Journal of Multidisciplinary Modeling and Optimization 2 (1) (2019), 16-26.

Three-Term Conjugate Gradient (TTCG) methods

Marwan S. JAMEEL[1] and Zeyad M. ABDULLAH[2]

[1]College of Environmental Science,
Mosul University, Mosul, Iraq
maroan88@ymail.com

[2]Department of Mathematics,
Tikrit University, Tikrit, Iraq
zeyaemoh1978@tu.edu.iq

Received: 02.05.2019, Accepted: 25.07.2019, Published: 02.08.2019

Abstract — *In this study, a comprehensive hybrid formula was developed for some known algorithms of "Three-Term Conjugate Gradient (TTCG) methods" for solving problems of unconstrained optimization by combining the three most important vectors (y_k, d_k, g_{k+1}) in an exceedingly new vector denoted by z_k defined in section two. The proposed vector z_k can also be considered as a special case or modified variant of the vector p_k within the general versions of Yasushi, Yabe, and Ford. As a theoretical aspect, global convergence, sufficient descend and conjugacy were studied in the presence of strong Wolfe condition. On the practical side, the proposed formula was compared with its counterpart to the researchers Yasushi, Yabe, and Ford. Where the results were encouraging and proved the efficiency of proposed algorithms than comparative algorithms using 35 nonlinear functions.*

Keywords Unconstrained optimization, Conjugate gradient, Global convergence.
Mathematics Subject Classification: 80C50, 30A40, 90C26.

1 Introduction

The discus problem is "unconstrained optimization":

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

(1)

where $f : R^n \to R$ is continuously differentiable and f gradient at x, which is represented by $g(x) = \nabla f(x)$ is existing. There exist many types of numerical methods to solve equation (1) including Steepest Descent (SD), Newton, CG and Quasi-Newton (QN) methods. Because of being simple and having requirement of very low memory, CG method plays a significant role, particularly when there is a large scale, the method of CG is very effec-

Cite as: *M. S. Jameel and Z. M. Abdullah, Three-Term Conjugate Gradient (TTCG) methods, Journal of Multidisciplinary Modeling and Optimization 2 (1) (2019), 16-26.*

tive. Let $x_0 \in R^n$ be a primary solution for problem (1). The nonlinear method of CG is generally planned by this iterative formula: [12]

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

where x_k the current is iterate point, $\alpha_k > 0$ represents a step length that is determined by a line search, and d_k refers to the search direction, defined by:

$$d_k = \begin{cases} -g_k & \text{if} \quad k = 0, \\ -g_k + \beta_k d_{k-1}, & \text{if} \quad k > 0, \end{cases} \tag{3}$$

where $\beta_k \in R$ is a parameter ($0 < \beta_k < 1$) and g_{k+1} denotes $g(x_{k+1})$. Some familiar formulas for β_k exist. They are shown below: [5]

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}, \qquad \text{(Fletcher-Reeves (FR), 1964)}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \qquad \text{(Hestenes -Stiefel (HS), 1952)}$$

$$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}, \qquad \text{(Polak- Ribiere (PR), 1969)}$$

where $\| \, . \, \|$ stands for the Euclidean norm of vectors and $y_k = g_{k+1} - g_k$. In this paper, three term CG methods were proposed, which are based on Yasushi, Yabe and Ford (2009) [10]. Generally, in the convergence analysis of CG methods, one hopes the ILS, such as the Strong Wolfe Conditions (SWC), which is shown as follows [6]:

Definition 1. Strong Wolfe Conditions (SWC) aims at finding line search α_k where:

$$\begin{aligned} & f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \ , \quad 0 \le \delta \le \frac{1}{2} \\ & \left| d_k^T g(x_k + \alpha_k d_k) \right| \le -\sigma d_k^T g_k \qquad , \quad \delta \le \sigma \le 1. \end{aligned} \tag{4}$$

Definition 2. Convex combination gives a finite number of points (which can be vectors, scalars) where all coefficients are non-negative and sum to 1 such that [13]

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots\dots\dots\dots + \alpha_n x_n \ \text{and} \ \alpha_1 + \alpha_2 + \dots\dots\dots\dots + \alpha_n = 1.$$

This article is organized as follows: in the second section, a general review is presented on three term CG algorithms. In section 3, new hybrid TTCG techniques are presented. In section 4, the properties of global convergence for the proposed new methods of CG are analysed. In section 5, some numerical comparisons were reported against general formula of Yasushi, Yabe and Ford 3TCG by substituting (y_k, d_k, g_k) instead of the vector

p_k in each case by using 35-test problems in the CUTE [7]; in addition, general conclusions are given in section 5.

