International Journal of Engineering & Technology, 7 (3.28) (2018) 54-57

International Journal of Engineering & Technology

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

Research paper

New Hybrid Conjugate Gradient Method with Global Convergence Properties under Exact Line Search

Yasir Salih[1], Mustafa Mamat[1*], Mohd Rivaie[2], Abdelrhaman Abashar[3], Mohamad Afendee Mohamed[1]

[1]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UNISZA), Kuala Terengganu, Terengganu, Malaysia*
[2]*Department of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM), Terengganu, Malaysia*
[3]*Faculty of Engineering, Red Sea University, Sudan*
**Corresponding author E-mail: must@unisza.edu.my*

Abstract

Conjugate Gradient (CG) method is a very useful technique for solving large-scale nonlinear optimization problems. In this paper, we propose a new formula for β_k, which is a hybrid of PRP and WYL methods. This method possesses sufficient descent and global convergence properties when used with exact line search. Numerical results indicate that the new formula has higher efficiency compared with other classical CG methods.
.

Keywords: *Nonlinear optimization; conjugate gradient coefficient; exact line search; global convergence; large scale.*

1. Introduction

In this paper, we focus on the unconstrained optimization problem defined by.

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

where $f: R^n \to R$ is a continuously differentiable function. The iterates of CG method for solving (1) are obtained by

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

where α_k is the step-size and computed by a line search procedure, that is

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

The search direction, d_k is given as

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

$$\beta_k$$

where β_k is the CG coefficient. There are many formulas for determining β_k, each one defines different CG method. Examples of the widely known formulas for β_k are

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

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

$$\beta_k^{DY} = \frac{\|g_k{}^2\|}{d_{k-1}^T(g_k - g_{k-1})}. \; \beta_k^{LS} = \frac{g_k^T(g_k - g_{k-1})}{-d_{k-1}^T g_{k-1}},$$

The above formulas are called Hestenes-Stiefel (HS) [1], Fletcher-Reeves (FR) [2], Polak,-Ribiere-Polyak (PRP) [3,4], Conjugate Descent (CD) [5], Liu-Storey (LS) [6] and Dai-Yuan (DY) [7] respectively. All of these methods have finite convergence properties under the exact line search, provided that $f(x)$ is strictly a convex quadratic function. For further information, readers can refer to [8-19].

2. New Parameter and Algorithm

In this section, we present a new formula for CG coefficient, referred here as β_k^{YHM} where YHM denotes Yasir, Hamoda and Mamat. The β_k^{YHM} is defined by

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} & if\ 0 \leq g_k^T g_{k-1} \leq \|g_k\|^2 \\ \dfrac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2} & \text{otherwise} \end{cases} \tag{5}$$

The following algorithm is implemented using the proposed hybrid CG method.

Algorithm 2.1: Hybrid CG Method

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

Step 2: Compute α_k by exact line search.

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

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

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

3. Global Convergence Properties

In this section, we study the global convergence properties of β_k^{YHM}. Firstly, we need to simplify β_k^{YHM} to ease the proving process. From (5), we have

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} & \text{if } 0 \leq g_k^T g_{k-1} \leq \|g_k\|^2 \\ \dfrac{g_k^T\left(g_k - \dfrac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2} & \text{otherwise} \end{cases}.$$

If $0 \leq g_k^T g_{k-1} \leq \|g_k\|^2$, then

$$\beta_k^{YHM} = \frac{g_k^T\ (g_k - g_{k-1})}{\|g_{k-1}\|^2} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2} \geq 0.$$

Otherwise,

$$\beta_k^{YHM} = \frac{g_k^T(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1})}{\|g_{k-1}\|^2} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2}.$$

Using the Cauchy-Schwartz inequalities implies that

$$\beta_k^{YHM} \geq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} \|g_k\| \|g_{k-1}\|}{\|g_{k-1}\|^2} = 0.$$

Hence, we get $\beta_k^{YHM} \geq 0$.

3.1. Sufficient descent condition

The sufficient descent condition is defined by:

$$g_k^T d_k \leq -c\|g_k\|^2 \text{ for } k > 0, c > 0. \tag{6}$$

The following theorem shows that our new CG coefficient with exact line search possesses sufficient descent property.

Theorem 1. Suppose that x_k and d_k are generated by the iterative CG method of the form (2), (4) and (5) and the step-size α_k is determined by exact line search, then condition (6) holds for all $k \geq 0$.

