Journal of Quality Measurement and Analysis
Jurnal Pengukuran Kualiti dan Analisis

JQMA 10(1) 2014, 75-85

GLOBAL CONVERGENCE ANALYSIS OF A NEW NONLINEAR CONJUGATE GRADIENT COEFFICIENT WITH STRONG WOLFE LINE SEARCH

(Analisis Penumpuan Sejagat bagi Pekali Kecerunan Konjugat Tak Linear Baharu dengan Gelintaran Garis Strong Wolfe)

A. ABDELRAHMAN ABDALLA, M. MAMAT, M. RIVAIE & O. OMER

ABSTRACT

Nonlinear conjugate gradient (CG) methods are the most important method for solving large-scale unconstrained optimisation problems. Many studies and modifications have been conducted recently to improve this method. In this paper, a new class of conjugate gradient coefficients β_k with a new parameter $m = \|g_k\| / \|g_{k-1}\|$ that possesses global convergence properties is presented. The global convergence and sufficient descent property is established using inexact line searches to determine that α_k is the step size of CG methods. Numerical result shows that the new formula is superior and more efficient when compared to other CG coefficients.

Keywords: unconstrained optimisation; conjugate gradient method; sufficient descent property; global convergence

ABSTRAK

Kaedah kecerunan konjugat (CG) tak linear adalah kaedah yang paling penting untuk menyelesaikan masalah pengoptimuman tak berkekangan yang berskala besar. Banyak kajian dan pengubahsuaian telah dijalankan baru-baru ini untuk meningkatkan kecekapan kaedah ini. Dalam makalah ini, suatu kelas baharu pekali kecerunan konjugat, β_k dengan parameter $m = \|g_k\| / \|g_{k-1}\|$ yang mempunyai sifat-sifat penumpuan sejagat dibentangkan. Sifat-sifat penumpuan sejagat dan penurunan yang mencukupi ditentukan daripada anggaran pencarian garis untuk menentukan bahawa saiz langkah kaedah CG, α_k . Hasil berangka menunjukkan bahawa rumus baharu ini adalah lebih baik dan lebih cekap berbanding dengan pekali CG yang lain.

Kata kunci: pengoptimuman tak berkekangan; kaedah kecerunan konjugat; sifat penurunan yang mencukupi; penumpuan sejagat

1. Introduction

The nonlinear conjugate gradient methods (CG) can be used to find the minimum value of function for unconstrained optimisation problems. In general, the method has the following form

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

where $f : R^n \to R$ is a continuously differentiable nonlinear function with gradient denoted by $g(x)$. The CG methods are given by an iterative method of the form

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

where x_k is the current iterative point. The $\alpha_k > 0$ is a step size and d_k is the search direction. The step size is obtained by carrying out a one dimensional search known as the 'line search'.

The most common technique is the inexact strong Wolfe line search, where α_k satisfies

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k, \left| g(x_k + \alpha_k d_k)^T \right| \le \sigma \left| g_k^T d_k \right|, \tag{3}$$

with $0 < \delta < \sigma < 1$ are both constants.

The search direction, d_k is defined by

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

where β_k is a scalar and g_k is the gradient of nonlinear function. There are at least six formulae for β_k, which are given as follows:

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \qquad \text{(Hestenes \& Stiefel 1952)},$$

$$\beta_k^{FR} = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}} \qquad \text{(Fletcher-Reeves \& Reeves 1964)},$$

$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}} \qquad \text{(Polak \& Ribierre 1969)},$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}} \qquad \text{(Fletcher 1987)},$$

$$\beta_k^{LS} = -\frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T g_{k-1}} \qquad \text{(Liu \& Storey 1992)},$$

$$\beta_k^{DY} = \frac{g_k^T g_k}{(g_k - g_{k-1})^T d_{k-1}} \qquad \text{(Dai \& Yuan 1999)}.$$

The convergence manner of the β_k s' formulae with some line search conditions has been studied by many authors (Dai & Yuan 1999; Fletcher 1987; Hestenes & Stiefel 1952; Polak & Ribierre 1969; Powell 1977).

Rivaie *et al.* (2012) presented a new modification of a conjugate gradient method. It has global convergence properties under exact line searches, but it did not prove the global convergence under an inexact line search.

This paper is organised as follows. In section 2, we present the underlying idea of modification and we present a new nonlinear conjugate gradient method and algorithm. In section 3, we establish sufficient descent property and global convergence property with the strong Wolfe line search for the case $\sigma = 0.1$. Lastly in section 4, we present our preliminary numerical results.

