IASJ
المجلات الاكاديمية العراقية

Tikrit Journal of Pure Science
ISSN: 1813 – 1662 (Print) --- E-ISSN: 2415 – 1726 (Online)

Journal Homepage: http://tjps.tu.edu.iq/index.php/j

Three Term Conjugate Gradient Technique and its Global Convergence based on the Zhang, Zhou and Li methods

Marwan S. Jameel

Department of Environmental Technology, College of Environmental, University of Mosul, Mosul, Iraq

https://doi.org/10.25130/tjps.v27i2.72

ARTICLE INFO.

Article history:

-Received: 22 / 10 / 2021

-Accepted: 3 / 11 / 2021

-Available online: / / 2022

Keywords:

Corresponding Author:

Name: Marwan S. Jameel

E-mail: :

marwanjameel3@gmail.com

Tel:

ABSTRACT

The optimal conjugation coefficient distinguishes conjugate gradient methods such as two-term, three-term, and conditional from other descent methods. A novel conjugation parameter formula is constructed from Zhang, Zhou, and Li's well-known formula to formulate a three-term conjugation gradient method in the unconstrained optimization domain. The conjugation parameter β_{k+1}^{M} and the third term parameter θ_{k+1}^{M} were constructed by incorporating the Perry conjugation condition into Shanno's memory-free strategy of a conjugate gradient. The approach demonstrated a steeply sloped search direction for each iteration by demonstrating stability, global convergence, and sufficient descent analysis in the presence of a strong Wolf case. The empirical results established that the proposed method is more efficient than Zhang et al.'s techniques. Through the use of a collection of nonlinear mathematical functions.

I- Introductions

Take the following issue of un-constrained optimisation:

$\min f(x), \quad x \in R^n$ … (1.1)

Real purpose function $f: \mathbb{R}^n \to \mathbb{R}$ is continuous and differentiable, $g(x) = \nabla f(x)$ is denoted by the gradient of f at x. In general, the mathematical discipline recognizes three main types of solutions: precise, approximation, and numerical approach. Additionally, various numerical approaches exist for solving Equation (1.1), including the steepest descent (S.D.) technique, the Newton technique, the C.G. technique, and the Quasi-Newton (Q.N.) technique. The C.G. approach is essential because to its simplicity and low memory requirement, especially when the scale is huge; the C.G. method is extremely efficient. The numerical type was classified by C.G. procedures. As a result, if is the initial assumption for solving Problem (1.1), the bracketing process entails starting with an initial guess, x_0, and descending downhill, computing f(x) at iterates x_1, x_2, x_3, x_4,..., until we reach an iterate x_n, at which the value of the objective purpose f(x) increases for the first time. Typically, a nonlinear C.G. approach is built in an iterative fashion to approximate the ideal solution by the use of

$$x_{k+1} = x_k + \alpha_k d_k \quad \text{…} \quad (1.2)$$

α_k is called the step length It is obtained through a series of line searches. To ensure a suitable decline in the function value without taking too short steps, the step length α_k can be set using the Wolfe, Goldstein, or Armijo criteria. And the direction d_k is in the case of two terms specified by

$$d_k = \begin{cases} -\nabla f_k & \text{if} \quad k = 0, \\ -\nabla f_k + \beta_k d_{k-1}, & \text{if} \quad k > 0, \end{cases} \quad \text{..(1.3)}$$

Alternatively, in the case of a three-term(T.T.C.G.) conjugate gradient as described by N. Andrei [4], one of the general forms can be used.

$$d_{k+1} = \begin{cases} -\nabla f_{k+1} & if \ k = 0 \\ -\nabla f_{k+1} - a_k s_k - b_k y_k & if \ k > 0 \end{cases} \quad \text{...(1.4)}$$

That is, the search direction will be the sum of $\{g_{k+1}, s_k, y_k\}$, also the $a_k \ and \ b_k$ take the mathematical form involved the terms$||y_k||^2$, $||g_k||^2$, $||g_{k+1}||^2$, $s_k^T g_k$ and $y_k^T g_k$ etc. commonalty $a_k = \beta_k$.where β_k is a parameter $(0 < \beta_k < 1)$ and g_{k+1} denotes $g(x_{k+1})$. There are some well-known formulas for β_k which are given as follows: [5]

$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}$ (Fletcher-Reeves (F.R.), 1964)

$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}$ (Hestenes -Stiefel (H.S.), 1952)

