The Scientific Journal of King Faisal University

Basic and Applied Sciences

A Three-Term Conjugate Gradient Method for Non-Convex Functions with Applications for Heat Transfer

Umar Audu Omesa[1,2], Ibrahim Mohammed Sulaiman [3], Maulana Malik [4], Basim A. Hassan [5], Waziri Muhammad Yusuf [2,6] and Mustafa Mamat [7]

[1] Department of Mathematics, Faculty of Sciences, Federal University of Agriculture, Zuru, Kebbi, Nigeria
[2] Numerical Optimization Research Group, Bayero University, Kano, Nigeria
[3] Department of Mathematics and Statistics, School of Quantitative Sciences, Universiti Utara Malaysia, Kedah, Malaysia
[4] Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia
[5] Department of Mathematics, Faculty of Sciences, University of Mosul, Mosul, Iraq
[6] Department of Mathematics, Faculty of Physical Sciences, Bayero University, Kano, Nigeria
[7] Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia

LINK	RECEIVED	ACCEPTED	PUBLISHED ONLINE	ASSIGNED TO AN ISSUE
https://doi.org/10.37575/b/sci/220053	02/05/2023	28/09/2023	28/09/2023	01/12/2023
NO. OF WORDS	**NO. OF PAGES**	**YEAR**	**VOLUME**	**ISSUE**
4632	7	2023	24	2

Abstract

The problem of unconstrained optimization (UOP) has recently gained a great deal of attention from researchers around the globe due to its numerous real-life applications. The conjugate gradient (CG) method is among the most widely used algorithms for solving UOP because of its good convergence properties and low memory requirements. This study investigates the performance of a modified CG coefficient for optimization functions, proof of sufficient descent, and global convergence of the new CG method under suitable, standard Wolfe conditions. Computational results on several benchmark problems are presented to validate the robustness and efficacy of the new algorithm. The proposed method was also applied to solve function estimations in inverse heat transfer problems. Another interesting feature possessed by the proposed modification is the ability to solve problems on a large scale and use different dimensions. Based on the theoretical and computational efficiency of the new method, we can conclude that the new coefficient can be a better alternative for solving unconstrained optimization and real-life application problems.

KEYWORDS
Computational efficiency, global convergence, inverse heat, low memory, optimization problems, theoretical efficiency

CITATION
Omesa, U.A., Sulaiman, I.M., Malik, M., Hassan, B.A., Yusuf, W.M. and Mamat, M. (2023). A three-term conjugate gradient method for non-convex functions with applications for heat transfer. *The Scientific Journal of King Faisal University: Basic and Applied Sciences*, 24(2), 12–8. DOI: 10.37575/b/sci/220053

1. Introduction

Consider the following optimization problem:

$$\min f(x), \qquad x \in R^n \tag{1.1}$$

where $f: R^n \to R$ is a smooth and convex function satisfying the condition:

$$f_i(\alpha x + \beta y) \le \alpha f_i(x) + \beta f_i(y),$$

for all $x, y \in R^n$, $\alpha, \beta \in R$, and $\alpha + \beta = 1$, $\alpha \ge 0$, $\beta \ge 0$. Convex optimization deals with minimizing specific convex functions over convex sets, while the non-convex optimization deals with problems where the objective function is non-convex. These problems are solved using local optimization methods that require an initial guess. This point significantly influences the performance of the objective value of the local solution.

One of the widely used methods for solving (1.1) is the conjugate gradient (CG) method, which computes as follows:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where x_{k+1} is the new iterate and $\alpha_k > 0$ is generated by a suitable line search process through the search direction d_k:

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \qquad d_k = -g_k \tag{1.3}$$

where β_k denotes the CG parameter that differentiates various CG formulas and $g_k = \nabla f(x)$ is the gradient of f (Yakubu *et al.*, 2020; Malik *et al.*, 2020). Generally, for $k = 0$, $d_0 = -g_0$, which represents the classical steepest descent direction. If α_k satisfies the exact line minimization condition and $f(x)$ is a strictly convex quadratic function, (1.2) and (1.3) will reduce to the linear CG method (Hager and Zhang, 2006). However, for the general nonlinear case, the parameter β_k is computed using algorithms that do not satisfy the conjugacy, such as:

$$\beta_k^{FR} = g_k^T g_k / \|g_{k-1}\|^2, \qquad \text{(Fletcher and Reeves, 1964)} \tag{1.4}$$

$$\beta_k^{PR} = (y_k^T g_{k+1}) / \|g_k\|^2, \qquad \text{(Polak and Ribiere, 1969)} \tag{1.5}$$

$$\beta_k^{DY} = g_{k+1}^T g_{k+1} / y_k^T s_k, \qquad \text{(Dai and Yuan, 2000)} \tag{1.6}$$

$$\beta_k^{DL} = g_{k+1}^T (y_k - t s_k) / y_k^T s_k, \qquad \text{(Dai and Liao, 2001)} \tag{1.7}$$

where $y_k = g_{k+1} - g_k$ and the parameter $t \ge 0$. For a detailed discussion on advances in the conjugate gradient method, refer to Hager and Zhang (2006) and (Sulaiman *et al.* (2022).