Proof. The proof is by induction. If $k = 0$, then $g_0^T d_0 = g_0^T(-g_0) = -\|g_0\|^2 < 0$. Hence, condition (6) holds true. Next, we show that condition (6) will also hold true for $k \geq 1$.

Multiply equation (4) by g_k^T, then

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

Note that for exact line search $g_k^T d_{k-1} = 0$. Thus, we get

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

Hence, condition (6) also holds true for all $k \geq 1$.

3.2. Global Convergence Properties

The following assumptions are necessary for proving the convergence of CG method.

Assumption 1.
$f(x)$ 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 2.
The gradient $g(x)$ is Lipschitz continuous in N and there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \leq L\|x - y\|, \quad \forall\, x, y \in N$.
After this assumption, we have the following lemma which was proved by [20].

Lemma 1. Suppose that Assumptions1 and 2 hold, let x_k be generated by Algorithm 2.1 and d_k satisfies $g_k^T d_k < 0$ for all k and α_k is obtained by (3), then

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{7}$$

The following theorem is based on Lemma 1.

Theorem 2. Suppose that Assumption 1 and 2 hold, the sequence $\{x_k\}$ is generated by Algorithm 2.1. If $\|s_k\| = \|\alpha_k d_k\| \to 0$ while $k \to \infty$, then

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{8}$$

Proof. Case 1: If $0 \leq g_k^T g_{k-1} \leq \|g_k\|^2$ then $\beta_k^{YHM} = \frac{g_k^T\ (g_k - g_{k-1})}{\|g_{k-1}\|^2}$ to proof this theorem, we use contradiction. If Theorem 1 is not true, then the constant $c > 0$ exists such that

$$\|g_k\| \geq c \tag{9}$$

From equation (4), we have $d_k = -g_k + \beta_k^{YHM} d_{k-1}$. Now, by squaring both sides of the equation, we obtain

$$\|d_k\|^2 = \|g_k\|^2 - 2\beta_k^{YHM} g_k d_{k-1} + (\beta_k^{YHM})^2 \|d_{k-1}\|^2 \tag{10}$$

For exact line search, $g_k^T d_{k-1} = 0$. Therefore,

$$\|d_k\|^2 = (\beta_k^{YHM})^2 \|d_{k-1}\|^2 + \|g_k\|^2 \tag{11}$$

Then, $\|d_k\|^2 \le \|g_k\|^2 + \frac{\|g_k\|^4}{\|g_{k-1}\|^2}\|d_{k-1}\|^2$

We divide the above equation by $\|g_k\|^4$ to get

$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} \qquad (12)$$

Utilizing (12) recursively and noting that

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

$$\frac{\|d_k\|^2}{\|g_k\|^4} \le \sum \frac{1}{\|g_k\|^2}$$

By (9), we know that $\frac{\|d_k\|^2}{\|g_k\|^4} \le \frac{k}{c^2}$. Therefore,

$$\frac{\|g_k\|^4}{\|d_k\|^2} \ge \frac{c^2}{k} \cdot \frac{(g_k^T d_k)^2}{\|d_k\|^2} \ge \frac{c^2}{k}$$

This contradicts the Zoutendijk condition, $\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$. The proof is completed.

Case 2: If $\beta_k^{YHM} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2}$, the proof of this theorem can be seen in [21].

4. Results and Discussion

In this section, we present the computational performance of the new hybrid CG method under exact line search. Thirty-one unconstrained optimization test problems are taken from [22] and fully listed in Table 1. We analyze the efficiency of our new formula by comparing its numerical performance with the FR [2], PRP [3, 4], WYL [21] and HRM3 [23] methods. All of the tested CG algorithms use $\|g_k\| < \varepsilon$ as the stopping criterion with $\varepsilon = 10^{-5}$ as suggested by [24]. Four initial points are selected for the numerical test, starting from a point close to the solution and then moving on to other points farther from it. The numerical testing process uses MATLAB version 8.3.0.532 (R2014a) software on a computer with Intel(R) Core™ i5-M520 (2.40GHz) CPU processor and 4GB RAM. In some cases, failure occurs when the line search is unable to find a positive step length.

