International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume-2, Special Issue 6 (Nov-Dec 2014), PP. 14-17

A NEW HYBRID WYL-AMRI CONJUGATE GRADIENT METHOD WITH SUFFICIENT DESCENT CONDITION FOR UNCONSTRAINED OPTIMIZATION

Ibrahim S. Mohammed[1], Mustafa Mamat[2], Kamil Uba Kamfa[3], Abdelrhaman Abashar[4]
[1, 2, 3,] Faculty of Informatics and computing
Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia
[4]Faculty of Engineering
Red Sea University, Sudan
kademi4u@yahoo.com

Abstract: Conjugate gradient method has played a significant role in solving large scale unconstrained optimization. Numerous survey and modifications have been done recently to improve this method. In this paper, we proposed a new hybrid method of Wei-Yao-Liu (WYL) method and the Abdelrhaman et al (AMRI) method, which possesses the sufficient descent condition under exact line search. The result of the numerical experiments show that the new proposed hybrid method perform better when compared with the WYL and AMRI methods. A set of test problems with different initial points are used, most of them are from Andrei (2008).

Keywords: Exact line search, Conjugate gradient Method, Unconstrained optimization, sufficient descent condition.

I. INTRODUCTION

Conjugate gradient (CG) method is very useful in solving large scale unconstrained optimization problems. This is due to its low memory requirement and simplicity in implementation. This method is of the form

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

where $f: R^n \to R$ is a continuously differentiable function. Generally, the conjugate gradient methods are iterative methods of the form,

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

where $\alpha_k \geq 0$ is the step length computed using exact line search by the formula given as

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

and d_k is the search direction computed as follow

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

where β_k is known as the conjugate gradient coefficient that characterizes different CG methods. Some classical methods such as the Fletcher Reeves (FR) [12], Dai and Yuan (DY) [21], and Conjugate Descent (CD), proposed by Fletcher [11], define as

$$\beta_k^{FR} = \frac{||g_k||^2}{||g_{k-1}||^2} \tag{5}$$

$$\beta_k^{DY} = \frac{||g_k||^2}{d_{k-1}^T(g_k - g_{k-1})} \tag{6}$$

$$\beta_k^{CD} = -\frac{||g_k||^2}{d_{k-1}^T g_{k-1}} \tag{7}$$

where proved to possess strong convergent properties, but they may not have modest practical performance due to jamming. On the other hand, the method of Polak Ribiere and polyak (PRP) [13, 17], Hestenes and Stiefel (HS) [10], and Liu and Storey (LS) [14], define below

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{||g_{k-1}||^2} \tag{8}$$

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{d_{k-1}^T(g_k - g_{k-1})} \tag{9}$$

$$\beta_k^{LS} = \frac{g_k^T(g_k - g_{k-1})}{-d_{k-1}^T g_{k-1}} \tag{10}$$

may not always be convergent, but they often have better computational performances. From equation (5 - 10), g_{k-1} and g_k denotes the gradient of $f(x)$ at the point x_{k-1} and x_k respectively. Also, $\|.\|$ denotes Euclidean norm of vectors.

Hybrid CG methods holds an important role in solving large scale unconstrained optimization. This is due to the part it plays in achieving better computational performance as well as retaining the strong global convergence properties of the various methods. Numerous researches and modifications have been done recently which focus mainly on hybridization of the different categories of the CG methods. These include Dai and Yuan [4], Touati-Ahmed and Storey [3], and Andrei [5-8]. The studies concentrate on the projection of various CG algorithms, usually with the aim of preventing the jamming phenomenon. Recently, Xiangrong LI and Xupei Zhao [16], suggested a hybrid method combining PRP and WYL methods. This method possesses some nice properties of the PRP method and the WYL method define as

$$\beta_k^{P-W} = \max\left\{\beta_k^{PRP}, \beta_k^{WYL}\right\} \tag{11}$$

Motivated by this idea, we proposed a new hybrid CG method between the WYL [20] method, and AMRI [19], where the WYL method is given as

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

and the AMRI method define as

$$\beta_k^{AMRI} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}|g_k g_{k-1}|}{\|d_{k-1}\|^2} \tag{13}$$

II. NEW HYBRID COEFFICIENT

In this section, we present the new coefficient as follows

$$\beta_k^{SW-A} = \max\left\{\beta_k^{WYL}, \beta_k^{AMRI}\right\} \tag{14}$$

where SW-A stands for Sulaiman, WYL method, and AMRI method. The new formula possesses some good properties of the WYL method and also the AMRI method.