The nonlinear CG algorithm plays a significant role in solving large-scale unconstrained, differentiable functions due to its simplicity and good convergence properties (Powell, 1984; Sulaiman *et al.*, 2020; Mamat *et al.*, 2020). Numerous studies have investigated the convergence of these classical CG methods. For instance, Zoutendijk (1970) studied the convergence of the Fletcher–Reeves (FR) method under exact line searches with the following condition:

$$g_k^T d_k \le c\|g_k\|^2 \qquad c > 0. \tag{1.8}$$

By restricting β_k^{PRP} to be non-negative, Gilbert and Nocedal (1992) confirmed the global convergence of the β_k^{PRP} method under suitable conditions. For more convergence results on CG methods, refer to Wolfe (1969), Sulaiman *et al.* (2021a), Grippo and Lucidi (1997), Malik *et al.* (2021), Sulaiman *et al.* (2021b), Hestenes and Stiefel (1952), Kamfa *et al.* (2020), Deng and Wan (2015), and Awwal *et al.* (2021).

Recently, Audu *et al.* (2020) developed a robust variant of the FR method by introducing a new term to the denominator for the classical FR method as follows:

$$\beta_k^{UMW} = \frac{\|g_k\|^2}{d_{k-1}^T\left(d_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|} g_k - g_k\right)}. \tag{1.9}$$

Corresponding Author: Ibrahim Mohammed Sulaiman 00601126322320, i.mohammed.sulaiman@uum.edu.my

The authors proved the decency property and established the global convergence under suitable conditions.

One of the efficient and widely recognized variants of the CG algorithm is the three-term gradient algorithm (Liu and Du, 2019; Maulana *et al.*, 2023), which is formulated by defining a new term as the classical CG search direction (1.3). Notes on the three-term CG method were first presented by Beale (1972). The author employed the classical β_k^{HS} method defined in Hestenes and Stiefel (1952) to construct the following d_k:

$$d_k = -g_k + \beta_k d_{k-1} + \gamma_k d_t, \quad (1.10)$$

where d_t denotes the restart direction and $\gamma_k = 0$ for $k = t+1$, and

$$\gamma_k = \frac{g_k^T y_t}{d_k^T y_t}, \qquad k > t+1. \quad (1.11)$$

In practice, this method is less effective and possesses finite termination properties (Hager and Zhang, 2006; Dai and Liao, 2001). In addition, Beale's method is not guaranteed to generate a descent direction under different line search procedures (Zhang *et al.*, 2007). As a result of these shortcomings, Powell (1984) and McGuire and Wolfe (1973) improved the performance of the Beale (1972) method using an efficient restart strategy by imposing the following condition,

$$g_k^T d_k \geq \varphi \|g_k\| \|d_k\|, \quad (1.12)$$

and the Powell Beale condition,

$$\left|g_{k-1}^T g_k\right| < 0.2 \|g_k\|^2, \quad (1.13)$$

on (1.10), which guarantees the descent condition.

A recent study on the three-term CG method was presented by Zhang *et al.* (2007). The authors utilized the good convergence properties of β_k^{PRP} to define a three-term CG method termed TTPRP, with the formula expressed as follows:

$$d_k = -g_k + \beta_k d_{k-1} + \theta_{k-1} y_{k-1}, \quad (1.14)$$

where $\theta_k = \frac{-g_k^T d_{k-1}}{\|g_{k-1}\|^2}$. The authors extended the idea to β_k^{HS} (TTHS) as follows:

$$d_k = \begin{cases} -g_k, & \text{if } g_k^T y_t < \varepsilon_1 \|g_{k-1}\|^r s_{k-1}^T s_{k-1} \\ -g_k + \beta_k d_{k-1} + \theta_{k-1} y_{k-1}, & \text{otherwise} \end{cases}$$

with $\theta_k = \frac{-g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}$ and $s_{k-1} = x_k - x_{k-1}, r \geq 0, \varepsilon_1 > 0$. Zhang *et al.* (2007) showed that both methods satisfy the descent condition,

$$g_k^T d_k \leq -\|g_k\|^2,$$

regardless of the line search procedure used, and further established the global convergence of TTPRP under modified Armijo conditions and the convergence of TTHS, which was studied under the standard Wolfe condition. Motivated by the idea of Zhang *et al.* (2007) and utilizing the efficacy modified FR parameter defined in (1.9), this study develops a new class of three-term CG methods for unconstrained optimization functions.

The remainder of the paper is structured as follows. Section 2 presents the method formulation and algorithm. In section 3, we establish the convergence of the proposed method under suitable Wolfe line conditions. Experimental results on a number of benchmark functions are presented in section 4 to demonstrate the robustness and efficacy of the new algorithm. In section 5, the new algorithm is applied to solve real-life application problems. Finally, the conclusion is presented in section 6.

2. Three-Term Conjugate Gradient Algorithm

This section presents the derivation process of the new three-term CG algorithm for optimization functions.

