Applied Mathematical Sciences, Vol. 10, 2016, no. 15, 721 - 734
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2016.56449

A Conjugate Gradient Method with Strong Wolfe-Powell Line Search for Unconstrained Optimization

Mohamed Hamoda [1], Mustafa Mamat [2], Mohd Rivaie [3] and Zabidin Salleh [1]

[1] School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu (UMT), 21030 Kuala Terengganu, Malaysia

[2] Department of Computer Science and Mathematics, Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, 22200 Terengganu, Malaysia

[3] Department of Computer Science and Mathematics, Univesiti Teknologi MARA (UITM), 23000 Terengganu, Malaysia

Abstract

In this paper, a modified conjugate gradient method is presented for solving large-scale unconstrained optimization problems, which possesses the sufficient descent property with Strong Wolfe-Powell line search. A global convergence result was proved when the (SWP) line search was used under some conditions. Computational results for a set consisting of 138 unconstrained optimization test problems showed that this new conjugate gradient algorithm seems to converge more stable and is superior to other similar methods in many situations.

Keywords: conjugate gradient coefficient, inexact line search, strong Wolfe–Powell line search, global convergence, large scale, unconstrained optimization

1. Introduction

Nonlinear conjugate gradient methods are well suited for large-scale problems due to the simplicity of their iteration and their very low memory requirements, that is, they are designed to solve the following unconstrained optimization problem:

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

where $f : R^n \to R$ is a smooth, nonlinear function, and its gradient is denoted by $g(x) = \nabla f(x)$. The iterative formula of the conjugate gradient methods is given by

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

where x_k is the current iterate point and α_k is the step length, which is computed by carrying out a line search, and d_k is the search direction defined by

$$d_k = \begin{cases} -g_k & \text{if} \quad k = 0, \\ -g_k + \beta_k d_{k-1} & \text{if} \quad k \geq 1, \end{cases} \tag{3}$$

where β_k is a scalar, and $g_k = g(x_k)$.

Various conjugate gradient methods have been proposed, and they mainly differ in the choice of the parameter β_k. Some well-known formulas for β_k are given below:

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, \quad \beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}, \quad \beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}, \quad \beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}},$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}}, \qquad \beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}}$$

where $\|.\|$ denotes the l_2 -norm. The corresponding method is respectively called, HS (Hestenes-Stiefel [1]), FR (Fletcher_Revees [2]), PRP (Polak_Ribiére_Polyak [3, 4]), CD (Conjugate Descent [5]), LS (Liu-Storey [6]), and DY (Dai_Yuan [7]) conjugate gradient method. The convergence behavior of the above formulas with some line search conditions has been studied by many authors for many years [5-17].

In the already-existing convergence analysis and implementations of the conjugate gradient method, the weak Wolfe–Powell (WWP) line search conditions are as follows:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \tag{4}$$

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k \tag{5}$$

where $0 < \delta < \sigma < 1$ and d_k is a descent direction.

The strong Wolfe–Powell conditions consist of (4) and,

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \leq \sigma \left| g_k^T d_k \right| \tag{6}$$

Furthermore, the sufficient descent property, namely,

$$g_k^T d_k \leq -c \|g_k\|^2 \tag{7}$$

Where c is, a positive constant, is crucial to ensure the global convergence of the nonlinear conjugate gradient method with the inexact line search techniques [12, 13].

2. New formula for β_k and its properties

During the last decade, much effort has been devoted to developing new modifications of conjugate gradient methods which do not only possess strong convergence properties, but they are also computationally superior to the classical methods. Such methods can be found in [18-30].

Recently, Wei et al. [31] gave a variant of the *PRP* method which is called the *WYL* method. Zhang studied and improved based on *WYL* a new conjugate gradient method, *NPRP*, and he proved that the *NPRP* method satisfied descent condition under strong Wolfe line search. Moreover, Zhang et al. proposed another modified method known as the *MPRP* method, where Dai and Wen [32] proposed a modified *NPRP* method known as the *DPRP* method. In this paper, enlightened by the above ideas, a modified *PRP* conjugate gradient method was proposed as follows:

$$\beta_k^{HRM} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} \tag{8}$$

where, HRM denotes Hamoda, Rivaie, and Mamat. According to the results obtained by [30], the value of the parameter u can be set to $0 < u < 1$, but in this paper, we will test our new method with an arbitrary value $u = 0.4$

Therefore, we first provide the following algorithm:

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 (SWP) line search.