In this paper, we present the sufficient descent condition of a new hybrid CG method. The algorithm of this new method presented in the next section. In section 3, we give the proof of the sufficient descent condition under exact line search. Numerical result and discussions are presented in section 4. Lastly, section 5 gives the conclusion.

The algorithm of β_k^{SW-A} is given as follows

ALGORITHM

Step1. Given an initial point $x_0 \in R^n$, $\varepsilon \in (0,1)$, Set $d_0 = -g_0$ if $||g_0|| \le \varepsilon$, then stop.

Step2. Compute β_k^{SW-A} based on (14).

Step3. Compute d_k based on (4). If $\| g_k \| \le \varepsilon$, then terminate,

Step4. Compute step size based on (3).

Step5. Update new point based on (2).

Step6. Convergent test and stopping criteria

If $f(x_k) \le f(x_{k+1})$ and $\| g_k \| \le \varepsilon$, then terminate, else, Set $k = k+1$ and go to Step 2.

III. CONVERGENT ANALYSIS

In this section, the convergent properties of β_k^{SW-A} will be studied. We only show the result of convergence for general CG method. To prove the convergence, we assumed that every search direction d_k should satisfy the descent condition

$$g_k^T d_k < 0 \tag{15}$$

for all $k \ge 0$. If there exist a constant $\lambda > 0$ for all $k \ge 0$ then, the search directions satisfy the following sufficient descent condition

$$g_k^T d_k \le -\lambda \| g_k \|^2 \tag{16}$$

The following Theorem is very important in establishing sufficient descent condition.

Theorem 1: Consider a CG method with the search direction (4) and β_k^{SW-A} given as (14), then condition (15) holds for all $k \ge 0$.

Proof. If $k = 0$, then it is clear that $g_0^T d_0 = -\lambda \| g_0 \|^2$. Hence, condition (15) holds true. We also need to show that for $k \ge 1$, condition (15) will also hold true

From (4), multiply both sides by g_{k+1}, we obtain

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

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

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

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

Therefore, it implies that d_{k+1} is a sufficient descent direction. Hence, $g_k^T d_k \le -\lambda \| g_k \|^2$ holds true. The proof is completed. ■

IV. NUMERICAL RESULT AND DISCUSSION

In this section, we present the result of our numerical experiments. We compared the performance of the new hybrid SWYL-AMRI method with the WYL method and the AMRI method base of number of iterations and CPU time. Most of the selected test problems considered are from Andrei [24], as presented in table 1. We considered $\| g_k \| < 10^{-6}$ to be stopping criteria. For each of the test problems, four initial points are used, starting with the point that is near the solution and moving to the point that is furthest from the solution. These four initial points will lead us to test the global convergence and the robustness of our method. All codes are written on MATLAB 7.6.0 (R 2008a) subroutine programming. The test results are run on an Intel® Core™ i5-2410M CPU @ 2.30 GHz processor, 4GB for RAM memory and Windows 7 Professional operating system.

Figure 1and 2 shows the performance results respectively, these were evaluated using the performance profile by Dolan and Moore [15]. Clearly, it shows that the new SWYL-AMRI hybrid method outperforms the WYL method and the AMRI method, as it was able to solve all the test problems successfully and reach 100%. However, the WYL method was able to solve about 97% of the test problems, and AMRI method was able to solve about 95% of the problems respectively. This shows that our new hybrid method is more effective than the WYL and AMRI methods\

