Malaysian Journal of Computing and Applied Mathematics 2018, Vol 1(1): 29-38

(Online)

Malaysian Journal of
Computing and Applied Mathematics

New Hybrid Conjugate Gradient Method with Global Convergence Properties for Unconstrained Optimization

***Yasir Salih[a], Mohamed Ahmed Hamoda[b] and Mohd Rivaie[c]**

[a]Department of Mathematics, Faculty of Education, Red Sea University, Sudan
[b]Department of Mathematics, Faculty of Science, Al-Asmarya Islamic University, Libya
[c]Department of Computer Science and Mathematics, Universiti Teknologi MARA, (UiTM)Terengganu Campus Kuala Terengganu, Malaysia

yasirsalih2015@gmail.com

Received: 11/04/2018, Accepted: 15/06/2018

Abstract

Nonlinear conjugate gradient (CG) method holds an important role in solving large-scale unconstrained optimization problems. In this paper, we suggest a new modification of CG coefficient β_k that satisfies sufficient descent condition and possesses global convergence property under strong Wolfe line search. The numerical results show that our new method is more efficient compared with other CG formulas tested.

Keywords: conjugate gradient method, large-scale, global convergence, strong Wolfe line search, unconstrained optimization.

Introduction

The general form of an unconstrained optimization problem is defined by

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

where $f: R^n \to R$is a continuously differentiable function and its gradient $g \equiv \nabla f(x)$is available. The iterative formula of the CG method is given by

$$X_{k+1} = x_k + \alpha_k d_k \quad k = 0,1,2,\dots, \tag{2}$$

where α_k the step-size computed by carrying out strong Wolfe line search procedure, defined as follows

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \tag{3}$$

and

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \le \sigma \left| g_k^T d_k \right| \tag{4}$$

where $0 < \delta < \sigma < 1$ The parameterd_k is the search direction defined by

$$d_k = \begin{cases} -g_k, & if \;\; k = 0 \\ -g_k + \beta_k d_{k-1}, & if \;\; k \ge 1 \end{cases} \tag{5}$$

where $\beta_k \in R$ a scalar known as the CG coefficient. Examples of most well-known classical formulas for β_k are Hestenes-Stiefel (HS) (Hestenes and Stiefel,1952), Fletcher-Reeves (FR) (Fletcher and Reeves, 1964), Polak-Ribiere-Polyak (PRP) (Polak and Ribiere, 1969), Conjugate Descent (CD) (Fletcher, 1980), Liu-Storey (LS) (Liu and Storey, 1991), and Dai-Yuan (DY) (Dai and Yuan, 1999). The parameters of these β_k are given as follows:

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

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

The convergence of CG method under different line searches has been studied by many authors such as Al-Baali (1985), Gilbert and Nocedal (1992), Liu and Storey (1991), Zoutendijk (1970), Touati-Ahmed and Storey (1990), and Andrei (2008). For further information, readers can refer to Abashar et al. (2017), Aini et al. (2017), Ghani et al. (2017a), Ghani et al. (2017b), Hamoda et al. (2016), Kamfa et al. (2017), Mohamed et al. (2017), Omer et al. (2015), Osman et al. (2017), Rivaie et al. (2015), and Zull et al. (2017).

Modified Formula and Algorithm

Recently, Wei et al. (2006) gave a variant of the PRP method which is called the WYL method, written as

$$\beta_k^{WYL} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2}. \tag{6}$$

The WYL method and PRP methods both come with restart properties. Zhang (2009) studied and improved WYL CG method and suggested the NPRP method, formulated as

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{\|g_{k-1}\|^2}. \tag{7}$$

Zhang (2009) proved that the NPRP method satisfies descent condition under strong Wolfe line search. Later, Dai and Wen (2012) proposed a modified NPRP method as follows:

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{\mu\left|g_k^T d_{k-1}\right| + \|g_{k-1}\|^2}, \qquad \mu > 1 \tag{8}$$

Based on the above ideas, we present a new $\beta_k$known as β_k^{YHM}, where YHM denotes Yasir, Hamoda and Mamat. The formula for β_k^{YHM} is defined by

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} & if \ \ 0 \le g_k^T g_{k-1} \le \|g_k\|^2 \\ \dfrac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2} & \text{otherwise} \end{cases} \tag{9}$$

The following algorithm is a general algorithm for solving optimization by CG methods.

