Bilangan : Jurnal Ilmiah Matematika, Kebumian dan Angkasa
Volume. 3, Nomor. 1 Tahun 2025
e-ISSN :3032-7113; p-ISSN :3032-6389, Hal 60-78
DOI: https://doi.org/10.62383/bilangan.v3i1.378
Available online at: https://journal.arimsi.or.id/index.php/Bilangan

Applying New Preconditioned Conjugated Gradient Algorithms to Unconstrained Optimization Problems

Sabreen M. Abbas[1], Abbas Y. Al-Bayati[2], Maysoon M. Aziz[3]

[1-2] Department of Mathematics, College of Computer Sciences and Mathematics, University of Mosul, Iraq, [3] University of Telafer, Iraq.

Email:mssabreen6@gmail.com[1], profabbasalbayati@yahoo.com[2], Maysoon M.Aziz@gmail.com[3]

Abstract: *In this paper, we study a new and improved preconditioned conjugate gradient (PCG) algorithm based on Dai and Liao's procedure to enhance the CG algorithm of (Maulana). The new PCG algorithm satisfies the coupling condition and the sufficient descent condition. This work proposes improved conjugate gradient methods to enhance the efficiency and robustness of classical conjugate gradient methods. The study changes the diagonal of the inverse Hessian approximation to quasi-Newton Broyden-Fletcher-Goldfarb-Shano (BFGS) updating to make a preconditioner for nonlinear conjugate gradient (NCG) methods used to solve large-scale optimization problems with no constraints. We will calculate the step size of this two-term algorithm by accelerating the Wolfe-Powell line searching technique. The proposed new PCG algorithms have proven their global convergence in certain specific conditions reported in this paper.*

Keywords: *Preconditioned Conjugate Gradient Algorithms, Unconstrained Optimization, Numerical Optimization Techniques.*

1. INTRODUCTION.

We examine the subsequent unconstrained optimization issue.[1]:

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

where $f: R^n \rightarrow R$ The function is continuously differentiable, and its gradient is accessible. We want to develop an approach for addressing large-scale problems when the Hessian of f is either unavailable or necessitates significant storage and processing resources, utilizing iterative techniques of the kind [2]:

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

x_k is the current iteration point and α_k is the step length being computed. The search direction d_k is determined by conducting a line search[3]:

$$d_{k+1} = \begin{cases} -g_{k+1}, & if\ k = 0 \\ -g_{k+1} + \beta_k d_k & if\ k \geq 1 \end{cases}$$

(3)

The procedure employs the conjugate gradient direction, whereby the notable parameter β_k is derived by equating the conjugate gradient direction with that of the Newton method [2], [3].

Any methodology of gradient has to update the point by the line search used. The Wolfe-Powell (WWP) customary search terms square measure usually employed

Received: Desember 07, 2024; Revised: Desember 21, 2024 ;Accepted: Januari 06, 2025;
Online Available : Januari, 09, 2025

in CG ways. The search terms within the Weak Wolfe-Powell (WWP) line square measure as follows:

$$f(x_{k+1}) - f(x_k) \leq \delta \alpha_k g_k^T d_k \quad (4)$$

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k \quad (5)$$

By this condition, d_k is a descent search direction such that $0 < \delta < \sigma < 1$. The Strong Wolfe-Powell (SWP) conditions defined in (4) and satisfies:

$$|g_{k+1}^T d_k| \leq \sigma g_k^T d_k \quad (6)$$

As a generalization of Strong Wolfe conditions, the absolute value opens in (6) with two disparities of inequality so that:

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

Moreover, from all of these previous conditions we obtain the characteristic of sufficient descent, namely:

$$g_k^T d_k \leq -c\|g_k\|^2 \quad (7)$$

as long as c > 0 and is a positive number.

The Preconditioned Conjugate Gradient (PCG) technique is an iterative methodology formulated to resolve extensive sparse systems of linear equations represented as Ax=b, A represent a symmetric positive-definite matrix. This approach is an augmentation of the Conjugate Gradient (CG) technique, improved by the application of a preconditioner to expedite convergence[4], [5].

An invertible matrix s is selected to accelerate convergence, and the concept of preconditioning is to adjust the variables x = sy. This iteration is obtained by first writing the conjugate gradient algorithm in the converted variable y and then translating it back to the x variable[6]:

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

$$d_{k+1} = -H_{k+1} g_{k+1} + H_k \beta_k d_k \quad (8)$$

Where $p = ss^T$. The update parameter β_k^* is equivalent to β_k (conjugacy parameter) except that the vectors g_k and d_k are replaced by $s^T g_k$ and $s^{-1} g_k$ respectively. As illustrations, we have for example β_k^{FR} and β_k^{PR}

$$\beta_k^{*FR} = \frac{z_{k+1}^T g_{k+1}}{g_k^T g_k} = \frac{g_{k+1}^T H_k g_{k+1}}{g_k^T g_k}, \quad \beta_k^{*PR} = \frac{g_{k+1}^T y_k}{g_k^T H_k g_k} \quad (9)$$

The preconditioned residual at iteration $k+1$, computed as $z_{k+1} = S^{-1} g_{k+1}$. We examine the convergence to attain a more profound comprehension of the effects of preconditioning. The velocity of the center of gravity is contingent upon the eigenvalues of the Hessian matrix. Assume that f is a quadratic function:

$$f(x) = \frac{1}{2}x^T Qx + b^T x, \quad (10)$$

Let Q be a symmetric matrix characterized by its eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ all $\lambda_i \geq 0$, $i = 1,2,3,\dots,n$. The error in the k-th conjugate gradient iterate, when employing an exact line search, adheres to the subsequent bound[7]

$$(x_k - x^*)^T Q(x_k - x^*) \leq \min_{p \in \rho_{k-1}} \max_{1 \leq i \leq n} \left(1 + \lambda_i p(\lambda_i)\right)^2 (x_0 - x^*)^T Q(x_0 - x^*) \quad (11)$$

P_k represents the collection of polynomials of degree k. For each integer $l \in [1, k]$, it follows that if $p \in P_k$ selected such that the degree k polynomial

