International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-12S2, October 2019

New Conjugate Gradient Method Addressing Large Scale Unconstrained Optimization Problem

UmmieKhalthumMohdYusof, MohdAsrulHery Ibrahim, MohdRivaie, Mustafa Mamat, MohamadAfendee Mohamed, Puspa Liza Ghazali

Abstract—An iterative conjugate gradient (CG) method is prominently known for dealing with unconstrained optimization problem. A new CG method which is modified by Wei Yao Liu (WYL) method is tested by standard test functions. Moreover, the step size is calculated using exact line search. Theoretical proofs on convergence analysis are shown. As a result, this new CG is comparable to the other methods in finding the optimal points by measuring the total iterations required as well as the computing time. Numerical results showed the execution between three CG methods in details.

Index Terms: Conjugate gradient (CG) method, global convergence, sufficient descent condition, unconstrained optimization.

I. INTRODUCTION

The conjugate gradient (CG) algorithm is an iterative method for unconstrained minimization that produces more appropriate approximation to the minimum of general unconstrained nonlinear problems at each iteration. The main advantages of this method is capable to solve large scale problems since it not necessitate to construct and store any matrix. Due to accomplish objective function, the performance of the conjugate coefficient plays an important rule. Starting in 1952, Hestenes-Stieffel develop this method by introduce conjugate coefficient namely β^{HS} method, then continues with β^{FR} by Fletcher -Reeves in 1965. Then, Polak- Ribiere introduced new alteration in 1969 as β^{PR} [1]. However β^{PR} is the most efficient among the class of conjugate directions methods. All the modification in this method based on conjugate coefficient, β_k. Recently, researchers keep develop their own β_k based on ideas from previous and make comparison in terms of their performance. In most recent studies, this method are simple and easy to implement compare Newton method because of Hessian matrix need more time consuming to perform [13], [20], [21]. Recent development, researchers come out with much effort to design and construct another model for CG methods which produce excellent numerical achievements that satisfy global convergence properties.

The objective function for unconstrained optimization problems is denoted as

$$\min_{x\in R^n} f(x) \qquad (1)$$

where $f: R^n \to R$ is continuously differentiable over n dimensional Euclidean space. To solve (1), we use iterative method termed as,

$$x_{k+1} = x_k + \alpha_k d_k \qquad (2)$$

with x_{k+1}, α_k, d_k are new iterate point, step size and search direction respectively. In (3), we found the search direction

$$d_k = \begin{cases} -g_k & \text{if } k = 0 \\ -g_k + \beta_k d_{k-1} & \text{if } k \geq 1 \end{cases} \qquad (3)$$

with g_k is a $f(x)$ gradient at x_k while β_k is CG coefficient.

Previous studies gives valuable intentions to latest researchers come out with new CG coefficient that gain ideas from the classical β_k as listed in Table 1.

Table 1: CG coefficients

CG Coefficients	Years
$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})}$	Hestenes-Steifel (1952)
$\beta_k^{FR} = \frac{g_k^T g_k}{\|g_{k-1}\|^2}$	Fletcher-Reeves (1964)
$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}$	Polak-Ribiere (1969)

Nowadays, many researchers develop new formula and some commonly used and also widely known are [2], [5], [6], [12], [15], [17]-[19], [23].

Line search is a technique to calculate the step size, α_k along a given search direction d_k.

To determine the values of the step size, either exact or inexact line search can be used. Armijo is the first method in inexact and followed by Goldstein, Wolfe and Strong Wolfe method. The ways to find α_k will effects both the convergence and the speed of convergence of the algorithm. However, this paper only studied CG algorithms with exact line search,

Revised Manuscript Received on September 14, 2019.

Ummie Khalthum Mohd Yusof, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

Mohd Asrul Hery Ibrahim, Faculty of Entrepneurship and Business, Universiti Malaysia Kelantan, Kelantan, Malaysia.

Mohd Rivaei, Faculty of Computer Sciences and Mathematics, Universiti Teknologi Mara (UiTM) Kuala Terengganu, Malaysia.