2.1 New Three-Term CG Method:

The proposed method is derived as follows:

$$d_0 = -g_0,$$
$$d_k = -g_k + \beta_k d_{k-1} + \theta_k^{(1)} y_{k-1}, \quad (2.1)$$

and

$$d_k = -g_k + \beta_k d_{k-1} + \theta_k^{(2)} y_{k-1}, \quad (2.2)$$

where

$$\beta_k^{UMW} = \frac{\|g_k\|^2}{d_{k-1}^T\left(d_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|} g_k - g_k\right)}, \quad (2.3)$$

and

$$\theta_k^{(1)} = -\frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2} \text{ and } \theta_k^{(2)} = -\frac{g_k^T d_{k-1}}{g_k^T (y_{k-1})}. \quad (2.4)$$

From (2.1)–(2.3), it is clear that

$$d_k^T g_k = -\|g_k\|^2, \quad (2.5)$$

which implies that d_k is a descent direction. It is also clear that $\theta_k = 0$ if the exact minimization condition is applied.

Algorithm 2.1

Step 1: Initialization: $x_0 \in R^n$, $d_0 = -g_0$, set $k = 0$. If $\|g_k\| \leq \varepsilon$, then stop. Otherwise continue.

Step 2: Compute β_k by (2.3)

Step 3: Compute θ_k by (2.4) i.e., $\theta_k^{(1)}$

Step 4: Determine α_k based on the following Wolfe conditions:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k, \quad (2.6)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \quad (2.7)$$

where $0 < \delta < \sigma < 1$.

Step 5: Update x_k based on (1.2)

Step 6: Check if $\|g_k\| = 0$, terminate. Otherwise, go back to step 2 with $k = k + 1$.

The proposed TTUMW method and the classic TTPRP formula have similar structures. However, the TTUMW contains a new coefficient, β_k, and a three-term parameter, θ_k, that differentiate the new method from other methods.

It is mandatory to show that the direction d_k, defined by (2.1) and (2.2), possess the descent properties before discussing the convergence results.

Lemma 2.0

Let β_k^{UMW} be defined by algorithm 2.1 where d_k follows from (2.1). If $g_k^T d_{k-1} = 0$, then

$$\beta_k^{UMW} \leq \frac{\|g_k\|^2}{\|d_{k-1}\|^2}.$$

Proof

Begin with the following simplification and let $g_k^T d_{k-1} = 0$ (Rivaie *et al.*, 2012).

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

$$= \frac{\|g_k\|^2}{\|d_{k-1}\|^2 - w - d_{k-1}^T g_k}, \quad w = \frac{\|g_k\|}{\|g_{k-1}\|} d_{k-1}^T g_k$$
$$\leq \frac{\|g_k\|^2}{\|d_{k-1}\|^2}.$$

This completes the proof.

Lemma 2.1

Let the sequence $\{x_k\}$ follow from algorithm (2.1), d_k from (2.1), and θ_k given as (2.4). Then,

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

Proof

From (2.1), (2.3), and (2.4), it follows that:

$$g_k^T d_k = -\|g_k\|^2 + g_k^T d_{k-1}\beta_k - g_k^T \cdot \frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2} \cdot y_{k-1}$$
$$\leq -\|g_k\|^2.$$

This completes the proof.

3. Convergence Analysis

The convergence result of the new method would be studied based on the following assumption.

Assumption 3.1

(1) The function f is bounded on $\Omega = \{x \in R^n: f(x) \leq f(x_0)\}$, where Ω defines the level set.

(2) f is smooth and bounded below on R^n, and its gradient, $g(x)$, is Lipchitz continuous on the neighborhood N of Ω. That is, for some $L > 0$, it implies:

$$\forall\, x, y \in N, \ \|g(x) - g(y)\| \leq L\|x - y\|. \tag{3.1}$$

For some positive constant μ, we have the following results, which come from Assumption 3.1:

$$\|g(x)\| \leq \mu \quad \forall x \in \Omega. \tag{3.2}$$

To prove Assumption 3.1, we begin by presenting the subsequent lemma with proof stemming from Zoutendijk (1970) and Wolfe (1969).

Lemma 3.1 Consider x_1 as the initial guess that Assumption 3.1 holds true. For a CG algorithm of the form (1.2), with α_k satisfying (2.6) and (2.7) and d_k is a descent direction, then,

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{3.3}$$

Proof

From (2.6), we have

$$g_k^T y_k = g_k^T(g_{k+1} - g_k) \geq (\sigma - 1) g_k^T d_k,$$

and after (3.1), it follows that

$$(g_{k+1} - g_k)^T d_k \leq \alpha_k L \|g_k\|^2.$$

Combining the two inequalities will give

$$\alpha_k \geq \frac{\sigma - 1}{L} \cdot \frac{g_k^T d_k}{\|d_k\|^2},$$

which reduces to

$$f_k - f_{k-1} \geq c \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2}, \tag{3.4}$$

where $c = \delta(1 - \sigma)/L$. Summing up (3.4) and noting that f is bounded below implies that (3.3) holds, and thus completes, the proof.

3.1 Sufficient Descent Condition:

The following results will be used to show that TTUMW satisfies the descent properties under the Wolfe line search.

Theorem 2.7

If supposed algorithm (1.2) holds where β_k is given as (2.3) and α_k is generated by (2.6) and (2.7), then

$$g_k^T d_k \leq -(1 - \sigma)\|g_k\|^2, \qquad \forall k \geq 0. \tag{3.5}$$

Proof

The proof of this theorem is induction. Suppose $k = 0$, then $g_0^T d_0 = \|g_0\|^2$. Hence, (3.5) holds. Assume that (3.5) holds true for $k - 1$, this implies that

$$g_{k-1}^T d_{k-1} < 0, \tag{3.6}$$

which suggests that (1.13) is true.