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 formula (8) and generate d_{k+1} by (3).

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

The following assumptions are often used in previous studies of the conjugate gradient methods:

Assumption A

$f(x)$ is bounded from below on the level set $\Omega = \{x \in R^n, f(x) \leq f(x_0)\}$, where x_0 is the starting point.

Assumption B

In some neighborhood N of Ω , the objective function is continuously differentiable, and its gradient is Lipschitz continuous, that is, there exists a constant $L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\| \; \forall x, y \in N . \tag{9}$$

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 the steepest direction automatically when a small step-size is generated, and the step-sizes are not produced successively [33].

Property (*)
Consider a CG method of the form (2) and (3). Suppose that, for all $k \geq 0$,

$$0 < \gamma \leq \|g_k\| \leq \bar{\gamma} \tag{10}$$

where γ and $\bar{\gamma}$ are two positive constants. We say that the method has property (*), if there exist constants $b > 1$, $\lambda > 0$ such that for all k, $|\beta_k| \leq b$, $\|S_k\| \leq \lambda$ implies $|\beta_k| \leq \frac{1}{2b}$, where $S_k = \alpha_k d_k$.

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

Lemma 2.1
Consider the method of form (2) and (3), Suppose that Assumptions A and B hold, then, the method β_k^{HRM} has property (*).

Proof

Set $b = \frac{5\bar{\gamma}^2(\gamma + \bar{\gamma})}{2\gamma^3} > 1$, $\lambda = \frac{\gamma^2}{10L\bar{\gamma}b}$. By (8) and (10) we have

$$\left|\beta_k^{HRM}\right| \leq \frac{\left|g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right)\right|}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} \leq \frac{\|g_k\|(\|g_k\| + \frac{\bar{\gamma}}{\gamma}\|g_{k-1}\|)}{0.4\|g_{k-1}\|^2} \leq \frac{5\bar{\gamma}(\bar{\gamma} + \frac{\bar{\gamma}^2}{\gamma})}{2\gamma^2} = \frac{5\bar{\gamma}^2(\gamma + \bar{\gamma})}{2\gamma^3} = b$$

From Assumption B, (9) holds. If $\|S_k\| \leq \lambda$ then,

$$\left|\beta_k^{HRM}\right| \leq \frac{\left(\|g_k - g_{k-1}\| + \left\|g_{k-1} - \frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1}\right\|\right)\|g_k\|}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} \leq \frac{(L\lambda + \|g_{k-1}\| - \|g_k\|)\|g_k\|}{u\|g_{k-1}\|^2} \leq \frac{(L\lambda + \|g_k - g_{k-1}\|)\|g_k\|}{u\|g_{k-1}\|^2}$$
$$\leq \frac{2L\lambda\|g_k\|}{0.4\|g_{k-1}\|^2} \leq \frac{5L\lambda\bar{\gamma}}{\gamma^2} = \frac{1}{2b}$$

The proof is finished.

3. The global convergence properties

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

Theorem 3.1
Suppose that the sequences $\{g_k\}$ and $\{d_k\}$ are generated by the method of form (2), (3) and (8), and the step length α_k is determined by the (SWP) line search (4) and

(6), if $g_k \neq 0$, then the sequence $\{d_k\}$ possesses the sufficient descent condition (7).

Proof

By the formulae (8), we have the following:

$$\beta_k^{HRM} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{u\|g_{k-1}\|^2 + (1-u)\|d_{k-1}\|^2} = 0$$

Thus we get, $\beta_k^{HRM} \geq 0$

Also

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

Hence, we obtain

$$0 \leq \beta_k^{HRM} \leq \frac{5\|g_k\|^2}{\|g_{k-1}\|^2} \tag{11}$$

Using (6) and (11), we get

$$\left|\beta_{k+1}^{HRM} g_{k+1}^T d_k\right| \leq \frac{5\|g_{k+1}\|^2}{\|g_k\|^2} \sigma \left|g_k^T d_k\right| \tag{12}$$