2 Three-term CG methods

Recently, researchers widely examined methods of three-term conjugate gradient for improving the classical conjugate gradient method efficiency. Beale presented the initial three-term nonlinear method of CG in [8], in which the following formula determines search direction:

$$d_{k+1} = -g_{k+1} + \beta_k + \gamma_k d_t.$$

In Beale's algorithm [8], the parameter $\beta_k = \beta_k^{FR}$ or $\{ \beta_k^{HS}, \beta_k^{PR}$ …, etc.$\}$. In [9], another method of "three-term conjugate gradient was proposed by Nazareth", in which the computation of search direction is done using this formula:

$$d_{k+1} = -y_k + \frac{y_k^T y_k}{y_k^T d_k} d_k + \frac{y_{k-1}^T y_k}{y_{k-1}^T d_{k-1}} d_{k-1}.$$

In [4], a descent modified algorithm of PRP conjugate gradient was developed, in which the following formula of three-term is used to obtain the search direction:

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

In [3], the modification of method of HS conjugate gradient was done by employing a method of descent three-term conjugate gradient, which is read

$$d_{k+1} = -g_{k+1} + \frac{g_{k+1}^T y_k}{s_k^T g_k} s_k - \frac{g_{k+1}^T s_k}{s_k^T g_k} y_k.$$

More recently in [10], a general form of three-term conjugate gradient methods, which always generate a sufficient descent direction by formula:

$$d_k = -g_k + \beta_k d_k - \beta_k \frac{g_k^T d_{k-1}}{g_k^T p_k} p_k.$$

The parameter β_k like Beale's form.

3 New Direction for TTCG

`The general idea of the hybrid was built by combining theoretical and reasonable advantages within completely different methods. As shown below, the motivation behind this hybrid is to select a set of good qualities for some known three-term conjugate gradient methods and generate new generic formulas for TTCG methods. Installing a method with high specifications is achieved by picking two parameters ϕ_1 *and* ϕ_2 with the created convex combination between typically used vectors in optimization (y_k, g_k, s_k) and place it rather than of P_k in Yasushi, Yabe and Ford and attached to the form of the most important features as following:

$$d_k = \begin{cases} -g_k, & \text{if } k=0 \text{ or } g_k^T z = 0, \\ -g_k + \beta_k d_{k-1} - \beta_k \dfrac{g_k^T d_{k-1}}{g_k^T z} z, & \text{otherwise}, \end{cases} \quad (5)$$

$$z = \phi_1 y_{k-1} + \phi_2 g_k + (1-\phi_1-\phi_2) d_{k-1}, \quad (6)$$

where ϕ_1 and ϕ_2 are scalars that take values in the interval [0, 1] and $\phi_2 < \phi_1$. The vector $z \in R^n$ as defined above is convex combination from type three which join the vectors y_{k-1} , g_k *and* d_{k-1}.This proposed method is reduced to the standard HS or PRP method in the case of exact line search since $g_k^T d_{k-1} = 0$. In this case, also it should be noticed that the proposed method includes "the three-term conjugate gradient methods proposed by Zhang et al." [1-3]. The methods (5),(6) with $\beta_k = \beta_k^{FR}, \phi_1 = \phi_2 = 0$ and $z = g_k$ becomes the method by [2] and if $g_k^T y_k \neq o$, the methods (5),(6) with $\beta_k = \beta_k^{PR}$, $\phi_1 = 1,\ \phi_2 = 0$ and $z = y_k$ becomes the method by [1]. If $g_k^T y_k \neq o$, the method (5)-(6) with $\beta_k = \beta_k^{HS}$ $\phi_1 = 1,\ \phi_2 = 0$ and $z = y_k$ becomes the method by [3]. In addition, the method (5)-(6) with $\beta_k = \beta_k^{PR}$ and $z = g_k$ becomes the method by [4]. More important, it can be considered as a key when selecting vectors y_{k-1} , g_k *and* d_{k-1} as switch from formula to formula and considered a special instance of the formula three-term conjugate gradient algorithm given by Narushima et al. [10], when putting $z = y_k$, it means $\phi_1 = 1,\ \phi_2 = 0$, as mentioned above.

