International Journal of Engineering & Technology, 7 (3.28) (2018) 92-96

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Solving a large scale nonlinear unconstrained optimization problems by using new coefficient of conjugate gradient method with exact line search direction

Talat Alkouli [1], Mustafa Mamat[1]*, Mohd Rivaie[2], Puspa Liza Ghazali[3]

[1]*Faculty of Informatics and Computing, University Sultan Zainal Abidin, 21030 Terengganu, Malaysia*
[2]*Department of Computer Science and Mathematics, Univesiti Technology MARA (UITM), 23000 Terengganu, Malaysia*
[3]*Faculty of Economy and Management Sciences, University Sultan Zainal Abidin, 21030 Terengganu, Malaysia*
**Corresponding author E-mail: must@unisza.edu.my*

Abstract

In this paper, an efficient modification of nonlinear conjugate gradient method and an associated implementation, based on an exact line search, are proposed and analyzed to solve large-scale unconstrained optimization problems. The method satisfies the sufficient descent property. Furthermore, global convergence result is proved. Computational results for a set of unconstrained optimization test problems, some of them from CUTE library, showed that this new conjugate gradient algorithm seems to converge more stable and outperforms the other similar methods in many situations.

Keywords: *Conjugate gradient; Exact line search; Global convergence; Sufficient descent condition; Unconstrained optimization.*

1. Introduction

The conjugate gradient method is an important and efficient method to solve the unconstrained optimization problems, especially for large scale problems. Its wide application in many fields such as economics, industry and engineering is due to low memory requirements and global convergence properties.

$$\min f(x)\ ,\quad x \in R^n \tag{1}$$

where $f: R^n \to R$ is continuously nonlinear differentiable function, bounded from below. A nonlinear conjugate gradient method generates a sequence x_k starting from an initial guess $x_0 \in R^n$, using the recurrence

$$x_{k+1} = x_k + \alpha_k d_k,\quad k = 0,1,2,\ldots \tag{2}$$

where α_k is the positive step size obtained by carrying out a one dimensional search, known as the 'line searches'. The most common is the exact line search which is

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha\, d_k) \tag{3}$$

and d_k is the search direction generated by

$$d_k = \begin{cases} -g_k & ,\ \ if\ k = 0\ ; \\ -g_k + \beta_k d_{k-1} & ,\ \ if\ k \geq 1\ , \end{cases} \tag{4}$$

where β_k is the CG update parameter and $g_k = \nabla f(x_k)^T$, where the gradient $\nabla f(x_k)$ of f at x_k is a row vector and g_k is a column vector.

Different CG method relates to different choices β_k. Table 1 organizes a chronological list of some choices for the classical CG parameter

$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}$	(Hestenses –Stiefel [1], 1952)
$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}$	(Fletcher –Reeves [2], 1964)
$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}$	(Polak-Ribiere–Polyak [3, 4], 1969)
$\beta_k^{CD} = \frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}$	(Conjugate Descent [5], 1987)
$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}$	(Liu–storey [6], 1991)
$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}$	(Dai–Yuan, [7], 1999)

Table 1: Various choices for the classical CG parameter

For strictly a convex quadratic objective functions, all methods in Table 1 under the exact line search have global convergence. But, for general non-quadratic functions or under the inexact line search, their behavior is relatively different [25, 26].

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The global convergence property of FR method was proved by [8] when exact line search was used. This was later refuted in a counter example introduced by [9]. Meanwhile, in [11] emphasis that FR is a superior method compared to others. In [10] has also shown that the PRP method has no global convergence under some traditional line searches. In [19] success to introduce the first global convergence result of the FR method under the Strong Wolfe condition. CD method and DY method were also proved to have global convergence under Strong Wolfe line search [12, 13]. Some convergent versions were proposed by using some new complicated line searches, or through restricting the parameter to a nonnegative number [20]. However, to the best of our knowledge, the global convergence of PRP, LS and HS methods have not been proven under all mentioned line searches. The main reason is that many conjugate gradient methods cannot guarantee the descent of objective function values at each iterative.

