ASM Sci. J., 12, Special Issue 6, 2019 for SKSM26, 1-8

A New Modification of Conjugate Gradient Method with Global Convergence and Sufficient Descent Properties for Unconstrained Optimization Problems

Mouiyad Bani Yousef[1], Mustafa Mamat[1] and Mohd Rivaie[2]

[1]*Department of Mathematics, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia*
[2]*Department of Computer Sciences and Mathematics, Universiti Technology MARA (UiTM) Terengganu, Campus Kuala Terengganu, Malaysia*

Conjugate gradient method (CG) is an important method for solving nonlinear unconstrained optimization problems. They are well-known for their global convergence properties and low memory requirement. This paper presents a modified CG method that is globally convergent under strong Wolfe-Powell (SWP) line search. The numerical results show that the new modification is more efficient than other CG methods tested.

Keywords: conjugate gradient method (CG); unconstrained optimization; large-scale optimization; sufficient descent property; Global convergence; strong Wolfe-Powell (SWP) line search

I. INTRODUCTION

The general unconstrained optimization problem is defined by the following,

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

where $f : R^n \to R$ is a continuously differentiable function. The CG method is a type of iterative algorithm that generates a sequence x_k by

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

where x_k is the current iteration point and x_{k+1} is the next iteration point. The parameter $\alpha_k > 0$ is the step length obtained by SWP line search conditions as follows:

$$f\left(x_k + \alpha_k d_k\right) - f\left(x_k\right) \le \delta \alpha_k\, g_k^T d_k, \quad (3)$$

$$|\, g\,(x_k + \alpha_k d_k)^T d_k \,| \le \sigma |g_k^T g_k|, \quad (4)$$

where $0 < \delta < \sigma < 1$, $g_k = g\left(x_k\right) = \nabla f(x)$ is the gradient of the function and d_k is the search direction. The formula of d_k for CG method is defined by

$$d_k = \begin{cases} -g_k & , \ k = 0 \\ -g_k + \beta_k d_{k-1} & , \ k \ge 1. \end{cases} \quad (5)$$

The parameter $\beta_k \in R$ is known as the CG coefficient that characterizes different CG methods. Examples of formulas for β_k include the Polak-Ribiere-Polyak (PRP) method (Polyak, 1969; Polak, 1969), the Fletcher-Reeves (FR) method(Fletcher *et al.*, 1964), the Hestenes-Steifel (HS) method (Hestenes & Stiefel, 1952), the 'Aini Rivaie-Mustafa (ARM) method (Aini, *et al.*, 2016), the Liu Storey (LS) method (Liu & Storey, 1991), the Conjugate Descent (CD) method

(Fletcher, 2013) and the Wei-Yao-Liu (WYL) method (Wei *et al.*, 2006). They are formulated as follows:

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

$$\beta_K^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}; \quad (7)$$

*Corresponding author's e-mail: md.magableh@yahoo.com

$$\beta_K^{HS} = \frac{g_k\left(g_k g_{k-1}\right)}{\left(g_k g_{k-1}\right)^T d_{k-1}}; \quad (8)$$

$$\beta_k^{ARM} = -\frac{m_k \|g_k\|^2 - \left|g_k^T g_{K-1}\right|}{m_k g_k^T g_{K-1}}$$, where

$$m_k = \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|} \quad (9)$$

$$\beta_K^{LS} = \frac{g_k^T\left(g_k - g_{k-1}\right)}{-d_{k-1} g_{k-1}}; \quad (10)$$

$$\beta_K^{CD} = \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}; \quad (11)$$

$$\beta_k^{WYL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2}, \quad (12)$$

where $\|.\|$ represents the Euclidean norm of vectors.

There are numerous studies centered on the convergence properties of CG methods, beginning with Zoutendijk (Zoutendijk, 1970) who proved the global convergence of FR method under exact line search. Then, Al-Baali (1985), Touti-Ahmed and Storey (1990) furthered the proof for Wolfe line search. Gilbert and Nocedal (1992) showed the global convergence of algorithm related to the FR method using the inexact line search with strong Wolfe conditions. Later, Al-Baali (1985)

extended the proof for $\sigma < 1/2$, while Guanghui *et al.*, (1995) proved the global convergence of FR for $\sigma \le 1/2$.