Algorithm 2.1:
Step1: Given an initial point$x_0 \epsilon R^n$, $\varepsilon > 0$, set $d_0 = -g_0$, $k = 0$
Step2: Compute$\beta_k$by formula (9)
Step3: Compute $d_k$based on (3).If $g_k = 0$, then stop.
Step4: Compute $\alpha_k$by inexact line search.
Step5: Update new point based on (2)

Step6: Convergence test and stopping criteria. If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| \le \epsilon$, then stop. Otherwise, set $k = k + 1$ and go to Step 1.

Global Convergence analysis

In this section, we study the global convergence properties of β_k^{YHM}, starting with the sufficient descent condition. Firstly, we need to simplify β_k^{YHM} so that the proving steps will be easier. From (9), we know that:

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

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

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

Otherwise,

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

By Cauchy - Schwartz inequality, it is implied that

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

Hence, we can deduce that for both cases of $0 \le g_k^T g_{k-1} \le \|g_k\|^2$ and otherwise,

$$\beta_k^{YHM} \ge 0 . \tag{10}$$

Sufficient descent condition

The sufficient descent condition is defined by:

$$g_k^T d_k \le -c\|g_k\|^2 \quad \text{for } k \ge 0 , c > 0 \quad . \tag{11}$$

The following theorem shows that YHM with inexact line search possesses the sufficient descent property.

Theorem 1. Suppose that the sequence $\{g_k\}$ and $\{d_k\}$ are generated by Algorithm (2.1) and the step-length α_k is determined by strong Wolfe line search. If $g_k \ne 0$, then the sequence $\{d_k\}$ satisfies the sufficient descent condition for all $k \ge 0$.

Proof. The proof of the descent property of $\{d_k\}$ is by induction. Firstly, we prove the theorem for the case of $k = 0$.

Case (1): If $0 \le g_k^T g_{k-1} \le \|g_k\|^2$, then

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

Since $g_0^T d_0 = -\|g_0\|^2 < 0$, then condition (11) is fulfilled for $k = 0$.

Now suppose that $d_i, i = 1,2,3, \dots , k$ are all descent directions that is $g_i^T d_i < 0$. From the strong Wolfe condition, and (10)

$$|\beta_{k+1}^{YHM} g_{k+1}^T d_k| \le \sigma \frac{\|g_{k+1}\|^2}{\|g_k\|^2} |g_k^T d_k|. \tag{12}$$

Now, we multiply $d_{k+1} = -g_{k+1} + \beta_{k+1}^{YHM} d_k$ with g_{k+1}^T to get

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

We divide both sides by $\|g_{k+1}\|^2$, which gives us

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1}^{YHM} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}. \quad (13)$$

Since $g_k^T d_{k+1} < 0$, then from (12), we have

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

Hence,

$$\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\sigma g_k^T d_k \le \beta_{k+1}^{YHM} g_{k+1}^T d_k \le -\frac{\|g_{k+1}\|^2}{\|g_k\|^2}\sigma g_k^T d_k \quad (14)$$

Substitute (13) into (14),then

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

By repeating this process and taking into account that $g_0^T d_0 = -\|g_0\|^2$, we get

$$-\sum_{j=0}^{k} \sigma^j \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \sum_{j=0}^{k} \sigma^j \quad (15)$$

Since $\sum_{j=0}^{k} \sigma^j < \sum_{j=0}^{\infty} \sigma^j = \frac{1}{1-\sigma}$, equation (15) can be written as

$$-\frac{1}{1-\sigma} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2 + \frac{1}{1-\sigma} \quad (16)$$

By making the restriction $\sigma \epsilon (0, \frac{1}{2})$, we can see that $g_{k+1}^T d_{k+1} < 0$. Therefore, by induction, $g_k^T d_k < 0$ holds for all $k \ge 0$. Substitute $c = 2 - \frac{1}{1-\sigma}$, $0 < c < 1$ into (16) and we get $(c-2)\|g_k\|^2 \le g_k^T d_k \le -c\|g_k\|^2$.