Many researchers have studied conjugate gradient method, but the most important properties they well-studied are sufficient condition, global convergence. For example:

In [15] proposed her conjugate gradient parameter method which is an extension of LS method and the resulting algorithm was proven to have both the sufficient descent and global convergence properties under inexact line search as follows

$$\beta_k^{ARM} = \frac{m\|g_k\|^2 - |g_k^T g_{k-1}|}{m d_{k-1}^T g_{k-1}} \quad \text{, where } m = \frac{\|d_{k-1}+g_k\|}{\|d_{k-1}\|}$$

Table 2 provides a chronological list of some new choices for the CG update parameter.

$$\beta_k^{RMIL} = \frac{g_k^T(g_k - g_{k-1})}{\|d_{k-1}\|^2}$$ (Rivaie et al. [14], 2012)

$$\beta_k^{KMAR} = \frac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T(g_k + g_{k-1})}$$ (Kamfa et al. [16], 2015)

$$\beta_k^{NRMI} = \frac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T(g_k - d_{k-1})}$$ (Shapiee et al. [17], 2016)

$$\beta_k^{MRM} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)}{\|g_{k-1}\|^2 + |g_k^T d_{k-1}|}$$ (Hamoda et al. [18], 2017)

Table 2: Various choices for the new CG update parameter

For non-quadratic objective functions, the global convergence property of all methods in Table 2 was proved when an exact line search was used.

2. New formula for β_k and its properties

During the last years, much effort has been devoted to develop new modifications of conjugate gradient methods, as we mention before, which do not only possess strong convergence properties, but they are also computationally superior to the classical methods. Since then a large number of variations of Conjugate Gradient algorithms have been suggested. A survey on their definition including 40 nonlinear Conjugate Gradient algorithms for unconstrained optimization is given by [21].

Lately, in [22] gave a variant of the *PRP* method which is called the *WYL* method. In [23] relied on *WYL* to suggest a new conjugate gradient method, *NPRP*, and he proved that the *NPRP* method satisfied descent condition and the Global convergence property under strong Wolfe line search. Subsequently, in [24] introduced an improved *NPRP* method known as the *DPRP* method.

In this section, enlightened by previous ideas [22-24], we propose our β_k which known as β_k^{TMR}. Where TMR denotes Tala't, Mustafa and Rivaie. The new β_k^{TMR} is a modification of PRP conjugate gradient method which is as follows:

$$\beta_k^{TMR} = \frac{g_k^T(n(g_k - g_{k-1}))}{n\|g_{k-1}\|^2}, where\ n = \frac{\|g_{k-1}\|}{\|g_k\|} \quad (5)$$

The algorithm is given as follows:

Algorithm 2.1

Step 1: Given $x_0 \in R^n, \varepsilon \geq 0,$ set $d_0 = -g_0$ if $\|g_0\| \leq \varepsilon$ then stop.

Step 2: Compute α_k by (3).

Step 3: Let $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = g(x_{k+1})$ if $\|g_{k+1}\| \leq \varepsilon$ then stop.

Step 4: Compute β_k by (5), and generate d_{k+1} by (4).

Step 5: set $k = k + 1$ go to Step 2.

3. Convergent analysis

In this section, the convergent properties of β_k^{TMR} will be studied. For an algorithm to be converge, it must satisfy the sufficient descent condition and the global convergence properties.

3.1. Sufficient descent condition

For the sufficient condition to hold,

$$g_k^T d_k \leq -C\|g_k\|^2 \text{ for } k \geq 0 \text{ and } C > 0 \quad (6)$$

Theorem 3.1.1.

Consider a CG method with the search direction (4) and β_k^{TMR} given as in (5), then condition (6) holds for all $k \geq 0$.

Proof:

If $k = 0$, then $g_0^T d_0 = -C\|g_k\|^2$

Hence, condition (6) holds true.

For $k \geq 1$, multiply (4) by g_{k+1}, then

$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1} + \beta_{k+1} d_k)$$

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

.