Multiplying (1.14) by g_k^T gives:

$$g_k^T d_k = -\|g_k\|^2 + \beta_k . g_k^T d_{k-1} + \theta_k^{(1)} g_k^T g_{k-1}.$$

Substituting (2.3) and (2.4), we have

$$= -\|g_k\|^2 + \frac{\|g_k\|^2}{d_{k-1}^T(d_{k-1} - m - g_k)} . g_k^T d_{k-1}$$
$$- \frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2} . g_k^T d_{k-1},$$

where $m = \frac{\|g_k\|}{\|g_{k-1}\|} g_k$.

From Lemma 2.0, it follows that $\beta_k^{UMW} = \frac{\|g_k\|^2}{\|d_{k-1}\|^2}$ when exact line search is used. Hence,

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

which follows from (3.5), and thus, completes the proof.

3.2. Global Convergence Property:

The global convergence analysis of the proposed method will be discussed under weak Wolfe conditions.

Theorem 2.8

Suppose Assumption 3.1 holds true, then Algorithm 2.1 produces an infinite sequence $\{x_k\}$ satisfying

$$\lim_{k\to\infty} \inf \|g_k\| = 0. \tag{3.7}$$

Proof

Assume (3.7) is not true. Then, there exists a constant $\vartheta > 0$ such that

$$\|g_k\| > \vartheta, \forall\, k. \tag{3.8}$$

From (2.1) and (2.4), we have:

$$\|d_k\| = \left\| -g_k + \beta_k^{UMW} d_{k-1} - \frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2} \ y_{k-1} \right\|,$$
$$\leq \|g_k\| + \left|\beta_k^{UMW}\right| \|d_{k-1}\| + \left|\frac{g_k^T d_{k-1}}{\|d_{k-1}\|^2}\right| \|y_{k-1}\|, \tag{3.9}$$
$$\leq \|g_k\| + \frac{\|g_k\|^2}{\|d_{k-1}\|^2} \|d_{k-1}\| + \frac{\|g_k\| \, \|d_{k-1}\|}{\|d_{k-1}\|^2} \|y_{k-1}\|, \tag{3.10}$$

$$= \|g_k\| + \frac{\|g_k\|^2}{\|d_{k-1}\|} + \frac{\|g_k\|}{\|d_{k-1}\|} L\|x_k - x_{k-1}\|, \quad (3.11)$$

$$\leq \|g_k\| + \frac{\|g_k\|^2}{\|g_{k-1}\|} + \frac{\|g_k\|}{\|g_{k-1}\|} L\,(\|x_k\| + \|x_{k-1}\|), \quad (3.12)$$

$$\leq \mu + \frac{\mu^2}{\vartheta} + \frac{\mu}{\vartheta} L(2\xi) \equiv T,$$

where (3.9) comes from the Cauchy–Schwarz inequality; (3.10) from Lemma 2.0; (3.11) from part 2 of Assumption 3.1; the fifth inequality from relation (2.6) in Liu and Du (2019); and (3.12) from (3.2), (3.8), and part 1 of Assumption (3.1).

Based on the above analysis, we can state that the sequence of $\{\|d_k\|\}$ has a common upper bound, that is,

$$\|d_k\| \leq T, \forall k. \quad (3.9)$$

By using (3.8) and (3.9), we have:

$$\sum_{k=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{\vartheta^4}{T^2} = +\infty,$$

which contradicts the Zoutendijk condition given in (3.3). Hence, the conclusion is that (3.7) is true. This completes the proof.

4. Results and Discussion

This section presents the performance results of the proposed method. The computational results are compared with that of other three-term CG methods from Zhang *et al.* (2007) to demonstrate the efficiency and robustness of our method. All test functions used for the experiments (Table 1) are from Andrei (2008). For each test function, we chose four initial points under the standard Wolfe conditions. All algorithms were coded on the MATLAB 2015a version, and the termination condition was set as $\|g_k\| \leq 10^{-6}$.

Table 1: Unconstrained Optimization Functions

Functions	Dim	Initial Points
Treccani	2	(0.5,0.5), (5,5), (10,10), (15,15)
Booth	2	(2,2), (9,9), (10,10), (13,13)
Three Hump Camel	2	(2,2), (9,9), (10,10), (13,13)
Sphere	2	(5,5), (15,15), (25,25), (50,50)
Ext DENCHNB	2	(5,5), (15,15), (25,25), (50,50)
Six Hump	2	(2,2), (10,10), (15,15), (20,20)
Hager function	2,4	(2,2), (10,10), (15,15), (20,20)
Quadratic QF2	2,4,10	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Power function	2,4,10	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Gen Tridiagonal	2,4,10,100	(2,2,...,2), (6,6,...,6), (9,9,...,9), (15,15,...,15)
Quadratic QF1	2,4,10,100	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Matyas function	2,4,10,100	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Dixon and price	2,4,10,100	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Gen Tridiagonal 2	2,4,10,100	(2,2,...,2), (6,6,...,6), (9,9,...,9), (15,15,...,15)
Gen Quartic	2,4,10,100,1000	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Gen Quartic	2,4,10,100,1000	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Sum square	2,4,10,100,1000	(2,2,...,2), (10,10,...,10), (15,15,...,15), (25,25,...,25)
Treccani	2,4,10,100,1000	(2,2,...,2), (9,9,...,9), (10,10,...,10), (13,13,...,13)
Ext freudestain andRoth	2,4,10,100,1000,10000	(5,5,...,5), (7,7,...,7), (9,9,...,9), (13,13,...,13)
Ext Beale	2,4,10,100,1000,10000	(2,2,...,2), (5,5,...,5), (15,15,...,15), (25,25,...,25)
Fletcher	2,4,10,100,1000,10000	(2,2,...,2), (3,3,...,3), (9,9,...,9), (15,15,...,15)