TABLE 1: A LIST OF PROBLEM FUNCTIONS

No	Function	Dim	Initial Points
1	Six hump	2	(1, 1), (2, 2), (5,5), (10, -10)
2	Three hump	2	(24, 24), (29, 29), (33, 33), (50, 50)
3	Booth	2	(10, -10), (20, 20), (50, 50), (100, 100)
4	Treccani	2	(5, 5), (10,10), (-20, 20), (-50, -50)
5	Matyas	2	(1, 1), (5, 5), (10, 10), (50, 50)
6	Extended Maratos	2, 4	(0,0,0,0), (0.5,5, 0.5, 5), (10, 0.5, 10, 0.5), (70, 70, 70, 70)
7	Ext FREUD & ROTH	2, 4	(13, 13, 13, 13), (21, 21, 21, 21), (25, 25, 25,25), (23, 23, 23, 23)
8	Generalized Trig	2, 4, 10	(0.5, 5, …, 5),(5, 10, …, 10),(7,7, …, 7),(50, 50, …, 50)
9	Fletcher	2, 4, 10	(23, 23, …, 23), (45, 45, …, 45), (50, 5,…, 5), (70, 70, …,70)
10	Extended Penalty	2, 4, 10, 100	(0.5,5, …,5),(10,-0.5…,-0.5), (105,105, …,105), (130,130, …,130)
11	Raydan 1	2, 4, 10, 100	(1, 1, …,1), (3, 3, …, 3), (5, 5, …, 5), (-10, -10, …, -10)
12	Hager	2, 4, 10, 100	(3, -3, …, -3),(21, 21, …, 21), (-23, 23, …, 23), (23, 23, …, 23)
13	Rosenbrock	2, 4, 10, 100, 500, 1000, 10000	(7, 7, …, 7), (13, 13, …, 13), (23, 23, …, 23), (35, 35, …, 35)
14	Shallow	2, 4, 10, 100, 500, 1000, 10000	(21, -21, …, -21), (21, 21, …, 21), (50,50, …, 50),(130, 130, …, 130)
15	Tridiagonal 1	2, 4, 10, 100, 500, 1000, 10000	(0, 0,…, 0), (1, -1, …, -1), (17, -17, …, -17), (30, 30, …, 30)
16	Ext White & Holst	2, 4, 10, 100, 500, 1000, 10000	(-5, -5, …, -5), (2, -2, …, -2), (3, -3, …, -3), (7, -7, …, -7)
17	Ext Denschnb	2, 4, 10, 100, 500, 1000, 10000	(8, 8, …, 8), (11, 11, …,11), (12, 12, …, 12),(13, 13, …, 13)
18	Diagonal 4	2, 4, 10, 100, 500, 1000, 10000	(2, 2, …, 2), (5, 5, …,5), (10, 10, …, 10), (15, 15, …, 15)

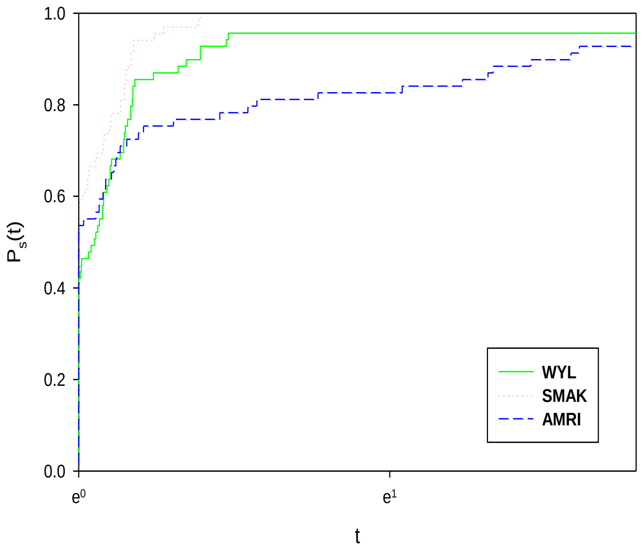

Fig.1: Performance profile based on the number of iteration

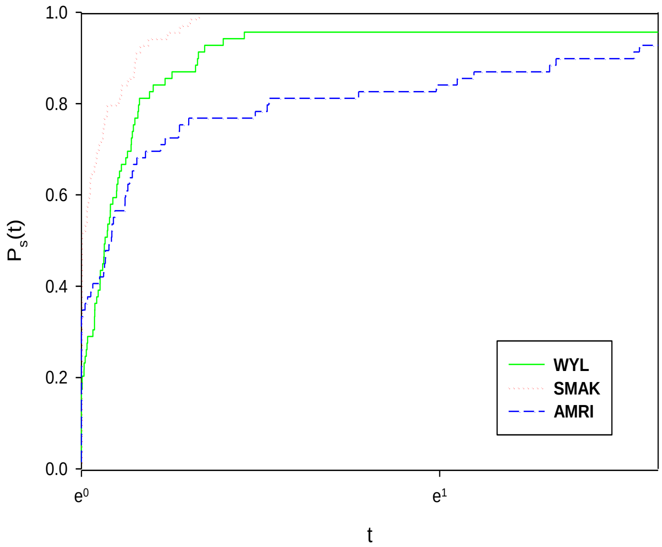

Fig.1: Performance profile based on CPU time

V. CONCLUSION

In this paper, we have examined a new hybrid method for solving unconstrained optimization. We showed that the new method satisfies the sufficient descent condition for all search direction under exact line search. The result of the numerical experiments show that the given method is competitive when compared to other classical conjugate gradient methods. In future, we hope to test this new method under different search rules.