For exact line search, we know that $g_{k+1}^T d_k = 0$. Thus,

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 \quad (7)$$

which implies that d_{k+1} is a sufficient descent direction. Hence,

$$g_k^T d_k \leq -C\|g_k\|^2 \text{ holds true.}$$

The proof is completed.

3.2. Global convergence properties

In this section, we will show that CG methods with β_k^{TMR} converge globally. However, we first need to simplify our new β_k^{TMR} so that our convergence proof will be significantly easier. In (5), we can see that

$$\beta_k^{TMR} = \frac{g_k^T\left(\frac{\|g_{k-1}\|}{\|g_k\|}(g_k - g_{k-1})\right)}{\frac{\|g_{k-1}\|}{\|g_k\|}\|g_{k-1}\|^2}$$

$$= \frac{\frac{\|g_{k-1}\|}{\|g_k\|}\|g_k\|^2 - \frac{\|g_{k-1}\|}{\|g_k\|}g_k^T g_{k-1}}{\frac{\|g_{k-1}\|}{\|g_k\|}\|g_{k-1}\|^2} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

Hence, we obtain

$$0 \leq \beta_k^{TMR} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \quad (8)$$

The following straightforward assumptions are always needed in the analysis of global convergence properties of CG methods.

Assumption A

$f(x)$ is bounded below on the level set R^n and is continuous and differentiable in a neighborhood N of the level set $\ell = \{x \epsilon R^n : f(x) \leq f(x_0)\}$ at the initial point x_0.

Assumption B

The gradient $g(x)$ is Lipschitz continuous in N, $i.e$:

$$\exists\, L > 0\ s.t\ \|g(x) - g(y)\| \leq L\|x - y\| \forall\ x, y \in N$$

Under these assumptions, we have the following lemma which was proven by [8].

Lemma 3.2.1.

Suppose that Assumptions A and B holds true for any iteration method of the form (1), where d_k is a descent search direction and α_k satisfies the exact minimization rule. Then, the following condition, known as the Zoutendijk condition holds

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

The proof of this lemma can be seen in [8].

Lemma 3.2.2

Suppose that Assumptions A and B hold, $\{x_k\}$ generated by the Algorithm 2.1, where the step size is determined by the exact line search (3). Then, Lemma 3.2.1 holds for all $k \geq 0$.

Proof.

Suppose that $\forall\ k,\ g_k \neq 0$. If $k = 0$, then

$$g_0^T d_0 = g_0^T(-g_0) = -\|g_{k0}\|^2.$$

If at a point x_k , d_k is not a descent direction then by the exact line search, we have

$x_{k+1} = x_k$, which implies $g_{k+1} = g_k$.

From (5), we have $\beta_k^{TMR} = 0$.

This means that at those points the directions will turn out to be the steepest descent directions. Those points are denoted by $N_1 = \{x_k | \beta_k^{TMR} = 0\}$ and the other points are denoted by $N_2 = \{x_k | \beta_k^{TMR} \neq 0\}$.

For all the points in N_1, since the directions are the steepest descent directions, from Lemma 3.2.1, we have

$$\sum_{x_k \in N_1} \frac{(g_k^T d_k)^2}{\|d_k\|^2}$$

The same as the above proof, for the points N_2 , we also have

$$\sum_{x_k \in N_2} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \sum_{x_k \in N_1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} + \sum_{x_k \in N_2} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

The proof is completed.

By using Lemma 3.2.1, we can obtain the following convergent theorem of the CG method using (8).

Theorem 2.

Suppose that Assumptions A and B holds true. For any CG method in the form of (1) and (4) combined with (8), where α_k is obtained by the exact minimization rule defined in (3). Also, the descent condition hold true. Then, either

$$\lim_{k \to \infty} \|g_k\| = 0 \quad or \quad \sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$$

Proof:

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

From (7) $\|d_k\|^2 = \|g_k\|^2 + \beta_k^2 \|d_{k-1}\|^2$

From (8) $\|d_k\|^2 = \|g_k\|^2 + \frac{\|g_k\|^4}{\|g_{k-1}\|^4}\|d_{k-1}\|^2$

Therefore $\frac{\|d_k\|^2}{\|g_k\|^4} - \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} = \frac{1}{\|g_k\|^2}$

So $\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} \leq \sum_{k=1}^{\infty} \|g_k\|^2$

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

Hence, the proof is completed.