Up to now, many researchers have come up with new formulas for CG coefficient with good numerical performance and globally convergent. Some current studies on nonlinear CG method can be seen in Hager and Zhang (2005), Andrei (2009), Wei *et al.* (2013), Abashar *et al.* (2017), Jusoh *et al.* (2014), and Rivaie *et al.* (2012; 2016).

This study is prepared as follows:

The first part consists of the new CG formula, followed by the CG algorithm and the convergence analysis. After that, the numerical results of the proposed method are presented and compared with other CG methods. Finally, this paper is concluded in the last section.

II. NEW FORMULA FOR CG METHOD

In this section, we propose a new formula for β_k known as β_k^{MMR}, where MMR denote Mouiyad, Mustafa and Rivaie. The new CG coefficient β_k^{MMR} is constructed by applying an alternative denominator in β_k^{AMR}. The formula is as follows:

$$\beta_k^{MMR} = \frac{m_k \|g_k\|^2 - \left(g_k^T g_{K-1}\right)}{m_k \|g_{k-1}\|^2}$$, where

$$m_k = \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|} \quad (13)$$

Note that

$$0 \le \frac{\|g_k\|^2 \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|} - \left(g_k^T g_{K-1}\right)}{\|g_{k-1}\|^2 \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|}} < \frac{\|g_k\|^2 \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|}}{\|g_{k-1}\|^2 \frac{\|d_{k-1} + g_k\|}{\|d_{k-1}\|}}$$

$$\le \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_k^{FR} \quad (14)$$

Therefore

$$0 \le \beta_k^{MMR} \le \beta_k^{FR}. \quad (15)$$

Hence, according to the argument presented in (Gilbert & Nocedal, 1992), β_k^{MMR} should obtain all of the advantages and properties of β_k^{FR}.

The algorithm is given as follows:

Algorithm 1

Step 1: Given $x_0 \in R^n$, Select $\varepsilon \ge 0$ and, $d_0 = -g_0$, $k = 0$. If $\|g_0\| \le \varepsilon$ then stop.

Step2: Compute α_k by (SWP) line search. (5) and (6)

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

Step 4: Compute β_k by (6) and generate d_{k+1} by (4).

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

III. CONVERGENCE ANALYSIS OF MMR METHOD

The following basic assumption are always necessary in the analysis to present the global convergence properties of CG methods with (SWP) line search.

Assumption 1.

(i) $f(x)$ is bounded from below on the level set $\Omega = \{ x \in R^n / f(x) \le f(x_0) \}$.

(ii)In some neighborhood N of Ω, f is continuously differentiable and its gradient $g(x)$ is Lipschitz continuous, hence there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \le L\|x - y\| \quad x, y \in N \text{ .(16)}$$

In 1992, Gilbert and Nocedal introduced the property (*) which plays an important role in the studies of CG methods. This property means that the next research direction approaches to the steepest direction automatically when a small step-size generated, and the step-sizes are not produced successively (Zhang *et al.*, 2012).

Property (*). Consider a CG method of the form (2) and (5). Suppose that, for all $k \ge 0$.

$$0 < \gamma \le \|g(x)\| \le \bar{\gamma} \text{ (17)}$$

Where γ and $\bar{\gamma}$ are two positive constants. We say that the method has the property (*) if
there exist constants $b > 1, \lambda > 1$ such that for

all k ,$|\beta_k| \le b$, $\|S_k\| \le \lambda$ implies $|g_k| \le \dfrac{1}{2b}$ where $S_k = \alpha_k \beta_k$.

The following lemma shows that the new method β_k^{MMR} has the property (*).

Lemma 1. Consider the method of form (2) and (5), suppose that Assumptions1 hold, then, the method β_k^{MMR} has the property (*).