By (3), we have $d_{k+1} = -g_{k+1} + \beta_{k+1} d_k$

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \tag{13}$$

We prove the descent property of $\{d_k\}$ by induction. Since $g_0^T d_0 = -\|g_0\|^2 < 0$, if $g_0 \neq 0$, now suppose that

$d_i, i = 1,2,....,k,$ are all descent directions, that is $g_i^T d_i < 0$

By (12), we get

$$\left|\beta_{k+1}^{HRM} g_{k+1}^T d_k\right| \leq \frac{5\|g_{k+1}\|^2}{\|g_k\|^2} \sigma(-g_k^T d_k) \tag{14}$$

That is,

$$\frac{\|g_{k+1}\|^2}{\|g_k\|^2} 5\sigma g_k^T d_k \leq \beta_{k+1}^{HRM} g_{k+1}^T d_k \leq -\frac{\|g_{k+1}\|^2}{\|g_k\|^2} 5\sigma g_k^T d_k \tag{15}$$

(13) and (15) deduce,

$$-1+\frac{5\sigma g_k^T d_k}{\|g_k\|^2} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1-\frac{5\sigma g_k^T d_k}{\|g_k\|^2}$$

By repeating this process and the fact $g_0^T d_0 = -\|g_0\|^2$, we have,

$$-\sum_{j=0}^{k}(5\sigma)^j \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2+\sum_{j=0}^{k}(5\sigma)^j \tag{16}$$

Since

$$\sum_{j=0}^{k}(5\sigma)^j < \sum_{j=0}^{\infty}(5\sigma)^j = \frac{1}{1-5\sigma}$$

(16) can be written as

$$-\frac{1}{1-5\sigma} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2+\frac{1}{1-5\sigma} \tag{17}$$

By making the restriction $\sigma \in (0,0.1)$, 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 $c = 2-\frac{1}{1-5\sigma}$ then, $0 < c < 1$, and (17) turns out to be

$$(c-2)\|g_k\|^2 \le g_k^T d_k \le -c\|g_k\|^2 \tag{18}$$

This implies that (7) holds. The proof is complete.

The following condition known as Zoutendijk condition was used to prove the global convergence of nonlinear CG methods [15, 34].

Lemma 3.1

Suppose that Assumptions A and B hold. Consider a CG method of the form (2) and (3), where d_k satisfies $g_k^T d_k < 0$, for all k, and α_k is obtained by (SWP) line search (4) and (6), Then,

$$\sum_{k=0}^{\infty}\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < \infty \tag{19}$$

The proof had been given in [35, 36]. In [10], Gilbert and Nocedal introduced the following important theorem:

Theorem 3.2

Consider any CG method of form (2) and (3), that satisfies the following conditions:

(1) $\beta_k \ge 0$

(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 (7), (9), (17) and Lemma 2.1, we found that the *HRM* method with the parameter $0<\delta<\sigma<1/10$ satisfies all four conditions in theorem 3.2 under the strong Wolfe-Powell line search, so the method is globally convergent.

4. Numerical Experiments

In the present numerical experiments, we selected thirty-two different functions which had been earlier considered in [37-39] for both small-scale and large-scale optimization problems. Each of these functions was tested with different variables which lie in the range from 2 to 10,000. We tested a set of 138 problems with strong Wolfe-Powell line search. The algorithm was implemented using MATLAB R2011b (7.13.0.564), applying the strong Wolfe-Powell line search. All of the numerical experiments were run on the same PC with an Intel (R) CoreTM i3-M350 (2.27GHz) CPU, 4GB of RAM, and the Windows 7 operating system. In order to assess the reliability of the new proposed method, *HRM*, we tested this method against the well-known classical and modified methods of the *FR*, *PRP*, *MPRP*, and *DPRP* methods using the same problems, and assumed that the best method should require fewer iterations and less CPU time. All of these algorithms terminated when $\|g_k\|<10^{-6}$. The step size α_k satisfies the strong Wolfe- Powell conditions, with $\delta=10^{-4}$, and $\sigma=0.001$. For the *MPRP* method, we chose $\mu_1=\mu_3=1$ and $\mu_2=3$, where $\mu=3$ in *DPRP* method. A list of test functions, Dimension, and the initial points used are shown in Table 1. 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 failure. In addition, we considered the search to have failed if the number of iterations exceeded 1,000 or CPU execution time exceeded 500 seconds. Numerical results were relatively compared with the CPU time and number of iterations. The performance results are shown in Figures 1 and 2 respectively, using a performance profile introduced by Dolan and More [40].