The performance results presented in Figures 1 and 2 were plotted using the performance profile tool introduced by Dolan and Moré (2002). This was achieved by recording the number of iterations (NOI) and CPU time for all solvers (*S*) on a set of problems (*P*). Suppose the set of solvers S consist of n_p problems and n_s solvers. Then, for every solver $s \in S$ and problem $p \in P$, we can define t_{ps} as the number of iterations or CPU time needed by solver $s \in S$ to solve problem $p \in P$.

For every algorithm, the Dolan and Moré tool plots a fraction *(P)* of the benchmark function to obtain a profile curve, as shown in Figures 1 and 2. The method with the curve lying the highest is regarded as the top performer.

Figure 1: Performance profile based on number of iterations

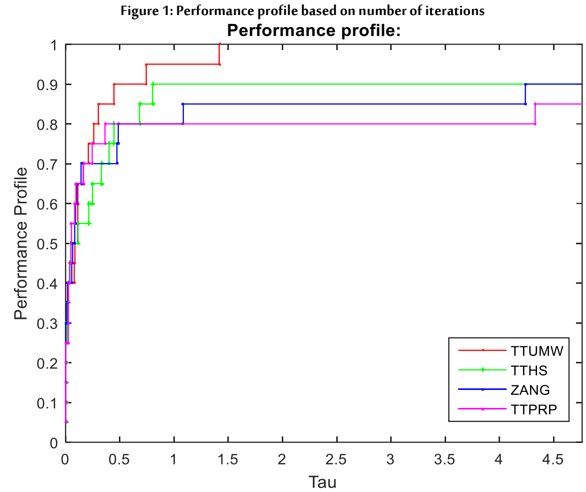
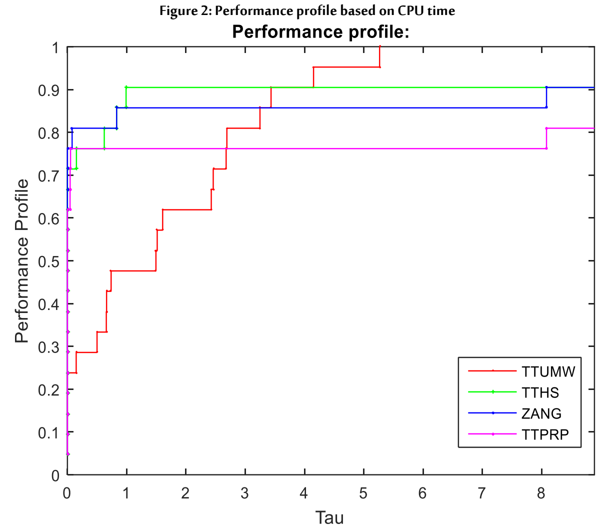


Figure 2: Performance profile based on CPU time



From Figure 1, it is clear that the convergence of all the methods follow a similar pattern. This can be attributed to the structure of the algorithms. Despite the close relationship among these algorithms, the proposed method demonstrated a better numerical performance because it was able to solve the majority of the test problems. In addition, based on Figure 2, it is clear that the proposed TTUMW method outperformed the three other methods considered for comparison. Based on these results, we can conclude that the proposed TTUMW method is both superior and promising.

5. Application of the TTUMW Method for Function Estimation in Inverse Heat Transfer Problems

The family of CG methods are known for their low memory requirement and global convergence properties when solving optimization functions. Most of these problems are traced to specific areas of engineering, namely, the sciences, economics, and social sciences. Recently, several studies have investigated the performance of CG methods on different application problems, such as robotic motion control, compressive sensing, fuzzy nonlinear problems, and inverse heat transformation problems (see Umar et al., 2020; Sulaiman et al., 2022b, 2022c; Razzaq *et al.*, 2020; Omesa *et al.*, 2023). This section examines the inverse heat transformation problem. The problem is transformed into a minimization problem and solved using certain stabilization techniques (Colaco and Orlande, 1999).

Consider the following inverse heat transformation problem:

$$\min Bi(X, Y, \tau) \quad (4.0)$$

where Bi is the dimensional heat transfer coefficient, X and Y are the dimensional coordinates, and τ denotes dimensional time. The method for solving (4.0) involves the iterative procedure of the CG method. Recent studies by Jarny *et al.* (1991) and Orlande *et al.* (1997) estimated the unknown heat transfer coefficient as follows:

$$Bi_{k+1}(X, Y, \tau) = Bi_k(X, Y, \tau) - b_k d_k(X, Y, \tau) \quad (4.1)$$

where k denotes the iteration number. The descent direction is a conjugation of the previous direction and gradient computed as:

$$d_k(X, Y, \tau) = j'_k(X, Y, t) + g_k d_{k-1}(X, Y, \tau). \quad (4.2)$$

The conjugation coefficient, β_k, utilized in this study is defined by equation (2.1), and the step size α_k is computed based on (2.6) and (2.7).