4. Results and discussion

In order to test the efficiency of TMR, we compare this new method with two classical methods (FR, PRP) and two of the recent modifications (WYL, RMIL). Table 4 shows the computational performance of a MATLAB program on a set of unconstrained optimization test problems. We select 34 test functions from [29].

In this test, we choose $\varepsilon = 10^{-6}$ and stopping criteria is set to $\|g_k\| \leq \varepsilon$ as in [28] suggested. Three initial points are chosen starting from a point closer to the solution point to a point further away from solution point, so that it can be used to test the global convergence of new CG coefficient. The dimensions *n* of 33 problems are 2, 4, 10, 100 and 1000.

In some cases, the computation stopped due to the failure of the line search to find the positive step size and thus it was considered as a fail. Numerical results are compared relative to the number iteration and CPU time. We use the performance profile intro-

duced by [27] to get the performance results that shown in Fig. 1 and 2.
The CPU processor used was Intel (R) Core TM i3-M350 (2.27GHz) with RAM 4 GB.

Table 1: List of Problem Functions

No.	Function	Dim	Initial Point
1	SIX HUMP	2	(0.5,0.5), (10,10), (40,40)
2	THREE HUMP	2	(-1,1),(1,-1),(2,-2)
3	LEON	2	(2,2),(4,4),(8,8)
4	MATYAS	2	(5,5),(10,10),(15,15)
5	BOOTH	2	(10,10),(25,25),(100,100)
6	RAYDAN	2	(3,3),(13,13),(23,23)
7	ZETTL	2	(5,5),(20,20),(50,50)
8	TRECANNI	2	(5,5),(10,10),(50,50)
9	EXTENDED WOOD	4	(5,..,5),(20,..,20),(30,..,30)
10	CLOVILLE	4	(2,..,2),(4,..,4),(10,..,10)
11	HAGER	2 4	(7,7),(15,15),(22,) (7,..,7),(15,..,15),(22,..,22)
12	DIXON & PRICE	2 4	(6,6),(30,30),(125,125) (6,..,6),(30,..,30),(125,..,125)
13	EXTENDED POWELL	100, 1000	(2,..,2),(4,..,4),(8,..,8)
14	GENERLIZED TRIDIAGONAL1	2 10,100	(3,3),(21,21),(90,90) (3,..,3),(21,..,21),(90,..,90)
15	GENERLIZED TRIDIAGONAL2	2 10,100	(15,15),(30,30),(150,150) (15,..,15),(30,..,30),(150,..,150)
16	EXTENDED PENALTY	2 10,100	(2,2),(20,20),(100,100) (2,..,2),(20,..,20),(100,..,100)
17	ROSENBROCK	2 10,100,1000	(3,3),(15,15),(75,75) (3,..,3),(15,..,15),(75,..,75)
18	SHALLOW	2 10,100,1000	(-2,-2),(12,12),(200,200) (-2,..,-2),(12,..,12),(200,..,200)
19	EXTENDED WHITE & HOLST	2 10,100,1000	(-3,-3),(7,7),(10,10) (-3,..,-3),(7,..,7),(10,..,10)
20	EXTENDED FREUDENSTEIN & ROTH	2 10,100,1000	(2,2),(5,5),(7,7) (2,..,2),(5,..,5),(7,..,7)
21	EXTENDED BEALE	2 10,100,1000	(-1,-1),(3,3),(10,10) (-1,..,-1),(3,..,3),(10,..,10)
22	PERTURBED QUADRATIC	2 10,100,1000	(1,1),(5,5),(10,10) (1,..,1),(5,..,5),(10,..,10)
23	EXTENDED TRIDIAGONAL1	2 10,100,1000	(6,6),(12,12),(17,17) (6,..,6),(12,..,12),(17,..,17)
24	DIAGONAL 4	2 10,100,1000	(1,1),(20,20),(40,40) (1,..,1),(20,..,20),(40,..,40)
25	EXTENDED HIMMELBLAU	2 10,100,1000	(10,10),(50,50),(125,125) (10,..,10),(50,..,50),(125,..,125)
26	FLETCHCR	2 10,100,1000	(3,3),(12,12),(15,15) (3,..,3),(12,..,12),(15,..,15)
27	EXTENDED DENSCHNB	2 10,100,1000	(5,5),(30,30),(50,50) (5,..,5),(30,..,30),(50,..,50)
28	EXTENDED BLOCK DIAG-ONAL BD1	2 10,100,1000	(1,1),(5,5),(10,10) (1,..,4),(5,..,5),(10,..,10)
29	GENERRAL-IZED QUARTIC	2 10,100,1000	(7,7),(70,70),(140,140) (7,..,7),(70,..,70),(140,..,140)
30	QUADRATIC QF2	2 10,100,1000	(4,4),(16,16),(40,40) (4,..,4),(16,..,16),(40,..,40)
31	EXTENDED QUADRATIC PENALTY QP2	2 10,100,1000	(10,10),(20,20),(30,30) (10,..,10),(20,..,20),(30,..,30)
32	QUARTC	2 10,100,1000	(8,8),(16,16),(32,32) (8,..,8),(16,..,16),(32,..,32)
33	SUM SQUARES	2 10,100,1000	(1,1),(5,5),(10,10) (1,..,1),(5,..,5),(10,..,10)
34	QUADRATIC QF1	2 10,100,1000	(3,3),(5,5),(10,10) (3,..,3),(5,..,5),(10,..,10)