Table 1. A list of problem functions

No	**Function**	**Dimension**	**Initial points**
1	Six Hump Camel	2	-10, -8, 8, 10
2	Booth	2	10, 25, 50, 100
3	Treccani	2	5, 10, 20, 50
4	Zettl	2	5, 10, 20, 30
5	Leon	2	2, 5, 8, 10
6	Three Hump	2	20, 50, 60, 150
7	Extended Wood	4	3, 5, 20, 30

8	Quartic	4	5, 10, 15, 20
9	Colville	4	2, 4, 7, 10
10	Extended Maratos	2, 4, 10, 100	1, 5, 8, 10
11	Fletcher	4, 10, 100, 500, 1000	7, 9, 11, 13
12	Perturbed Quadratic	2, 4, 10, 100, 500, 1000	1, 5, 10, 15
13	Extended Himmelblau	100, 500, 1000, 10000	50, 70, 100, 125
14	Extended Rosenbrock	2, 4, 10, 100, 500, 1000, 10000	13, 25, 30, 50
15	Shallow	2, 4, 10, 100, 500, 1000, 10000	10, 25, 50, 70
16	Extended Tridiagonal 1	2, 4, 10,100, 500, 1000, 10000	12, 17, 20, 30
17	Generalized Tridiagonal 1	2, 4,10, 100	25, 30, 35, 50
18	Extended white & Holst	2, 4,10,100, 500, 1000, 10000	3, 10, 30, 50
19	Generalized Quartic	2, 4,10,100, 500, 1000, 10000	1, 2, 3, 5
20	Extended Powell	4, 8, 20, 100, 500, 1000	4, 5, 7, 30
21	Extended Denschnb	2, 4, 10, 100, 500, 1000, 10000	8, 13, 30, 50
22	Hager	2, 4, 10, 100	1, 3, 5, 7
23	Extended Penalty	2, 4, 10, 100	10, 50, 75, 100
24	Quadrtic QF2	2, 4, 10, 100 ,500, 1000	10, 30, 50, 100
25	Extended Quadratic Penalty QP2	2, 4, 10, 100, 500, 1000, 10000	17, 18, 19, 20
26	Extended Beale	2, 4,10, 100, 500, 1000, 10000	1, 3, 13, 30
27	Diagonal 2	2, 4, 10, 100, 500, 1000	-1, 1, 2, 3
28	Raydan1	2, 4, 10,100	1, 3, 5, 7
29	Sum Squares	2, 4, 10,100, 500, 1000	1, 10, 20, 30
30	Generlized Tridiagonal 2	2, 4, 10, 100	1, 10, 20, 30
31	Quadratic QF1	2, 4, 10,100, 500, 1000	1, 2, 3, 4
32	Dixon & Price	2, 4, 10, 100	100, 125, 150, 175

Under a strong Wolfe-Powell line search, the performance profile of all methods measured by the number of iterations required is shown in Figure 1, and the performance profile based on the CPU time used is in Figure 2. The shapes of the profile plots in both Figures 1 and 2 are almost alike. A thorough inspection of the left side of both figures indicates that the lowest curve represents the *FR* method. Therefore, this method possesses the lowest performance. The top left side curve indicates that the *PRP* method is the best performer. The curves for methods *DPRP*, *MPRP*, and *HRM* fall in between the two extreme curves. Thus, the performance of this set of methods is in the middle of the sets based the number of iterations and CPU time.

From results shown in Figures 1 and 2, it is evident that the *FR* method achieved a success rate of only 0.65, while the *PRP* method scored 0.79, and the *DPRP* method recorded 0.88. Furthermore, the *MPRP* method achieved 0.94, and our new method, *HRM* achieved 1, that is, our new method scored a 100% success rate. Such result indicates that *HRM* method is the best among the four methods in the perspectives of the number of iterations and the CPU time. Hence, our new method successfully solved all the test problems, and it is competitive with the well-known conjugate gradient methods for unconstrained optimization.

