International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:8, No:1, 2014

A New Modification of Nonlinear Conjugate Gradient Coefficients with Global Convergence Properties

Ahmad Alhawarat, Mustafa Mamat, Mohd Rivaie, Ismail Mohd

***Abstract*—**Conjugate Gradient (CG) method has been enormously used to solve large scale unconstrained optimization problems due to the number of iteration, memory, CPU time, and convergence property, in this paper we proposed a new class of nonlinear conjugate gradient coefficient with global convergence properties proved by exact line search. The numerical results for our CG method new present an efficient numerical result when it compared with well-known formulas.

***Keywords*—**Conjugate gradient method, conjugate gradient coefficient, global convergence.

I. Introduction

NONLINEAR conjugate gradient method (CG) is useful method to find the minimum value of function for unconstrained optimization problems. Let us consider the following form:

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

where $f:R^n\to R$ is continuously differentiable and its gradient is denoted by $g(x)=\nabla f(x)$, the method to find a sequence of points $\{x_k\}$ starting from initial point $\{x\in R^n\}$ is given by iterative formula:

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

where x_k is the current iteration point and $\alpha_k>0$ is the step size obtained by some line search. In this paper we used exact line search which is,

$$f(x_k+\alpha_k d_k)=\min f(x_x+\alpha d_k),\alpha\geq 0.\tag{3}$$

Many researchers do not prefer to study this method, because it is very slow especially when the initial point is far away from the optimal solution point, so it is very expensive when compared with inexact line search, but our formula gives a good numerical result by using fast computer processors which is an advantage for exact line search method.

The search direction d_k is defined by:

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

where $g_k=g(x_k)$ and β_k is scalar known as the conjugate gradient coefficient. The most well know formulas for β_k are as follows: (Hestenses–Stiefel (HS) [1]), (Fletcher–Reeves (FR) [2]), (Polak-Ribiere–Polyak (PR) [3]), (Conjugate Descent (CD) [4]), (Liu–Storey (LS) [5]), (Dai–Yuan, (DY) [6]), (Wei et al. (WYL) [7]). (Mohd Rivaie, Mustafa Mamat, Ismail Mohd (RMIL) [17]).

$$\beta_k^{FR}=\frac{g_k^T g_k}{g_{k-1}^T\,g_{k-1}},$$

$$\beta_k^{PRP}=\frac{g_k^T(g_k-g_{k-1})}{g_{k-1}^T g_{k-1}},$$

$$\beta_k^{CD}=-\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}},$$

$$\beta_k^{HS}=\frac{g_{k+1}^T(g_{k+1}-g_k)}{d_k^{\,T}(g_{k+1}-g_k)},$$

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

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

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

The global convergence of FR method with exact line search was achieved by [8]; its behavior on numerical

Ahmad Alhawarat is with School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Terengganu, Malaysia (e-mail: abadee2010@yahoo.com).
Mustafa Mamat is with Fakulti Informatik dan Komputeran, Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia (e-mail: must@unisza.edu.my).
Ismail Mohd is with Department of Mathematics, Faculty of Science and Technology, Universiti Malaysia Terengganu (UMT), Malaysia (e-mail: ismail@umt.edu.my).
Mohd Rivaie is with Department of Computer Science and Mathematics, Univesiti Teknologi Mara (UiTM) Terengganu, Campus Kuala Terengganu, Malaysia (e-mail: rivaie75@yahoo.com).

computation is unpredictable. Sometimes it is as efficient as PRP method, nevertheless most time it is very slow, also DY and CD have the same performance with exact line search.

Global convergence of PRP method for convex objective function under exact line search was proved by Polak and Ribiere in 1969 [3], in the other hand Powell gave out a counter example which shows that there exist non convex function, which PRP method does not converge globally even though the exact line search is used. After that Powell 1986 suggested that it is very important to achieve the global convergence of β_k should not be negative. Gilbert and Nocedal [9] proved that nonnegative PRP method is globally convergent with the Wolfe-Powel line search, but it is still open for Strong Wolf condition. HS method and LS method have the same performance as PRP method with exact line search. Therefore, PRP method is the most efficient method compare to the other conjugate gradient methods, there has been much research on convergence of these methods you can see [10]-[14].