2.1 The New Three-Term CG-Algorithms:

Step1 (Initializing): Given an initial point $x_0 \in R^n$ and positive parameters, $0 < \phi_2 < \phi_1 \leq 1$, $\psi = 0.2$, $0 \leq \delta \leq 0.5$ and $\delta \leq \sigma \leq 1$. Set the initial search direction $d_0 = -g_0$ and let k = 0.

Step2 (Criterion of Termination): When $\|g_k\| \leq \varepsilon$, after that stop.

Step 3 (Line search): Determine step length $\alpha_k > 0$ satisfies "the Strong Wolfe condition" (4) with Acceleration scheme [5]: "compute $z = x_k + \alpha_k d_k$, $y_k = g_k - g_z$, $g_z = \nabla f(z)$, and Compute $a_k = \alpha_k g_k^T d_k$, $b_k = -\alpha_k y_k^T d_k$, if $b_k \neq 0$, then compute $\varphi_k = -\dfrac{a_k}{b_k}$ and update the variables as $x_{k+1} = x_k + \varphi_k \alpha_k d_k$; otherwise update the variables as $x_{k+1} = x_k + \alpha_k d_k$.

Step4 (Finding the direction): Compute the new search direction (5),(6), where the scalar parameter β_k is indecently chosen (in practice, β_k^{FR} is substituted).

Step5 (Restart procedure): If $\left|g_{k+1}^T g_k\right| \geq \psi \|g_{k+1}\|^2$, then go to Step (1) else continue (this is Powell restart).

Step6 (Loop): Let k = k +1 and go to Step (2).

4 Convergence Analysis

Now, the basic global convergence property of the new three-term CG-Algorithms must be proved under the condition that the following assumption is held.

Assumption (A):

(i) The level set $S = \{x : x \in R^n, f(x) \le f(x_0)\}$ is bounded, where x_0 is the starting point, and there exists a positive constant such that, for all: $B > 0$ and defined below.

(ii) In a neighbourhood Ω of S, f is differentiable continuously and its gradient g is continuously Lipchitz, namely, a constant $L \ge 0$ exists, where

$$\| g(x) - g(x_k) \| \le L \| x - x_k \|, \forall x, x_k \in \Omega \tag{7}$$

Obviously, Assumption (A, i) results in “a positive constant D, where:

$$B = \max\{\|x - x_k\|, \forall x, x_k \in S\}. \tag{8}$$

Here B refers to Ω diameter. From Assumption (A, ii), it is also known that a constant $\gamma \ge 0$ exists, where:

$$\|g(x)\| \le \gamma, \forall x \in S \tag{9}$$

In a number of studies on methods of CG, the descent condition or sufficient descent has a significant role; however, this condition is sometimes difficult to be achieved [1]

Theorem 4.1. (Descent property)[1: Suppose that the assumption (A) is held, independently of choice the parameter β_k and line search, consider the search directions d_k generated from (5-6), it is proved that the search direction easily satisfies the sufficient method with $c = 1$,

$$d_k^T g_k \le -c\|g_k\|^2.$$

Proof. Start with multiplying the direction d_k in (5-6) by the gradient g_k

$$d_k^{\,T} g_k = -\|g_k\|^2 + \beta_k\, d_{k-1}^T g_k - \beta_k \frac{g_k^{\,T} d_{k-1}}{g_k^{\,T} z_k} z_k^T g_k, \tag{10}$$

$$d_k^{\,T} g_k = -\|g_k\|^2 + \beta_k\, d_{k-1}^T g_k - \beta_k \frac{g_k^{\,T} d_{k-1}}{g_k^{\,T} z_k} g_k^T z_k, \tag{11}$$

$$d_k^{\,T} g_k = -\|g_k\|^2.$$

Hence, by comparing the result with standard sufficiently descent condition, the proposed direction held this condition by the value of $c = 1$.

