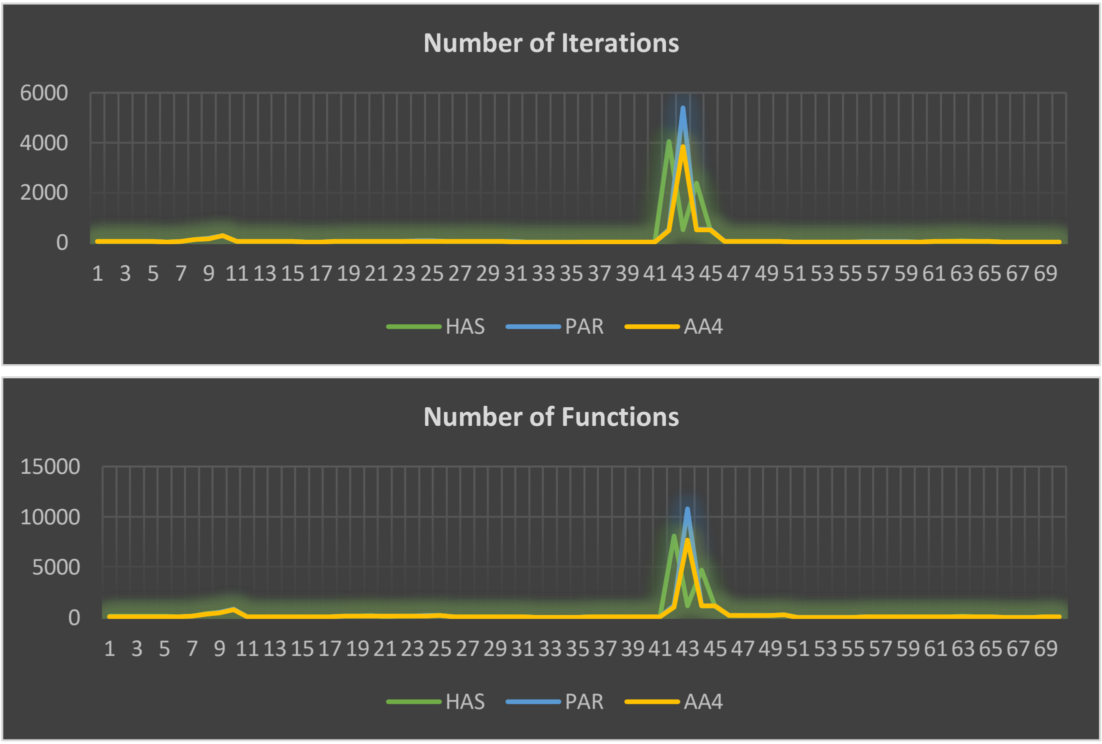


Figure (1): shows clearly that the new method AA4 is better than standard methods (HAS) and (PAR) according to the (NOIS) and the (NOFS).

5. Conclusion

We suggested a novel hybrid β_k^{AA4} for CJG that possesses global convergent property in this part. According to numerical data, this novel β_k^{AA4} outperforms (HAS) and (PAR). We can and have proposed several new hybrids for CJG of no constrained optimization in the future.

References

[1] A.Y. Al - Bayati, and N.H. AL-Assady, *Conjugate gradient method*, Technical Research, school of computer studies, Leeds University, 1(1986).

[2] N. Andrei, *New Hybrid Conjugate Gradient Algorithms for Unconstraint Optimization*, Encyclopedia of Optimization 2nd Edition, C.A. Floudas, and P. Pardalos (Eds.), Entry 76, (2008c).

[3] Y.H. Dai & Y. Yuan, A nonlinear conjugate gradient with a strong global convergence property, *SIAM J. Optim.*, 10(1) (1999), 177–182. https://doi.org/10.1137/s1052623497318992

[4] Y. H. Dia, and Y. Yuan, An Efficient Hybrid Conjugate Gradient Method for Unconstraint Optimization, *Ann. Operation Research*, (2001), 33-47.

[5] R. Fletcher, and C. Reeves, Function minimization by conjugate gradients, Comput. J., 7(2) (1964), 149-154. https://doi.org/10.1093/comjnl/7.2.149

[6] R., Fletcher, *Practical Methods of Optimization vol.1: Unconstrained Optimization*. Jhon Wiley & Sons, New York, (1987).

[7] M.R. Hestenes and E. Steifel, Method of the conjugate gradient for solving linear equations, *J. Res. Nat. Bur. Stand*, 6(49) (1952), 409-436.

[8] I. Jusoh, M. Mamat, and M. Rivaie, A New Edition of Conjugate Gradient Methods for Large-Scale Unconstrained Optimization, *International Journal of Mathematical Analysis*, 46(8) (2014), 2277 -2291. https://doi.org/10.12988/ijma.2014.44115

[9] Y. Liu and C. Storey, *Efficient generalized conjugate gradient algorithms part 1: Theory*, J. Comput. Appl. Math., 69(1) (1991), 129-137. https://doi.org/10.1007/bf00940464

[10] I. Mohammed S., M. Mamat, K. Kamfa, and M. Danlami, A Descent Modification of Conjugate Gradient Method for Optimization Models, *Iraqi Journal of Science*, 61 (7) (2020), 1745-1750.

[11] J. Nocedal, Theory of Algorithms for Unconstraint Optimization, *Acta Numerica*, (1992), 199-242.

[12] E. Polak and G. Ribiere, Note sur la convergence de directionsconjugees, Rev. Francaise Informat Recherche Operationelle. 3E Annee, (16) (1969), 35-43.

[13] M. Rivaie, M. Mamat, L. W. June and I. Mohd, A new conjugate gradient coefficient for large scale nonlinear unconstrained optimization, *International Journal of Mathematical Analysis*, 23(6) (2012), 1131-1146.

[14] Z. Salleh, G. Alhamzi, I. Masmali, and A. Alhawarat, A Modified Liu and Storey Conjugate Gradient Method for Large Scale Unconstrained Optimization Problems, *Algorithms*, 14(8) (2021), 227. https://doi.org/10.3390/a14080227

[15] H. Yan, L. Chen, and B. Jiao, , *HH-LS-SD Hybrid Conjugate Gradient algorithm for Unconstraint Optimization*, Second International Workshop on Computer Science and Engineering, (2009), 264-268.

[16] M. B. Yousef, M. Mamat and M. Rivaie, A New Modification of Conjugate Gradient Method with Global Convergence and Sufficient Descent Properties for Unconstrained Optimization Problems, *ASM Science Journal*, 12(6) (2019).

[17] L. Zhang, A derivative-free conjugate residual method using secant condition for general large-scale nonlinear equations, Numer. Algorithms, 83(4) (2020), 1277–1293. https://doi.org/10.1007/s11075-019-00725-7

[18] G. Zoutendijk, *Nonlinear Programming Computational Methods, Integer and Nonlinear Programming*, J. Ababie (Ed), North-Holand. Amsterd, (1970), 37-86