To generate the simulated measurement, we needed to rewrite $Bi(X, Y, \tau)$ as follows:

$$Bi = 6 f_t(t) f_X(X) f_Y(Y). \quad (4.3)$$

Different functional forms were tested for $f_t(t), f_X(X),$ and $f_Y(Y)$ (Lally *et al.*, 1990).

By generating the following simulated measurements with the functional form, this study demonstrates how to estimate the spatial changes and temperature shifts of the heat transfer coefficient as follows:

$$Bi(X, Y, \theta) = 6e^{-(0.5X)^2} f_\theta(\theta) \quad (4.4)$$

where

$$f_\theta(\theta) = 51{,}338 - 893.278\theta + 8566.53\theta^2 - 46{,}264.9\theta^3 + 154{,}280\theta^4 - 333{,}217\theta^5 + 47{,}539\theta^6 - 445{,}692\theta^7 + 264{,}504\theta^8 - 90{,}201.7\theta^9 + 13{,}476.3\theta^{10} \quad 0.103 < \theta < 1. \quad (4.5)$$

The function given by equation (4.5) comes from the data of Stewart *et al.* (1996), with the initial guess set as $Bi^0(X, Y, \theta) = 1$. By considering the final time of 0.125 and applying algorithm 2.1, the solution (6.9161, 0.0168) was arrived at after one iteration with a CPU time of 0.7155.

6. Conclusion

This study investigated the performance of new three-term CG algorithms for non-convex optimization functions. The new algorithm was extended to solve the problem of heat transfer. An interesting feature of the proposed algorithm is that it possesses the descent property $g_k^T d_k \leq -\|g_k\|^2$, irrespective of the line search condition used. The global convergence analysis of the method was discussed under suitable conditions. Numerical computation on a set of benchmark problems was presented to determine the performance of the method. The experimental results showed that the new algorithm outperformed the classical three-term CG in terms of number of iterations and CPU time.

Acknowledgment

This research was supported by Ministry of Higher Education (MoHE) of Malaysia through Fundamental Research Grant Scheme (FRGS/1/2021/STG06/UUM/02/4) with S/O code 20123.

Biographies

Umar Audu Omesa

Department of Mathematics, Faculty of Sciences, Federal University of Agriculture, Zuru, Kebbi, Nigeria, 002349075784600, umarabdul64@gmail.com

Dr. Umar Audu Omesa is from Kebbi state, Nigerian and currently serve as a lecturer in Kebbi, Nigeria. He holds a PhD in Applied Mathematics from Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia in 2019. He is an author of more than 10 Scopus scholarly research papers and has presented his research work in international conferences. His work has been cited by 19 counties (Including Brazil, Russia, Taiwan, Thailand, and United states). His research interest include: Fuzzy nonlinear system, Optimization methods, Nonlinear analysis. Scopus ID: 57216225815.

Ibrahim Mohammed Sulaiman

Department of Mathematics and Statistics, School of Quantitative Sciences, Universiti Utara Malaysia, Kedah, Malaysia, 00601126322320, i.mohammed.sulaiman@uum.edu.my

Dr. Sulaiman is from Kano State, Nigeria, but currently serve as an International Senior Lecturer in Malaysia. He holds a PhD degree in Applied Mathematics from UniSZA, Malaysia in 2018. He is an author of more than 80 ISI/Scopus scholarly research papers published by largest global publishers (Elsevier, Springer, Sage, PLOS One, Willey, & Nature). His work has been cited by 49 counties (Including United states, Jordan, Canada, Australia, Brazil, China, Thailand, and United Kingdom) His main research interest includes; optimization methods, fuzzy nonlinear problems, and Fluid. https://orcid.org/0000-0001-5246-6636

Maulana Malik

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia, 009628988072292, m.malik@sci.ui.ac.id

Mr. Malik is an Assistant Professor from Indonesia. He is currently pursuing a PhD degree in Applied Mathematics at Universiti Sultan Zainal Abidin (UniSZA), Terengganu, Malaysia. He is an author of 35 ISI/Scopus scholarly research papers published by largest global publishers (Elsevier, Springer, Willey, & Frontiers). His work has been cited by 119 documents from 33 counties (Including Russia, Italy, Morocco, United states, Jordan, Australia, China, Nigeria, Thailand). His research interest includes: Portfolio selection, Unconstrained optimization and application, Numerical methods, Nonlinear problems. ORCID: 0000-0003-3060-0624

Basim A. Hassan

Department of Mathematics, Faculty of Sciences, University of Mosul, Mosul, Iraq, 009647518095345, basimah@uomosul.edu.iq

Dr. Basim A. Hassan from Iraq and currently a professor at department of Mathematics, Faculty of Sciences, University of Mosul in Iraq. He is an author of more than 52 ISI/Scopus scholarly research papers published by largest global publishers (Elsevier, Springer, Sage, PLOS One, Willey, AIMS) and presented some of his research in international conferences. His work has been cited by 22 counties (Including Austria, Russia, Kuwait, China, Indonesia, Thailand). His research interest includes: Unconstrained optimization and Applications, Numerical methods, Nonlinear problems.