Table 1: A list of problem functions

No.	Function	Dimension	Initial Points
1	Six Hump Camel	2	-10, -8, 8, 10
2	Three Hump Function	2	-10, 10, 20, 30
3	Booth	2	10, 25, 50, 100
4	Treccani	2	5, 10, 20, 50
5	Zettl	2	5, 10, 20, 30
6	Ex – Rosenbrock	2,4,10,100,500,1000,10000	13, 25, 30, 50
7	Diagonal 4	2,4,10,100,500,1000	1, 3, 6, 12
8	Shalow	2,4,10,100,500,1000,10000	10, 25, 50, 70
9	Ex - Tridiagonal1	2,4,10,100,500,1000,10000	6, 12, 17, 20
10	Ex- white and Holst	2,4,10,100,500,1000,10000	3, 5, 7, 10
11	Perturbed Quadratic Function	2,4,10,100,500,1000	1, 3, 5, 10
12	Ex- Denschnb	2,4,10,100,500,1000,10000	8, 13, 30, 50
13	Ex- Beale	2,4,10,100,500,1000,10000	-1, 3, 7, 10
14	Ex – Himmelblau	10,100,500,1000,10000	50, 70, 100, 125
15	Generalized Quartic	2,4,10,100,500,1000,10000	1, 2, 5, 7
16	Hager	2,4,10,100,500,1000,10000	5, 10, 15, 20
17	Ex – Penalty	2,4,10,100	80, 100. 111, 120
18	Quadratic QF2	2,4,10,100,500,1000	5, 20, 50, 100
19	Ex - Quadratic Penalty qp2	2,4,10,100,500,1000	10, 20, 30, 50
20	Diagonal 2	2,4,10,100,500,1000	1, 5, 10, 15
21	Raydan1 Function	2,4,10,100	1, 3, 7, 10
22	Sum Squares Function	2,4,10,100,500,1000	1, 3, 7, 10
23	Generalized Tridiagonal 1	2,4,10,100	7, 10, 13, 20
24	Fletcher	4,10,100,500,1000,10000	3, 5, 8. 10
25	Quadratic QF1	2,4,10,100,500,1000	3, 5, 8, 10
26	Generalized Tridiagonal 2	2,4,10,100	15, 18, 20, 22
27	Leon Function	2	2, 5, 8, 10
28	Ex-Wood	4	3, 5, 20,30
29	Quartic Function	4	5, 10, 15, 20
30	Matyas Function	2	1, 5, 10, 15
31	Colville Function	4	2, 4, 7, 10

We use the performance profile introduced by [25] to assess the efficiency of each tested method. Figure 1 and 2 show the numerical performances of those methods relative to the number of iterations and CPU time respectively. From the results, we can see that YHM solves 100% of the test problems, followed by HRM3 which solves 99.5%. PRP solves 99.4%, WYL solves 98% and FR solves 72% of the test problems. Since only YHM manage to solve all of the test functions, we can say that it has the best overall performance.

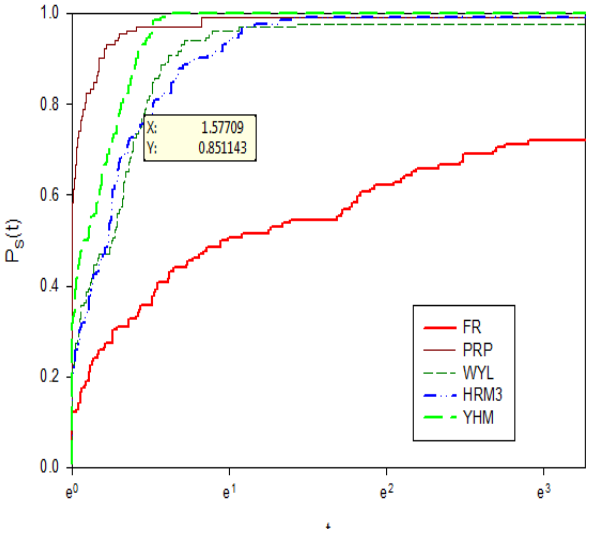

Fig. 1: Performance profile relative to the number of iteration.

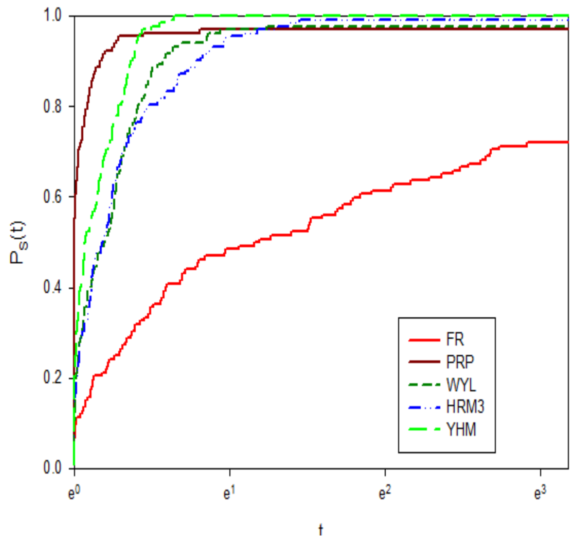

Fig. 2: Performance profile relative to the CPU time.