Mustafa Mamat, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

Mohamad Afendee Mohamed, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

Puspa Liza Ghazali, Faculty of Economics and Management Sciences, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

$$\min_{\alpha\geq 0} f\left(x_k + \alpha_k d_k\right) \quad (4)$$

II. THE NEW COEEFICIENT

A. Propose New Coefficient

In this study, a new CG method with namely β_k^{UAM} (Ummie, Asrul and Mustafa) proposed by pursuing the excellent performance of coefficient proposed by [22], [7] in term of successful to solve the test problem.

$$\beta_{k+1}^{UAM} = \max\left\{0, \frac{g_{k+1}^T\left(g_{k+1} - \frac{\|g_k\|}{\|g_{k+1}\|}g_k\right)}{\|g_k\|^2}\right\} \quad (5)$$

Hence, the complete algorithms for the CG methods are arranged by the following pseudo-code,

1. Set k to 0 with initial point, x_0 an element of R
2. Compute β_k as (5)
3. Calculate the d_k using (3). If $\|g_k\| = 0$, then exit.
4. Compute the α_k by (4)
5. Compute x_{k+1} based on (2)
6. Evaluate convergence property, followed by stopping condition such that if $f\left(x_{k+}\right) < f\left(x_k\right)$ and $\|g_{k+1}\| \leq \varepsilon$, then exit. Else, jump to step 1 and increase k by one.

B. Convergence Analysis

The convergence analysis of β_k^{UAM} are presented. Furthermore, we establish the sufficient descent condition and global convergent to ensure an algorithm satisfy both conditions.

C. Sufficient Descent Condition

For $k \geq 0$, each d_k should satisfy the descent condition by the following assumption

$$g_k^T d_k < 0 \quad (6)$$

Suppose there exists a constant, $c_1 > 0$, then

$$g_k^T d_k \leq c_1 \|g_k\|^2 \quad (7)$$

which dictates the fulfillment of this condition by search direction.

Theorem 1. Assume CG method as in (5), search direction as (3) with the condition (7) also hold for $k \geq 0$.

Proof. If $k \geq 0$ then we get (7). When $k \geq 1$, we prove by induction. By multiplying our search direction (3) with g_{k+1}^T we get,

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_{k+1}^{UAM} g_{k+1}^T d_k$$

For exact line search, $g_{k+1}^T d_k = 0$. Hence $g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2$. Thus, sufficient descent direction holds and completes the proof.

D. Global Convergence Properties

In making sure there are remain not less than zero to examine the global convergence properties, β_k^{UAM} is simplified as follow:

$$\beta_{k+1}^{UAM} = \begin{cases} 0 & \text{for } \|g_{k+1}\|^2 < \frac{\|g_k\|}{\|g_{k+1}\|} g_{k+1}^T g_k \\ \frac{g_{k+1}^T\left(g_{k+1} - \frac{\|g_k\|}{\|g_{k+1}\|}g_k\right)}{\|g_k\|^2} & \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \end{cases} \quad (8)$$

Thus,

$$0 \leq \beta_{k+1}^{UAM} \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad (9)$$

In studying of global convergence properties, the following assumptions are deemed necessary.

Assumption 1

1) At the initial point x_0, f is bounded below objective function and is continuously differentiable on $\ell = \left\{x \in R^n \left| f(x) \leq f(x_0)\right.\right\}$ for some neighborhood N.

2) For f having gradient $g(x)$ given by Lipschitz continuous in N, then for any given constant L > 0, the expression $\|g(x) - g(y)\| \leq L\|x - y\|$ holds for any $x, y \in N$.

From here, we reserve to the subsequent lemma (Zoutendijk).

Lemma 1. Let Assumption 1 be true. An iterative CG scheme having search direction d_k as well as step size α_k fulfils the exact line search. Consequently, Zoutendjik condition holds such that

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

Theorem 2

Assume both Assumption 1 and Theorem 1 is true. Let any β_k^{UAM} as in (5) and also (3) such that α_k is obtained using (4). Consequently, expression (10) representing sufficient descent condition holds such that

$$\lim_{i\to\infty} \|g_i\|^2 = 0 \text{ or } \sum_{i=0}^{\infty} \frac{\left(g_i^T d_i\right)^2}{\|d_i\|^2} < \infty \quad (10)$$

Proof. Theorem 2 is verified by contradiction. If incorrect, then a constant $c > 0$ holds, whereas

$$\|g_k\| \geq c \quad (11)$$

Claiming (3) and squaring both sides, we come to

$$\|d_{k+1}\|^2 = \left(\beta_{k+1}^{UAM}\right)^2 \|d_k\|^2 - 2g_{k+1}^T d_{k+1} - \|g_{k+1}\|^2 \quad (12)$$

Consuming $\left(g_{k+1}^T d_{k+1}\right)^2$ and dividing both sides,

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} = \frac{\left(\beta_{k+1}^{UAM}\right)^2 \|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} - \frac{2}{g_{k+1}^T d_{k+1}} - \frac{\|g_{k+1}\|^2}{g_{k+1}^T d_{k+1}} \quad (13)$$

Completing the square,

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

$$\leq \frac{\left(\beta_{k+1}^{UAM}\right)^2 \|d_k\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} + \frac{1}{\|g_{k+1}\|^2}$$