Waziri Muhammad Yusuf

Department of Mathematics, Faculty of Physical Sciences, Bayero University, Kano, Nigeria, 02348036364455, mywaziri.mth@buk.edu.ng

Prof. Waziri is a UPM graduate, Nigeria Professor from Yobe state of Northern Nigeria. He obtained is PhD from department of Mathematics, faculty of science Universiti Putra Malaysia, Serdang, Malaysia. He is an author of more than 122 ISI/Scopus scholarly research papers published by largest global publishers (Elsevier, Springer, Sage, PLOS One, Willey, AIMS). His work has been cited by

351 documents from counties (Including Brazil, China, Japan, Indonesia, Taiwan, and Thailand). His research interest includes: Jacobian matrices, Newton method, approximation theory, large-scale systems, nonlinear equations. ORCID: 0000-0001-7112-659X

Mustafa Mamat

Faculty of Informatics and Computing, University Sultan Zainal Abidin, Terengganu, Malaysia, 0060198955027, must@unisza.edu.my

Prof. Mamat is a graduate of UPM, Malaysian Professor. He is an author of more than 475 ISI/Scopus scholarly research papers published by largest global publishers (Elsevier, Springer, Sage, PLOS One, Willey, & Nature). His work has been cited 3,368 times from countries (including United states, South Africa, Jordan, Australia, Sudan, China, Thailand, and United Kingdom). To date, he has successfully supervised more than 60 postgraduate students. His research interests include conjugate gradient methods, steepest descent methods, Broydens family, and quasi-Newton methods.

References

Andrei, N. (2008). An unconstrained optimization test functions collection. *Advanced Modeling and Optimization*, **10**(1), 147–61.

Audu, O.U., Mamat, M., Sulaiman, I.M., Waziri, M.Y. and Abba, V.M. (2020). A new modification of conjugate gradient parameter with efficient line search for nonconvex function. *International Journal of scientific and Technology Research*, **9**(3), 1–4.

Awwal, A.M., Sulaiman, I.M., Malik, M., Mamat, M., Kumam, P. and Sitthithakerngkiet, K. (2021). A spectral RMIL+ conjugate gradient method for unconstrained optimization with applications in portfolio selection and motion control. *IEEE Access*, **9**(n/a) 75398–414.

Beale, E.M.L. (1972). A derivation of conjugate gradients. In: F.A. Lootsma (Ed.), *Numerical Methods for Nonlinear Optimization*. London: Academic Press.

Colaco, M.J. and Orlande, H.R. (1999). Comparison of different versions of the conjugate gradient method of function estimation. *Numerical Heat Transfer: Part A: Applications*, **36**(2), 229–49.

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

Dai, Y.H. and Yuan, Y. (2000). *Nonlinear Conjugate Gradient Methods*. Shanghai: Shanghai Science and Technology Publisher.

Deng, S. and Wan, Z. (2015). A three-term conjugate gradient algorithm for large-scale unconstrained optimization problems. *Applied Numerical Mathematics*, **92**(n/a), 70–81.

Dolan, E.D. and Moré, J.J. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming*, **91**(2), 201–13.

Fletcher, R. and Reeves, C.M. (1964). Function minimization by conjugate gradients. *The Computer Journal*, **7**(2), 149–54.

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

Grippo, L. and Lucidi, S. (1997). A globally convergent version of the Polak-Ribiere conjugate gradient method. *Mathematical Programming*, **78**(3), 375–91.

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

Hestenes, M.R. and Stiefel, E. (1952). Methods of conjugate gradients for solving. *Journal of Research of the National Bureau of Standards*, **49**(6), 409.

Jarny, Y., Ozisik, M.N. and Bardon, J.P. (1991). A general optimization method using adjoint equation for solving multidimensional inverse heat conduction. *International journal of heat and mass transfer*, **34**(11), 2911–9.

Kamfa, K., Waziri, M.Y., Sulaiman, I.M., Ibrahim, M.A.H., Mamat, M. and Abas, S.S. (2020). An efficient hybrid bfgs-cg search direction for solving unconstrained optimization problems. *Journal of Advanced Research in Dynamical and Control Systems,* **12**(2), 1035–41.

Lally, B., Biegler, L. and Henein, H. (1990). Finite difference heat-transfer modeling for continuous casting. *Metallurgical Transactions B*, **21**(4), 761–70.

Liu, J. and Du, S. (2019). *Modified Three-Term Conjugate Gradient Method and Its Applications*. Available at: https://www.hindawi.com/journals/mpe/2019/5976595/ (accessed on 10/09/2022).

Malik, M., Abubakar, A.B., Ibrahim, S.M., Mamat, M., Abas, S.S. and Firman, S. (2021). A new three-term conjugate gradient method for unconstrained optimization with applications in portfolio selection and robotic motion control. *IAENG International Journal of Applied Mathematics*, **51**(2021), 471–86.

Malik, M., Mamat, M., Abas, S.S., Ibrahim, S.M. and Sukono, F. (2020). A new spectral conjugate gradient method with descent condition and global convergence property for unconstrained optimization. *Journal of Mathematical and Computational Science*, **10**(5), 2053–69.