Recently [8] gave a new β_k which is a variant of PRP method. It seems like original PRP method which has been studied in both exact line search and inexact line search, and many modifications appeared as;

$$\beta_K^{VHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \text{ [15]},$$

$$\beta_K^{NRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} \left\|g_k^T g_{k-1}\right\|}{\|g_{k-1}\|^2} \text{ [16]},$$

$$\beta_K^{NHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})} \text{ [16]},$$

Without loose of generating in Sections II and III we will present our new formula, the algorithm, sufficient descent condition, and the global convergence properties with its proof. The numerical results and discussion will be presented in Section IV. Finally, the conclusions are presented in Section V.

II. The New Formula

In this section we present our new β_k^{AMR*}, where $AMR*$ denotes to Ahmad, Mustafa, and Rivaie which is extended for β_k^{WYL} method with coefficient, that is,

$$\beta_K^{AMR*} = \frac{g_k^T (m.g_k - g_{k-1})}{m.(g_{k-1}^T g_{k-1})}, \tag{5}$$

where $m_k = \frac{\|g_{k-1}\|}{\|g_k\|}$, and $\|.\|$ means the Euclidean norm.

The algorithm is given as follows:

1st Step: Initialization. Given $x_0 \in R^n$, set $k = 0$.

2nd Step: Compute β_k based on (5).

3rd Step: Compute d_k based on (4), If $\|g_k\| = 0$, then stop.

4th Step: Compute α_k based on (3).

5th Step: Updating new point based on (2).

6th Step: Convergent test and stopping criteria.

If $\|g_k\| \le \varepsilon$ then stop. Otherwise go to Step2 with $k = k+1$

III. Convergent Analysis of AMR* Method

In this section, the convergent properties of β_k^{AMR*} is studied, for above algorithm to be convergent, it should have fulfilled the sufficient descent condition and the global convergence properties.

A. Sufficient Descent Condition

For the Sufficient descent condition to hold,

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

Theorem 1. Consider a CG method with the search direction (4) and β_k^{AMR*} given as (5), then condition (6) holds for all $k \ge 0$ and $c > 0$.

Proof. From (4) we have if $k = 0$, the $g_0^T d_0 = -c\|g_0\|^2$. for $k \ge 1$, we need to multiply (4) by g_{k+1}^T and set $k = k+1$ then we have,

$$g_{k+1}^T d_{k+1} = g_{k+1}^T(-g_{k+1}^T + \beta_{k+1} d_k) = -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k. \tag{7}$$

for exact line search easy to know $g_{k+1}^T d_k = 0$. Thus

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2. \tag{8}$$

Thus, sufficient descent direction holds. The proof is completed.

B. Global Convergence Properties

We need to show that β_k^{AMR*} is globally convergent under exact line search, before we start we need to simplify β_k^{AMR*} to

$$\beta_k^{AMR*} = \frac{g_k^T \left(\frac{\|g_{k-1}\|}{\|g_k\|} g_k - g_{k-1}\right)}{\frac{\|g_{k-1}\|}{\|g_k\|} (g_{k-1}^T g_{k-1})}$$

show that $\beta_k^{AMR*} \ge 0$ and $\beta_K^{AMR*} \le \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}$, by using Cauchy - Schwartz inequality, we have

$$\beta_K^{AMR*} = \frac{g_k^T\left(\frac{\|g_{k-1}\|}{\|g_k\|}g_k - g_{k-1}\right)}{\frac{\|g_{k-1}\|}{\|g_k\|}(g_{k-1}^T g_{k-1})} \geq \frac{\|g_k\|^2\|g_{k-1}\| - \|g_k\|\|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^3} = 0.$$