Theorem 4.2. (Conjugacy Property): Suppose that the step-size α_k satisfies the standard Wolfe conditions, consider the search directions d_k generated from (5-6), then the search directions d_{k+1} are conjugate for all k that is

$$d_{k+1}^T y_k = -c_0\, g_{k+1}^T s_k,$$

where c_0 positive constant.

Proof. Begin by multiplying the proposed direction by the vector y_k

$$y_k^T d_{k+1} = -y_k^T g_{k+1} + \beta_k y_k^T d_k - \beta_k \frac{g_{k+1}^T d_k}{g_{k+1}^T z} y_k^T z \tag{12}$$

By using the following reality to get

$y_k = g_{k+1} - g_k \;\Rightarrow\; y_k^T z = g_{k+1}^T z - g_k^T z \;\Rightarrow\; y_k^T z \le g_{k+1}^T z$

Taking the last one and put it in (12)

$$\begin{aligned}
&\le -y_k^T g_{k+1} + \beta_k y_k^T d_k - \beta_k \frac{g_{k+1}^T d_k}{g_{k+1}^T z} g_{k+1}^T z \\
&\le -y_k^T g_{k+1} + \beta_k y_k^T d_k - \beta_k g_{k+1}^T d_k \\
&= -y_k^T g_{k+1} + \beta_k (g_{k+1}^T d_k - g_k^T d_k - g_{k+1}^T d_k) \\
&\le -y_k^T g_{k+1} - \beta_k g_k^T d_k \\
&= -(\|g_{k+1}\|^2 - g_{k+1}^T g_k) - \beta_k g_k^T d_k \\
&= -\|g_{k+1}\|^2 + g_{k+1}^T g_k - \beta_k g_k^T d_k \\
&\le -\|g_{k+1}\|^2 + \|g_{k+1}\|^2 - \beta_k g_k^T d_k
\end{aligned}$$

using curvature inequality in (4)

$$\le -\beta_k g_k^T d_k \quad \le -\frac{\beta_k}{\sigma\alpha} g_{k+1}^T s_k.$$

Hence the conjugacy condition $d_{k+1}^T y_k = -c_0\, g_{k+1}^T s_k$ is done with $c_0 = \beta / \sigma\alpha$.

Property 4.1. Consider a general CG method and suppose that [11]

$$0 < \varsigma \le \|g_k\| \le \gamma, \quad \forall k \ge 0$$

(13)

It can be said that a CG method has the property (4.1) if there exists two constants b>1 and $\lambda > 0$ such that for all k,

$$|\beta_k| \le b \tag{14}$$

$$\text{If } \|s_k\| \le \lambda \text{ then } |\beta_k| \le \frac{1}{2b} \text{ for all } \lambda > 0. \tag{15}$$

Lemma 4.2. Assume that d_{k+1} is a descent direction and g_k satisfies the Lipchitz condition $\|g(x) - g(x_k)\| \le L\|x - x_k\|$ for all x on the line segment connecting x and x_k, where L is constant if the line search direction satisfies Strong Wolfe condition, then[6]:

$$\alpha_k \geq \frac{(1-\sigma)\left|d_k^{\ T} g_k\right|}{L\|d_k\|^2}. \tag{16}$$

Proof. Using curvature inequality in (4)

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

$$\Rightarrow \sigma d_k^{\ T} g_k \leq d_k^{\ T} g_{k+1}. \tag{17}$$

Subtracting $d_k^{\ T} g_k$ from both sides of (35) and using Lipchitz condition yields:

$$(1-\sigma) d_k^{\ T} g_k \leq d_k^{\ T}(g_{k+1} - g_k) \leq L\alpha_k \|d_k\|^2 \tag{18}$$

As d_k is a descent direction” and $\sigma \leq 1$, so (16) holds:

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

The conclusion of the following Lemma, often called the Zoutendijk condition, is used to prove the global convergence of any nonlinear CG method. Zoutendijk [18] originally gave it under the Strong Wolfe line search (4). In the following Lemma, this condition will be proved.