Applying (7) results

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

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} \leq \frac{\|g_{k+1}\|^4}{\|g_k\|^4} \frac{\|d_k\|^2}{\left(g_{k-1}^T d_{k-1}\right)^2} + \frac{1}{\|g_{k+1}\|^2}$$

$$\frac{\|d_{k+1}\|^2}{\left(g_{k+1}^T d_{k+1}\right)^2} - \frac{\|g_{k+1}\|^2 \|d_k\|^2}{\|g_k\|^4 \|d_{k-1}\|^2} \leq \frac{1}{\|g_{k+1}\|^2}$$

Since $\beta_k^{UAM} \geq 0$ it becomes

$$\frac{\|d_{i+1}\|^2}{\left(g_{i+1}^T d_{i+1}\right)^2} \leq \frac{1}{\|g_{i+1}\|^2}$$

Therefore,

$$\sum_{i=0}^{\infty} \frac{\left(g_i^T d_i\right)^2}{\|d_i\|^2} = \infty$$

However, it is contradictory to Lemma 1. Hence, condition (8) is true and the proof is concluded.

III. RESULTS AND DISCUSSION

An overview numerical result from many different conjugate gradient methods gives various results in how performance profile looks. By using test problem based on Andrei [3], the methods are fully tested with dimension between 2 to 500 variables [14].

In total, we have 180 test problems where the initial points are subtracted from the minimum point. The stopping criteria, $\|g_k\| \leq 0$ if the iteration number exceeds the limit of 10,000.

The outcomes are displayed in Fig. 1 and 2 utilizing performance profile improved by [16]. The performance profile aims to show performance of the new coefficient in order to meet optimal solutions compared with another two coefficients. Regarding on Fig. 1, the β^{UAM} yields the best performance in terms of test problem solved with the fastest method. However, in Fig. 2, it can be summarized that β^{UAM} is as good as β^{WYL} at the starting movement, but by referring to the right hand side of graph, the percentage of test problem solved by the new method is the highest compared to others. In conclusion, it's shown our coefficient is better and able to deal with the entire test problem [4].

Table 2:Numerical results for each methods

Coefficient	Number of Iterations	CPU Times
β^{UAM}	10416	304.37
β^{WYL}	15906	384.2969
β^{FR}	110311	551.0469

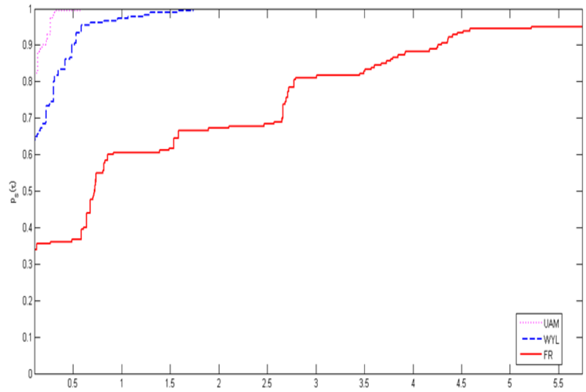

Fig. 1: Performance profile - number of iterations

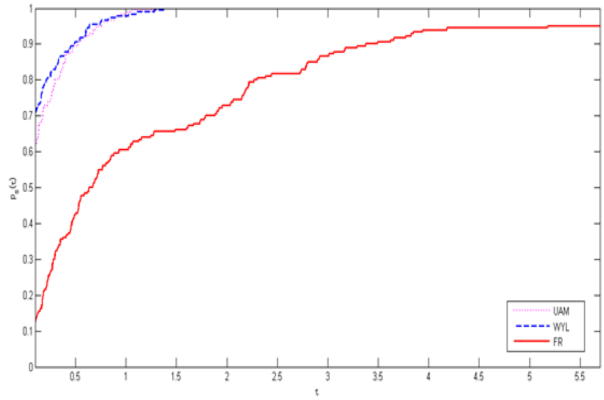

Fig. 2: Performance profile - CPU time

IV. CONCLUSION

A new coefficient, β^{UAM} with impressive performance is introduced in order to achieve optimal solutions. For the next interest, this new coefficient can be extended in hybrid between CG and BFGS methods with the search direction proposed by [8]-[11].

V. ACKNOWLEDGMENT

The research is partially funded by the Malaysian government through FRGS/1/2017/STG06/Unisza/01/1 and Universiti Sultan ZainalAbidin, Malaysia.

REFERENCES

1 M. Al-Baali, "Numerical experience with a class of self-scaling quasi-newton algorithms," Journal of Optimization Theory and Applications, 3(96), 1998, pp. 533-553.