$\beta_k^{PR} = \frac{g_k^T y_{k-1}}{g_{k-1}^T g_{k-1}}$ (Polak- Ribiere (P.R.), 1969)

Where $\| \cdot \|$ represents the Euclidean norm and g_{k+1} denotes $g(x_{k+1})$ and $y_k = \nabla f_{k+1} - \nabla f_k$ of vectors. To ensure that the function value drops appropriately without taking too few steps, the Wolfe, Goldstein, or Armijo criterion can be used to set the step length α_k. Generally, one looks for the ILS, such as the strong wolfe conditions (SWC), in the theoretical convergence analysis of CG-approach, as illustrated in the following [6]:

The purpose of the strong wolfe line search is to α_k locate s.t.:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k , \quad 0 \le \delta \le \frac{1}{2} \quad \ldots(1.5)$$

$$|d_k^T g(x_k + \alpha_k d_k)| \le -\sigma d_k^T g_k \quad , \quad \delta \le \sigma \le 1$$

This article has been disseminated as follows: The next part provides background information and a predecessor to a novel method: three-term CG-algorithms and Shanno's C.G. methods as memoryless Q.N. methods. Section 3 will have detailed the newly generated algorithm. We examined the global convergence properties of the suggested novel C.G. methods in Section 4. We published various numerical comparisons versus hybridization formulations in Section 5 by depending on them Zhang et al. Section 6 provide basic conclusions based on [1-3] approaches employing 35-test problems in the CUTE [7].

II- Background and Preliminary to New Method

2.1. Three-term CG methods

Recently, considerable research has been conducted on a three-term conjugate gradient technique in order to increase the effectiveness of the standard conjugate gradient method. Beale developed the first nonlinear three-term CG method in [8], specifying the search direction as:

$$d_{k+1} = -g_{k+1} + \beta_k d_k + \gamma_k d_r \ldots(2.1)$$

The parameter $\beta_k = \beta_k^{FR}$ or, { β_k^{HS} , β_k^{PR} …, etc.}, in the Beale algorithm [9]. Nazareth developed a variant of the three-term formula approach in [10], in which the search path is calculated via:

$$d_{k+1} = -y_k + \frac{y_k^T y_k}{y_k^T d_k} d_k + \frac{y_{k-1}^T y_k}{y_{k-1}^T d_{k-1}} d_{k-1} \quad \ldots(2.2)$$

[4] described a descending modified PRP conjugate gradient method in which the search path is determined by the following three-term formula:

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

The H.S. conjugate gradient approach was refined in [3] using a downward three-term formulas. it says:

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

Additionally, in [10], a typical structure of three-term conjugates gradient approaches is described, which typically generates a sufficient descent direction using the 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 is β_k similar to Beale's form.

2.2. Shanno's Conjugate Gradient strategies as memoryless Quasi~Newton approaches:

Quasi-Newton approach for function minimization is strategies for minimizing (1.1) of the following structural elements:

$x_{k+1} = x_k + p_k$ and $p_k = -\alpha_k H_k g_k$ …(2.4)

At each step, a preliminary approximation matrix H_k to the inverse Hessian is placed to ensure that the secant condition is satisfied. Typically, updates are chosen from the Broyden type described by the skill of

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \theta v_k v_k^T + \frac{p_k p_k^T}{p_k^T y_k} \ldots(2.5)$$

$v_k = \sqrt{y_k^T H_k y_k} \left(\frac{p_k}{p_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right)$, θ a scalar.

These strategies have been the subject of a great deal of current research, particularly concerning the desire for the parameter θ. The $BFGS$ replacement corresponding to $\theta = 1$ and described by:

$$H_{k+1} = H_k - \frac{H_k y_k p_k^T + p_k y_k^T H_k}{p_k^T y_k} + \left(1 + \frac{y_k^T H_k y_k}{p_k^T y_k}\right) \frac{p_k p_k^T}{p_k^T y_k}, \quad \ldots (2.6)$$

Benefits from computational solid and theoretical proof. As the first-class update of the Broyden type[6].

The matrix H_k This is the main distinction between conjugate gradient and Quasi-Newton iterative approach. For large variables, it is frequently hard to store an approximation to the inverse Hessian in the available PC memory. As a result, desired strategies to the Quasi-Newton methods must be found. Conjugate gradient techniques were initially being developed for problems of this type.