2. Modification

Our motivation mainly comes from Rivaie *et al.* (2012) and Hestenes and Stiefel (1952), where

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

We insert the parameter $m = \|g_k\| / \|g_{k-1}\|$ in the numerator and the denominator of β_k^{RMIL}. We define the new CG coefficient as β_k^{AMRO}. Hence,

$$\beta_k^{AMRO} = \frac{g_k^T\left(g_k - mg_{k-1}\right)}{d_{k-1}^T\left(d_{k-1} - mg_k\right)}, \tag{5}$$

where $m = \|g_k\| / \|g_{k-1}\|$.

Algorithm 2.1

Step 1: Initialisation. Given $x_0 \in R^n$, $d_0 = -g_0$, $k := 0$, if $g_0 = 0$ then stop.

Step 2: Compute β_k based on the Eq. (5).

Step 3: Compute d_k based on (4). If $\|g_k\| \le \varepsilon$, then stop. Otherwise, go to the next step.

Step 4: Compute $\alpha_k > 0$ based on (3).

Step 5: Updating new point based on (2). If $\|g_k\| \le \varepsilon$, then stop. Otherwise, set $k := k+1$ and go to step 3.

3. Convergence Analysis

In this section, we analyse and study the convergence properties of β_k^{AMRO}. We assume that $g_k \neq 0$ for all k, otherwise a stationary point has been found. The following assumptions are often used to prove the convergence of the nonlinear conjugate gradient methods.

3.1. *Sufficient Descent Conditions*

Before giving the sufficient descent conditions, we need the following assumptions.

Assumption 3.1. (*i*) *The function f*(*x*) *is bounded below on the level set* R^n *and is continuous and differentiable in a neighbourhood N of the level set* $\ell(x_0)=\{x\in R^n \mid f(x)\le f(x_0)\}$ *at the initial point* x_0. (*ii*) *The gradient* $g(x)=\nabla f(x)$ *is Lipschitz continuous in N, so a constant* $L>0$ *exists, such that*

$$\|g(x)-g(y)\|\le L\|x-y\|, \qquad \text{for any } x,y\in N \tag{6}$$

From (5), *we have*

$$\left|\beta_k^{AMRO}\right|\le\begin{cases}\dfrac{2\|g_k\|^2}{\|d_{k-1}\|^2} & \text{if } g_k^T d_{k-1}\le 0\\[2ex] \dfrac{\sqrt{2}\|g_k\|}{\|d_{k-1}\|} & \text{if } g_k^T d_{k-1}>0\end{cases} \tag{7}$$

Theorem 3.1. *Consider any method of Eqs.* (2) *and* (4) *be generated by Algorithm* 2.1 *and let the step-size* α_k *be determined by the strong Wolfe line search* (3), *where* $\beta_k=\beta_k^{AMRO}$ *is given in* (5), *then for all* $k\ge 0$,

$$g_k^T d_k\le -c\|g_k\|^2, \tag{8}$$

holds.

Proof. If $k=0$, then $g_0^T d_0\le -c\|g_0\|^2<0$. Hence, condition (8) holds. We also need to show that for $k\ge 1$, the condition (8) also holds. By (7), if $g_k^T d_{k-1}\le 0$ and the second inequality in (3), then we get

$$\left|\beta_k^{AMRO} g_{k+1}^T d_k\right|\le \frac{2\|g_{k+1}\|^2}{\|d_k\|^2}\sigma\left|g_k^T d_k\right|. \tag{9}$$

From (4), multiply by g_{k+1}, we get

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2}=-1+\beta_{k+1}^{AMRO}\frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}. \tag{10}$$

We prove the descent property of $\{d_k\}$ by induction. Since $g_0^T d_0\le -c\|g_0\|^2<0$, if $g_0\ne 0$, now we suppose that $d_i, i=1,2,...,k$ are all descent direction, for example, $g_i^T d_i<0$.
By (10), we get

$$\left|\beta_{k+1}^{AMRO} g_{k+1}^T d_k\right| \le \frac{2\|g_{k+1}\|^2}{\|d_k\|^2}\sigma\left|-g_k^T d_k\right|. \tag{11}$$

Then,

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

From (10) and together with (12), we get

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

Repeating this process and using the fact $g_0^T d_0 = -\|d_0\|^2$ imply

$$-\sum_{i=0}^{k}\left[2\sigma\right]^i \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2+\sum_{i=0}^{k}\left[2\sigma\right]^i. \tag{14}$$

By the restriction $\sigma \in (0,1)$, so

$$\sum_{i=0}^{k}\left[2\sigma\right]^i < \sum_{i=0}^{\infty}\left[2\sigma\right]^i = \frac{1}{1-2\sigma}. \tag{15}$$

Then (14) can be written as

$$-\frac{1}{1-2\sigma} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -2+\frac{1}{1-2\sigma}. \tag{16}$$