Thus we get

$$\beta_k^{AMR*} \geq 0. \quad (9)$$

and,

$$\beta_k^{AMR*} = \frac{g_k^T\left(\frac{\|g_{k-1}\|}{\|g_k\|}g_k - g_{k-1}\right)}{\frac{\|g_{k-1}\|}{\|g_k\|}(g_{k-1}^T g_{k-1})}$$

$$\leq \frac{\|g_k\|^2\|g_{k-1}\| + \|g_k\|\|g_k^T\|\|g_{k-1}\|}{\|g_{k-1}\|^2} \leq \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}$$

which implies

$$\beta_K^{AMR*} \leq \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}. \quad (10)$$

The following assumptions are needed to use in our proof.

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 and Ω is bounded.

Assumption B. In some neighborhood N of Ω, f is continues and differentiable, and its gradient is Lipchitz continues, that is, for any $x, y \in N$, there exists a constant $L \geq 0$ such that:

$$\|g(x) - g(y)\| \leq L\|x - y\|.$$

Under the above assumptions, we have the following Lemma:

Lemma 1. Suppose the Assumptions A and B hold true, consider any form of (4), for all k and α_k satisfied (3) the following condition, known as the Zoutendijk condition holds $\sum_{k=0}^{\infty}\frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$, the proof of this Lemma can be seen [8] .

By substitute (8) in Zoutendijk condition then it is equivalent to,

$$\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2} < \infty. \quad (11)$$

C. Angle Condition

Theorem 2. Suppose that Assumptions A and B hold, and the sequence $\{x_x\}$ is generated by aforementioned Algorithm, if $\|x_{k+1} - x_k\| \to 0$ while $k \to \infty$, then

$$\liminf_{k\to\infty}\|g_k\| = 0. \quad (14)$$

Proof. We use proof by contradiction, firstly consider θ_k is the angle between d_k and the steepest descent search direction $-g_k$, where

$$\cos\theta_k = \frac{-g_k^T d_k}{\|g_k\|\|d_k\|}. \quad (15)$$

By using (4) and (15) we indicate the following relations,

$$\|d_k\| = \sec\theta_k\|g_k\| \quad (16)$$

$$\beta_{k+1}\|d_k\| = \tan\theta_{k+1}\|g_{k+1}\| \quad (17)$$

Combining (16) and (17), indicate

$$\tan\theta_{k+1} = \frac{g_{k+1}^T\left(\frac{\|g_k\|}{\|g_{k+1}\|}g_{k+1} - g_k\right)}{\frac{\|g_k\|}{\|g_{k+1}\|}\|g_k\|^2}\sec\theta_k\frac{\|g_k\|}{\|g_{k+1}\|} \quad (18)$$

$$\tan\theta_{k+1} \leq \frac{\|g_{k+1}\|\left\|\frac{\|g_k\|}{\|g_{k+1}\|}g_{k+1} - g_k\right\|}{\|g_k\|^2}\sec\theta_k$$

$$\tan\theta_{k+1} \leq \frac{\|g_{k+1}\|\left\|\frac{\|g_k\|}{\|g_{k+1}\|}g_{k+1} - g_{k+1} + g_{k+1} - g_k\right\|}{\|g_k\|^2}\sec\theta_k$$

$$\tan\theta_{k+1} \leq \frac{\|g_{k+1}\|\|\|g_k - g_{k+1}\| + \|g_{k+1} - g_k\|}{\|g_k\|^2}\sec\theta_k$$

$$\tan\theta_{k+1} \leq \frac{2\|g_{k+1}\|\|g_{k+1} - g_k\|}{\|g_k\|^2}\sec\theta_k$$