The purpose of this study is to determine the link between conjugate gradient and quasi-Newton methods. To begin, make a matrix.

$$G_{k+1} = I - \frac{p_k y_k^T}{y_k^T p_k} - \frac{y_k p_k^T}{y_k^T p_k} + \frac{p_k p_k^T}{p_k^T y_k}.$$

and G_{k+1} to maintain the quasi-Newton Equation, resulting in symmetric replacement.

$$G_{k+1}^* = I - \frac{p_k y_k^T + y_k p_k^T}{y_k^T p_k} + \left(1 + \frac{y_k^T y_k}{p_k^T y_k}\right) \frac{p_k p_k^T}{p_k^T y_k}. \quad \ldots(2.7)$$

Note G_{k+1}^* Satisfying (2.7) still reduces precisely to the P.R. technique under exact searches. We now are aware that if we alternative I for H_k in (2.6), receives exactly G_{k+1}^*

Then C.G. strategies are equivalent to the quasi-Newton BFGS approach. As the identification matrix, the approximation to the inverse Hessian is reset at each step. As no storing is employed to enhance the approximation to the inverse Hessian's accuracy. It is worth noting that the C.G. method specified the following criteria, namely

$d_{k+1} = -G^*_{k+1} g_{k+1}$

Does now no longer, besides a doubt, require the matrix G^*_{k+1} Rather,

$$d_{k+1} = -g_{k+1} + \left[\frac{y_k^T g_{k+1}}{p_k^T y_k} - \left(1 + \frac{y_k^T y_k}{p_k^T y_k}\right)\frac{p_k^T g_{k+1}}{p_k^T y_k}\right] p_k - \frac{p_k^T g_k}{p_k^T y_k} y_k \quad ...(2.8)$$

Memoryless Q.N. approach and no extra data beyond that required through ordinary C.G. algorithm are needed.

III- New Direction for TTCG.

In this part, we discuss the derived approach for three-term (T.T.C.G.)conjugate gradients, which is comparable to the generic platform proposed by Zhang, Zhou, and Li [1-3].

$$d_{k+1} = \begin{cases} -g_k & if\ k = 0 \\ -g_{k+1} + \beta_{k+1}\, d_k - \theta_k y_k & if\ k > 0 \end{cases} \quad ...(3.1)$$

Based on a formula similar to the last Equation, we give our method by new derived parameters for β_k and θ_k .Starting with (2.8) by taking a minor modification to

$$d_{k+1} = -g_{k+1} + \left[\frac{y_k^T g_{k+1}}{p_k^T y_k} - \left(1 + \frac{y_k^T y_k}{p_k^T y_k}\right)\frac{p_k^T g_{k+1}}{p_k^T y_k}\right] p_k - \frac{p_k^T g_{k+1}}{p_k^T y_k} y_k$$

Using the hyperplane

$$\times_{k=} \{d \in R^n : \ y_k^T d = -p_k^T g_{k+1}\} \ ... (3.3)$$

Since $p_k = \alpha d_k$, so we can write

$$y_k^T p = -\alpha\, p_k^T g_{k+1} \ .....(3.4)$$

Here by importing of Equation (3.4) is placed in the second term inside the large arc of equation (3.1),

$$d_{k+1} = -g_{k+1} + \left[\frac{y_k^T g_{k+1}}{p_k^T y_k} - \left(1 + \frac{y_k^T y_k}{p_k^T y_k}\right)\frac{p_k^T g_{k+1}}{-\alpha p_k^T g_{k+1}}\right] p_k - \frac{p_k^T g_{k+1}}{p_k^T y_k} y_k \ ... (3.5)$$

Here we neglected the term close to interior brackets with equality of eq(3.4)

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

The final version of the trend is presented as in (3.1) and as given by Andia (1.4) . But with new parameters

$$d_{k+1}^{new} = \begin{cases} -g_k & if\ k = 0 \\ -g_{k+1} - a_k d_k - b_k y_k & if\ k > 0 \end{cases} \quad ....(3.6)$$

Denoting the conjugate parameter $a_k = \beta_{k+1}^M$ and spectral scaling for the gradient $b_k = \theta_{k+1}^M$as follows

$$\beta_{k+1}^M = \frac{y_k^T g_{k+1}}{d_k^T y_k} + \left(1 + \frac{y_k^T y_k}{p_k^T y_k}\right) \text{ and } \theta_{k+1}^M = \frac{p_k^T g_{k+1}}{p_k^T y_k} \text{ or } = \frac{d_k^T g_{k+1}}{d_k^T y_k} \ .... (3.7)$$

Furthermore, we obtain new coefficients β_{k+1}^M and θ_{k+1}^M. The processes of the technique is given in the flowchart illustrated in figure.1.