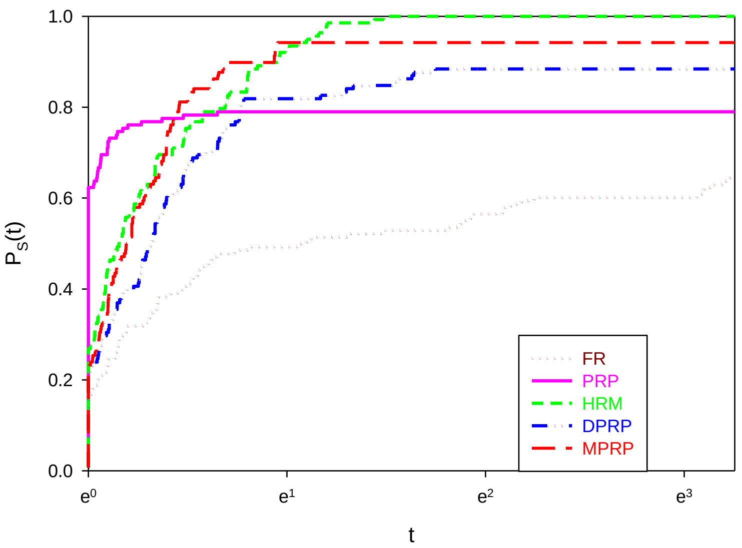

Figure 1: Performance profile relative to the number of iterations

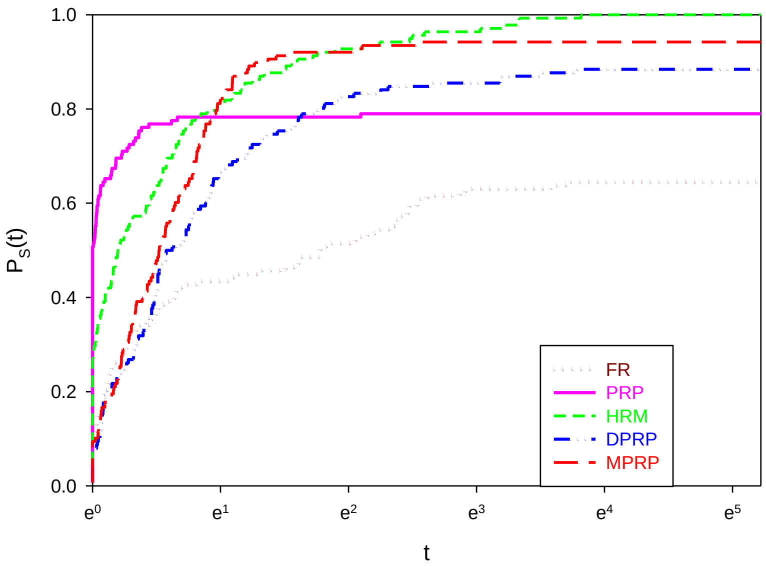

Figure 2: Performance profile relative to the CPU time

5. Conclusion

In this paper, we proposed a new conjugate gradient method for unconstrained optimization. Results showed that it could satisfy the sufficient descent condition and converge globally if the strong Wolfe-Powell line search was used. Numerical results show that the *HRM* method is efficient for the addressed problems.

Acknowledgements. The authors would like to thank the University of Malaysia Terengganu (Grant no FRGS Vot 59256)

References

[1] M. R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Bureau of Standards*, **49** (1952), 409-436. http://dx.doi.org/10.6028/jres.049.044

[2] R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, *The Computer Journal*, **7** (1964), 149-154. http://dx.doi.org/10.1093/comjnl/7.2.149

[3] E. Polak and G. Ribiere, Note Sur la convergence de directions conjuge`es, *ESAIM: Mathematical Modelling and Numerical Analysis*, **3E** (1969), 35-43.