Suppose (14) does not hold true, then for all k, there exist $\varepsilon > 0$, such that

$$\|g_k\| \geq \varepsilon \quad (19)$$

By $\|x_{k+1} - x_k\| \to 0$ and Lipschitz condition

$$\|g_{k+1} - g_k\| \leq \varepsilon.$$

Since

$$\|g_{k+1}\| - \|g_k\| \leq \|g_{k+1} - g_k\|$$

This imply that

$$\|g_{k+1}\| \le 2\varepsilon .$$

So

$$\tan\theta_{k+1} \le 4\sec\theta_k .$$

Since $\sec\theta_k \ge \tan\theta_k$, for all $\theta_k \in [0, \frac{\pi}{2})$, we have $\tan\theta_{k+1} \le \frac{4}{\tan\theta_k}$

Therefore the angle between d_k and the steepest descent direction $-g_k$ is bounded away from $\frac{\pi}{2}$, so from (11), (12), and (17) we have,

$$\sum_{k=0}^{\infty}\frac{\|g_k\|^4}{\|d_k\|^2} = \sum_{k=0}^{\infty}\|g_k\|^2(\cos\theta_k)^2 < \infty$$

This implies $\lim_{k\to\infty}\inf\|g_k\| = 0$, which contradicts (14). The proof is complete.

D. Linear Convergence Rate

In this section, we shall discuss linear convergence rate for AMR* method under this convergence, we need the following necessary assumption to prove.

Assumption C. The sequence (2) where α_k generated by the exact line search d_k and β_k generated by (4) and (5) respectively, converges to x^*. In addition, $\nabla^2 f(x^*)$ is a symmetric positive definite matrix and twice continuously differentiable on $N(x^*, \varepsilon_0) = \{x \,|\, \|x - x^*\| < \varepsilon_0\}$.

The conclusion of the following Lemma is used to prove the linear convergence of nonlinear conjugate gradient methods. The proof can be seen given in [18], [19].

Lemma 2. If Assumption 2 holds true, then m, M and ε_1 exist with $0 < m \le M$ and $\varepsilon < \varepsilon_1$ such that,

$$\begin{aligned}
& m\|y\|^2 \le y^T\nabla f^2(x)y \le M\|y\|^2 \\
& \forall x, y \in N(x^*, \varepsilon), \\
& \frac{1}{2}m\|x - x^*\|^2 \le f(x) - f(x^*) \le \frac{1}{2}M\|x - x^*\|^2 \\
& \forall x, \in N(x^*, \varepsilon), \\
& m\|x - y\|^2 \le (g(x) - g(y))^T(x - y) \le M\|x - y\|^2 \\
& \quad \forall x, y \in N(x^*, \varepsilon) . \qquad (20)
\end{aligned}$$

Thus we get

$$m\|x - y\|^2 \le g(x)^T(x - x^*) \le M\|x - x^*\|^2$$
$$\forall x \in N(x^*, \varepsilon) . \qquad (21)$$

Using Cauchy–Schwartz inequality, (20) and (21) we obtain

$$m\|x - x^*\| \le \|g(x)\| \le M\|x - x^*\|$$
$$\forall x, y \in N(x^*, \varepsilon) .$$

and

$$\|g(x) - g(y)\| \le M\|x - x^*\|$$
$$\forall x, y \in N(x^*, \varepsilon) .$$

Lemma 3. Supposed Assumption C holds true, and let θ_k be the angle between $-g_k$ and d_k, the sequence x_k is generated by the exact line search and d_k is a descent direction. If a constant $\eta > 0$ exist for which,

$$\prod_{i=0}^{k-1}\cos\theta_i \ge \eta^k \qquad (22)$$

then a constant a > 0 and $r \in (0,1)$ exist, such that ,

$$\|x_{k+1} - x^*\| \le ar^{k+1} .$$

Hence x_k will converge to x^* at least R-linearly. The proof for this Lemma can be seen from [20].

Theorem5. If Assumption C holds true, then constants $a > 0$ and $r \in (0,1)$ exists such that the sequence generated by (2), (4) and (1) using the exact line search satisfies,

$$\|x_k - x^*\| \le ar^k , \qquad (23)$$

hence, x_k will converge to x^* at least R-linearly.