Proof. Set $b = \dfrac{\bar{\gamma}^2(\gamma+\bar{\gamma})}{\gamma^3} > 1$, $\lambda = \dfrac{\gamma^2}{4L\bar{\gamma}b}$ by (13) and (17)

we have

$$\left|\beta_k^{MMR}\right| \le \frac{\|g_k\|^2 \frac{\|d_{k-1}+g_k\|}{\|d_{k-1}\|} - \left(g_k^T g_{K-1}\right)}{\|g_{k-1}\|^2 \frac{\|d_{k-1}+g_k\|}{\|d_{k-1}\|}} \le \frac{\|g_k\|^2 - \left(g_k^T g_{K-1}\right)}{\|g_{k-1}\|^2}$$

$$\le \frac{\left|\|g_k\|^2 - \left(g_k^T g_{K-1}\right)\right|}{\|g_{k-1}\|^2}$$

$$\le \frac{\|g_k\|^2 + \|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^2} \le \frac{\bar{\gamma}^2 + (\gamma\bar{\gamma})}{\gamma^2} = \frac{\gamma\bar{\gamma}^2 + \gamma\bar{\gamma}^2}{\gamma^3}$$

$$= \frac{\bar{\gamma}^2(\gamma + \bar{\gamma})}{\gamma^3} = b$$

From the Assumption ii, (16) holds. If $\|S_k\| \le \lambda$ then,

$$\left|\beta_k^{MMR}\right| \le \frac{\left|\|g_k\|^2 - \left(g_k^T g_{K-1}\right)\right|}{\|g_{k-1}\|^2} \le \frac{\left|g_k\left(g_k^T - g_{k-1}\right)\right|}{\|g_{k-1}\|^2} \le \frac{2L\lambda\|\bar{\gamma}\|}{\gamma^2} = \frac{1}{2b}$$

The proof is finished.

IV. SUFFICIENT DESCENT

The sufficient descent condition should be satisfied as follows:

$$g_k^T d_k \le -c\|g_k\|^2 \text{ for } k \ge 0 \text{ and } c > 0 \text{ .(18)}$$

Note that the CG coefficient, β_k^{MMR} satisfies

$$0 \le \beta_k^{MMR} \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

$$0 \le 0 \le \beta_{k+1}^{MMR} \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_k^{FR} \quad (19)$$

The following theorem shows that the formula MMR with SWP line search possess the sufficient descent condition

Theorem 1

Suppose that the sequences $\{g_k\}$ and $\{d_k\}$ are generated by the method of form (2), (5) and (13), and the step length α_k is determined by the (SWP) line search (3) and (4), then the sequence $\{d_k\}$ possesses the sufficient descent condition (18).

Proof. By the formulae (13), we have

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

$$\le \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

Hence, we obtain

$$0 \le \beta_k^{MMR} \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

Using (4) and (19), we get

$$\left|\beta_{k+1}^{MMR} g_{k+1} d_k\right| \le \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \sigma \left|g_k d_k\right| \quad (20)$$

By (5), we have $d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \quad (21)$$

We prove the descent property of $\{d_k\}$ by induction. Since

$g_0^T d_0 = -\|g_0\|^2 < 0$ if $g_0 \ne 0$,

Now suppose d_i $i = 1, 2, ..., k$ are all descent directions,

that is $g_i^T d_i < 0$

By (20), we get

$$\left|\beta_{k+1}^{MMR} g_{k+1} d_k\right| \le \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \sigma(-g_k d_k) \quad (22)$$

That is,

$$\frac{\|g_{k+1}\|^2}{\|g_k\|^2} \sigma g_k^T d_k \le \beta_{k+1}^{PRP} g_{k+1} d_k \le -\frac{\|g_{k+1}\|^2}{\|g_k\|^2} \sigma g_k^T d_k \quad (23)$$

(21) and (23) deduce,

$$-1 + \frac{\sigma g_k^T d_k}{\|g_k\|^2} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1 - \frac{\sigma g_k^T d_k}{\|g_k\|^2}$$

By repeating this process and the fact

$g_0^T d_0 = -\|g_0\|^2$ we have,

$$-\sum_{j=0}^{k} (\sigma)^j \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \sum_{j=0}^{k} (\sigma)^j \quad (24)$$

Since

$$\sum_{j=0}^{k} (\sigma)^j < \sum_{j=0}^{\infty} (\sigma)^j = \frac{1}{1-\sigma}$$

(24) Can be written as

$$-\frac{1}{1-\sigma} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \frac{1}{1-\sigma} \quad (25)$$

By making the restriction $\sigma \in (0, \frac{1}{4})$ we have

$g_{k+1}^T d_{k+1} < 0$.