This implies that condition (11) holds. The proof is completed.

Case (2): when $\beta_k^{YHM} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2}$, the proof of this theorem can be seen in (Wei et al., 2006).

Global Convergence Properties

The following assumptions are often used in the studies of the CG method.

Assumption 1

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

B. In some neighbourhood N of Ω, the objective function is continuously differentiable and its gradient is Lipchitz continuous, that is, there exists constant $L > 0$ such that:

$$\|g(x) - g(y)\| \le L\|x - y\|, \qquad \forall x, y \in N \quad (17)$$

In 1992, Gilbert and Nocedal introduced property (*) which plays an important role in the studies of CG method. This property means that the following search direction automatically approaches the steepest direction when a small step-length is generated, and the step-length are not produced successively (Zhang et al., 2012).

Property (*)

Consider a CG method of the form (2) and (3). Suppose that for all $k \ge 1$,

$$0 < \gamma \le \|g_k\| \le \gamma^- \quad (18)$$

where γ and γ^- are two positive constants. The method has property (*) if there exist constants $b > 1$ and $\lambda > 0$ such that for all k: $|\beta_k| \le b$, $\|s_k\| \le \lambda$ implies $\|\beta_k\| \le \frac{1}{2b}$ where $s_k = \alpha_k d_k$.

The following lemma shows that the new parameter β_k^{YHM} possesses property (*).

Lemma1. Consider the method of form (2) and (3), and suppose that Assumption 1 holds, then the CG method withβ_k^{YHM} has property (*).

Proof: Case 1: If$0 \leq g_k^T g_{k-1} \leq \|g_k\|^2$, then $\beta_k^{YHM} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2}$.

Set $b = \frac{\gamma^{-2}}{\gamma^2} > 1$, $\lambda = \frac{\gamma^2}{2L\gamma^{-2}b}$.

By (8) and (18) $\left|\beta_k^{YHM}\right| = \frac{|g_k^T(g_k - g_{k-1})|}{\|g_{k-1}\|^2} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \leq \frac{\gamma^{-2}}{\gamma^2} = b$.

By Assumption 1,if$\|s_k\| \leq \lambda$, then$\left|\beta_k^{YHM}\right| = \frac{\|g_k\|\|g_{k-1}\|}{\|g_{k-1}\|^2} \leq \frac{L\lambda\gamma^{-2}}{\gamma^2} = \frac{1}{2b}$.

The proof is complete.

Case 2: When$\beta_k^{YHM} = \frac{g_k^T\left(g_k - \frac{\|g_k\|}{\|g_{k-1}\|} g_{k-1}\right)}{\|g_{k-1}\|^2}$,the proof of this theorem can be seen in(Wei et al., 2006).

Lemma 2. Suppose that Assumption 1holds and$x_k$is generated by Algorithm 2.1 where$d_k$satisfies $g_k^T d_k < 0$ for all k. The step sizeα_k is obtained by (SWP) line search (4) and (5), then,

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{19}$$

Proof. By Assumption 1 and the strong Wolfe line search, we obtain

$$(1-\sigma) g_k^T d_k \leq (g_{k+1} - g_k)^T d_k \leq L\alpha_k \|d_k\|^2$$

Hence,

$$\alpha_k \geq \frac{-(1-\sigma) g_k^T d_k}{L\|d_k\|^2} \tag{20}$$

We combine (20) with (12), which then results to

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \frac{L}{1-\sigma} \sum_{k=1}^{\infty} \left(-\alpha_k g_k^T d_k\right) < \infty$$

The proof is complete.

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

(1) $\beta_k \geq 0$
(2) The search directions fulfil the sufficient descent condition.
(3) The Zoutendijk condition holds.
(4) Property(*) holds.

If Assumptions1 and2 hold, then the iteration are globally convergent. From equations (11), (16), and (17) and Lemma 2, we found that the YHM method satisfies all four conditions in Theorem 2 under the strong Wolfe line search, so the method is globally convergent.