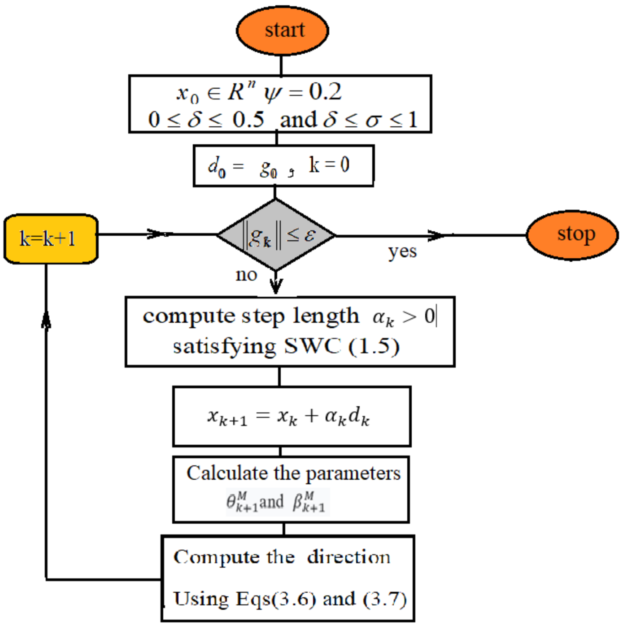

Fig. 1: Flowchart of our M-New (C.G.) conjugate gradient method

IV- Convergence Analysis.

We must demonstrate the M-New three-term CG-Algorithms' fundamental global convergence characteristic under the following premise.

Hypothesis (A):

(i)The level set $S = \{x : x \in \mathbb{R}^n, f(x) \le f(x_0)\}$ 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 neighborhood Ω of S, f is continuously differentiable, and its gradient g is Lipschitz continually; namely, there exists a constant $L \geq 0$ such that

$\|\nabla f_k - \nabla f_y\| \leq L\|x - y\|$ for all x,y belog to Ω …(4.1)

Obviously, from the Assumption (A, i), There is such a positive constant D that:

$$B = \max \{\|x - y\|, \forall y, x \in S\} \quad \text{…..} \quad (4.2)$$

Where B denotes the diameter of Ω, we also know from assumption (A, ii) that there exists a constant $\gamma \geq 0$, such that:

$\|\nabla f(x)\| \leq \gamma, \forall x \in S$ …..(4.3)

In certain studies of the C.G. techniques, the adequate sufficient descent or descent requirement is critical, although this condition is not always easy to maintain. [1]

Theorem: (Descent property)

Suppose that the assumption (A) hold, independently of choice the parameter β_k and line search with SWC(1.5), consider the search directions d_k generated from (1.2 and 1.4 We demonstrate that the direction of the search easily fulfills the sufficient method with $c = 1$:

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

Proof:

Starting with multiply the direction d_{k+1} in (3.2) by the gradient $g = g_{k+1}$

$d_{k+1}{}^T g = -\|g\|^2 - \beta_{k+1}^M\, d_k^T g + \theta_{k+1}^M\, y_k^T g$ ….(4.4)

Setting amount of parameters β_{k+1}^M and θ_{k+1}^M

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

$$= -\|g\|^2 - \left[\frac{y_k{}^T g}{d_k{}^T y_k} + (1 + \frac{y_k{}^T y_k}{\alpha\, d_k{}^T y_k})\right] d_k^T g + \frac{y_k^T g}{d_k{}^T y_k} d_k{}^T g$$

Since $d_k^T g_{k+1} < d_k^T y_k$ we get the inequality

$$\leq -\|g\|^2 + \left[\frac{y_k{}^T g}{d_k{}^T y_k} + (1 + \frac{y_k{}^T y_k}{\alpha\, d_k{}^T y_k})\right] d_k^T y_k - \frac{y_k^T g}{d_k{}^T y_k} d_k{}^T y_k$$

$$\leq -\|g\|^2 - (1 + \frac{y_k{}^T y_k}{\alpha\, d_k{}^T y_k}) d_k{}^T y_k$$

Implies that

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

For last term is increase the force of inequality by neglecting. With optimal constant $c = 1$.