$1 + \lambda p(\lambda)$ vanishes with multiplicity 1 at λ_i, $1 \leq i \leq l - 1$, and with multiplicity $k - l + 1$ at $(\lambda_l + \lambda_n)/2$ then we have

$$(x_k - x^*)^T Q(x_k - x^*) \leq \left(\frac{\lambda_l - \lambda_n}{\lambda_l + \lambda_n}\right)^{2(k-l+1)} (x_0 - x^*)^T Q(x_0 - x^*) \quad (12)$$

Upon substituting the variable x with sy in equation (4.3), we derive

$$f(Sy) = \frac{1}{2}y^T \mathrm{s}^T Q \mathrm{s} y + b^T \mathrm{s} y, \quad (13)$$

The matrix $Q\mathrm{s}\mathrm{s}^T = QP$ is analogous to the matrix $\mathrm{s}^T Q \mathrm{s}$, which is related to the quadratic in y. As a consequence, the optimal preconditioner is $P = Q^{-1}$, which achieves convergence in a single iteration, as the eigenvalues of $\mathrm{s}^T Q \mathrm{s}$ are uniformly equal to 1. Any matrix that approximates the inverse of the Hessian $\nabla^2 f(x^*)^{-1}$ is an effective preconditioner for a generic nonlinear function f. For instance, to illustrate the selection of P (see[8] [9], [10]).

Our aim in this paper in brief, accelerating the speed of the algorithms by the Precondition Conjugate Gradient method and the Variable matrix.

2. NEW PRECONDITION APPROACH

In this section, we will provide a conjugate gradient method to improve Maulana methods, based on our previous exploration of the preconditional conjugate gradient and the equations presented improve Maulana methods in the third chapter. To formulate it, we use the equations that are listed below.

$$d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M1} d_k \tag{14}$$

$$\beta_k^{M1} = \frac{{g_{k+1}}^T \left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k - g_k \right)}{{g_k}^T (g_{k+1} - d_k)} \tag{15}$$

$$d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M2} d_k \tag{16}$$

$$\beta_k^{M2} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|} |{g_{k+1}}^T g_k| - |{g_{k+1}}^T g_k|}{(1-\mu)\|d_k\|^2 + \mu \|g_k\|^2} \tag{17}$$

Where $\mu = 0.6$

A. Outline of New Precondition CG Algorithms ($M1$)

a. For $x_0 \in R^n$ initial point of minimum , $0 < \varepsilon < 1$, $0 < \delta < \frac{1}{2}$, $and\ \delta < \sigma < 1$, $H_0 = I\ (identity\ matrix)$.

b. set $d_0 = -H_0 g_0$, $k = 0$.

c. If $\|g_k\| < \varepsilon$, then stop, otherwise continue to the next step.

d. Compute step size α_k by Wolfe line search (4),(6).

e. Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| < \varepsilon$, then stop.

f. Calculate the new search directions PCG by:

$$d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M1} d_k$$

$$\beta_k^{M1} = \frac{{g_{k+1}}^T \left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k - g_k \right)}{{g_k}^T (g_{k+1} - d_k)}$$

g. Set $k = k + 1$, and go to step 3.

B. Property (Gilbert and J. Nocedal, 1992)

Suppose that the general conjugate gradient method is used and that $0 < \delta \leq \|g_k\| \leq \bar{\delta}$ is achieved in it, then this method has property (2.1.1) if the constants $b > 1$ and $p > 1$ are found, for example, for every k,

$$|\beta_k| \leq b \tag{18}$$

$$\|s_k\| \leq p \Rightarrow |\beta_k| \leq \frac{1}{2b} \tag{19}$$

C. Assumption A.

Let f(x) have its lower bound defined on the set

S={x∈R^n,f(x)≤f(x_0)}, (20)

where x_0 is the initial point.

D. Assumption B.

The objective function is continuously differentiated and the gradient of the objective function is Lipchitz continuous in some neighborhood N of S, such that:

‖g(x)-g(y)‖≤L‖x-y‖ ∀x,y∈N. (21)

The sequence $\{x_k\}$ produced by NEW2 and NEW2 is included in S since $\{f(x_k)\}$ is decreasing. Furthermore, from Assumption A, we can deduce that for any constant B and any B and $\gamma_1 >$, we have:

$$\|x\| \le B, \|g(x)\| \le \gamma_1, \forall x \in S \quad (22)$$

Assumption A's conditions are taken as given throughout the paper's subsequent sections. Then, there is a handy lemma that was first presented in [4,10].

E. Theorem (2.1.3) (Zoutendijk condition)

Suppose that Assumption (2.1.2) (2.1.3) holds. Any CG method of the form (10), where d_k is a descent search direction and α_k satisfies the (SWP) in (4),(6). Then the following holds[11]:

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \quad (23)$$

From **Theorem (2.1.3)** and from (20)-(21) for the New Precondition CG Algorithms with the Wolfe line search, we can easily obtain the following condition:

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < +\infty \quad (24)$$

F. Theorem

Suppose that Assumption (2.1.2) (2.1.3) holds. Any CG technique of the forms (2) and (3) with d_k constitutes a descent search direction, and α_k adheres to the Strong Wolfe Conditions (SWP) in (4) and (6). If Any CG method of the

$$\sum_{k\ge 0} \frac{1}{\|d_k\|^2} = \infty \Rightarrow \lim_{k\to\infty} \inf\|g_k\| = 0 \quad (25)$$

[12]

3. THE DESCENT PROPERTY OF A CG NEW METHOD (M1)

The descent property for our suggested conjugate gradient technique, designated as PCG, must be shown below. Subsequently, we assert the adequate decline.