Numerical results and discussions

In this section, we present the results of the numerical tests conducted on our new parameter. The test problems used are taken from Andrei (2008), as shown in Table 1.We measure the performance of the proposed method by comparing it with other, well-established CG methods; FR, PRP, WYL and DPRP. A laptop with Intel(R) Core™ i5-M520 (2.40GHz) CPU processor and 4GB RAM in addition to MATLAB software version 8.3.0.532 (R2014a)are used to execute the optimization algorithms. We consider$\|g_k\| \leq \epsilon$ as the stopping criteria as suggested by Hillstrom (1977) with $\epsilon = 10^{-6}$. The dimensions of the test problems lay in the range of 2 to 10000. For each test function, we use four initial points, starting from a point close to the solution to another point far from it. In some cases, the computation is stopped due to the line search failing to find a positive step-size, thus it is considered a failure. The performance results are shown in Figures 1 and 2, respectively, based on the performance profile introduced by Dolan and More (2002).

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	Ex –rosenbrock	2,4,10,100,500,1000,10000	13, 25, 30, 50
6	Extended penalty	2,4,10,100	50, 60, 70, 80
7	Generalized Tridiagonal 1	2,4,10,100	30, 35, 40, 45
8	Shalow	2,4,10,100,500,1000,10000	10, 25, 50, 70
9	Ex-Tridiagonal1	2,4,10,100,500,1000,10000	12, 17, 20, 30
10	Extended White and Holst	2,4,10,100,500,1000,10000	3, 10, 30, 50
11	Quadrtic qf2	2,4,10,100,500,1000	10, 30, 50, 100
12	Extended Denschnb	2,4,10,100,500,1000,10000	8, 13, 30, 50
13	Hager	2,4,10,100	1, 3, 5, 7
14	Ex-Powell	4,8,20,100,500,1000	-1, 1, 7, 11
15	Extended Beale	2,4,10,100,500,1000,10000	-19, 1, 13, 23
16	Ex–Himmelblau	100,500,1000,10000	50, 70, 100, 125
17	Diagonal 2	2,4,10,100,500,1000	-1, 1, 2, 3
18	Perturbed quadratic	2,4,10,100,500,1000	1, 5, 10, 15
19	Sum Squares function	2,4,10,100,500,1000	1, 10, 20, 30
20	Ex- quadratic penalty QP2	4,10,100,500,1000,10000	17, 18, 19, 20
21	Raydan1 function	2,4,10,100	1, 3, 5, 7
22	Generalized Tridiagonal 2	2,4,10,100	1, 7, 8, 14
23	Quadratic QF1	2,4,10,100,500,1000	1, 2, 3, 4
24	Dixon and Price	2,4,10,100	100, 125, 150, 175
25	Fletcher	4,10,100,500,1000	7, 11, 13, 15
26	Ex-Maratos	2,4,10,100	5, 10, 12, 15
27	Leon function	2	2, 5, 8, 10

28	Extended wood	4	3, 5, 20, 30
29	Quartic function	4	5, 10, 15, 20
30	Matyas function	2	5, 10, 15, 20
31	Colville function	4	2, 4, 7, 10