So, by induction, $g_k^T d_k < 0$ holds for all $k \ge 0$.

Denote $c = 2 - \frac{1}{1-\sigma}$ then, $0 < c < 1$, and (25) turns out to be

$$(c-2)\|g_k\|^2 \le g_k d_k \le -c\|g_k\|^2 \quad (26)$$

this implies that (18) holds. The proof is complete.

V. GLOBAL CONVERGENCE

The following condition known as Zoutendijk condition is used to prove the global convergence of nonlinear CG methods (Zoutendijk, 1970; Wolfe, 1969).

Lemma 2. Suppose that Assumptions 1 hold. Consider a CG method of the form (2) and (5), where d_k satisfies $g_k^T d_k < 0$, for all k, and α_k is obtained by (SWP) line search (3) and (4),

Then,

$$\sum_{K=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \quad (27)$$

The proof had been given by Wei *et al.* 2006 and Yuan *et al.* 2010. Gilbert and Nocedal (1992) introduced the following important theorem.

Theorem2

Consider any CG method of form (2) and (5), that satisfies the following conditions:

(1) $\beta_k \geq 0$

(2) The search directions satisfy the sufficient descent condition.

(3) The Zoutendijk condition holds.

(4) Property (*) holds. If the Lipschitz and boundedness Assumptions hold, then the iterates are globally convergent.

From (16), (18), (27) and Lemma 1, we found that the MMR method with the parameter $0 < \delta < \sigma < 1/4$ satisfies all four conditions in theorem 2 under the strong Wolfe-Powell line search, so the method is globally convergent.

VI. NUMERICAL RESULTS

In this section, we compare the computational performance of MMR method with PRP, FR, WYL and ARM under SWP line search. These comparisons are based on number of iteration and CPU time. To perform the test, a set of twenty-five test functions with varying number of variables $(2 \leq n \leq 10000)$ have been selected from Andrei (2008).

For each test problems, we take four different initial points in order to study the global convergence properties of the new CG formula. All functions and initial points used are listed in Table 1.

Table 1. A list of problem functions

No Functions		Initial points
1 Six Hummp	2	(-3, -3) (21,21) (-43, -43) (77,77)
2 Three Hump	2	(21,21) (25,25) (71,71) (67,67)
3 Zettl	2	(7,7) (18,18) (-116, -116)
4 FELETCHCR	2	(-2, -27) (-112, -112) (101,101) (17,17)
5 Colivlle	4	(28,..,28) (199,..,199)
6 Dixon and Price	2,4	(12,..,12) (19,..,19) (-108,..,-108)
7 Hager	2,4	(6,...,6) (16,...,16) (-78,...,-78)
8Raydan1	2,4	(7,..,7) (18,..,18) (-89,.., -89)
9Raydan 2	2,4	(-7,.., -7) (-75,.., -75)
10ARWAHEAD	2,4,10	(4,...,4) (21,...,21) (80,...,80)
11 Freuidenstein and Roth	2,4,10	(-19,..., -19) (18,..,18) (4.5,..,4.5)
12 Extended Maratos	2,4,10	(-4,..., - 4) (5,....,5) (18,..,18) (-84,.., -84)
13 Generalized Tridiagonal 1	2,4,10	(3,...,3) (14,...,14) (70,...,70)
14 Generalized Quartic 1	2,4,10,100,500,1000	(11,,...,11) (29,...,29) (87,...,87) (-80,...,-80)
15 Extended Beale	2,4,10,100,500,1000	(107,...,107) (-1. 3,..., -1.3) (72,...,72)
16 Extended	2,4,10,100,500,1000	(5,..,5)