Starting by the direction of precondition(14)

$$d_{k+1}^{\mathrm{NEWi}} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{\mathrm{M1}} d_k \tag{26}$$

Where

$$H_{k+1}^{BFGS} = \left(H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k}\right)$$

$$H_{k+1}^{BFGS} = (a_k + b_k) \tag{27}$$

$$\beta_k^{M1} = \frac{{g_{k+1}}^T \left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_{k-1}\|} g_k - g_k\right)}{{g_k}^T (g_{k+1} - d_k)} \tag{28}$$

Multiply (26) by g_{k+1}

$$d_{k+1}^T g_{k+1} = -H_{k+1}^{BFGS} \|g_{k+1}\|^2 + H_k^{BFGS} \beta_k^{\mathrm{M1}} d_k^T g_{k+1} \tag{29}$$

By using (IEL)

$$d_k^T g_{k+1} = d_k^T g_{k+1} - d_k^T g_k + d_k^T g_k$$

$$= d_k^T (g_{k+1} - g_k) + d_k^T g_k = d_k^T y_k + d_k^T g_k < d_k^T y_k \tag{30}$$

Substituting (30) in (29)

$$d_{k+1}^T g_{k+1} = -H_{k+1}^{BFGS} \|g_{k+1}\|^2 + H_k^{BFGS} \beta_k^{\mathrm{M1}} d_k^T y_k$$

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

$$d_{k+1}^T g_{k+1} = -(a_k + b_k) \|g_{k+1}\|^2$$
$$+ \left(\frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_k\|} g_{k+1}^T g_k - g_{k+1}^T g_k}{{g_k}^T g_{k+1} - {g_k}^T d_k}\right) H_k^{BFGS} d_k^T y_k$$

$H_k = H_k^{BFGS}$ $+ive$ definite always, so that

And by Powell condition

$$g_{k+1}^T g_k = 0.2 \|g_{k+1}\|^2$$

$$d_{k+1}^T g_{k+1} = -(a_k + b_k) \|g_{k+1}\|^2$$
$$+ \left(\frac{\|g_{k+1}\|^2 + \frac{0.2 \|g_{k+1}\|^3}{\|g_k\|} + 0.2 \|g_{k+1}\|^2}{{g_k}^T g_{k+1} - {g_k}^T d_k}\right) (a_{k-1} + b_{k-1}) d_k^T y_k$$

By using descent condition

$$g_k{}^T d_k \leq 0$$

$$d_{k+1}^T g_{k+1} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{\frac{\|g_k\|\|g_{k+1}\|^2 + 0.2\|g_{k+1}\|^3 + 0.2\|g_k\|\|g_{k+1}\|^2}{\|g_k\|}}{g_k{}^T g_{k+1}}\right)(a_{k-1} + b_{k-1})d_k^T y_k$$

$$d_{k+1}^T g_{k+1} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{\|g_k\| + 0.2\|g_{k+1}\| + 0.2\|g_k\|}{\|g_k\|g_k{}^T g_{k+1}}\right)\|g_{k+1}\|^2(a_{k-1} + b_{k-1})d_k^T y_k$$

$$d_{k+1}^T g_{k+1} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{1.2\|g_k\| + 0.2\|g_{k+1}\|}{\|g_k\|g_k{}^T g_{k+1}}\right)\|g_{k+1}\|^2(a_{k-1} + b_{k-1})d_k^T y_k \qquad (31)$$

$$d_{k+1}^T g_{k+1} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{1.2\|g_k\| + 0.2\|g_{k+1}\|}{\|g_k\|g_k{}^T g_{k+1}}\right)\|g_{k+1}\|^2(a_{k-1} + b_{k-1})d_k^T y_k$$

$$d_{k+1}^T g_{k+1} \leq -\left((a_k + b_k) + \left(\frac{1.2\|g_k\| + 0.2\|g_{k+1}\|}{\|g_k\|g_k{}^T g_{k+1}}\right)(a_{k-1} + b_{k-1})d_k^T y_k\right)\|g_{k+1}\|^2$$

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

where

$$c > 0$$

4. GLOBAL CONVERGENCE (M1)

For the same assumptions and preliminaries in previous chapters, we complete our theoretics analysis.

A. Theorem(4.1):

Let the **Property** (2.1.2) (2.1.3) be fulfilled, and the CG algorithm in (8) and (13), since d_{k+1} is a sufficient descent direction, for every $k \geq 0$, then $\|s_k\|$ approaches zero, and if the constants are found $\delta, \bar{\delta}, \gamma, \text{and} \bar{\gamma}$ are in this form $(0 < \delta \leq \|g_k\| \leq \bar{\delta})$ $(0 < \gamma \leq \|g_{k+1}\| \leq \bar{\gamma})$, and that the function f is a general function with the Lipschitz condition, then

$$\lim_{k\to\infty} \inf\|g_k\| = 0. \qquad (32)$$

Proof:

The sequence of approximate solution generated by the precondition direction$\{d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M1} d_k\}$

$|d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M1} d_k|$

$$\|d_{k+1}\| = |H_{k+1}^{BFGS}| \|g_{k+1}\| + |H_k^{BFGS}| |\beta_k^{M1}| \|d_k\| \tag{33}$$

We take each part separately and deduce its value:

$$H_{k+1}^{BFGS} = \left(H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k}\right)$$

$$|H_{k+1}^{BFGS}| = \left|\left(H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T H_k y_k}{y_k^T s_k}\right) \frac{s_k s_k^T}{y_k^T s_k}\right)\right|$$

$$|H_{k+1}^{BFGS}| \leq \left(|H_k| - \frac{|H_k| \|y_k\| \|s_k\| + \|s_k\| \|y_k\| |H_k|}{\|y_k\| s_k}\right) + \left(\left(1 + \frac{\|y_k\| |H_k| \|y_k\|}{\|y_k\| \|s_k\|}\right) \frac{\|s_k\| \|s_k\|}{\|y_k\| \|s_k\|}\right)$$

From Lipschitz condition we have

$\|y_k\| \leq l \|s_k\|$

$$H_{k+1}^{BFGS} \leq \left(|H_k| - \frac{|H_k| l \|s_k\|^2 + l \|s_k\|^2 |H_k|}{l \|s_k\|^2}\right) + \left(\left(1 + \frac{l^2 \|s_k\|^2 |H_k|}{l \|s_k\|^2}\right) \frac{\|s_k\|^2}{l \|s_k\|^2}\right) \tag{34}$$

