Applied Mathematical Sciences, Vol. 9, 2015, no. 37, 1823 - 1832
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2015.411995

A Conjugate Gradient Method with Inexact Line Search for Unconstrained Optimization

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

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

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

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

Copyright © 2014 Mohamed Hamoda et al. 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.

Abstract

In this paper, an efficient nonlinear modified PRP conjugate gradient method is presented for solving large-scale unconstrained optimization problems. The sufficient descent property is satisfied under strong Wolfe-Powell (SWP) line search by restricting the parameter $\sigma < 1/4$. The global convergence result is established under the (SWP) line search conditions. Numerical results, for a set consisting of 133 unconstrained optimization test problems, show that this method is better than the PRP method and the FR method.

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 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 current iterate point and α_k is a 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 being 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 [11]), FR (Fletcher_Revees [8]), PRP (Polak_Ribiére_Polyak [18, 19]), CD (Conjugate Descent [7]), LS (Liu-Storey [15]), and DY (Dai_Yuan [5]) conjugate gradient method. The convergence behavior of the above formulas with some line search conditions has been studied by many authors for many years (e.g.[1, 3-5, 7, 9, 10, 12, 13, 15-17, 20-24]).

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

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

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

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 insure the global convergence of the nonlinear conjugate gradient method with the inexact line search techniques [1, 9, 21].

2. New formula for B_k and its properties

Therefore, many of the variants of the PRP method had been widely studied. In this paper, a variant of the PRP method is known as β_k^{MRM}, where MRM denotes Mohamed, Rivaie and Mustafa, β_k^{MRM} is defined by

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

Now we give the following algorithm firstly.

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 the 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

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

In [9], 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 [24].

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

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

where γ and $\bar{\gamma}$ are two positive constants. We say that the method has the 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^{MRM} 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^{MRM} has the property (*).

Proof. Set $b = \frac{\bar{\gamma}^2(\gamma+\bar{\gamma})}{\gamma^3} > 1$, $\lambda = \frac{\gamma^2}{4L\bar{\gamma}b}$. By (8) and (10) we have

$$\left|\beta_k^{MRM}\right| \le \frac{\left|g_k^T(g_k - \frac{\|g_k\|}{\|g_{k-1}\|}g_{k-1})\right|}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \le \frac{\|g_k\|(\|g_k\| + \frac{\bar{\gamma}}{\gamma}\|g_{k-1}\|)}{\|g_{k-1}\|^2} \le \frac{\bar{\gamma}(\bar{\gamma} + \frac{\bar{\gamma}^2}{\gamma})}{\gamma^2} = \frac{\bar{\gamma}^2(\gamma+\bar{\gamma})}{\gamma^3} = b$$

From the Assumption B, (9) holds. If $\|s_k\| \le \lambda$ then,

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

$$\le \frac{2L\lambda\|g_k\|}{\|g_{k-1}\|^2} \le \frac{2L\lambda\bar{\gamma}}{\gamma^2} = \frac{1}{2b}$$

The proof is finished.

3. The global convergence properties

The following theorem shows that the formula MRM 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, then the sequence $\{d_k\}$ possesses the sufficient descent condition (7).

Proof. By the formulae (8), we have

$$\beta_k^{MRM} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}g_k^T g_{k-1}}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \ge \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \ge \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} = 0$$

Thus we get, $\beta_k^{MRM} \ge 0$

Also

$$\beta_k^{MRM} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}g_k^T g_{k-1}}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \le \frac{\|g_k\|^2 + \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \le \frac{\|g_k\|^2 + \frac{\|g_k\|}{\|g_{k-1}\|}\|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^2 + \left|g_k^T d_{k-1}\right|} \le \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}$$

Hence we obtain

$$0 \le \beta_k^{MRM} \le \frac{2\|g_k\|^2}{\|g_{k-1}\|^2} \tag{9}$$

Using (6) and (9), we get

$$\left|\beta_{k+1}^{MRM} g_{k+1}^T d_k\right| \le \frac{2\|g_{k+1}\|^2}{\|g_k\|^2}\sigma\left|g_k^T d_k\right| \tag{10}$$

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

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1}\frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2} \tag{11}$$

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 that
$d_i, i = 1,2,....,k,$ are all descent directions, that is $g_i^T d_i < 0$
By (10), we get

$$\left|\beta_{k+1}^{MRM} g_{k+1}^T d_k\right| \le \frac{2\|g_{k+1}\|^2}{\|g_k\|^2}\sigma(-g_k^T d_k) \tag{12}$$