Lemma 4.3. Suppose Assumption (A) holds. Consider the iteration process of the form (5),(6), where d_{k+1} satisfies the descent condition ($d_k^{\ T} g_k \leq 0$) for all $k \geq 1$ and α_k satisfies (4). Then

$$\sum_{k\geq 1} \frac{(g_k^{\ T} d_k)^2}{\|d_k\|^2} < +\infty. \tag{19}$$

Proof: From the first inequality in (4), the following equation can be obtained:

$$f_{k+1} - f_k \leq \delta \alpha_k g_k^{\ T} d_k.$$

Combining this with the results in Lemma (4.2), yields

$$f_{k+1} - f_k \leq \frac{\delta(1-\sigma)}{L} \frac{(g_k^{\ T} d_k)^2}{\|d_k\|^2}. \tag{20}$$

Using the bound-ness of function f in Assumption (A), hence

$$\sum_{k\geq 1} \frac{(g_k^{\ T} d_k)^2}{\|d_k\|^2} < +\infty. \tag{21}$$

Theorem 4.3. Suppose that assumption A holds and consider the new algorithm obtained by (3-1,3-2) where α_k is computed by Wolf Line Search , then

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

Proof. The proof is well done by contradiction, so it is supposed that the conclusion is not true, then $\|g_k\| \neq 0$, as mentioned above, there exists a constants $\varsigma, \gamma > 0$ such that

$$0 < \varsigma \leq \|g_k\| \leq \gamma, \quad \forall k \geq 0$$

Now by taking the square norm of both sides of the proposed new direction

$$\|d_{k+1}\| = \left\| -g_{k+1} + \beta_{k+1} d_k - \beta_{k+1} \frac{g_{k+1}^T d_k}{g_{k+1}^T z} z \right\|$$
$$\leq \|g_{k+1}\| + \beta_{k+1}\|d_k\| + \beta_{k+1} \frac{|(g_{k+1}^T d_k)|}{|(g_{k+1}^T z)|}\|z\|$$
$$\leq \|g_{k+1}\| + \beta_{k+1}\|d_k\| + \beta_{k+1} \frac{\|g_{k+1}\|\|d_k\|}{\|g_{k+1}\|\|z\|}\|z\| \qquad \text{(By Cauchy Schwarz)}$$
$$= \|g_{k+1}\| + 2\ \beta_{k+1}\|d_k\| \ < \ \gamma + 2\ \beta_{k+1} B \qquad \{C = \gamma + 2\ \beta_{k+1} B\}$$

So that $\|d_{k+1}\|^2 < C^2$, dividing by the quality $\|g_{k+1}\|^4$ to get

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} < \frac{C^2}{\|g_{k+1}\|^4}$$
$$\sum_{k=1}^{\infty} \frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} > C^2 \gamma^{-2} = \infty.$$

Which is in contrary to Lemma 4.3, then $Lim \inf \|g_k\| = 0$.

5 Numerical Results

To evaluate the reliability of the proposed methods, they were tested against the Yasushi, Yabe and Ford 3TCG methods with different options of p_k such as (y_k, d_k, g_{k+1}) using the same test problems as shown in Table (1). The comparison includes some of the known test functions that contributed to CUTE [7] in different dimensions (100, 400, 600, 1000). The program was written with a double precision account using Fortran 6.6. The comparative performance of the algorithm is evaluated by considering both the total number of function evaluations that is normally assumed as the usually factor in each iteration, total number of iterations and the time. The standard of convergence criterion was

$$\|g_{k+1}\| \leq 1 \times 10^{-6}. \qquad (22)$$

Percentage Performance of each New algorithm was against 100% Yasushi,Yabe and Ford 3TCG algorithms with different choice of p_k by y_k, d_k, g_{k+1} respectively, as shown in Table 4.2.

Based on the above tables, it can be concluded that the new algorithm beats Yasushi, Yabe and Ford 3TCG methods in all NOI; NOFG and Time. NOI is about (35-70) % percentages. However, the new algorithm also beats Yasushi, Yabe and Ford 3TCG methods in all NOFG about (8-52) % percentages. In addition, the new algorithm also beats Yasushi, Yabe and Ford 3TCG methods time about (27-73) % percentages.