And simplistically (25)

$$H_{k+1}^{BFGS} \leq (|H_k| - 2|H_k|) + \left(\frac{1}{l} + |H_k|\right)$$

$$H_{k+1}^{BFGS} \leq \frac{1}{l} \tag{35}$$

Or we can say since H_{k+1}^{BFGS} +ive definite then directly we obtain the global convergence requirements, if we continue by brevis context of the theorem

$$\beta_k^{M1} = \frac{{g_{k+1}}^T \left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k - g_k\right)}{{g_k}^T (g_{k+1} - d_k)}$$

$$\beta_k^{M1} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|^2}{\|g_k\|} g_k - g_{k+1}^T g_k}{{g_k}^T g_{k+1} - {g_k}^T d_k}$$

$$|\beta_k^{M1}| = \left|\frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|^2}{\|g_k\|} g_k - g_{k+1}^T g_k}{{g_k}^T g_{k+1} - {g_k}^T d_k}\right|$$

$$|\beta_k^{M1}| \le \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|^2}{\|g_k\|}\|g_k\| - \|g_{k+1}\|\|g_k\|}{\|g_k\|\|g_{k+1}\| - \|g_k\|\|d_k\|}$$

$$|\beta_k^{M1}| \le \frac{-\|g_{k+1}\|\|g_k\|}{\|g_k\|\|g_{k+1}\| - \|g_k\|\|d_k\|}$$

$$|\beta_k^{M1}| \le \frac{-\bar{\gamma}\bar{\delta}}{\bar{\delta}\,\bar{\gamma} - \bar{\delta}\,\|d_k\|}$$

$$|\beta_k^{M1}| \le \frac{-\bar{\gamma}\bar{\delta}}{-\bar{\delta}\,(\|d_k\| - \bar{\gamma})}$$

$$|\beta_k^{M1}| \le \frac{\bar{\gamma}}{\|d_k\| - \bar{\gamma}} = \mathrm{E}_1 \qquad (36)$$

Substituting the (35) and(36) in(33)

$$\|d_{k+1}\| = \frac{1}{l}\bar{\gamma} + \mathrm{E}_1\|\mathrm{d_k}\|$$

$$\|d_{k+1}\| = \frac{1}{l}\bar{\gamma} + \mathrm{E}_1\|\mathrm{d_k}\|$$

Then we get

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

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \le \sum_{k=0}^{\infty} \frac{1}{c^2}\frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

Then

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

5. Outline of New Precondition CG Algorithms ($M2$)

1. For $x_0 \in R^n$ initial point of minimum , $0 < \varepsilon < 1$, $0 < \delta < \frac{1}{2}$, $and\ \delta < \sigma < 1$, $H_0 = I$ $(identity\ matrix)$.
2. set $d_0 = -H_0 g_0$, $k = 0$.
3. If $\|g_k\| < \varepsilon$, then stop, otherwise continue to the next step.
4. Compute step size α_k by Wolfe line search (4), (6).
5. Let $x_{k+1} = x_k + \alpha_k d_k$, if $\|g_{k+1}\| < \varepsilon$, then stop.
6. Calculate the new search directions PCG by:

$$d_{k+1}^{NEWi} = -H_{k+1}^{BFGS} g_{k+1} + H_k^{BFGS} \beta_k^{M2} d_k$$

$$\beta_k^{M2} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|} |g_{k+1}{}^T g_k| - |g_{k+1}{}^T g_k|}{(1-\mu)\|d_k\|^2 + \mu\|g_k\|^2}$$

7. Increment k by 1, and go to step 3.

6. The Descent Property of a CG New Method(M2)

The descent property for our proposed new conjugate gradient scheme must be demonstrated below, referred to as PCG. In the next, we argue the sufficient descent.

Starting by the direction of precondition(16)

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{NEWi}} = -H_{k+1}^{BFGS}\mathrm{g}_{\mathrm{k+1}} + H_k^{BFGS}\beta_{\mathrm{k}}^{\mathrm{M2}}\mathrm{d}_{\mathrm{k}} \quad (37)$$

Where

$$H_{k+1}^{BFGS} = \left(\mathrm{H_k} - \frac{\mathrm{H_k} y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T \mathrm{H_k} y_k}{y_k^T s_k}\right)\frac{s_k s_k^T}{y_k^T s_k}\right)$$

$$H_{k+1}^{BFGS} = (a_k + b_k) \quad (38)$$

$$\beta_k^{M2} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|} |g_{k+1}{}^T g_k| - |g_{k+1}{}^T g_k|}{0.4\|d_k\|^2 + 0.6\|g_k\|^2} \quad (39)$$

where $\mu = 0.6$

Multiply (37) by $\mathrm{g}_{\mathrm{k+1}}$

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} = -H_{k+1}^{BFGS}\|g_{k+1}\|^2 + H_k^{BFGS}\beta_{\mathrm{k}}^{\mathrm{M2}} d_k^T \mathrm{g}_{\mathrm{k+1}} \quad (40)$$

By using (IEL)

$$d_k^T g_{k+1} = d_k^T g_{k+1} - d_k^T g_k + d_k^T g_k$$
$$= \mathrm{d_k^T}(\mathrm{g_{k+1}} - \mathrm{g_k}) + \mathrm{d_k^T g_k} = \mathrm{d_k^T y_k} + \mathrm{d_k^T g_k} < \mathrm{d_k^T y_k} \quad (41)$$

Substituting (41) in (40)

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} = -H_{k+1}^{BFGS}\|g_{k+1}\|^2 + H_k^{BFGS}\beta_{\mathrm{k}}^{\mathrm{M1}} d_k^T \mathrm{y}_{\mathrm{k}}$$

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} = -H_{k+1}^{BFGS}\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|}|{g_{k+1}}^T g_k| - |{g_{k+1}}^T g_k|}{0.4\|d_k\|^2 + 0.6\|g_k\|^2} H_k^{BFGS} d_k^T \mathrm{y}_{\mathrm{k}}$$

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} = -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|}|{g_{k+1}}^T g_k| - |{g_{k+1}}^T g_k|}{0.4\|d_k\|^2 + 0.6\|g_k\|^2}\right) H_k^{BFGS} d_k^T \mathrm{y}_{\mathrm{k}}$$