In [27] offered a model to evaluate and compare the performance of the set solvers S on a test set P. Supposing n_s solvers and n_p problems exists, for each problem p and solver s, they defined $t_{p,s}$ = computing time (the number of iterations, CPU time or others) required to solve problems p by solver s.
Lacking a standard form for comparisons, they compared the performance of problem p by solver s with the best performance for any solver to the same problem using the performance ratio

$$r_{p,s} = \frac{1}{min\{t_{p,s} : s \in S\}}$$

Thus, $p_s(t)$ is the probability for solver $s \in S$ that a performance ratio $r_{p,s}$ was within a factor $t \in R$ of the best possible ration. Then, function p_s is the cumulative distribution function for the performance ratio. The performance profile $p_s: R \rightarrow [0,1]$ for a solver was a non-decreasing, piecewise, and continuous from the right. The value of $p_s(1)$ is the probability that the solver will win over the rest of the solvers. In general, a solver with high values of $P(\tau)$ or at the top right of the figure is better or represent the best solver.

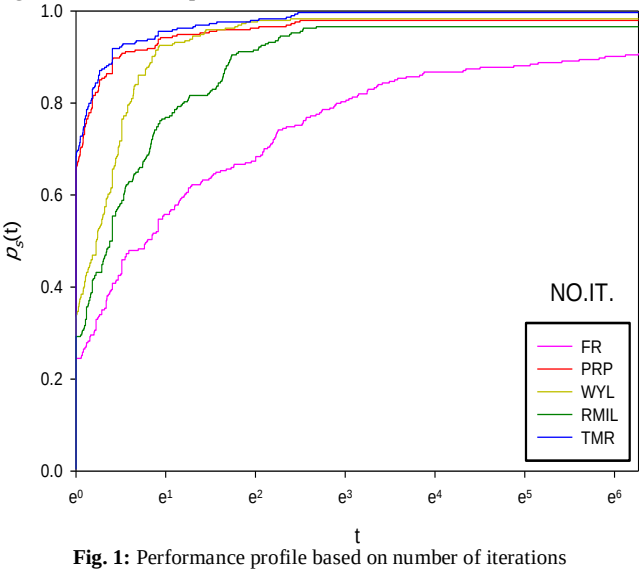


Fig. 1: Performance profile based on number of iterations

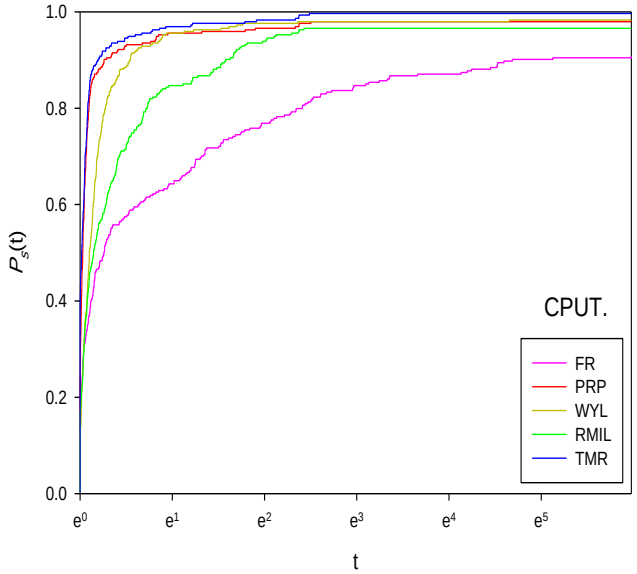


Fig. 2: Performance profile based on CPU time

Figures show the performance profile of the five methods we used based on number of iterations and CPU time.
Both figures depict that TMR outperforms on the other methods, since it can solve all of the test problems and reach 100% percent-

age. In contrast, FR, PRP, WYL and RMIL only reach 90%, 97%, 98% and 96% respectively in solving the given test problems.
It can be seen that TMR possesses restart properties and clearly dominates the other methods. To sum up, our numerical results suggest a new efficient conjugate gradient method. The applications of CG can be refer to [30, 31].

5. Conclusion and further research

Even though, the CG methods are appropriate and efficient for solving nonlinear optimization problems. Many studies on CG methods have led to a new variety of these methods. Meanwhile, their performance has been shown to be slightly better than the previous methods; they seem to be more difficult and complicated. In this paper, we have proposed a new and simple β_k. Besides that, our numerical results show that the TMR method is competitive to the standard conjugate gradient methods and possesses good convergence properties with exact line search. For further research, we intend to study the convergence of the TMR using other line search.