	DENSCHNB		(25,...,25) (200,..,200)
17	Extended Himmelblau	2,4,10,100,500,1000	(40,...,40) (17.8,...,17.8) (115,...,115)
18	Extended Rosenbrock	2,4,10,100,500,100	(10,...,10) (18,...,18) (-111,..., -111)
19	Extended Shallow	2,4,10,100,500,1000	(11,...,11) (41,...,41) (120,...,120)
20	Extended Strait	2,4,10,100,500,1000	(4,...,4) (236,...,236) (136,...,136)
21	Extende Tridiagonal	2,4,10,100,500,1000	(13,..,13,) (24.7,..,24.7)(120,..,120)
22	Sum Sqaures	2,4,10,100,500,1000	(5,...,5) (-3.8,...,-3.8) (22,...,22)
23	QUARTC	2,4,10,100,500,1000	(5,...,5) (17,...,17) (97,...,97)
24	Both	2,4,10,100,500,1000	(9,...,9) (-9,..., -9) (11,...,11) (-11,..., -11)
25	Extended White and Holst	2,4,10,100,500, 1000	6,...,6) (9,...,9) (3,...,3) (-3,...,-3)

The algorithm is set to stop when $\|g_k\| \le 10^{-6}$. All codes are written in MATLAB version R2015a subroutine programming and run on a PC computer with Intel(R) Core™ i3-4005U CPU @ 1.70GHz processor, 4GB RAM and Windows 10 Professional operating system. For result analysis, we use Sigma Plot 10 program to graph the data based on the performance profile proposed by Dolan and More (2002).

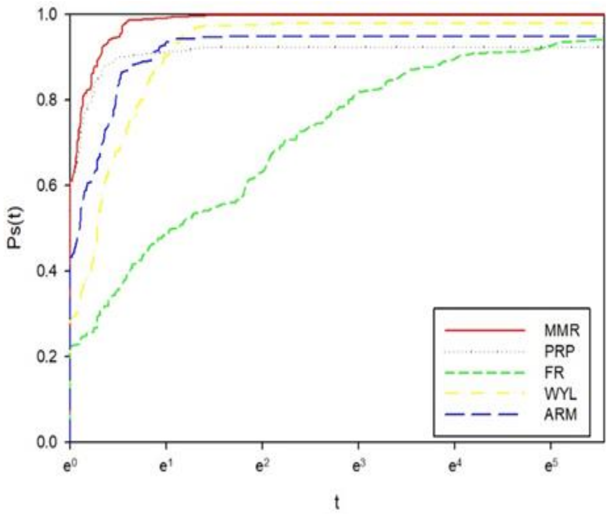

Figure 1. Number of iterations Performance profile graph

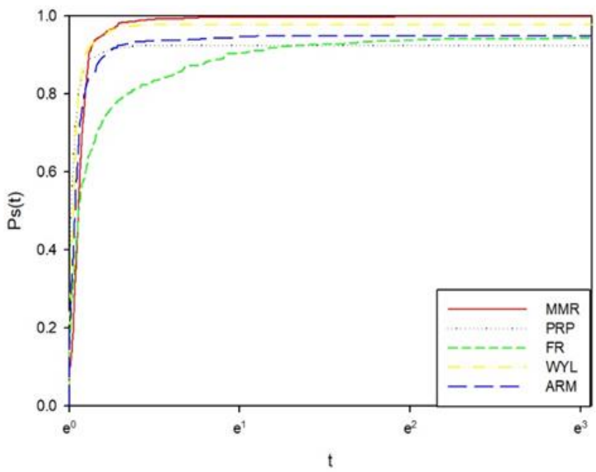

Figure 2. CPU time Performance profile graph

Figures 1 and 2 show that MMR has the best efficiency in terms of number of iteration and CPU time and plots the performance of our new method with respect to other four other CG methods, PRP, FR, WYL, and AMR. The curve at the top left signifies the fastest solver, while the curve at the top left is the most robust. The best method should be the one at the top left and right of the performance profile.

In Figure 1, the curve for MMR method is at the top of other curves, while in Figure 2, it is mostly at the top. In addition, MMR manages to solve 100% of the test problems. Compared to that, the PRP, FR, WYL, and ARM methods only solve 92%, 94%, 97%, and 94% of the test problems, respectively. Hence, our new formula is the most robust amongst all the methods tested.