That is,

$$\frac{\|g_{k+1}\|^2}{\|g_k\|^2} 2\sigma\, g_k^T d_k \le \beta_{k+1}^{MRM} g_{k+1}^T d_k \le -\frac{\|g_{k+1}\|^2}{\|g_k\|^2} 2\sigma\, g_k^T d_k \tag{13}$$

(11) and (13) deduce,

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

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

$$-\sum_{j=0}^{k}(2\sigma)^j \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \sum_{j=0}^{k}(2\sigma)^j \tag{14}$$

Since

$$\sum_{j=0}^{k}(2\sigma)^j < \sum_{j=0}^{\infty}(2\sigma)^j = \frac{1}{1-2\sigma}$$

(14) can be written as

$$-\frac{1}{1-2\sigma} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \frac{1}{1-2\sigma} \tag{15}$$

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 $c = 2 - \frac{1}{1-2\sigma}$ then, $0 < c < 1$, and (15) turns out to be

$$(c-2)\|g_k\|^2 \le g_k^T d_k \le -c\|g_k\|^2 \quad (16)$$

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

The following condition known as Zoutendijk condition is used to prove the global convergence of nonlinear CG methods[23, 25].

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,

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \quad (17)$$

The proof had been given in [14, 22]. In[9], 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 MRM method with the parameter $0 < \delta < \sigma < 1/4$ satisfies all four conditions in theorem 3.2 under the strong Wolfe-Powell line search, so the method is globally convergent.

4. Numerical results and discussion

In this section, we selected 27 test functions considered in Andrei [2]. For each test function we have considered from 1 to 7 numerical experiments with number of variables lay in the range from 2 to 10000, shown in table1, also for each test function, we used four initial points, starting from a closer point to the solution and moving on to the one that is furthest from it. We performed a comparison with two CG methods FR and PRP. The step size α_k satisfies the strong Wolfe-Powell conditions, with $\delta = 10^{-4}$, $\sigma = 0.001$ and $\|g_k\| < 10^{-6}$. A list of functions and the initial points used are shown in table1, where all the problems are solved by MATLAB program. We used the strong Wolfe Powell line search to compute the step size. The CPU processor used was Intel (R) Core^{TM} i3-M350 (2.27GHz), with RAM 4 GB. 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 a failure. In addition, we considered a failure if the number of iterations exceeds 1000 or CPU time exceeds 500 (Sec). Numerical results are compared relative on the CPU time and number of iterations. The performance results are shown in Figs.1 and 2 respectively, using a performance profile introduced by Dolan and More [6].

Table 1. A list of problem functions

No	Function	Dimension	Initial points
1	Six Hump	2	-10, 10, -8, 8
2	Booth	2	10, 25, 50, 100
3	Treccani	2	5, 10, 20, 50
4	Zettl	2	5, 10, 20, 30
5	Extended Maratos	2, 4,10, 100	1, 5, 8, 10
6	Fletcher	4, 10, 100, 500, 1000	7, 9, 11, 13
7	Perturbed Quadratic	2, 4, 10, 100, 500, 1000	1, 5, 10, 15
8	Extended Himmelblau	100, 500, 1000, 10000	50, 70, 100, 125
9	Extended Rosenbrock	2, 4, 10, 100, 500, 1000, 10000	13, 25, 30, 50
10	Shallow	2, 4, 10, 100, 500, 1000, 10000	10, 25, 50, 70
11	Extended Tridiagonal 1	2, 4, 10,100, 500, 1000, 10000	12, 17, 20, 30
12	Generlyzed Tridiagonal 1	2, 4,10, 100	25, 30, 35, 50
13	Extended white & Holst	2, 4,10,100, 500, 1000, 10000	3, 10, 30, 50
14	Generalized Quartic	2, 4,10,100, 500, 1000, 10000	1, 2, 3, 5
15	Extended Powell	4, 8, 20, 100, 500, 1000	4, 5, 7, 30
16	Extended Denschnb	2, 4, 10, 100, 500, 1000, 10000	8, 13, 30, 50
17	Hager	2, 4, 10, 100	1, 3, 5, 7
18	Extended Penalty	2, 4, 10, 100	10, 50, 75, 100
19	Quadrtic QF2	2, 4, 10, 100 ,500, 1000	10, 30, 50, 100
20	Extended Quadratic Penalty QP2	2, 4, 10, 100, 500, 1000, 10000	17, 18, 19, 20
21	Extended Beale	2, 4,10, 100, 500, 1000, 10000	1, 3, 13, 30
22	Diagonal 2	2, 4, 10, 100, 500, 1000	-1,1, 2, 3
23	Raydan1	2, 4, 10,100	1, 3, 5, 7
24	Sum Squares function	2, 4, 10,100, 500, 1000	1, 10, 20, 30
25	Generlized Tridiagonal 2	2, 4, 10, 100	1, 10, 20, 30
26	Quadratic QF1	2, 4, 10,100, 500, 1000	1, 2, 3, 4
27	Dixon and Price	2, 4, 10, 100	100, 125, 150, 175