$H_k = H_k^{BFGS}$ $+ive$ definite always, so that

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} = -(a_k + b_k)\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|y_k\|}|{g_{k+1}}^T g_k| - |{g_{k+1}}^T g_k|}{0.4\|g_k\|^2 + 0.6\|g_k\|^2}(a_{k-1} + b_{k-1}) d_k^T \mathrm{y}_{\mathrm{k}}$$

By using descent condition

$${g_k}^T d_k \leq 0$$

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \frac{\frac{\|y_k\|\|g_{k+1}\|^2 - \|g_{k+1}\||{g_{k+1}}^T g_k| - \|y_k\||{g_{k+1}}^T g_k|}{\|y_k\|}}{\|g_k\|^2}(a_{k-1} + b_{k-1}) d_k^T \mathrm{y}_{\mathrm{k}}$$

By Powell condition we get

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \frac{\frac{l\|s_k\|\|g_{k+1}\|^2 + 0.2\|g_{k+1}\|^3 + 0.2l\|s_k\|\|g_{k+1}\|^2}{l\|s_k\|}}{\|g_k\|^2}(a_{k-1} + b_{k-1}) d_k^T \mathrm{y}_{\mathrm{k}}$$

$$\mathrm{d}_{\mathrm{k+1}}^{\mathrm{T}}\mathrm{g}_{\mathrm{k+1}} \leq -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{l\|s_k\| + 0.2\|g_{k+1}\| + 0.2l\|s_k\|}{l\|s_k\|\|g_k\|^2}\right)\|g_{k+1}\|^2(a_{k-1} + b_{k-1}) d_k^T \mathrm{y}_{\mathrm{k}} \qquad (42)$$

$$\mathrm{d}_{k+1}^{T}\mathrm{g}_{k+1} \le -(a_k + b_k)\|g_{k+1}\|^2 + \left(\frac{1.2l\|s_k\| + 0.2\|g_{k+1}\|}{l\|s_k\|\|g_k\|^2}\right)\|g_{k+1}\|^2\,(a_{k-1} + b_{k-1})\,d_k^T\mathrm{y}_k$$

$$\mathrm{d}_{k+1}^{T}\mathrm{g}_{k+1} = -\left((a_k + b_k) - \left(\frac{1.2l\|s_k\| + 0.2\|g_{k+1}\|}{l\|s_k\|\|g_k\|^2}\right)(a_{k-1} + b_{k-1})\,d_k^T\mathrm{y}_k\right)\|g_{k+1}\|^2$$

$$\mathrm{d}_{k+1}^{T}\mathrm{g}_{k+1} \le -c\|g_{k+1}\|^2$$

where

$c > 0.$

7. Global convergence(M2)

For the same assumptions and preliminaries in previous chapters, we complete our theoretics analysis.

Theorem:

Let the **Property** (2.1.2) (2.1.3)be fulfilled, and the CG algorithm in (8) and (16), since d_{k+1} is a sufficient descent direction, for every $\mathrm{k} \ge 0$, then $\|\mathrm{s}_k\|$ approaches zero, and if the constants are found $\delta, \bar{\delta}, \gamma, \text{and} \bar{\gamma}$ are in this form $(0 < \delta \le \|\mathrm{g}_k\| \le \bar{\delta})$ $(0 < \gamma \le \|\mathrm{g}_{k+1}\| \le \bar{\gamma})$, and that the function f is a general function with the Lipschitz condition, then

$$\lim_{k\to\infty} \inf\|\mathrm{g}_k\| = 0. \tag{43}$$

Proof:

The sequence of approximate solution generated by the precondition direction$\{\mathrm{d}_{k+1}^{\mathrm{NEWi}} = -H_{k+1}^{BFGS}\mathrm{g}_{k+1} + H_k^{BFGS}\beta_k^{\mathrm{M2}}\mathrm{d}_k\}$

$$\left|\mathrm{d}_{k+1}^{\mathrm{NEWi}} = -H_{k+1}^{BFGS}\mathrm{g}_{k+1} + H_k^{BFGS}\beta_k^{\mathrm{M2}}\mathrm{d}_k\right|$$

$$\|d_{k+1}\| = \left|H_{k+1}^{BFGS}\right|\|\mathrm{g}_{k+1}\| + \left|H_k^{BFGS}\right|\left|\beta_k^{\mathrm{M2}}\right|\|\mathrm{d}_k\| \tag{44}$$

We take each part separately and deduce its value:

$$H_{k+1}^{BFGS} = \left(\mathrm{H}_k - \frac{\mathrm{H}_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T \mathrm{H}_k y_k}{y_k^T s_k}\right)\frac{s_k s_k^T}{y_k^T s_k}\right)$$

$$\left|H_{k+1}^{BFGS}\right| = \left|\left(\mathrm{H}_k - \frac{\mathrm{H}_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k}\right) + \left(\left(1 + \frac{y_k^T \mathrm{H}_k y_k}{y_k^T s_k}\right)\frac{s_k s_k^T}{y_k^T s_k}\right)\right|$$

$$\left|H_{k+1}^{BFGS}\right| \le \left(|\mathrm{H_k}| - \frac{|\mathrm{H_k}|\|y_k\| * \|s_k\| + \|s_k\| * \|y_k\||\mathrm{H_k}|}{\|y_k\|\|s_k\|}\right) + \left(\left(1 + \frac{\|y_k\||\mathrm{H_k}|\|y_k\|}{\|y_k\|\|s_k\|}\right)\frac{\|s_k\|\|s_k\|}{\|y_k\|\|s_k\|}\right)$$

From Lipschitz condition we have

$$\|y_k\| \le l\|s_k\|$$

$$H_{k+1}^{BFGS} \le \left(|\mathrm{H_k}| - \frac{|\mathrm{H_k}|l\|s_k\|^2 + l\|s_k\|^2|\mathrm{H_k}|}{l\|s_k\|^2}\right) + \left(\left(1 + \frac{l^2\|s_k\|^2|\mathrm{H_k}|}{l\|s_k\|^2}\right)\frac{\|s_k\|^2}{l\|s_k\|^2}\right)$$