Table 1: Show the Comparison between New and Yasushi,Yabe and Ford 3TCG methods for the total of n different dimensions n= 100, 400, 700,1000, for each test problem ($\phi_1 = 0.3, \phi_2 = 0.4$, $\varepsilon = 1x10^6$).

Number of Problem	New method by Z vectors NOI/NOFG/TIME	3TCG by $\mathbf{y_k}$ vectors **(Yasushi,Yabe, Ford)** NOI/NOFG/TIME	3TCG by $\mathbf{g_{k+1}}$vectors **(Yasushi,Yabe, Ford)** NOI/NOFG/TIME	3TCG by d_k vectors **(Yasushi,Yabe, Ford)** NOI/NOFG/TIME
1	235 365 0.17	265 388 0.22	348 480 0.28	425 556 0.30
2	84 161 0.02	84 161 0.01	84 161 0.02	84 161 0.00
3	35 78 0.02	35 78 0.02	35 78 0.01	35 78 0.01
4	348 389 0.19	388 429 0.22	494 536 0.26	2305 2351 1.73
5	228 265 0.04	302 339 0.05	236 273 0.03	373 406 0.04
6	176 188 0.14	214 226 0.17	75 87 0.06	179 189 0.13
7	250 260 0.03	60 70 0.02	2176 2188 0.06	6016 6107 0.46
8	66 79 0.08	66 79 0.06	66 79 0.08	66 79 0.08
9	126 174 0.01	141 187 0.02	90 138 0.01	151 203 0.02
10	75 116 0.05	83 118 0.05	75 94 0.05	75 94 0.04
11	178 233 0.05	228 283 0.06	165 220 0.03	248 300 0.07
12	95 135 0.01	2069 2113 0.17	95 135 0.01	2069 2113 0.18
13	14 28 0.00	14 28 0.01	14 28 0.00	14 28 0.02
14	516 548 0.05	450 493 0.07	478 514 0.06	2909 2970 0.55
15	373 402 0.07	348 377 0.06	1420 1449 0.28	435 464 0.07
16	49 117 0.03	43 112 0.01	47 55 0.01	2033 2122 0.35
17	36 48 0.00	36 47 0.02	36 48 0.00	36 48 0.02
18	43 52 0.00	43 54 0.01	43 52 0.00	43 52 0.02
19	353 388 0.06	43 54 0.01	294 329 0.03	424 459 0.06
20	29 38 0.02	29 38 0.01	29 38 0.00	29 38 0.01
21	274 307 0.03	320 353 0.05	247 277 0.03	364 393 0.05
22	174 209 0.03	183 198 0.06	251 273 0.10	191 219 0.04
23	220 263 0.08	219 261 0.08	183 209 0.05	308 387 0.09
24	49 77 0.00	45 82 0.02	51 79 0.02	47 84 0.00
25	101 127 0.02	101 127 0.01	101 127 0.01	101 127 0.03
26	216 227 0.06	207 218 0.06	234 245 0.08	333 352 0.08
27	49 80 0.00	49 80 0.02	55 84 0.00	55 84 0.00
28	49 80 0.00	32 66 0.02	32 66 0.01	**32 66 0.02**
29	35 43 0.01	35 43 0.02	35 43 0.03	35 43 0.03
30	18 30 0.00	18 30 0.00	18 30 0.02	18 30 0.02
31	75 94 0.03	75 94 0.03	75 94 0.03	75 94 0.03
32	285 333 0.04	287 331 0.03	294 342 0.02	364 448 0.06
33	8 28 0.00	11 34 0.00	11 34 0.00	11 34 0.00
34	32 44 0.00	32 44 0.00	32 44 0.01	32 44 0.02
35	2064 2100 0.04	4082 4135 0.46	4036 4064 0.19	2138 2245 0.29
Total	6958 8106 2:18	10637 11770 3:33	11955 12993 3:08	22053 23468 8:18

Table 2: Performance of the new algorithm against 100% of Yasushi, Yabe and Ford algorithm, as followed in Table 1.