In figures 1-2, the horizontal axis of the figure gives the percentage of the test problems for which a method is the fastest, while the vertical axis gives the percentage of the test problems that were successfully solved by each method. Fig.1 presents the performance profiles of MRM, FR and PRP relative to the number of iterations. Fig.2 presents the performance profiles of the three methods relative to the CPU time. The interpretation in Figures 1-2 shows that the new method outperform the other two methods relative to both performance, number of iterations and CPU time, since MRM can solve all the test problems and reach 100%, while PRP can solve only 79% of the problems and FR solved only 65%, the performance of MRM lies between FR and PRP and we can say that MRM near PRP. Hence we considered that MRM method is computationally efficient.

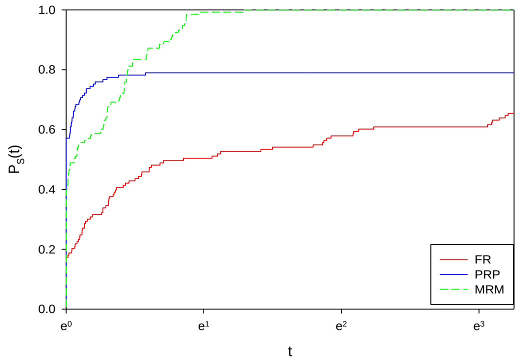

Figure 1. Performance profile relative to the number of iterations.

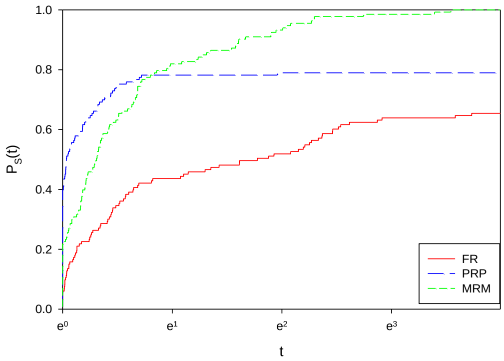

Figure 2. Performance profile relative to the CPU time.

5. Conclusion and future research

In this paper, we proposed a new β_k for unconstrained optimization, we prove that it is a global convergence with strong Wolfe Powell line search. Based on our numerical experiments, we concluded that the new method more efficient and more robust than the classical methods FR and PRP.

Our future work is concentrated on studying the convergence properties of our new method using different inexact line searches.

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

References

[1] 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

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

[3] Y. 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 (2000), 345-358. http://dx.doi.org/10.1137/s1052623494268443

[4] Y. H. Dai and Y. Yuan, "Convergence properties of the Fletcher-Reeves method," *IMA Journal of Numerical Analysis*, 16 (1996), 155-164. http://dx.doi.org/10.1093/imanum/16.2.155

[5] Y. H. Dai and Y. 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

[6] 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

[7] R. Fletcher, *Practical Method of Optimization*, 2 ed. I. New York, 1987. http://dx.doi.org/10.1002/9781118723203

[8] 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

[9] 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

[10] 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

[11] 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

[12] Y. F. Hu and C. Storey, "Global Convergence Result for Conjugate-Gradient Methods," *Journal of Optimization Theory and Applications*, 71 (1991), 399-405. http://dx.doi.org/10.1007/bf00939927

[13] S. Jie and Z. Jiapu, "Global Convergence of Conjugate Gradient Methods without Line Search," *Annals of Operations Research*, 103 (2001), 161–173. http://dx.doi.org/10.1023/a:1012903105391

[14] 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

[15] 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

[16] J. J. More, B. S. Garboww, and K. E. Hillstrom, "Testing Unconstrained Optimization Software," *ACM Transactions on Mathematical Software* 7 (1981), 17-41. http://dx.doi.org/10.1145/355934.355936

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

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

[19] 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

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

[21] 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

[22] Z. 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

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

[24] Y. Q. Zhang, H. Zheng, and C. L. Zhang, "Global Convergence of a Modified PRP Conjugate Gradient Method," in *International Conference on Advances in Computational Modeling and Simulation,* (2012), 986-995. http://dx.doi.org/10.1016/j.proeng.2012.01.1131

[25] G. Zoutendijk, "Nonlinear programming, computational methods," in *Integer and nonlinear programming*, ed North-Holland, Amsterdam, 1970, pp. 37-86.

Received: December 10, 2014; Published: March 9, 2015