5. Conclusion

In this paper, we propose a new formula for β_k under exact line search. The proposed method is shown to possess sufficient descent and global convergence properties. Based on the numerical results, our new method has the best performance compared to other tested CG methods.

Acknowledgement

The authors would like to thank the Universiti Sultan Zainal Abidin (Unisza), and Red Sea University of Sudan for financial support.

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.

[2] R. Fletcher, and C.M. Reeves, Function minimization by conjugate gradients. Computer Journal, 7(1964), 149-154.

[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, USSRComputational Mathematics and Mathematical Physics, 9(1969), 94-112.

[5] R. Fletcher, Practical Method of Optimization, New York, 2000.

[6] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, Part 1: Theory, Journal of Optimization Theory and Applications, 69(1991), 129-137.

[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.

[8] M. Mamat, M. Rivaie, I. Mohd and M. Fauzi (2010). A new conjugate gradient coefficient for unconstrained optimization, Int. J. Contemp. Math. Sciences, 5(29), 1429-1437.

[9] 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.

[10] M. Rivaie, A. Abashar, M. Mamat and I. Mohd (2014). The convergence properties of a new type of conjugate gradient methods, 8(1), 33-44.

[11] Aini, N., Hajar, N. Mamat, M., Zull, N., and Rivaie, M. (2017). Hybrid quasi-newton andconjugate gradient method for solving unconstrained optimization problems, Journal of Engineering and Applied Sciences, 12(8), 4627-4631.

[12] Ghani, N. H. A., Mamat, M., Susilawati, F., Zull, N., Mohamed, N. S., and Reivaie, M. (2017). Another descent conjugate gradient methodwith strong-Wolfe line search. Journal of Engineering and Applied Sciences, 12(18), 4632-4636.

[13] Ghani, N. H. A., Mohamed, N. S. Zull, N., Shoid, S., Rivaie, M., and Mamat (2017). Performance comparison of a new hybrid conjugate gradient method under exact and inexact line searches. Journal of Physics: Conference Series, 890(1).

[14] Ghani, N. H. A., Rivaie, M., Mamat, M. A modified form of conjugate gradient method for unconstrained optimization problems. (2016), AIP Conference Proceedings, 1739, 020076.

[15] Hager, W. W. and H. Zhang, A survey of nonlinear conjugate gradient methods. Pacific Journal of Optimization, 2006, 2(1). 35-58.

[16] Mohamed, N. S., Mamat, M., and Rivaie, M. (2017). A new nonlinear conjugate gradient coefficient under strong Wolfe-Powell line search. AIP Conference Proceedings, 1870

[17] Wan Osman, W. F. H., Hery Ibrahim, M. A., and Mamat, M. (2017). Hybrid DFP-CG method for solving unconstrained optimization problems. Journal of Physics: Conference Series, 890(1).

[18] Zull, N., Aini, N., Mamat, M., Ghani, N. H. A., Mohamed, N. S., Rivaie, M. and Sukono. (2017). An alternative approach for conjugate gradient method. Journal of Engineering and Applied Sciences, 12(18), 4622-4626.

[19] I. S. Mohammed, M. Mamat, A. Abashar, M. Rivaie and Z. Salleh (2017). The Global Convergence Properties of an Improved Conjugate Gradient Method. Journal of Applied Mathematical Sciences, 9(38), 1857-1868.

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

[21] 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.

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

[23] M. Hamoda, M. Mamat, M .Rivaie and Z. Salleh (2016). A Conjugate Gradient Method with Strong Wolfe-Powell Line Search for Unconstrained Optimization. Journal of Applied Mathematical Sciences, 10(15), 721-734.

[24] K. E. Hillstrom, A simulation test approach to the evaluation of nonlinear optimization algorithm, ACM Transactions on Mathematical Software, 3(4) (1977), 305-315.

[25] E. D. Dolan and J. J. Mor, Benchmarking optimization software with performance profiles, Mathematical Programming, 91(2002), 201-213.