References

[1] Stiefel, E. "Methods of conjugate gradient for solving linear equation." *J. Res. Nat. Bur. Standards* 49 (1952): 409-436.
[2] Fletcher, Reeves, and Colin M. Reeves. "Function minimization by conjugate gradients." *Computer Journal* 7(2), (1964): 149-154.
[3] Polyak, Boris T. "The conjugate gradient method in extremal problems." *USSR Computational Mathematics and Mathematical Physics* 9(4), (1969): 94-112.
[4] Polak, Elijah, and Gerard Ribiere. "Note sur la convergence de méthodes de directions conjuguées." *Revue française d'informatique et de recherche opérationnelle. Série rouge* 3(16), (1969): 35-43.
[5] R.Fletcher, practical method of optimization, Unconstrained optimization, Wiley, New York, (1997): 214-242.
[6] Liu, Y., and C. Storey. "Efficient generalized conjugate gradient algorithms, part 1: Theory." *Journal of Optimization Theory and Applications* 69(1) (1991): 129-137.
[7] Dai, Yu-Hong, and Yaxiang Yuan. "A nonlinear conjugate gradient method with a strong global convergence property." *SIAM Journal on Optimization* 10, 1 (1999): 177-182
[8] Zoutendijk, G. "Nonlinear programming, computational methods." *Integer and Nonlinear Programming* (1970): 37-86.
[9] Powell, Michael James David. "Restart procedures for the conjugate gradient method." *Mathematical Programming* 12, 1 (1977): 241-254.
[10] Powell, Michael JD. "Nonconvex minimization calculations and the conjugate gradient method." In *Numerical Analysis*, pp. Springer, Berlin, Heidelberg, (1984). 122-141.
[11] Powell, M. JD. "Convergence properties of algorithms for nonlinear optimization." *Siam Review* 28(4), (1986): 487-500.
[12] Wei, Zengxin, Guoyin Li, and Liqun Qi. "New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems." *Applied Mathematics and Computation* 179(2), (2006): 407-430.
[13] Dai, Zhi-Feng. "Two modified HS type conjugate gradient methods for unconstrained optimization problems." *Nonlinear Analysis: Theory, Methods and Applications* 74(3) (2011): 927-936.
[14] Rivaie, Mohd, Mustafa Mamat, Leong Wah June, and Ismail Mohd. "A new class of nonlinear conjugate gradient coefficients with global convergence properties." *Applied Mathematics and Computation* 218(22), (2012): 11323-11332.
[15] Aini, N., Rivaie, M., and Mamat, M. "A modified conjugate gradient coefficient with inexact line search for unconstrained optimization". *AIP Conference Proceedings*, 1739(1), 020082, 2016.
