An Extension of Polak-Ribière-Polyak Method using Exact Line Search

Mahmoud Dawahdeh, Mustafa Mamat, Mohd Rivaie, Mohamad Afendee Mohamed, Puspa Liza Ghazali

ABSTRACT--- Lately, many large-scale unconstrained optimization problems rely upon nonlinear conjugate gradient (CG) methods. Many areas such as engineering, and computer science have benefited because of its simplicity, fast and low memory requirements. Many modified coefficients have appeared recently, all of which to improve these methods. This paper considers an extension conjugate gradient method of Polak-Ribière-Polyak using exact line search to show that it holds for some properties such as sufficient descent and global convergence. A set of 113 test problems is used to evaluate the performance of the proposed method and get compared to other existing methods using the same line search.

Index Terms — Conjugate gradient (CG) method, exact line search, global convergence, sufficient descent property.

I. INTRODUCTION

The problem of unconstrained optimization is represented as a function for which the nonlinear CG methods are used to minimize. Generally, the form of the optimization problem is:

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

such that $f : R^n \rightarrow R$ is continuously differentiable. The gradient of the function f(x) is available and denoted by $g(x) = \nabla f(x)$. Now, the iterative formula of CG method and the search direction d_k are respectively defined by:

$$x_{k+1} = x_k + \alpha_k d_k \qquad , k = 0,1,2,3, ... \qquad (2)$$

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

where $\alpha_k > 0$ is the step length which can be obtained through some methods (line search, etc.) and β_k a CG coefficient that different CG methods are defined based on it. In addition, g_k a gradient at point x_k of *f*(*x*). There exist two types of line searches namely exact and inexact line search. We use exact line search in this study as a tool to compute the step length, and it can be given as:

$$f(x_k + \alpha_k d_k) = \min f(x_k + \alpha d_k) \ , \ \alpha \geq 0 . \qquad (4)$$

This method is not preferred to be studied by many researchers, because they believe that it is quite slow. They prefer to use the other line search which is inexact. But when fast computer processors are used with our method, it gives good numerical results and so, this will be an advantage for exact line search method. The most common β_k formulas are:

Hestenes-Stiefel (HS) [1], Polak-Ribi`re-Polyak (PRP) [2], Gilbert and Nocedal (PRP+) [3], Fletcher-Reeves (FR) [4], Liu-e Storey (LS) [5], Dai-Yuan (DY) [6], Wei et al. (WYL) [7], Conjugate Descent (CD) [8] which are given by:

$$\beta_K^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \qquad (5)$$

$$\beta_K^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \qquad (6)$$

$$\beta_K^{PRP+} = max \{\beta_K^{PRP} , \ 0\}, \qquad (7)$$

$$\beta_K^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} , \qquad (8)$$

$$\beta_K^{LS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \qquad (9)$$

$$\beta_K^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})}, \qquad (10)$$

$$\beta_k^{WYL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2} , \qquad (11)$$

$$\beta_K^{CD} = \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} , \qquad (12)$$

In 1952, Hestenes-Stiefel [1] proposed the earliest formula for solving the quadratic functions. Accordingly, a formula by Fletcher and Reeves [4] came later in 1964 for nonlinear functions. Based on FR method, some investigations on convergence properties were initiated by Zoutendijk [9], subsequently followed by Al-Baali [10] with the proof of global convergence using SWP line search having $\sigma < 1/2$. After that, Guanghui et al. [11] made an extension to the result with $\sigma \leq 1/2$. With an exact line search, Elijah and Ribiere [2] showed that PRP method finally globally converged. However, Powell [12] disproved the global convergence property of PRP method by showing that there exists a nonconvex function. As a result, Powell introduced nonnegative PRP method. Later, Gilbert and Nocedal [3] proved the global convergence property of PRP+ for $\beta_K^{PRP+} = max \{\beta_K^{PRP}, 0\}$ using some line search. Unfortunately, using SWP line search with general nonlinear functions, PRP+ failed to converge. Moreover, Wei et al. [7] discovered a new positive CG method that behaves very similar to the old PRP method. This is followed by several modifications which appeared in [13], [14] much later.