ACKNOWLEDGMENT

The Author would like to thank the Government of Kano State for funding this study. We are also grateful to UNISZA for their consideration and comment.

REFERENCES

[1] J. C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization. SIAM J. Optim. 2(1), 1992, pp. 21–42

[2] Y. F. Hu and C. Storey, Global convergence result for conjugate gradient methods. J. Optim. Theory Appl. 71(2), 1991, pp. 399–405

[3] D. Touati-Ahmed, C. Storey, Efficient hybrid conjugate gradient techniques. J. Optim. Theory Appl.64(2), 1990, pp. 379–397

[4] Y. H. Dai, and Y. Yuan, An efficient hybrid conjugate gradient method for unconstrained optimization.Ann. Oper. Res. 103(1-4), 2001, pp. 33–47

[5] N. Andrei, Another hybrid conjugate gradient algorithm for unconstrained optimization. Numer. Algorithms.47(2), 2008, pp. 143–156

[6] N. Andrei, A hybrid conjugate gradient algorithm for unconstrained optimization as a convex combination of Hestenes–Stiefel and Dai–Yuan. Stud. Inform. Control 17(1), 2008, pp. 55–70

[7] N. Andrei, Hybrid conjugate gradient algorithm for unconstrained optimization. J. Optim. Theory Appl. 141(2), 2009, pp. 249–264

[8] N. Andrei, Accelerated hybrid conjugate gradient algorithm with modified secant condition for unconstrained optimization. Numer. Algorithms 54(1), 2010, pp. 23–46

[9] G. Zoutendijk, Nonlinear programming, computational methods, in Integer and Nonlinear Programming, J. Abadie, ed., North-Holland, Amsterdam, 1970, pp. 37-86..

[10] M. R. Hestenes and E. L. Stiefel, Methods of conjugate gradients for solving linear systems, J. Research Nat. Bur. Standard 49, 1952,pp. 409-436.
[11] R. Fletcher, Practical Method of Optimization, Vol. 1, Unconstrained Optimization, John Wiley & Sons, New York, 1987.
[12] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, Comput. J. 7, 1964, pp. 149-154.
[13] E. Polak and G. Ribière, Note sur la convergence de directions conjuguée, Rev. Francaise Informat Recherche,Operationell, 3e Année 16, 1969, pp35-43.
[14] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms part 1: Theory, J. Comput. Appl. Math. 69, 1992, pp. 129-137.
[15] E. Dolan, J.J. More, Benchmarking optimization software with performance profile, Math. Prog. 91, 2002, pp. 201–213
[16] X. Li, X. Zhao, A hybrid conjugate gradient method for optimization problems, Natural Science, 3, 2011, pp. 85-90.
[17] B. T. Polyak, The conjugate gradient method in extreme problems, USSR Comp. Math. Math. Phys. 9, 1969, pp. 94-112.
[18] M. Al-Baali, Descent property and global convergence of the Fletcher-Reeves method with inexact line search, IMA J. Numer. Anal. 5, 1985, pp. 121-124.
[19] A. Abashar et al, Global convergence properties of a new class of conjugate gradient method for unconstrained optimization, Appl. Math. Sci. 8, 2014, pp. 3307-3319
[20] Z. Wei, S. Yao, and L. Liu, The convergence properties of some new conjugate gradient methods. Applied Math and Comp, 183, 2006, pp. 1341-1350.
[21] Y. Dai, and Y. Yuan, A nonlinear conjugate gradient with a strong global convergence properties. SIAM Journal on Optimization, 10, 2000, pp. 177-182.
[22] Y. Dai, and Y. Yuan, Nonlinear conjugate gradient methods. Shanghai Scientific and Technical Publishers, Shanghai. 1998
[23] G. L. Yuan, Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. Optim Letters, 3, 2009, pp. 11-21.
[24] N. Andrei, An unconstrained optimization test functions collection, Advanced Modelling and optimization 10(1), 2008, pp. 147-161.