Thus, by induction, $g_k^T d_k < 0$ holds for all $k \ge 0$, denote $c = 2-\frac{1}{1-2\sigma}$, then $c \in (0,1)$ and (16) becomes

$$c-2 \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -c. \tag{17}$$

Also from (7), if $g_k^T d_k > 0$ and the second inequality in (3), then we get

$$\left|\beta_{k+1}^{AMRO} g_{k+1}^T d_k\right| \le \frac{2\|g_k\|}{\|d_{k-1}\|}\sigma\left|g_k^T d_k\right|. \tag{18}$$

By (18), we get

$$\left|\beta_{k+1}^{AMRO} g_{k+1}^T d_k\right| \le \frac{2\|g_k\|}{\|d_{k-1}\|}\sigma\left(-g_k^T d_k\right), \tag{19}$$

$$\frac{2\|g_k\|}{\|d_{k-1}\|}\sigma\left(g_k^T d_k\right) \le \beta_{k+1}^{AMRO} g_{k+1}^T d_k \le -\frac{2\|g_k\|}{\|d_{k-1}\|}\sigma\left(g_k^T d_k\right). \tag{20}$$

From (10) and (20), we have

$$-1+\sqrt{2}\sigma\frac{g_k^T d_k}{\|g_{k+1}\|\|d_k\|} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1-\sqrt{2}\sigma\frac{g_k^T d_k}{\|g_{k+1}\|\|d_k\|}. \tag{21}$$

The second inequality of (3) implies

$$-1+\sqrt{2} \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -1-\sqrt{2}. \tag{22}$$

Thus, by induction, $g_k^T d_k < 0$ holds for all $k \ge 0$, denote $c = 1-\sqrt{2}$, then $c \in (0,1)$ and (22) becomes

$$c-2 \le \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \le -c. \tag{23}$$

Both (17) and (23) imply that (8) holds. □

3.2. *Global Convergence Properties*

Theorem 3.2. *Suppose that Assumption* 3.1 *holds. Consider any CG method in the form of* (2) *and* (4), *where* α_k *is obtained by the strong Wolfe inexact line search* (3). *Also, the descent condition holds. Then*

$$\lim_{k\to\infty} \inf\|g_k\| = 0. \tag{24}$$

Proof. To prove Theorem 3.2, we use a contradiction method. That is, if Theorem 3.2 is not true, then a constant $\varepsilon > 0$ exists, such that

$$\|g_k\| \ge \varepsilon. \tag{25}$$

Rewriting (4) as $d_{k+1} + g_{k+1} = \beta_{k+1} d_k$ and squaring both sides of the equation, we obtain

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

From (7), if $g_k^T d_{k-1} \le 0$, then we can get

$$\|d_{k+1}\|^2 = -\|g_{k+1}\|^2 - 2g_{k+1}^T d_{k+1} + \frac{4\|g_{k+1}\|^4}{\|d_{k+1}\|^2}. \tag{26}$$

We can divide (26) by $\|g_{k+1}\|^4$ and from (8), we get

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \frac{4}{\|g_{k+1}\|^2} + \frac{2c}{\|g_{k+1}\|^2}.$$

Suppose that (24) does not hold. Then, there exists $\varepsilon > 0$, such that (25) holds for all $k \ge 0$,

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \frac{4}{\varepsilon^2} + \frac{2c}{\varepsilon^2} = \frac{2c+4}{\varepsilon^2},$$

$$\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \frac{\varepsilon^2}{2c+4},$$

$$\sum_{k=0}^{\infty} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \infty \cdot \tag{27}$$

Also, from Eq. (7), if $g_k^T d_{k-1} > 0$, then we get

$$\|d_{k+1}\|^2 \le \|g_{k+1}\|^2 - 2g_{k+1}^T d_{k+1}. \tag{28}$$

We divide (28) by $\|g_{k+1}\|^4$ and from (8), we get

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \le \frac{1}{\|g_{k+1}\|^2} + \frac{2c}{\|g_{k+1}\|^2} \le \frac{2c}{\varepsilon^2} + \frac{1}{\varepsilon^2} \cdot \tag{29}$$

$$\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \frac{\varepsilon^2}{2c+1},$$

$$\sum_{k=0}^{\infty} \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \ge \infty \cdot \tag{30}$$

From (27) and (30), we get (24) and this shows that (24) holds. The proof is completed.□

4. Numerical Results

In this section, most of the problems from Andrei (2008) have been used to test and analyse the efficiency of AMRO compared to FR, PRP and CD. The stopping criterium is set to $\|g_k\| \geq \varepsilon$, where $\varepsilon = 10^{-6}$. As suggested by Hilstrom (1977) for each test problem, four or five initial points are used. All runs are performed on a PC ACER (Intel® Core™ i3-3217u CPU @ 1.8 GHZ, with 4.00 GB RAM, Windows 7 Home Premium). Numerical results are compared based on the number of iterations and CPU time. Every problem mentioned in Table 1 is solved using Matlab10 subroutine programming. We used the inexact strong Wolfe line search as to give the inexact value of the step-size. The performance results are shown in Figures 1 and 2, respectively, using a performance profile introduced by Dolan and More (2002).