Property (4.1): Assume that there is a general CG-method and that [11]:

$\varsigma \leq \|g_k\| \leq \gamma, \quad \forall k \geq 0$ …..(4.5)

ς is positive. We say that a CG-method has the Property (4.1) if there exist two constants b> 1 and $\lambda > 0$ such that for all k,

We define a CG-method as having the Property (4.1) if two constants b>1 exist and $\lambda > 0$ such that for all k,

$|\beta_k^M| \leq b$ (4.6) ….

If $\|p_k\| \leq \lambda$ then $|\beta_k^N| \leq \frac{1}{2b}$ for all $\lambda > 0$ …. (4.7)

Corollary (4.2): $|\theta_{k+1}^M| < \bar{b}$

Proof: by using Cauchy Schwarz property and Leibnitz inq. implies

$$|\theta_{k+1}^M| = \left|\frac{p_k^T g_{k+1}}{p_k^T y_k}\right| = \frac{|p_k^T g_{k+1}|}{|p_k^T y_k|} \leq \frac{\|g_{k+1}\|}{\|y_k\|} < \frac{\gamma}{B} = \bar{b}$$

Lemma(4.3): Assume that is a descending direction d_k and g satisfy, $\|\nabla f(y) - \nabla f(x)\| \leq L\|y - x\|$, the Lipschitz condition, where L is constant that all points within the line segment connecting y and x. If the direction of the line search fits the Strong Wolfe criterion, then[6]:

$$\alpha_k \geq \frac{(1-\sigma)|d_k{}^T g_k|}{L\|d_k\|^2} \quad (4.8) \quad \text{….}$$

Proof: Using curvature inequality in (1.5)

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

$\Rightarrow \sigma\ d_k{}^T g_k \leq d_k{}^T g_{k+1}$ …..(4.9)

Subtracting $d_k{}^T g_k$ from both of (4.9) and applying the Lipschitz condition results:

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

(4.10)

Since d_k is descent direction and $\sigma \leq 1$, then (34) holds:

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

The following is the conclusion: The Lemma, which is frequently referred to as the Zoutendijk condition, is used to establish the global convergence of any nonlinear CG~method. Zoutendijk [15] first mentioned it in connection with the Strong Wolfe line search (1.5). This condition will be established in the following Lemma.

Lemma (4.4): Assume hypothesis (A) is true. Consider the iteration process of the type (1.2)-(1.4), where the descent condition ($d_k{}^T g_k \leq 0$) satisfies for all $k \geq 1$ and with α_k condition(1.5). Then

$$\sum_{k \geq 1} \frac{(g_k{}^T d_k)^2}{\|d_k\|^2} < +\infty \quad \text{….(4.11)}$$

Proof: using first inequality in (1.5), we can get:

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

Combining this with the results in Lemma **(4.3)**, yields

$$f_{k+1} - f_k \leq \frac{\delta(1-\sigma)}{L} \frac{(g_k{}^T d_k)^2}{\|d_k\|^2} \quad \text{…… (4.12)}$$

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

$$\sum_{k\geq1}\frac{(g_k^T d_k)^2}{\|d_k\|^2}<+\infty \qquad \ldots(4.13)$$

Theorem

Assuming A is true and considering the new algorithm produced by (1.2,1.4) and (3.6, 3.7) where α_k is computed using wolf Line Search,

$$\underset{k\to\infty}{L.i.m}\ \inf\|g_k\|=0$$

Prove

Contradiction manner utilize to prove, so we suppose the conclusion is not proper, then $\|g_k\|\neq 0$, as mentioned above, there exist constants $\varsigma,\gamma>0$, so

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

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

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

$$\|d_{k+1}\|=\left\|-g_{k+1}-\beta_{k+1}^M d_k+\theta_{k+1}^M y_k\right\|$$

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

$$\leq\|g_{k+1}\|+\beta_{k+1}^M\|d_k\|+\theta_{k+1}^M\|y_k\|$$

$$\leq\|g_{k+1}\|+\beta_{k+1}^M\|d_k\|+\theta_{k+1}^M\|y_k\|$$

(By Cauchy Schwarz)

$$<\ \gamma+b\lambda+\bar{b}B=C\quad\{C=\gamma+b\lambda+\bar{b}B\}$$

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

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4}<\frac{C^2}{\|g_{k+1}\|^4}$$