Proof. If Assumption C hold true, then we assume, $\forall x_0 \in N(x^*, \varepsilon)$, therefore from (4) and (10) we indicate

$$\begin{aligned}
\|d_k\| &\le \|g_k\| + \beta_k\|d_{k-1}\| \le \|g_k\| + \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}\|d_{k-1}\| \\
&\le \|g_k\| + \frac{2\|g_k\|^2}{\|d_{k-1}\|} \le \left(1 + \frac{2\|g_k\|}{\|d_{k-1}\|}\right)\|g_k\|
\end{aligned}$$

thus,

$$\begin{aligned}
\cos\theta_k &= \frac{-g_k^T d_k}{\|g_k\|\cdot\|d_k\|} = \frac{\|g_k\|^2}{\|g_k\|\|d_k\|} \\
&\ge \left(1 + \frac{2\|g_k\|}{\|d_{k-1}\|}\right)^{-1} > 0
\end{aligned}$$

Hence, Lemma 3 holds true. By (22) we can obtain (23). The proof is completed.

IV. Numerical Results and Discussions

In this section, we use some test problems to find the computational results to analyze the efficiency of AMR*. We performed a comparison with other CG methods, including PR and WYL. The tolerance ε is selected as equal to 10^{-5} for all algorithms to investigate how rapidly the iteration of these algorithm towards the optimal solution, also the gradient value as the stopping criteria. Hence the stopping criteria are set $\|g_k\| \le 10^{-5}$, the test functions can be found on many trusted web sites with a lot of cods for exact and inexact line search made by Fortran, Matlab, C, and C++, Hager and Andrei as an example and others. We used Maple 13 subroutine program, with CPU processor Intel (R) Core (TM), i7 CPU, and 4GB RAM memory. The performance results are shown in Figs. 1 and 2, respectively, using a performance profile introduced by Dolan and More [20]. In this performance profile, they introduced the notion of a means to evaluate and compare the performance of the set solvers S on a test set P Assuming n_s solvers and n_p problems exists for each problem P and solver S, they define $t_{p,s}$ = computing time (the number of iterations or CPU time or others) required to solve problems p by solver s

Requiring a baseline for comparisons, they compared the performance on problem p by solver s with the best performance by any solver on this problem using the performance ratio

$$t_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in S\}}.$$

Suppose that a parameter $r_M \ge r_{p,s}$ for all $r_{p,s} = r_M$ is chosen, and if and only if solver s does not solve problem p. The performance of solvers s on any given problem might be of interest, but because we would like to obtain an overall assessment of the performance of the solver, then it was defined

$$\rho_s(t) = \frac{1}{n_p} size\{p \in P : r_{p,s} \le t\},$$

Thus $\rho_s(t)$ was the probability for solver $s \in S$ that a performance ratio $r_{p,s}$ was within a factor $t \in R$ of the best possible ration, function p_s was the cumulative distribution function for perform- ance ratio, the performance profile $p_s : R \to [0,1]$ for a solver was a non-decreasing piecewise, and continuous from the right. The value of $p_s(1)$ is the probability that the solver will win over the rest of the solvers. In general, a solver with high values of $p(t)$ or at the top right of the figure are preferable or represent the best solver.

Figs. 1 and 2 show that AMR*is best performance, since it can solve all test problems and reach 100%. Although the performance of PRP seems to be much better than AMR*, but it can solve only 90% and WYL solved 87% and it seems less than AMR* for both performance. Hence we considered AMR* as the best and superior method above since that can solve all problems.

V. Conclusion

In this paper, we have proposed a new and simple β_k that is easy to implement. Our numerical results have shown that, our new method has the best performance compared to the other standard CG methods. We have also provide proof showing that this method converges globally with a linear convergence rate, in future we intend to test β_k^{AMR*} using inexact line search under strong Wolf condition, Armijo line search and Grippp-Lucidi line search, also we can make some modification to improve β_k^{AMR*} to get results characterized by speed, accuracy, and less space in memory using various types line search.