Tools	3TCG By y	New	3TCG By g	New	3TCG By s	New
NOI	100%	65.41%	100%	58.20%	100%	31.55%
NOFG	100%	68.87%	100%	62.38%	100%	34.54%
Time	100%	64.78%	100%	73.40%	100%	28.04%

6 Conclusions

In this paper, a three-term conjugate gradient method was projected. The good property of these methods is that this algorithm can produce sufficient descent with conjugancy direction, under a few assumptions. The proposed CG methods are shown to be globally convergent for uniformly convex and general functions, respectively. Some numerical results are reported against Yasushi, Yabe and Ford 3TCG algorithm which demonstrated the viability of the new proposed CG algorithms with the scalars ϕ_1 and ϕ_2.

6 Appendix

The details of the 35-test functions used are:

1-Extended Trigonometric Function. 2-Extended Penalty Function. 3-Raydan2 Function.

4-Hager Function. 5-Generalized Tridiagonal-1 Function. 6-Extended Three Exponential Function. 7-Diagonal 4 Function. 8-Diagonal5 Function. 9-Extended Himmelblau Function. 10-Generalized PSC1 Function. 11- Extended Block Diagonal BD1 Function. 12-Extended Quadratic Penalty QP1 Function. 13-Extended Quadratic QF2 Function. 14-Extended EP1 Function.15-Extended Tri-diagonal 2 Function. 16- DIXMAANA Function. 17-DIXMAANB Function. 18- DIXMAANC Function. 19-EDENSCH Function. 20-DIAGONAL 6 Function. 21-ENGVALI Function. 22-DENSCHNA Function. 23-DENSCHNC Function. 24-DENSCHNB Function. 25-DENSCHNF Function. 26-Extended Block–Diagonal BD2 Function. 27-Generalized quadratic GQ1 Function. 28-DIAGONAL 7 Function. 29- DIAGONAL 8 Function. 30- Full Hessian Function.
31-SINCOS Function. 32- Generalized quadratic GQ2 Function. 33-ARGLINB Function. 34-HIMMELBG Function. 35-HIMMELBH Function

References:

[1] L. Zhang, W. Zhou and D.H. Li, A descent modified Polak-Ribi`ere-Polyak conjugate gradient method and its global convergence, IMA J. Numer. Anal., 26(4) 2006, 629-640.

[2] L. Zhang, W. Zhou and D.H. Li, Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search, Numer. Math., 104(4) 2006, 561-572.

[3] L. Zhang, W. Zhou and D.H. Li, Some descent three-term conjugate gradient methods and their global convergence, Optim. Method. Softw., 22(4) 2007, 697-711 .

[4] W. Cheng, A two-term PRP-based descent method, Numer. Func. Anal. Opt., 28(11-12) 2007, 1217-1230.

[5] N. Andrei, Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization, J. Optim. Theory Appl., 141(2) 2009, 249-264.

[6] H. Yabe and Takano M., Global Convergence Properties of Nonlinear Conjugate Gradient Methods with Modified Secant Condition, Comput. Optim. Appl., 28(2) 2004, 203-225.

[7] K. E Bongartz, A.R Conn, N. I. M. Gould and P. L. Toint., CUTE: constrained and unconstrained testing environments. ACM T. Math. Software., 21(1) 1995, 123-160.

[8] E. M. L. Beale, A derivative of conjugate gradients, in: F.A. Lootsma (Ed.), Numer. Meth. Part D E. 1972, 39-43.

[9] L. Nazareth, A conjugate direction algorithm without line search. J. Optimiz. Theory. App. ,23(3) 1977, 373-387.

[10] Y. Narushima, H. Yabe and J. A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, SIAM J. Optimiz., 21(1) 2011, 212-230.

[11] W. W. Hager, and Z. Hongchao, A survey of nonlinear conjugate gradient methods, Pac. J. Optim., 2(1) 2006, 35-58.

[12] A. Y. Al-Bayati and M. S. Al-Jameel, New Scaled Proposed Formulas for Conjugate Gradient Methods in Unconstrained Optimization, J. Comp. Sci. Math., 11(2) 2014, 25-46,.

[13] R. T. Rockafellar, Convex analysis. Princeton University Press, New York, 1970.

Marwan S. JAMEEL, ORCID: https://orcid.org/0000-0003-3397-2749

Zeyad M. ABDULLAH, ORCID: https://orcid.org/0000-0002-7863-1860