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

This is in contradiction with Lemma (4.4), then $Lim\inf\|g_k\|=0$

V- Numerical Results

To determine the reliability of our newly proposed approaches, we compared them to Zhang. et al. 's [1-3], which are 3TCG-methods using the same test problems as showing in figures (1), (2) and (3). The comparison involves some well-known test functions contributed in CUTE [7] with different dimensions (100) to (1000) variety increasing number. The program is written in double-precision arithmetic using Fortran 66. The algorithm's comparative performance is determined by the total number of function evaluations that normally assume the most expensive element in each iteration, as well as the total number of iterations. The actual condition for convergence was

$$\|g_{k+1}\|\leq 1\times10^{-6}\quad\ldots(5.1)$$

We plot the fraction P of issues for which the process takes less than a factor of the ideal time for each technique. The figure's left side represents the percentage of test issues for which a method is the fastest; the right side indicates the proportion of test problems successfully answered using each method. The top curve denotes the approach that resolved the most problems in the least amount of time.

Comparison between M-New TT, Zhange as in [1], Zhange[2], Zhange[3] CG~methods for the total of n different dimensions $\boldsymbol{n}$ =[100, 1000] for each test problem are given in figures 2, figures 2, figures 3 and figures 4 according for measure of preferability in numerical optimization the iteration and function evaluation number the last one is cpu. As follows

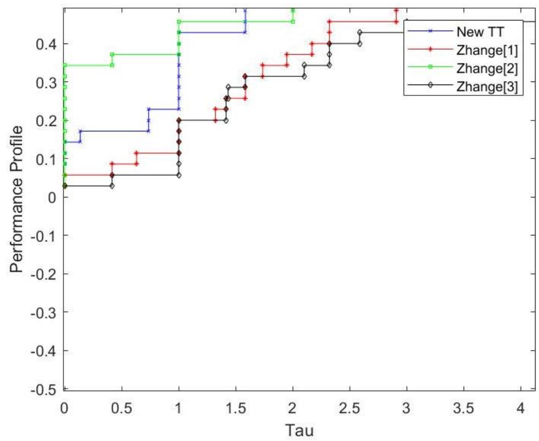

Fig. 2: The time performance between compared algorithms

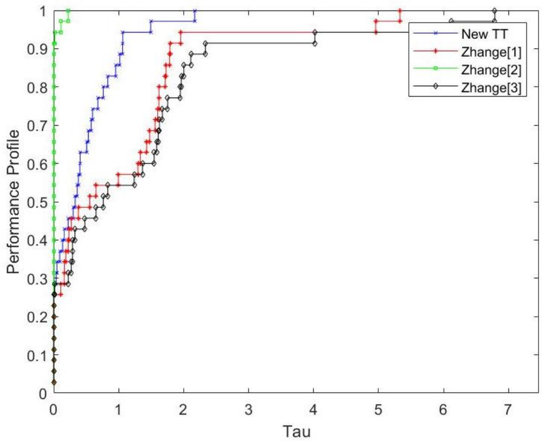

Fig. 3: The number of function gradient evaluation performance between compared algorithms.

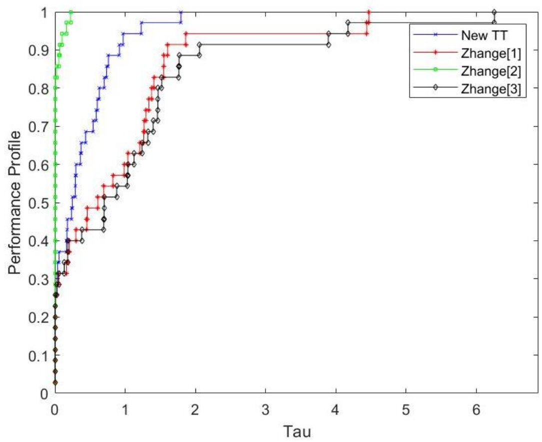

Fig. 4: The number of iteration performance between compared algorithms.

Effectiveness of each proposed technique as a percentage of Zhang [1], Zhang [2] and Zhang [3] algorithms, as shown in table 1

Table 1: proposed technique against 100% compared algorithms according to mentioned factors

Tools	Zhange as in [1]	M-New	Zhange as in [3]	M-New
Iter	100%	40%	100%	19%
Fg	100%	46%	100%	27%
Time	100%	45%	100%	26%

We inferred from the preceding table that the M-New method outperforms Zhange, as described in [1], and Zhange as in [3] CG-algorithms in all iter; fg and time in about (55-74)% percentages. However, Zhange as in [2] algorithm equivalent of M-New in all iter; fg and time.