And simplistically (43)

$$H_{k+1}^{BFGS} \le (|\mathrm{H_k}| - 2|\mathrm{H_k}|) + \left(\frac{1}{l} + |\mathrm{H_k}|\right)$$

$$H_{k+1}^{BFGS} \le \frac{1}{l} \qquad (45)$$

Or we can say since H_{k+1}^{BFGS} +ive definite then directly we obtain the global convergence requirements, if we continut by brevise context of the theorem

$$\beta_k^{M2} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|}|g_{k+1}{}^T g_k| - |g_{k+1}{}^T g_k|}{(1-\mu)\|d_k\|^2 + \mu\|g_k\|^2}$$

where μ =0.6

$$\left|\beta_k^{M2} = \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|}|g_{k+1}{}^T g_k| - |g_{k+1}{}^T g_k|}{0.4\|d_k\|^2 + 0.6\|g_k\|^2}\right|$$

$$|\beta_k^{M2}| \le \frac{\|g_{k+1}\|^2 - \frac{\|g_{k+1}\|}{\|g_{k+1} - g_k\|}\|g_{k+1}\| * \|g_k\| - \|g_{k+1}\| * \|g_k\|}{0.4\|d_k\|^2 + 0.6\|g_k\|^2}$$

$$|\beta_k^{M2}| \le \frac{\frac{\|y_k\|\|g_{k+1}\|^2 - \|g_{k+1}\|^2\|g_k\| - \|y_k\|\|g_{k+1}\| * \|g_k\|}{\|y_k\|}}{\|g_k\|^2}$$

$$|\beta_k^{M2}| \le \frac{\|y_k\|\|g_{k+1}\|^2 - \|g_{k+1}\|^2\|g_k\| - \|y_k\|\|g_{k+1}\| * \|g_k\|}{\|y_k\|\|g_k\|^2}$$

$$|\beta_k^{M2}| \le \frac{l\|s_k\| * \|g_{k+1}\|^2 - \|g_{k+1}\|^2\|g_k\| - l\|s_k\|\|g_{k+1}\|\|g_k\|}{l\|s_k\|\|g_k\|^2}$$

Let $\|s_k\| = D$

$$|\beta_k^{M2}| \leq \frac{lD\bar{\gamma}(\bar{\gamma} - \bar{\delta}) - \bar{\gamma}^2\bar{\delta}}{lD\bar{\delta}^{\,2}} = E_2 \qquad (46)$$

Substituting the (39) and(38) in(37)

$$\|d_{k+1}\| = \frac{1}{l}\bar{\gamma} + E_2\|d_k\|$$

$$\|d_{k+1}\| = \frac{1}{l}\bar{\gamma} + E_2\|d_k\|$$

Then we get

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

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \leq \sum_{k=0}^{\infty} \frac{1}{c^2} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

then

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

8. NUMERICAL EXAMPLES

In this paper, we test and compare the new algorithm Maulana1 (M1) and Maulana2 (M2), with legacy algorithm BFGS, on unconstrained problems. We apply this to problems with dimensions 100 and 1000. We execute all computations on an HP laptop, outfitted with the Win10 operating system, 4 GB of RAM, and a Core i7, utilizing the Fortran program. We then utilize MATLAB programming to compare the results in the form of curves the Algorithms are based on the number of iterations (NOI), the number of function evaluations(NOFG), and(TIME). Experimental results and discussion are as follows:

In Table (7.1) we compare (BFGS70 Full VM) with both (Maulana1 with BFGS and Maulana2 with BFGS) methods. Total number of occurrences NOI, total job evaluations NOFG, and total TIME to solve 40 test problems.

Table(7.1). Comparison between (Maulana1with BFGS) against (Maulana2 with BFGS) PCG-Algorithms

Table(7.2) Percentage performance of (Maulana1with BFGS) against (Maulana2 with BFGS)

Test Problems	**BFGS70 Full VM-algorithm**	**Maulana1with BFGS NOI/NOFG/TIME**	**Maulana2with BFGS BNOI/NOFG/TIME**
1- Trigonometric	102/198/0.03	86/173/0.18	85/164/0.02
2- Extended Rosenbrock (CUTE)	157/364/0.01	157/364/0.04	126/289/0.03
3- Extended White & Holst	162/363/0.01	162/363/ 0.04	133/299/0.03
4- Extended Beale	50/102/0.01	48 /98/ 0.01	47/104/0.01
5- Raydan 2	12/36/0.00	12/ 36/0.01	12/36/0.02
6- Extended Tridiagonal 1	42/93/0.00	37/83/0.01	41/89/0.00
7- Extended Three Expo Terms	306/8089/1.29	68/529/0.46	68/529/0.43
8- Generalized Tridiagonal 2	198/310/0.01	175/278/0.05	179/292/0.05
9- Diagonal 4	16/40/0.00	16/40/0.01	16/40/0.00
10- Diagonal 5	12/36/0.02	11/34/0.02	11/34/0.03
11- Extended Himmelblau	72/124/0.00	66/115/0.02	66/115/0.01
12- Extended PSC1	25/63/0.00	23/58/0.03	23/58/0.03
13- Extended BD1	252/402/0.01	220/354/0.08	89/174/0.03
14- Extended Hiebert	324/752/0.02	324/752/0.08	273/676/0.08
15- Extended EP1	2/16/0.00	1/14/ 0.00	1/14/0.00
16- Extended Tridiagonal 2	143/234/0.00	113/183/0.03	114/177/0.04
17- ARROWHEAD (CUTE)	41/318/0.02	24/57/0.01	21/67/0.01
18- NONDIA (CUTE)	44/93/0.02	36/77/0.01	36/76/0.00
19- DQDRTIC (CUTE)	20/52/0.00	20/ 52/0.01	20/52/0.00
20- DIXMAANA (CUTE)	24/56/0.02	21/50/0.02	20/48/0.01
21- DIXMAANB (CUTE)	39/75/0.01	36/69/0.02	36/69/0.02
22- DIXMAANC (CUTE)	51/99/0.00	48/93/0.02	48/93/0.03
23- Broyden Tridiagonal	144/245/0.02	125/214/0.03	126/242/0.04
24- Tridiagonal Perturbed Quadratic	53/153/0.00	53/153/0.02	53/153/0.02
25- LIARWHD (CUTE)	77/172/0.00	71/159/0.02	65/145/0.03
26- DIAGONAL 6	12/36/0.00	12/36/0.01	12/36/0.01
27- DENSCHNA (CUTE)	34/74/0.00	33/72/0.02	33/72/0.02
28- DENSCHNC (CUTE)	46/96/0.02	46/96/0.04	43/94/0.04
29- DENSCHNB (CUTE)	24/60/0.00	24/60/0.01	20/52/0.01
30- DENSCHNF (CUTE)	80/148/0.02	76/142/0.03	84/157/0.03
31- Extended Block-Diagonal BD2	40/86/0.01	37/81/0.03	34/75/0.03
32- Generalized quartic GQ1	25/70/0.00	23/66/0.01	23/66/0.01
33- DIAGONAL 7	11/42/0.00	11/ 42/ 0.02	11/42/0.01
34- Full Hessian	8/28/0.00	8 /28/ 0.01	8/28/0.02
35- SINCOS	25/63/0.02	23/58/0.03	23/58/0.03
36- Generalized quartic GQ2	131/215/0.01	115/188/0.03	116/189/0.04
37- ARGLINB (CUTE)	0/12/0.00	0/12/0.00	0/12/0.00
38- FLETCHCR (CUTE)	97/209/0.02	88/191/ 0.03	93/197/0.03
39- HIMMELBG (CUTE)	32/44/0.00	28/40/0.01	28/40/0.01
40- HIMMELBH (CUTE)	24/56/0.00	24/56/0.00	24/56/0.01
Total	**2.957/13.724/1.6**	**2.501/5.566/1.51**	**2.261/5.209/1.27**
Total Work= NOI+NOFG+TIME	**18.281**	**9.577**	**8.74**