Revised Manuscript Received on July 10, 2019.

Mahmoud Dawahdeh, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia.

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

Mohd Rivaie, Department of Computer Science and Mathematics, Universiti Technologi MARA (UiTM) Terengganu, Kuala Terengganu Campus, 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.

$$\beta_k^{VHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T y_{k-1}} \tag{13}$$

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{w|g_k^T d_{k-1}| + \|g_{k-1}\|^2} \tag{14}$$

where $w \geq 1$.

Zhang [15] introduced another method, a variant of β_k^{WYL},

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \tag{15}$$

Generally, CG methods are aiming at robustness and efficiency [16]–[19]. This paper is structured as follows. Our proposed CG method is presented in section 2. Then, the convergence analysis presented in section 3. Numerical results are shown in Section 4. The Last Section is 5, where the conclusion is.

II. NEWLY PROPOSED FORMULA AND ALGORITHM

In this section, our new β_k is presented which is known as β_k^{DMAR} and it is extended to β_k^{WYL} and β_k^{NPRP} method. It is known as (Dawahdeh, Mamat, and Rivaie) DMAR method. The idea of the newly proposed formula mainly comes from Zhang [15] and is given in (16).

$$\beta_k^{DMAR} = \begin{cases} \frac{\|g_k\|^2 - \mu_k |g_k{}^T g_{k-1}|}{\|g_{k-1}\|^2}, & \|g_k\|^2 \geq \mu_k |g_k{}^T g_{k-1}| \\ 0 & , \quad otherwise \end{cases}. \tag{16}$$

where $\|.\|$ is the Euclidean norm. We define μ_k as

$$\mu_k = \frac{\|g_k\|}{\|y_{k-1}\|^2} = \frac{\|g_k\|}{\|g_k - g_{k-1}\|^2} \quad .$$

Noteworthy

$$0 \leq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_k - g_{k-1}\|^2} |g_k{}^T g_{k-1}|}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_K^{FR} \text{ and so,}$$

$$0 \leq \beta_k^{DMAR} < \beta_K^{FR}. \tag{17}$$

Hence, according to the argument presented in [3], β_k^{DMAR} will inherit all properties belong to β_K^{FR}.

Algorithm 1

Step 1: Choose an initial point x_0, and set initial search direction $d_0 = -g_0$ and $k = 0$.

Step 2: Compute d_k from (3) using (19).

Step 3: Compute α_k using (4).

Step 4: Update x_{k+1} following to (2).

Step 5: If $\|g_k\| \leq 10^{-6}$ then stop; otherwise let $k = k + 1$ and go to Step 2.

III. CONVERGENCE ANALYSIS

We begin the study of convergence properties of β_k^{DMAR} with sufficient descent condition which will be needed later to prove the global convergence property. In CG methods, one important rule is that of descent condition. By proving that

$$g_k^T d_k \leq 0$$

then, the expression $f(x_{k+1}) < f(x_k)$ holds. We could extend to the definition of sufficient descent condition by the following,

$$g_k^T d_k \leq -C\|g_k\|^2, \text{ when } k \geq 0, C > 0. \tag{18}$$

Prior proceeding to the proof of sufficient descent direction condition for DMAR, noteworthy β_k^{DMAR} satisfies

$$0 \leq \beta_k^{DMAR} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

$$0 \leq \beta_{k+1}^{DMAR} \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \tag{19}$$

Theorem 1. If $g_k \neq 0$, suppose the direction d_k is generated by Algorithm 1, then we have the condition (18) holds for any $k \geq 0$.

Proof. Clearly, the result is true for $k = 0$, $d_0 = -g_0$. Now for $k \geq 1$, from (3), and (19) we have

$$d_k = -g_k + \beta_k^{DMAR} d_{k-1}. \tag{20}$$

Multiplying both side by g_k^T, then we get

$$g_k^T d_k = -\|g_k\|^2 + \beta_k^{DMAR} g_k^T d_{k-1}.$$

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

Specific to exact line search, we have $g_k^T d_{k-1} = 0$. Hence,

$$g_k^T d_k = -\|g_k\|^2.$$

And so, (18) holds true which implies that d_k is a sufficient descent direction. This completes the proof.