VI- Conclusions

A M-New form of the (T.T.C.G.) conjugate gradient method has been proposed. The parameters β_{k+1}^{M} and θ_{k+1}^{M} specify the new direction. These algorithms can provide sufficient descent under specific assumptions, which is an appealing feature of these approaches. Uniformly convex and generic functions have been shown to be globally convergent. Several numerical outcomes vs Zhang et al.[1-3]-algorithms were shown, illustrating the usefulness of our new proposed C.G. algorithms with the specifications $\theta_{k+1}^{M}$and β_{k+1}^{M} .

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 Journal of Numerical Analysis, 26 (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**"**, Numerische Mathematik, 104 (2006),561–572.

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

[4] Andrei, N. (2013). A simple three-term conjugate gradient algorithm for unconstrained optimization. Journal of Computational and Applied Mathematics, 241, 19-29.

[5] Shanno, D. F. (1978). Conjugate gradient methods with inexact searches. Mathematics of operations research, 3(3), 244-256.

[6] W. Cheng, A two-term PRP-based descent method, Numerical Functional Analysis and Optimization, 28 (2007), 1217–1230.

[7] Theory Appl. Andrei N., "Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization", J. Optim. 141, pp. 249–264, (2009).

[8] Deng, S., & Wan, Z. (2015). A three-term conjugate gradient algorithm for large-scale unconstrained optimization problems. Applied Numerical Mathematics, 92, 70-81.

[9] Yabe H. and Takano M., "Global Convergence Properties of Nonlinear Conjugate Gradient Methods with Modified Secant Condition", Computational Optimization and Applications, 28, 203–225, (2004).

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

[11] Beale, E.M.L. A derivative of conjugate gradients, in: F.A. Lootsma (Ed.), Numerical Methods for Nonlinear Optimization, Academic Press, London, 1972, pp. 39–43.

[12] Nazareth, L. (1977), A conjugate direction algorithm without line search. Journal of Optimization Theory and Applications, 23, 373-387.

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

[14] Hager W.W. and Zhang H. "A survey of nonlinear conjugate gradient methods", Department of Mathematics, University of Florida, Gainesville, FL, pp. 32611- 8105, (2005).

[15] A. Y. Al-Bayati and M. S. Al-Jameel, "New Scaled Proposed Formulas for Conjugate Gradient Methods in Unconstrained Optimization", AL-Rafidain J. of Computer Science and Mathematics, Mosul, Iraq, vol. 11, no. 2, pp.25-46, 2014.

[16] Zoutendijk, G. (1970). Some algorithms based on the principle of feasible directions. In Nonlinear programming (pp. 93-121). Academic Press.

طريقة التدرج المترافق الثلاثية الحدود وتقاربها الشامل بالاستناد لطرائق Zhang و Zhou و Li

مروان صالح جميل

قسم تقانات البيئة ، كلية علوم البيئة وتقاناتها ، جامعة الموصل ، الموصل ، العراق

الملخص

معامل الترافق الأمثل تعتبر الأساس التي تعتمد عليها طرائق التدرج المترافق مثل ثنائية الحدود وثلاثية الحدود والمشروطة وتميزها عن طرائق الانحدار الأخرى. من خلال الصيغة المعروفة لـ Zhang, Zhou and Li تم اشتقاق صيغة جديدة لمعلمة الترافق لصياغة طريقة تدرج ترافق ثلاثية الحدود في مجال الامثلية غير المقيدة. بإدخال الشرط الترافق لـPerry الى طريقة Shanno الخالية من الذاكرة للتدرج المترافق تم استحداث معلمة الترافق β_{k+1}^{M} ومعلمة الحد الثالث θ_{k+1}^{M}.وقد اثبتت الطريقة إن لها الاتجاه بحث ذو انحدار شديد عند كل تكرار من خلال إثبات الاستقرارية والتقارب الشامل وتحليل الانحدار الكافي مع وجود حالة الوولف القوية. وأثبتت النتائج عمليًا أن الخوارزمية المقترحة أكثر كفاءة من خوارزميات المقارنة الخاصة بـ .Zhang et al باستخدام مجموعة من الدوال الرياضية غير الخطية.