algorithms

TOOLS	BFGS	Maulana1with BFGS	Maulana2with BFGS
NOI	100%	84.6%	76.5%
NOFG	100%	40.6%	38%
TIME	100%	94.4%	79.4%

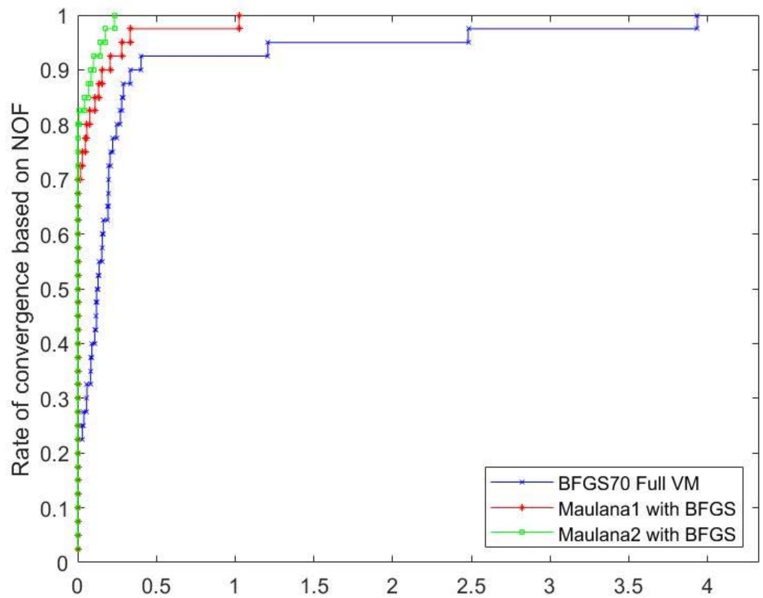
Table(7.3) No. of best NOFG test problems

Tools	No. of best NOFG BFGS	No. of best NOFG Maulana1 with BFGS	No. of equal NOFG In both
NOI	0	26	14
NOFG	0	26	14
TIME	30	3	7

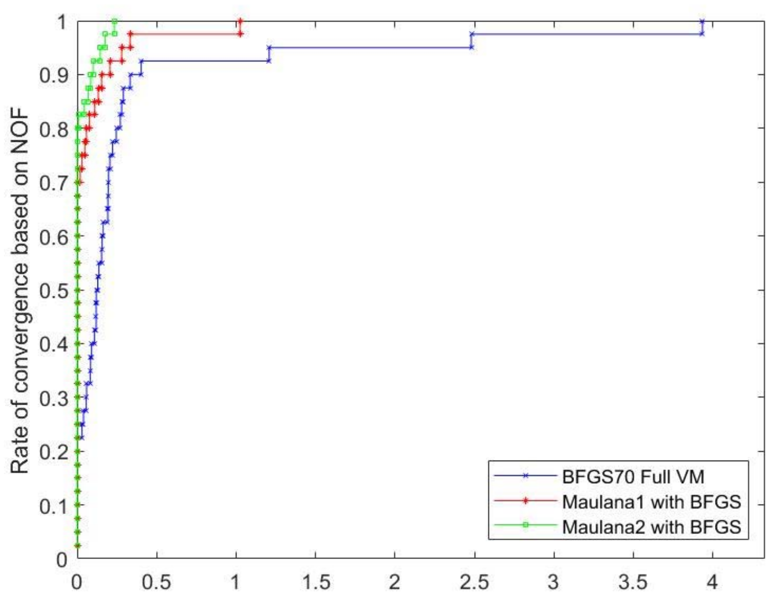
Table(7.4) No. of best NOFG test problems

Tools	No. of best NOFG BFGS	No. of best NOFG Maulana2 with BFGS	No. of equal NOFG In both
NOI	1	30	9
NOFG	2	29	9
TIME	29	5	6