[16] Kamfa, Kamilu Uba, Mustafa Mamat, Abdelrhaman Abashar, Mohd Rivaie, Puspa Liza Binti Ghazali, and Zabidin Salleh. "Another modified conjugate gradient coefficient with global convergence properties." (2017). 9(37): 1833–1844.
[17] Shapiee, Norrlaili, Mohd Rivaie, and Mustafa Mamat. "A new classical conjugate gradient coefficient with exact line search." *AIP Conference Proceedings*, 1739(1), 020082, 2016.
[18] Hamoda, M., M. Rivaie, M. Mamat, and Z. Salleh. "A new nonlinear conjugate gradient coefficient for unconstrained optimization." *Applied Mathematical Sciences* 9 (2015): 1813-1822.
[19] Al-Baali, Mehiddin. "Descent property and global convergence of the Fletcher—Reeves method with inexact line searches." *IMA Journal of Numerical Analysis* 5(1), (1985): 121-124
[20] Li, Min, and Heying Feng. "A sufficient descent LS conjugate gradient method for unconstrained optimization problems." *Applied Mathematics and Computation* 218(5), (2011): 1577-1586.
[21] N. Andrei, "40 conjugate gradients algorithms for unconstrained optimization", *Bull. Malay. Math. Sci. Soc.* 34 (2011): 319–330.
[22] Wei, Zengxin, Shengwei Yao, and Liying Liu. "The convergence properties of some new conjugate gradient methods." *Applied Mathematics and computation* 183(2), (2006): 1341-1350.
[23] Zhang, Yueqin, Hao Zheng, and Chuanlin Zhang. "Global convergence of a modified PRP conjugate gradient method." *Procedia Engineering* 31 (2012): 986-995.
[24] Dai, Zhifeng, and Fenghua Wen. "Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property." *Applied Mathematics and Computation* 218, 14 (2012): 7421-7430.
[25] Dai, Yu-Hong, and Yaxiang Yuan. "Nonlinear conjugate gradient methods." *Shanghai Science and Technology Publisher, Shanghai* (2000).
[26] Yuan, Yaxiang, and Wenyu Sun. "Theory and methods of optimization." (1999).
[27] Dolan, Elizabeth D., and Jorge J. Moré. "Benchmarking optimization software with performance profiles." *Mathematical Programming* 91(2), (2002):201-213.
[28] Kenneth E. "A simulation test approach to the evaluation of nonlinear optimization algorithms." *ACM Transactions on Mathematical Software*, 3(4), (1977): 305-315.
[29] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, 10 (2008), 147-161.
[30] Mamat, M., Deraman, S.K., Noor, N.M.M., Rokhayati, Y. Diet problem and nutrient requirement using fuzzy linear programming approach (2012*) Asian Journal of Applied Sciences*, 5(1): 52-59.
[31] Abidin, N.Z., Mamat, M., Ganderfield, B., Zulkefli, J.H., Baten, Md. A., Wibowo, A. Combating obesity through healthy eating behavior: A call for system dynamics optimization, *Plos One*, 9(12), e114135.