[4] B. T. Polyak, The conjugate gradient method in extreme problems, *USSR Computational Mathematics and Mathematical Physics*, **9** (1969), 94-112. http://dx.doi.org/10.1016/0041-5553(69)90035-4

[5] R. Fletcher, *Practical Method of Optimization*, 2 ed., vol. I, New York, 2000. http://dx.doi.org/10.1002/9781118723203

[6] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, Part 1: Theory, *Journal of Optimization Theory and Applications*, **69** (1991), 129-137. http://dx.doi.org/10.1007/bf00940464

[7] Y.-H. Dai and Y.-X. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Journal on Optimization*, **10** (1999), 177-182. http://dx.doi.org/10.1137/s1052623497318992

[8] M. J. D. Powell, Restart procedures for the conjugate gradient method, *Mathematical Programming*, **12** (1977), 241–254. http://dx.doi.org/10.1007/bf01593790

[9] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 1999. http://dx.doi.org/10.1007/b98874

[10] J. C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM Journal on Optimization*, **2** (1992), 21-42. http://dx.doi.org/10.1137/0802003

[11] L. Guanghui, H. Jiye, and Y. Hongxia, Global convergence of the Fletcher-Reeves algorithm with inexact linesearch, *Applied Mathematics-A Journal of Chinese Universities*, **10** (1995), 75-82. http://dx.doi.org/10.1007/bf02663897

[12] D. Touati-Ahmed and C. Storey, Efficient Hybrid Conjugate Gradient Techniques, *Journal of Optimization Theory and Applications*, **64** (1990), 379-397. http://dx.doi.org/10.1007/bf00939455

[13] M. Al-Baali, Descent Property and Global Convergence of the Fletcher-Reeves Method with Inexact Line Search, *IMA Journal of Numerical Analysis*, **5** (1985), 121-124. http://dx.doi.org/10.1093/imanum/5.1.121

[14] Y.-H. Dai, J. Han, G. Liu, D. Sun, H. Yin and Y.-X. Yuan, Convergence Properties of Nonlinear Conjugate Gradient Methods, *SIAM Journal on Optimization*, **10** (1998), 345-358. http://dx.doi.org/10.1137/s1052623494268443

[15] P. Wolfe, Convergence conditions for ascent methods, *SIAM Review*, **11** (1969), 226-235. http://dx.doi.org/10.1137/1011036

[16] H. Liu, A new conjugate gradient method for unconstrained optimization, *Far East Journal of Mathematical Sciences (FJMS)*, **40** (2010), 145-152.

[17] M. Rivaie, M. Fauzi and M. Mamat, New modifications of conjugate gradient coefficient with global convergence properties, *2012 IEEE Symposium on Humanities, Science and Engineering Research (SHUSER)*, Kuala Lumpur, (2012), 625-629. http://dx.doi.org/10.1109/shuser.2012.6268897

[18] A. Y. Al-Bayati, M. S. Al-Salih and M. M. M. Ali, A Modified Family of CG-Algorithm with a New Closed-Form Line-Search Procedure, *Australian Journal of Basic & Applied Sciences*, **7** (2013), 214-220.

[19] A. Abashar, M. Mamat, M. Rivaie, M. Fauzi and Z. Salleh, A modified DPRP conjugate gradient method for unconstrained optimization, *Far East Journal of Mathematical Sciences (FJMS)*, **97** (2015), 31-44.

http://dx.doi.org/10.17654/fjmsmay2015_031_044

[20] M. Hamoda, A. Abashar, M. Mamat and M. Rivaie, A comparative study of two new conjugate gradient methods, *AIP Conference Proceedings*, **1643** (2015), 616-621. http://dx.doi.org/10.1063/1.4907502

[21] A. Abashar, M. Mamat, M. Rivaie and I. Mohd, Global convergence properties of a new class of conjugate gradient method for unconstrained optimization, *Applied Mathematical Sciences*, **8** (2014), 3307 - 3319. http://dx.doi.org/10.12988/ams.2014.43246

[22] M. Hamoda, M. Rivaie, M. Mamat and Z. Salleh, A New Nonlinear Conjugate Gradient Coefficient for Unconstrained Optimization, *Applied Mathematical Sciences*, **9** (2015), 1813-1822. http://dx.doi.org/10.12988/ams.2015.411994

[23] M. Hamoda, M. Rivaie, M. Mamat and Z. Salleh, A Conjugate Gradient Method with Inexact Line Search for Unconstrained Optimization, *Applied Mathematical Sciences*, **9** (2015), 1823-1832. http://dx.doi.org/10.12988/ams.2015.411995

[24] R. Jaafar, M. Mamat, M. F. Embong and M. Rivaie, A Comparative Study of Modified BFGS and Scale Modified BFGS, *Applied Mathematical Sciences*, **5** (2011), 3981-3989.