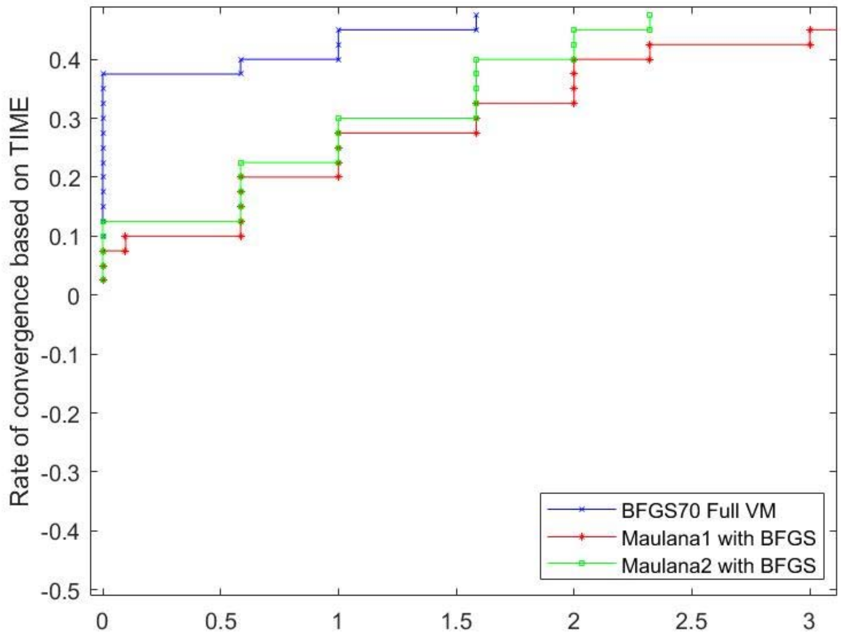
Tables(7.1-7.4) show that compared to the baseline BFGS70 Full VM-algorithm, the (PCG-Maulana1with BFGS) algorithm improves upon it by a factor of (**84.6%,40.6%**, and **94.4%)** in terms of NOI, NOFG, and TIME. Comparing the (PCG-Maulana2with BFGS) algorithm improves upon it by a factor of **(76.5%, 38%, 79.4%)** in terms of NOI, NOFG, and CPU, as shown in Table(6.3-6.4), reveals that the (PCG -Maulana1and PCG -Maulana2) algorithm achieves the strongest results in NOI, NOFG, and TIME under the accelerated Wolfe-Powell line search, demonstrating that the (PCG -Maulana1and PCG -Maulana2) algorithm is significantly more effective than the (BFGS70 Full VM-algorithm) CG-algorithm. The (PCG -Maulana1and PCG -Maulana2) algorithm achieves its best performance by making full use of all available resources (including NOI, NOFG, and TIME.



Figure(7.1): Profile of performance (BFGS against Maulana1 with BFGS) relative to the NOI



Figure(7.2): Profile of performance (BFGS against Maulana1 with BFGS) relative to the NOF



Figure(7.3): Profile of performance (BFGS against Maulana1 with BFGS) relative to the TIME

Tables ((7.1)-(7.4))demonstrate that the combinations of (Maulana1 and Maulana2 with BFGS), exhibit significantly greater efficiency compared to the individual method (BFGS). Our numerical analysis reflects the most recent findings. The technological mechanism in this context is noteworthy. Figures (7.1) through (7.3) illustrate the performance profiles of our approach in comparison to other methods. The data indicates that PCG The process demonstrates superior performance compared to identical iterative methods. Assessment of employment and duration. Based on the analysis of figures (7.1) - (7.3), we determined that the new algorithms outperform the algorithm used for comparison in our study**.**

9. CONCLUSION

In conclusion, this study introduces a novel and enhanced Preconditioned Conjugate Gradient (PCG) algorithm grounded in Dai and Liao's procedure, aiming to improve the efficiency and robustness of classical conjugate gradient methods, including Maulana's approach. By satisfying the coupling and sufficient descent conditions, the proposed algorithm demonstrates significant advancements in addressing large-scale unconstrained optimization problems. The integration of a modified quasi-Newton BFGS preconditioner and an accelerated Wolfe-Powell line search for step size calculation further enhances its performance. Theoretical analysis confirms the global convergence of the algorithm under specific conditions, highlighting its potential for practical optimization tasks. Future work may explore its application in diverse problem domains and further refine its preconditioning strategies**.**

REFERENCES

Andrei, N. (2007). Scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization. Optim Methods Softw, 22(4), 561–571.

Andrei, N. (2020). Nonlinear conjugate gradient methods for unconstrained optimization. Springer.

Aziz, R. F., Younis, M. S., & Jameel, M. S. (2024). An investigation of a new hybrid conjugate gradient approach for unconstrained optimization problems. Bilangan: Jurnal Ilmiah Matematika, Kebumian dan Angkasa, 2(4), 11–23.

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

Dener, A., Denchfield, A., & Munson, T. (2019). Preconditioning nonlinear conjugate gradient with diagonalized quasi-Newton. In Proceedings of the Platform for Advanced Scientific Computing Conference (pp. 1–7).

Hager, W. W., & Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. Pacific Journal of Optimization, 2(1), 35–58.

Jameel, M. S., Abdullah, Z. M., Fawzi, F. A., & Hassan, B. A. (2021). A new shifted conjugate gradient method based on shifted quasi-Newton condition. Journal of Physics: Conference Series, IOP Publishing, 12105.

Jameel, M. S., Basheer, G. T., Al-Bayati, A. Y., & Algamal, Z. Y. (2021). Parameter estimation of a truncated regression model based on improving numerical optimization algorithms. Journal of Physics: Conference Series, IOP Publishing, 012059.

Jameel, M., Al-Bayati, A., & Algamal, Z. (2023). Scaled multi-parameter (SMP) nonlinear QN-algorithms. AIP Conference Proceedings, AIP Publishing.

Knyazev, A. V., & Lashuk, I. (2008). Steepest descent and conjugate gradient methods with variable preconditioning. SIAM Journal on Matrix Analysis and Applications, 29(4), 1267–1280.

Stiefel, E. L. (1958). Kernel polynomial in linear algebra and their numerical applications. In Further contributions to the determination of eigenvalues (NBS Applied Math. Ser., vol. 49, pp. 1–22).

Zoutendijk, G. (1970). Nonlinear programming, computational methods. In Integer and nonlinear programming (pp. 37–86).