Mamat, M., Sulaiman, I.M., Malik, M. and Zakaria, Z.A. (2020). An efficient spectral conjugate gradient parameter with descent condition for un-constrained optimization. *Journal of Advanced Research in Dynamical and Control Systems*, **12**(2), 2487–93.

Maulana, M., Sulaiman, I.M., Auwal, B.A., Gianinna, A. and Sukono. (2023). A new family of hybrid three-term conjugate gradient method for unconstrained optimization with application to image restoration and portfolio selection. *AIMS Mathematics*, **8**(1), 1–28.

McGuire, M.F. and Wolfe, P. (1973). *Evaluating a Restart Procedure for Conjugate Gradients*. Available at: https://dominoweb.draco.res.ibm.com/reports/rc4382.pdf (accessed on 10/09/2022).

Omesa, U.A., Sulaiman, I.M., Yusuf, W.M., Hassan, B.A., Moyi, A.U., Abdul-Rahman, A. and Mamat, M. (2023). An efficient hybrid Conjugate gradient algorithm for solving intuitionistic fuzzy nonlinear equations. *The Scientific Journal of King Faisal University: Basic and Applied Sciences*, **24**(1), 8–13. DOI: 10.37575/b/sci/220041

Orlande, H.R., Colaço, M.J. and Malta, A.A. (1997). Estimation of the heat transfer coefficient in the spray cooling of continuously cast slabs. *Asme-Publications-Htd, United States*, **340**(n/a), 109–16.

Polak, E. and Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, **3**(1), 35–43.

Powell, M.J. (1984). Nonconvex minimization calculations and the conjugate gradient method. In: D.F. Griffiths (eds.), *Numerical Analysis*. Berlin: Springer.

Razzaq, A., Seadawy, R. and Raza, N. (2020). Heat transfer analysis of viscoelastic fluid flow with fractional Maxwell model in the cylindrical geometry. *Physica Scripta*, **95**(11), 115220.

Stewart, I., Massingham, J.D. and Hagers, J.J. (1996). Heat transfer coefficient effects on spray cooling. *Iron and Steel Engineer*, **73**(7), n/a.

Sulaiman, I.M. and Mamat, M. (2020). A new conjugate gradient method with descent properties and its application to regression analysis. *Journal of Numerical Analysis, Industrial and Applied Mathematics*, **12**(1-2), 25–39.

Sulaiman, I.M., Awwal, A.M., Malik, M., Pakkaranang, N. and Panyanak, B. (2022c). A derivative-free MZPRP projection method for convex constrained nonlinear equations and its application in compressive sensing. *Mathematics*, **10**(16), 2884.

Sulaiman, I.M., Malik, M., Awwal, A.M., Kumam, P., Mamat, M. and Al-Ahmad, S. (2022a). On three-term conjugate gradient method for optimization problems with applications on COVID-19 model and robotic motion control. *Advances in Continuous and Discrete Models*, **2022**(1), 1. DOI:10.1186/s13662-021-03638-9

Sulaiman, I.M., Malik, M., Giyarti, W., Mamat, M., Ibrahim M.A.H. and Ahmad, M.Z. (2022b). The application of conjugate gradient method to motion control of robotic manipulators. In: I.M. Khairuddin, M.A. Abdullah, A.F. Ab-Nasir, J.A.M. Jizat, M.A.M. Razman, A.S. Abdul Ghani, M.A. Zakaria, W.H.M. Isa and A.P.P. Abdul Majeed (eds.) *Enabling Industry 4.0 Through Advances in Mechatronics*. Singapore: Springer.

Sulaiman, I.M., Mamat, M., Owoyemi, A.E., Ghazali, P.L., Rivaie, M. and Malik, M. (2021b). The convergence properties of some descent conjugate gradient algorithms for optimization models. *Journal of Mathematics and Computer Science*, **22**(3), 204–15.

Sulaiman, I.M., Mamat, M., Waziri, M.Y., Yakubu, U.A. and Malik, M. (2021a). The performance analysis of a new modification of conjugate gradient parameter for unconstrained optimization models. *Mathematics and Statistics*, **9**(1), 16–23. DOI:10.13189/ms.2021.090103

Umar, A.O., Sulaiman, I.M., Mamat, M., Waziri, M.Y., Foziah, H.M., Altien, J.R. and Deiby, T.S. (2020). New Hybrid conjugate gradient method for solving fuzzy nonlinear equations. *Journal of Advanced Research* *in Dynamical and Control Systems*, **12**(2), 585–90.

Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM review*, **11**(2), 226–35.

Yakubu, U.A., Sulaiman, I.M., Mamat, M., Ghazali, P. and Khalid, K. (2020). The global convergence properties of a descent conjugate gradient method. *Journal of Advanced Research in Dynamical and Control Systems*, **12**(2), 1011–6.

Zhang, L., Zhou, W. and Li, D. (2007). Some descent three-term conjugate gradient methods and their global convergence. *Optimization Methods and Software*, **22**(4), 697–711.

Zoutendijk, G. (1970). Nonlinear Programming, Computational methods. In: J. Abadie (Ed.) *Integer and Nonlinear Programming*. Amsterdam: North-Holland.