[25] I. S. Mohammed, M. Mamat, A. Abashar, Mohd Rivaie and Zabidin Salleh, A Modified Nonlinear Conjugate Gradient Method for Unconstrained Optimization, *Applied Mathematical Sciences*, **9** (2015), 2671-2682. http://dx.doi.org/10.12988/ams.2015.5141

[26] M. Rivaie, A. Abashar, M. Mamat and I. Mohd, The convergence properties of a new type of conjugate gradient methods, *Applied Mathematical Sciences*, **8** (2014), 33-44. http://dx.doi.org/10.12988/ams.2014.310578

[27] M. Rivaie, M. Mamat, W. J. Leong and I. Mohd, A new class of nonlinear conjugate gradient coefficients with global convergence properties, *Applied Mathematics and Computation*, **218** (2012), 11323-11332. http://dx.doi.org/10.1016/j.amc.2012.05.030

[28] J. Wang and X. Chi, CG Global Convergence Properties with Goldstein Linesearch*, *Bulletin of the Brazilian Mathematical Society*, **36** (2005), 197-204. http://dx.doi.org/10.1007/s00574-005-0036-0

[29] M. Mamat, M. Rivaie, I. Mohd and M. Fauzi, A New Conjugate Gradient Coefficient for Unconstrained Optimization, *Int. J. Contemp. Math. Sciences*, **5** (2010), 1429 - 1437.

[30] O. Omer, M. Mamat and M. Rivaie, The global convergence properties of a family of conjugate gradient method under the strong wolfe line search, *Abstract and Applied Analysis*, 2015.

[31] Z.-X. Wei, S. W. Yao and L. Y. Liu, The convergence properties of some new conjugate gradient methods, *Applied Mathematics and Computation*, **183** (2006), 1341-1350. http://dx.doi.org/10.1016/j.amc.2006.05.150

[32] Z.-F. Dai and F. Wen, Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property, *Applied Mathematics and Computation*, **218** (2012), 7421-7430. http://dx.doi.org/10.1016/j.amc.2011.12.091

[33] Y.-Q. Zhang, H. Zheng and C.-L. Zhang, Global Convergence of a Modified PRP Conjugate Gradient Method, *Procedia Engineering*, **31** (2012), 986-995. http://dx.doi.org/10.1016/j.proeng.2012.01.1131

[34] G. Zoutendijk, Nonlinear Programming, Computational Methods, Chapter in *Integer and Nonlinear Programming*, North-Holland, Amsterdam, 1970, 37-86.

[35] G.-Y. Li, C.-M. Tang and Z.-X. Wei, New conjugacy condition and related new conjugate gradient methods for unconstrained optimization, *Journal of Computational and Applied Mathematics*, **202** (2007), 523-539. http://dx.doi.org/10.1016/j.cam.2006.03.005

[36] Z.-X. Wei, G. Li and L. Qi, New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems, *Applied Mathematics and Computation*, **179** (2006), 407-430. http://dx.doi.org/10.1016/j.amc.2005.11.150

[37] N. Andrei, An unconstrained optimization test functions collection, *Advanced Modeling and Optimization*, **10** (2008), 147-161.

[38] M. Molga and C. Smutnicki, *Test Functions for Optimization Needs*, 2005.

[39] S. Mishra, Some new test functions for global optimization and performance of repulsive particle swarm method, *Munich Personal RePEc Archive*, Apr 13 2007.

[40] E. D. Dolan and J. J. Mor, Benchmarking optimization software with

performance profiles, *Mathematical Programming*, **91** (2002), 201-213.
http://dx.doi.org/10.1007/s101070100263

Received: July 4, 2015; Published: March 8, 2016