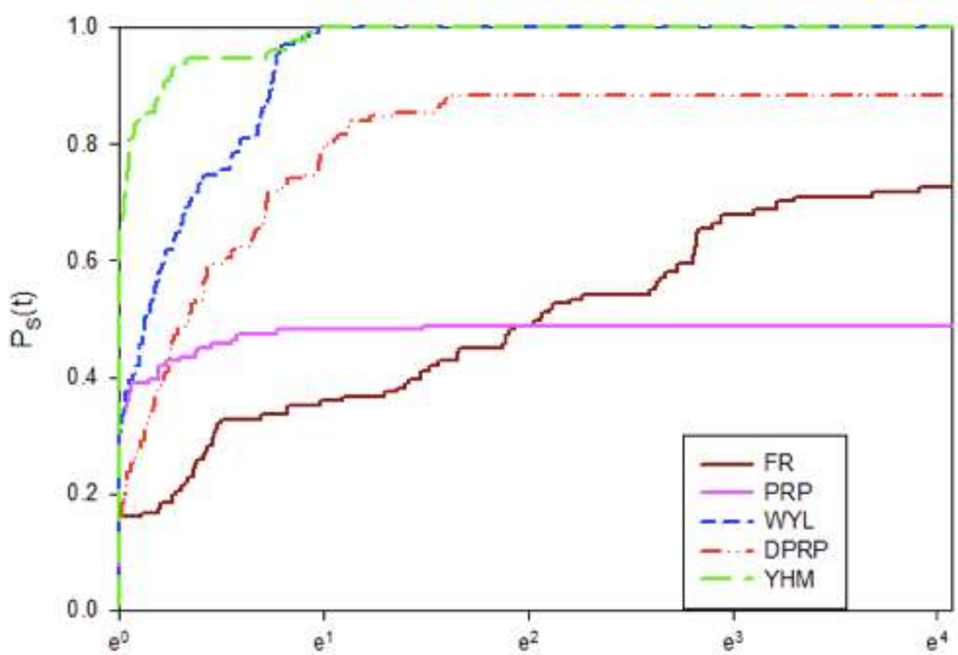

Figure 1. Performance profile relative to the number of iteration.

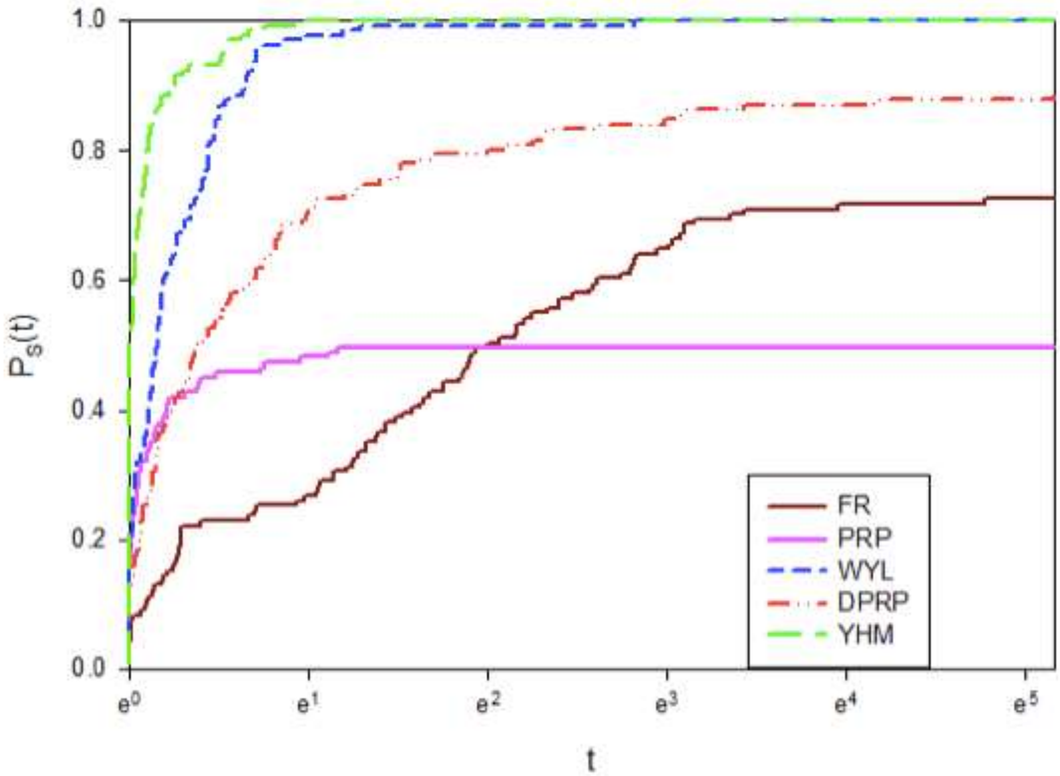

Figure 2. Performance profile relative to the CPU Time

From Figures 1 and 2, we found that our proposed algorithm solves 100% of the test problems, followed by WYL which solves 99.4% and DPRP with 88.4%of problems solved. Older CG methods like FR and PRP solve about 57% and 49 % of the test functions, respectively.