Now, we will show that our CG method with β_k^{DMAR} converges globally. The following basic assumption is frequently required for completeness.

Assumption 1.

(i) The level set $L = \{ x \in R^n : f(x) \leq f(x_0) \}$ is bounded.

(ii) For some neighborhood N of L, f is continuously differentiable function, having Lipschitz continuous gradient, and there exists a constant $K > 0$ such that

$$\|g(x) - g(y)\| \leq K\|x - y\|, \forall x, y \in N.$$

Lemma 1. Let Assumption 1 hold. x_k is given by Algorithm 1, then we have

$$\sum_{K=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \tag{21}$$

Go back to [9] for the proof of Lemma 1.

The following theorem proves the global convergence property of DMAR, depending on the above Assumption and Lemma.

Theorem 2. Consider the sequence x_k is generated by Algorithm 1, and suppose that Assumption 1 holds. Then, we obtain

$$\lim_{K\to\infty} \inf \|g_k\| = 0. \tag{22}$$

Retrieval Number: B10630782S319/19©BEIESP
DOI : 10.35940/ijrte.B1063.0782S319

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Proof. Suppose the contradiction, there exists a positive constant $\epsilon > 0$ such that

$$\|g_k\| \geq \epsilon, \text{ for all } k \geq 0, \quad (23)$$

which means

$$\frac{1}{\|g_k\|^2} \leq \frac{1}{\epsilon^2}, \text{ for all } k \geq 0 \text{ and } \|g_k\| \neq 0. \quad (24)$$

Rewriting (20) as $d_k + g_k = \beta_k^{DMAR} d_{k-1}$, and squaring both sides we obtain

$$\|d_k\|^2 + \|g_k\|^2 + 2g_k^T d_k = (\beta_k^{DMAR})^2 \|d_{k-1}\|^2,$$

then

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + (\beta_k^{DMAR})^2 \|d_{k-1}\|^2. \quad (25)$$

Dividing both sides of (25) by $(g_k^T d_k)^2$, we have

$$\frac{\|d_k\|^2}{\|g_k\|^4} = \frac{\|d_k\|^2}{(g_k^T d_k)^2} = (\beta_k^{DMAR})^2 \frac{\|d_{k-1}\|^2}{(g_k^T d_k)^2} - \frac{2g_k^T d_k}{(g_k^T d_k)^2} - \frac{\|g_k\|^2}{(g_k^T d_k)^2}$$

By applying the condition that, $g_k^T d_k = -\|g_k\|^2$, we find that:

$$\frac{\|d_k\|^2}{\|g_k\|^4} = \frac{\|d_k\|^2}{(g_k^T d_k)^2} = (\beta_k^{DMAR})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} - \frac{2(-\|g_k\|^2)}{\|g_k\|^4} - \frac{\|g_k\|^2}{\|g_k\|^4}$$

$$= (\beta_k^{DMAR})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} + \frac{2}{\|g_k\|^2} - \frac{1}{\|g_k\|^2}$$

$$= (\beta_k^{DMAR})^2 \frac{\|d_{k-1}\|^2}{\|g_k\|^4} + \frac{1}{\|g_k\|^2}$$

By applying (19) yields

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

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

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

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

Note that ($\|d_0\|^2 = \|g_0\|^2$), so we get :

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{1}{\|g_0\|^2} + \frac{1}{\|g_k\|^2}$$

Then, we have

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=0}^{k-1} \frac{1}{\|g_i\|^2} \leq \frac{k}{\epsilon^2}$$

And so, we get

$$\sum_{K=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \epsilon^2 \sum_{K=1}^{\infty} \frac{1}{k} = +\infty.$$

which is contradictory to Lemma 1, and hence complete the proof.

IV. NUMERICAL RESULTS

To analyze the performance of our method, we use some standard test functions as given in Table 1, which is adopted from Adorio and Diliman [20], Andrei [21], and CUTEr [22]. We conducted comparison studies between our method DMAR against other existing CG methods such as WYL [7] and NPRP [15] methods which are from the most well-known conventional and recently CG methods. Measurements are based on CPU time and the number of iterations.