TABLE I
A List of Problem Functions

No.	Function	Number of variables	Initial points
1	Beale	10, 100, 500, 1000	(13,13,…,13), (30,30…,30), (50,50…,50)
2	Colville	4	(-10,-10,-10,-10), (-5,-5,-5,-5), (5,55,5), (10,10,10,10)
3	Extended Himmelblau	10, 100, 500, 1000	(1,1,…,1), (5,5,…,5), (10,10…,10),(100,100,…,100)
4	Generalize Quadratic	10, 100, 500, 1000	(1,1,…,1), (5,5,…,5), (10,10…,10),(100,100,…,100)
5	Generalized Tridiagonal	10, 100, 500, 1000	(2,2,…,2), (5,5,…,5), (10,10…,10),(100,100,…,100)
6	Goldstein-Price's	2	(3,-3), (5,-5), (10,-10), (25,-25)
7	liarwhd	10, 100	(2,2,…,2), (4,4,…,4), (10,10,…,10),(100,100,…,100)
8	Rosenbrock	10, 100, 500, 1000	(2,2…,2), (5,5…,5), (10,10…,10),(100,100…,100)
			(2,2…,2), (5,5…,5), (10,10…,10)
9	Three-hump	2	(-1,1), (5,5), (-5,5)
10	White-Holst	10, 100, 500	(2,2…,2), (5,5…,5), (10,10…,10)
11	Fletcher	10, 100, 500, 1000	(2,2…,2), (5,5…,5), (10,10…,10)
12	Extended Freudenstein and Roth	10, 100, 500, 1000	(2,2…,2), (5,5…,5), (10,10…,10)
13	Powell	10, 100,500	(5,5…,5), (10,10…,10), (15,15…,15)
14	Extended Tridiagonal 1	10, 100, 500, 1000	(1,1…,1), (5,5…,5), (10,10…,10)
15	Extended Tridiagonal 2	10, 100, 500, 1000	(1,1…,1), (5,5…,5), (10,10…,10)
16	Extended wood	10, 100, 500, 1000	(-1,-1,…,-1), (5,5,…,5), (10,10…,10)
17	Extended denschnf	10, 100, 500, 1000	(1,1,…,1), (5,5,…,5), (10,10…,10), (100,100,…,100)

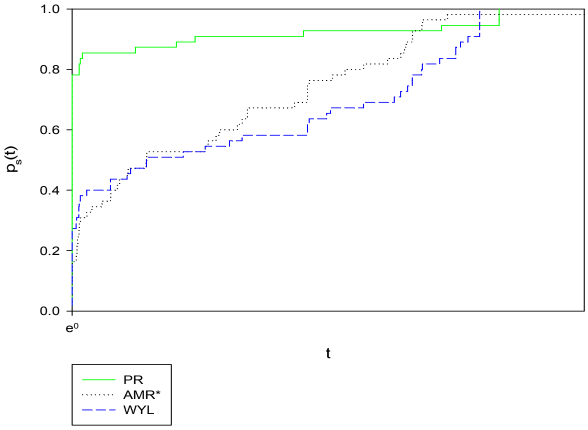

Fig. 1 Performance profile based on the number of iteration

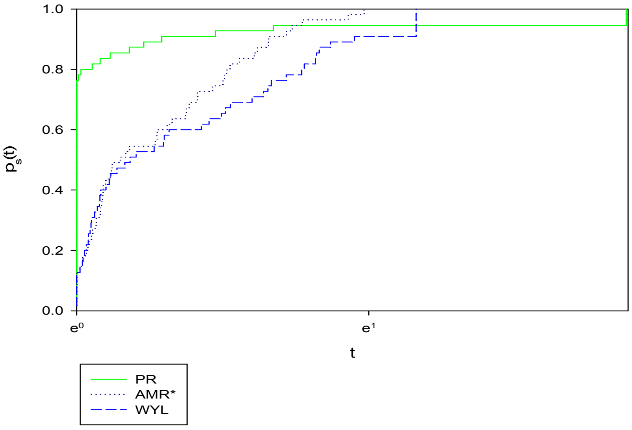

Fig. 2 Performance profile based on the CPU time

ACKNOWLEDGMENT

The authors are grateful to the editor and the anonymous reviewers for their comments and suggestions which improved this paper substantially. They would also like to thank The Ministry of Education Malaysia (MOE) for funding this research under The Fundamental Research Grant Scheme (Grant no. 59256).