2 N. Aini, M. Rivaie, and M. Mamat, "A modified conjugate gradient coefficient with inexact line search for unconstrained optimization," AIP Conference Proceedings, 1787(1), 2016, pp. 1-6.

3 N. Andrei, "An unconstrained optimization test functions collection," Advance Modelling and Optimization, 1(10), 2008, pp. 147-161.

4 E. Dolan, and J. J. More, "Benchmarking optimization software with performance profile," Mathematic Programming, 91(2), 2002 pp. 201-213.

5 N. H. A. Ghani, M. Rivaie, and M. Mamat M, “A modified form of conjugate gradient method for unconstrained optimization problems,” AIP Conference Proceedings, 1739(1), 2016, pp. 1-8.
6 N. Hajar, M. Mamat, M. Rivaie, and I. Jusoh, “A new type of descent conjugate gradient method with exact line search,” AIP Conference Proceedings, 1739(1), 2016, pp. 1-8.
7 M. Hamoda, M. Rivaie, M. Mamat M, and Z. Salleh, “A conjugate gradient method with inexact line search for unconstrained optimization,” Applied Mathematical Sciences, 9(37), 2015, pp. 1823-1832.
8 M. A. H. Ibrahim, M. Mamat, and L.W. June, “The hybrid BFGS-CG method in solving unconstrained optimization problems,” Hindawi Publishing Corporation, 2014, 2014, pp. 1-6.
9 M. A. H Ibrahim, M. Mamat, and L. W. June, “BFGS method: A new search direction,” Sains Malaysiana, 10(43), 2014, pp. 1591-1597.
10 M. A. H Ibrahim, M. Mamat, L. W. June, and A. Z. M. Sofi, “The algorithms of Broyden-CG for unconstrained optimization problems,” International Journal of Mathematical Analysis, 8(52), 2014, pp. 2591-2600.
11 M. A. H. Ibrahim, M. Mamat, A. Z. M. Sofi, I. Mohd, and W. M. A. W. Ahmad, “Alternative algorithm of Broyden FAMILY for unconstrained optimization,” AIP Conference Proceedings, 1309(1), 2008, pp. 670-680.
12 W. Khadijah, M. Rivaie, M. Mamat, and I. Jusoh, “A spectral KRMI conjugate gradient method under the strong-wolfe line search,”AIP Conference Proceedings, 1739(1), 2016, pp. 1-8.
13 Long, X. Hu, and L. Zhang L, “Improved Newton’s method with exact line searches to solve quadratic matrix equation,” Journal of Computational and Applied Mathematics, 222(2), 2008, pp. 645-654.
14 Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs. Berlin: Springer, 1996.
15 N. S. Mohamed, M. Mamat, F. S. Mohamad, and M. Rivaie, “A new coefficient of conjugate gradient methods for nonlinear unconstrained optimization,” Jurnal Teknologi, 78(6-4), 2016, pp. 131-136.
16 J. L. More, B. S. Garbow, and K. E. Hillstrom, “Testing unconstrained optimization software,” ACM Transaction on Mathematical Software, 7(1), 1981, pp. 17-41.
17 M. Rivaie, A. Abashar, M. Mamat, and I. Mohd I, “The convergence properties of a new type of conjugate gradient methods,” Applied Mathematical Sciences, 8(1), 2014, pp. 33-44.
18 N. Shapiee, M. R. M. Ali, M. Mamat, and Z. Salleh, “A new simple conjugate gradient coefficient for unconstrained optimization,” Applied Mathematical Sciences, 9(63), 2015, pp. 3119-3130.
19 S. Shoid, M. Rivaie, M. Mamat, and Z. Salleh, “A new conjugate gradient method with exact line search,” Applied Mathematical Sciences, 9(96), 2015, pp. 4799-4812.
20 A. Z. M. Sofi, M. Mamat, I. Mohd, and M. A. H Ibrahim, “Fletcher Reeves Like CG Formula approach on Broyden Family Update,” 3rd International Conference on Mathematical Sciences, 2014, pp. 527-532.
21 A. Z. M. Sofi, M. Mamat, I. Mohd, and Y. Dasril, “An alternative hybrid search direction for unconstrained optimization,” Journal of Interdisciplinary Mathematics, 11(5), 2008, pp. 731-739.
22 Z. Wei, S. Yao and L. Liu, “The convergence properties of some new conjugate gradient method,” Applied Mathematics and Computation, 183(2), 2006, pp. 1341-1350.
23 N. Zull, M. Rivaie, M. Mamat, Z. Salleh, and Z. Amani, “Global convergence of a new spectral conjugate gradient by using strong wolfe line search,” Applied Mathematical Sciences, 9(63), 2015, pp. 3105-3117.