The tolerance ε is selected to be 10^{-6}, it is used to study the efficiency of these methods towards the optimal. Moreover, we chose $\|g_k\| \leq 10^{-6}$ as the stopping criteria. In case the iteration counts grows bigger than 1000 times, the method is considered as failing.

All methods were simulated on MATLAB R2015a subroutine running on a PC with Intel R Core TM, i5-2410 having 3 GB RAM. Fig. 1 and 2 respectively show the measurement taken for both parameters. The performance measurements were based on that of Dolan and Mor´ [23]. Observably, the best method composes of high values of $p_s(t)$ and is normally located on the upper most right section with within both figures.

Table 1: Test cases

N Function	Dimension	Initial Points
1 Extended Beale	2	(1.8, 1.8)
2 Perturbed Quadratic	2	(0.5, 0.5)
3 Raydan 1	2	(1, 1), (10, 10)
4 Diagonal 4	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1), (10, … , 10)
5 Extended Himmelblau	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1), (10, 10)
6 Extended Powell	1000, 5000	(1, 1, . . ., 1)
7 FLETCHCR	2	(0.5, 0.5)
8 NONSCOMP	2	(3,3), (10, 10)
9 Extended DENSCHNB	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1), (10, 10)
10 Extended Quadratic Penalty QP1	2	(1, 1)
11 Extended Penalty	2	(1, 1), (2, 2)
12 Hager	2	(1, 1), (10, 10)
13 BIGGSB1	2	(0, 0)
14 Six Hump	2	(10, 10)
15 Three Hump	2	(1, 1), (-1, -1)
16 Booth	2	(1, 1), (3, 3)
17 Zettl	2	(0, 0), (1, 1)
18 Shallow	2	(-2, -2)
19 Generalized Quartic	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1), (10, 10)
20 Quadratic QF2	2	(0.5, 0.5), (1, 1)
21 Price	2	(-10, -10), (1, 1)
22 Generalized Tridiagonal 1	500	(2, 2, . . ., 2)
23 Generalized Tridiagonal 2	2	(1, 1)
24 Quadratic QF1	2	(1, 1)
25 Extended Quadratic Penalty QP2	500	(1, 1, . . ., 1)
26 Extended White and Holst	2	(3, 3), (9, 9)
27 ARWHEAD	2	(1, 1), (10, 10)
28 QUARTC	2, 500, 1000, 5000, 10000	(2, 2, . . ., 2)
29 Diagonal 4	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1)
30 Extended Block Diagonal BD1	2, 500, 1000	(0.1, 0.1, . . ., 0.1)
31 Extended Three Exponential Terms	2	(0.1, 0.1), (1, 1)
32 Extended Cliff	2, 500, 1000	(0, 0, . . ., 0)
33 DE jongs	2, 500, 1000, 5000, 10000	(10, 10, . . ., 10)
34 BIGGSB1	2	(0, 0)
35 POWER	2	(1, 1)
36 Dixon and Price	2	(1, 1)
37 Sphere	2, 500, 1000, 5000, 10000	(1, 1, . . ., 1)
38 Sum Squares	2	(1, 1)
39 TRIDIA	2	(1, 1)
40 DIXON3DQ	2	(-1, -1)
41 Generalized Quartic GQ1	2, 500, 1000, 5000, 10000	(1, . . ., 1), (10, . . ., 10)
42 Generalized Quartic GQ2	2	(1, 1), (10, 10)
43 Extended Trigonometric		(0.2, 0.2), (1, 1)

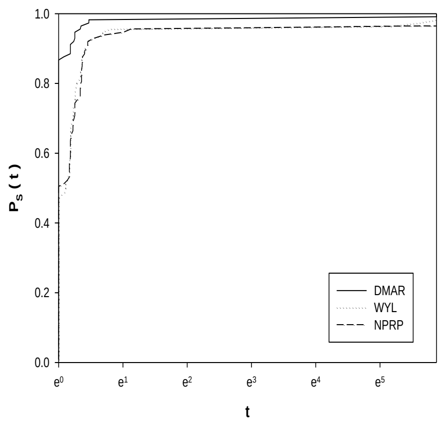

Fig. 1: Number of iterations

Each method is measure by the percentage of success in solving the test problem, represented by, $p_s(t)$ on vertical axis., whereas, the horizontal axis determines the fastest methods.