Conclusion

This paper gives anew$\beta_k$formula for solving unconstrained optimization problems. Under strong Wolfe line search, this newβ_k possesses global convergence properties. Numerical results show that the YHM method is very efficient and has the best performance when compared with other tested CG methods.

References

Abashar, A., Mamat, M., Rivaie, M., and Salleh, Z. (2017). Amodified dprp conjugate gradent method for unconstrained optimazation.

Aini, N., Mamat, M., Zull, N., and Rivaie, M. (2017). *Hybrid quasi-newton and conjugate gradient method for solving unconstrained optimization problems*: Journal of engineering and apllied sciences.

Al-Baali, M. (1985). Descent property and global convergence of the Fletcher—Reeves method with inexact line search. *IMA Journal of Numerical Analysis, 5*(1), 121-124.

Andrei, N. (2008). An unconstrained optimization test functions collection. *Adv. Model. Optim, 10*(1), 147-161.

Dai, Y.-H., and Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization, 10*(1), 177-182.

Dai, Z., and Wen, F. (2012). Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property. *Applied Mathematics and Computation, 218*(14), 7421-7430.

Dolan, E. D., and Moré, J. J. (2002). Benchmarking optimization software with performance profiles. *Mathematical programming, 91*(2), 201-213.

Fletcher, R. (1980). Practical methods of optimization. Vol. 1, Unconstrained optimization.

Fletcher, R., and Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal, 7*(2), 149-154.

Ghani, N., Mohamed, N., Zull, N., Shoid, S., Rivaie, M., and Mamat, M. (2017). *Performance comparison of a new hybrid conjugate gradient method under exact and inexact line searches.* Paper presented at the Journal of Physics: Conference Series.

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.

Gilbert, J. C., and Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on optimization, 2*(1), 21-42.

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

Hestenes, M. R., and Stiefel, E. (1952). *Methods of conjugate gradients for solving linear systems* (Vol. 49): NBS.

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

Kamfa, K. U., Mamat, M., Abashar, A., Rivaie, M., Ghazali, P. L. B., and Salleh, Z. (2017). Another modified conjugate gradient coefficient with global convergence properties.

Liu, Y., and Storey, C. (1991). Efficient generalized conjugate gradient algorithms, part 1: theory. *Journal of optimization theory and applications, 69*(1), 129-137.

Mohamed, N. S., Mamat, M., and Rivaie, M. (2017). *A new nonlinear conjugate gradient coefficient under strong Wolfe-Powell line search.* Paper presented at the AIP Conference Proceedings.

Omer, O., Rivaie, M., Mamat, M., and Amani, Z.(2015). *A new conjugate gradient method with sufficient descent without any line search for unconstrained optimization.* Paper presented at the AIP Conference Proceedings.

Osman, W. F. H. W., Ibrahim, M. A. H., and Mamat, M. (2017). *Hybrid DFP-CG method for solving unconstrained optimization problems.* Paper presented at the Journal of Physics: Conference Series.

Polak, E., and Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue française d'informatique et de recherche opérationnelle. Série rouge, 3*(16), 35-43.

Rivaie, M., Mamat, M., and Abashar, A. (2015). A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches. *Applied Mathematics and Computation, 268*, 1152-1163.

Touati-Ahmed, D., and Storey, C. (1990). Efficient hybrid conjugate gradient techniques. *Journal of optimization theory and applications, 64*(2), 379-397.

Wei, Z., Yao, S., and Liu, L. (2006). The convergence properties of some new conjugate gradient methods. *Applied Mathematics and Computation, 183*(2), 1341-1350.

Zhang, L. (2009). An improved Wei–Yao–Liu nonlinear conjugate gradient method for optimization computation. *Applied Mathematics and Computation, 215*(6), 2269-2274.

Zhang, Y., Zheng, H., and Zhang, C. (2012). Global convergence of a modified PRP conjugate gradient method. *Procedia Engineering, 31*, 986-995.

Zoutendijk, G. (1970). Nonlinear programming, computational methods. *Integer and nonlinear programming, 143*(1), 37-86.

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 Engeneering and Applied Sciences, 12(18), 4622-4626.