VII. CONCLUSION

In this paper, we have considered a new CG method called the MMR method. We proved that it satisfies the sufficient descent condition and possesses global convergence properties when used with SWP line search. The numerical results demonstrate promising results for the MMR method, in which it has shown higher efficiency and robustness than other tested CG parameters.

VIII. ACKNOWLEDGEMENTS

The authors are grateful to Universiti Sultan Zainal Abidin, for financial support under the fundamental research Grant Scheme (FRGS)/1/2017/STG06/UniSZA/01/1.

IX. REFERENCES

Abashar, A., Mamat, M., Rivaie, M.&Mohd, I. (2017). Global convergence properties of a new class of conjugate gradient method for unconstrained optimization.

Aini, N., Rivaie, M. &Mamat, M. (2016). A modified conjugate gradient coefficient with inexact line search for unconstrained optimization. In AIP Conference Proceedings (Vol. 1787, No. 1, p. 080019). AIP Publishing.

Al-Baali, M. (1985). Descent property and global convergence of the Fletcher—Reeves method with inexact line search. IMA Journal of Numerical Analysis, 5(1), pp. 121-124.

Andrei, N. (2008). An unconstrained optimization test functions collection. Adv. Model. Optim, 10(1), pp. 147-161.

Andrei, N. (2009). Accelerated conjugate gradient algorithm with finite difference Hessian/vector product approximation for unconstrained optimization. Journal of Computational and Applied Mathematics, 230(2), pp. 570-582.

Dolan, E. D.&Moré, J. J. (2002). Benchmarking optimization software with performance profiles. Mathematical programming, 91(2), pp. 201-213.

Fletcher, R. & Reeves, C. M. (1964). Function minimization by conjugate gradients. The computer journal, 7(2), pp. 149-154.

Fletcher, R. (2013). Practical methods of optimization. John Wiley & Sons.

Gilbert, J. C.&Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. SIAM Journal on optimization, 2(1), pp. 21-42.

Guanghui, L., Jiye, H.&Hongxia, Y. (1995). Global convergence of the Fletcher-Reeves algorithm with inexact linesearch. Applied Mathematics-A Journal of Chinese Universities, 10(1), pp. 75-82.

Hager, W. W.& Zhang, H. (2005). A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM Journal on optimization, 16(1), pp. 170-192.

Hestenes, M. R.&Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems (Vol. 49, No. 1). Washington, DC: NBS.

Jusoh, I., Mamat, M.&Rivaie, M. (2014). A new edition of conjugate gradient methods for large-scale unconstrained optimization. Intl. J. Math. Anal, 8, pp. 2277-2291.

Liu, Y.& Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: theory. Journal of optimization theory and applications, 69(1), pp. 129-137.

Polyak, B. T. (1969). The conjugate gradient method in extremal problems. USSR Computational Mathematics and Mathematical Physics, 9(4), pp. 94-112.

Polak, E.&Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. Revue françaised'informatiqueet de rechercheopérationnelle. Série rouge, 3(16), pp. 35-43.

Rivaie, M.,Mamat, M., Leong W. J. & Ismail, M. (2012). A new class of nonlinear Conjugate gradient coefficients with global convergence properties, Applied Mathematics and Computation, 218, pp. 11323-11332.

Rivaie, M., Abashar, A., Mamat, M.&Mohd, I. (2016). The convergence properties of a new type of conjugate gradient methods.

Touati-Ahmed, D.& Storey, C. (1990). Efficient hybrid conjugate gradient techniques. Journal of optimization theory and applications, 64(2), pp. 379-397.

Wei, Z., Yao, S.& Liu, L. (2006). The convergence properties of some new conjugate gradient methods. Applied Mathematics and computation, 183(2), pp. 1341-135.

Wolfe, P. (1969). Convergence conditions for ascent methods. SIAM review, 11(2), pp. 226-235.

Yuan, G., Lu, S.& Wei, Z. (2010). A line search algorithm for unconstrained optimization. Journal of Software Engineering and Applications, 3(05), 503.

Zhang, Y., Zheng, H.& Zhang, C. (2012). Global convergence of a modified PRP conjugate gradient method. Procedia Engineering, 31, pp. 986-995.

Zoutendijk, G. (1970). Nonlinear programming, computational methods. Integer and nonlinear programming, pp. 37-86.