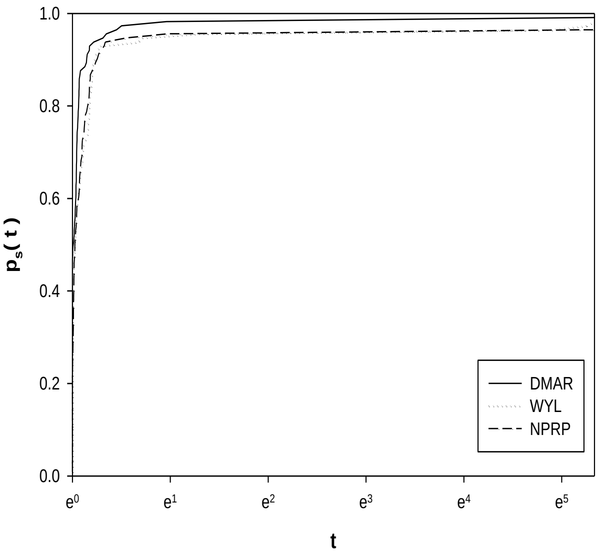

Fig. 2: CPU time

From both figures, it is clear that DMAR method has outclassed both WYL and NPRP in terms of the two parameters in most of the tested functions. In addition, DMAR has successfully solved all tested functions while WYL and NPRP have failed to do so.

V. CONCLUSION

This paper considers the CG methods that have been famously applied in many areas such as computer science, and engineering [24]–[29]. By proposing a new formula for CG method, the numerical simulation shows that the proposed coefficient exhibited superb performance in comparison to WYL and NPRP. Also, the global convergence is proved for DMAR method with the exact line search besides, the sufficient descent property.

ACKNOWLEDGMENT

This research is funded by the Malaysian Ministry of Education via Fundamental Research Grant Scheme (FRGS/1/2017/STG06/Unisza/01/1).