REFERENCES

[1] M.R. Hestenes, E.L. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Na- tl. Bur. Stand. Sec. B 49 (1952), 409–432.
[2] R. Fletcher, C. Reeves, Function minimization by conjugate gradients, comput. J. 7 (1964), 149–1 -54.
[3] B. Polak, G. Ribiére, Note surla convergence des méthodes de directions conjuguées, Rev. Fr. Aut- om. Inform. Rech. Oper.,3e Année. 16 (1969), 35–43.
[4] R. Fletcher, Practical method of optimization, vol 1, unconstrained optimization, John Wiley & Sons, New York, 1987.
[5] Y.L. Liu, C.S. Storey, Efficient generalized conjugate gradient algorithms, part 1: theory, J. Optim. Theory Appl. 69 (1991), 129–137.
[6] R. Fletcher, and C. Reeves, Function minimization by conjugate gradients, Comput. J. 7(1964), 149-154.
[7] Z. Wei, S. Yao, L. Liu, The convergence properties of some new conjugate gradient methods, App-l. Math. Comput. 183 (2006), 1341–1350.
[8] G. Zoutendijk, Nonlinear progrAMRing computational methods, in: J. Abadie (Ed.), Integer and N- onlinear ProgrAMRing, North-Holland, Amsterdam (1970), 37–86.
[9] J.C. Gilbert, J. Nocedal, Global convergence properties of conjugate gradient methods for optimi- zation, SIAM J. Optimizat. 2 (1) (1992), 21–42.
[10] Cheng, W.Y.: A two-term PRP-based descent method. Numer. Funct. Anal. Optim. 28(2007)1217–1230
[11] Dai, Z.F., Tian, B.S.: Global convergence of some modified PRP nonlinear conjugate gradient methods. Opt. Lett. (2010), doi:10.1007/s11590-010-0224-8
[12] Yu, G.H., Zhao, Y.L., Wei, Z.X.: A descent nonlinear conjugate gradient method for largescale unconstrained optimization. Appl. Math. Comput. 187 (2007), 636–643
[13] Zhang, L., Zhou, W., Li, D.: A descent modified Polak-Ribi-re-Polyak conjugate gradient method and its global convergence. IMA J. Numer. Anal. 26(2006), 629–640
[14] Zhang, L., Zhou, W., Li, D.: Some descent three-term conjugate gradient methods and their global convergence. Optim. Methods Softw. 22 (2007), 697–711
[15] Z. Wei, S. Yao, L. Liu, The convergence properties of some new conjugate gradient methods, Appl. Math. Comput. 183 (2006), 1341–1350.
[16] L. Zhang, An improved Wei–Yao–Liu nonlinear conjugate gradient method for optimization computation, Appl. Math. Comput. 215 (2009), 2269–2274.
[17] Mohd Rivaie, Mustafa Mamat, Ismail Mohd, Leong Wah June, A new class of nonlinear conjugate gradient coefficients with global convergence properties, Appl.Math. Comput218 (2012), 11323-11332.
[18] Z.F. Dai, Two modified HS type conjugate gradient methods for unconstrained optimization problems, Nonliner Anal. 74 (2011), 927–936.
[19] G. Yuan, X. Lu, and Z. Wei, *A conjugate gradient method with descent direction for unconstrained optimization*, J. Comp. App. Maths. 233(2009), 519-530.
[20] E. Dolan, J.J. More, Benchmarking optimization software with performance profile, Math. Prog. 91 (2002), 201–213.