REFERENCES

1. M. R. Hestenes, and E. Steifel, “Method of conjugate gradient for solving linear equations,” Journal of Research National Bureau Standards, 49(6), 1952, pp. 409-436.
2. E. Polak, and G. Ribiere, “Note sur la convergence de directions conjugees,” Francaise Informat Recherche Operationelle, 3E Annee, 16, 1969, pp. 35-43.
3. J. C. Gilbert, and J. Nocedal, “Global convergence properties of conjugate gradient methods for optimization,” SIAM Journal on Optimization, 2(1), 1992, pp. 21-42.
4. R. Fletcher, and C. Reeves, “Function minimization by conjugate gradients,” Computer Journal, 7(2), 1964, pp. 149-154.
5. Y. Liu, and C. Storey, “Efficient generalized conjugate gradient algorithms, part 1: Theory”, Journal of Optimization Theory and Applications, 69(1), 1991, pp. 129-137.
6. Y. H. Dai, and Y. Yuan, Nonlinear Conjugate Gradient Methods. Shanghai Science and Technology Publisher, 2000.
7. Z. Wei, S. Yao, and L. Liu, “The convergence properties of some new conjugate gradient methods,” Applied Mathematics and Computation, 183(2), 2006, pp. 1341–1350.
8. R. Fletcher, Practical Methods of Optimization. New York: Wiley-Interscience John Wiley and Sons, 2001.
9. G. Zoutendjik, “Nonlinear programming computational methods,” in Integer and Nonlinear Programming, J. Abadie, Ed. Amsterdam: North-Holland, 1970, pp. 37-86.
10. M. Al-Baali, “Descent property and global convergence of the Fletcher–Reeves method with inexact line search,” IMA Journal of Numerical Analysis, 5(1), 1985, pp. 121-124.
11. L. Guanghui, H. Jiye, and Y. Hongxia, “Global convergence of the Fletcher-Reeves algorithm with inexact line search,” Applied Mathematics-A Journal of Chinese Universities, 10(1), 1995, pp. 75–82.
12. M. J. D. Powell, Nonconvex minimization calculations and the conjugate gradient method,” Lecture Notes in Mathematics, 1066, 1964, pp. 122-141.
13. Y. Shengwei, Z. Wei, and H. Huang, “A note about WYL’s conjugate gradient method and its applications,” Applied Mathematics and Computation, 191(2), 2007, pp. 381–388.
14. Z. Dai, and F. Wen, “Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property,” Applied Mathematics and Computation, 218(14), 2012, pp. 7421–7430.
15. L. Zhang, “An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation,” Applied Mathematics and Computation, 215(6), 2009, pp. 2269–2274.
16. M. Rivaie, M. Mamat, L. W. June, and I. Mohd, “A new class of nonlinear conjugate gradient coefficients with global convergence properties,” Applied Mathematics and Computation, 218(22), 2012, pp. 11323–11332.
17. A. Alhawarat, M. Mamat, M. Rivaie, and I. Mohd, “A new modification of nonlinear conjugate gradient coefficients with global convergence properties,” International Journal of Mathematical, Computational, Statistical, Natural and Physical Engineering, 8(1), 2014, pp. 54-60.
18. M. Rivaie, M. Mamat, and A. Abashar, “A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches,” Applied Mathematics and Computation, 268, 2015, pp. 1152-1163.
19. M. Dawahdeh, M. Mamat, A. Alhawarat, and M. Rivaie, “A new modification of conjugate gradient method with global convergence properties,” AIP Conference Proceedings, 2013(1), 2018, pp. 1-13.
20. E. P. Adorio, and U. P.Diliman, Mvf-multivariate test functions library in C for unconstrained global optimization. 2005, Available: http://www.geocities.ws/eadorio/mvf.pdf.
21. N, Andrei, “An unconstrained optimization test functions collection,” Advanced Modeling and Optimization, 10(1), 2008, pp. 147-161.
22. I. Bongartz, A. R. Conn, N. Gould, and P. L. Toint, “CUTE: Constrained and unconstrained testing environment,” ACM Transactions on Mathematical Software, 21(1), 1995, pp. 123–160.
23. E. D. Dolan, and J. M. Jorge, “Benchmarking optimization software with performance profiles,” Mathematical Programming, 91(2), 2002, pp. 201-213.
24. A. Abashar, M. Mamat, M. Rivaie, I. Mohd, and O. Omer, “The proof of sufficient descent condition for a new type of conjugate gradient methods,” AIP Conference Proceedings, 1602, 2014, pp. 296-303.

25. N. Z. Abidin, M. Mamat, B. Dangerfield, J. H. Zulkepli, M. A. Baten, and A. Wibowo, "Combating obesity through healthy eating behavior: A call for system dynamics optimization," Plos One, 9(12), 2014, pp. 1-17.
26. M. Mamat, Subiyanto, and A. Kartono, "Mathematical model of cancer treatments using immunotherapy, chemotherapy and biochemotherapy," Applied Mathematical Sciences, 7(5-8), 2013, pp. 247-261.
27. M. Mamat, S. K. Deraman, N. M. M. Noor, and Y. Rokhayati, "Diet problem and nutrient requirement using fuzzy linear programming approach," Asian Journal of Applied Sciences, 5(1), 2012, pp. 52-59.
28. M. Rivaie, M. Mamat, I. Mohd, and M. Fauzi, "Comparative study of conjugate gradient coefficient for unconstrained optimization," Australian Journal of Basic and Applied Sciences, 5(9), 2011, pp. 947-951.
29. M. Mamat, Y. Rokhayati, N. M. M. Noor, and I. Mohd, "Optimizing human diet problem with fuzzy price using fuzzy linear programming approach," Pakistan Journal of Nutrition, 10(6), 2011, pp. 594-598.