We use the performance profile to introduce 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 exist, for each problem *p* and solver *s*, they defined $t_{p,s}$ as computing time (the number of iterations or CPU time or others) required to solve problems *p* by solver *s*. They compared the performance on problem *p* by solver *s* with the best performance by any solver on this problem using the performance ratio $r_{p,s} \leq t_{p,s} / \min\left\{t_{p,s} : s \in S\right\}$. Suppose that a parameter $r_M \geq r_{p,s}$ for all *p* and *s* are chosen, and $r_M = r_{p,s}$ if and only if solver *s* does not solve problem *p*. The performance solver *s* of given problems have to be the best, but we would like to obtain all evaluation performance of the solver, then it was defined $P(t)_s = \frac{1}{n_p} size\left\{p \in P : r_{p,s} \leq t\right\}$. The $P(t)_s$ was probability for solver $s \in S$ that a performance ratio $r_{p,s}$ was within a factor $t \in R$ of efficient ratio. Then, function P_s was the cumulative distribution function for the performance ratio. The performance profile $P : R \rightarrow \left[0,1\right]$ for solver was a non-decreasing, piecewise, and continuous from the right. The value $P(1)_S$ is the probability that the solver will win over the rest of the solvers. In all, a solver with high values of $P(t)_s$ or at the top right of the figure are preferable or represent the best solver.

Table 1: List of problem functions

No	Functions	n	Initial points
1	Zettl	2	(-10,-10),(-3,-3),(8,8),(20,20),(30,30)
2	Six-hump camel back	2	(-7,-7),(-2,-2),(3,3),(10,10)
3	Three-hump Camel back	2	(-3,-3),(2,2),(8,8),(13,13)
4	Trecanni	2	(-10,-10),(-5,-5),(7,7),(20,20),(30,30)
5	Hager	(2,4,10,100)	(-5,…,-5),(-1,…,-1),(3,…,3),(7,…,7)
6	Raydan1	(2,4,10,100)	(-3,…,-3),(2,…,2),(5,…,5),(10,…,10)
7	Shallow	(6,10,100,500,1000)	(8,…,8),(20,…,20),(40,…,40),(100,..,100)
8	Extended Tridiagonal2	(4,10)	(-6,…,-6),(-1,…,-1),(2,…,2),(5,…,5)
9	Extended Maratos	(2,4,10,100,500,1000)	(-9,…,-9),(-5,…,-5),(2,…,2),(6,…,6),(10,…,10)
10	Extended Tridiagonal1	(4,6,10,100,500,1000)	(2,…,2),(8,…,8),(14,…,14),(28,…,28)
11	Himmelblau	(4,10,100,500,1000)	(5,…,5),(11,…,11),(17,…,17),(31,…,31)
12	Generalised Quartic	(2,10,100,500,1000)	(-5,…,-5),(2,…,2),(7,…,7),(11,…,11)
13	Extended Rosenbrock	(4,10,100,500,1000)	(3,…,3),(6,…,6),(11,…,11),(23,…,23)
14	Extended Denschnb	(4,10,10,100)	(-5,…,-5),(-1,…,-1),(3,…,3),(10,…,10)
15	Arwhead	(2,4,100,500,1000)	(10,…,10),(50,…,50),(100,…,100),(200,…,200)
16	Freudenstein & Roth	(4,10,100,500,1000,5000)	(-6,…,-6),(-3,…,-3),(1,…,1),(7,…,7)
17	Flethcr	(4,10,100,500)	(-7,…,-7),(-2,…,-2),(5,…,5),(17,…,17)
18	Extended White & Holst	(4,10,100,500,1000)	(-8,..,-8),(-3,…,-3),(2,…,2),(10,…,10)
19	Powell	(4,100,500,1000)	(-5,…,-5),(-1,…,-1),(3,…,3),(10,…,10)
20	Extended Penalty	(2,4,6,10,100,500,1000,5000)	(-3,…,-3),(-1,…,-1),(2,…,2),(8,…,8)

5. Results and Discussions

Figures 1 and 2 show that AMRO CG, β_k^{AMRO} is capable of solving all the test problems and reach 100% accuracy. The performance of the conjugate gradient coefficients FR, PRP and CD can be divided into two groups. The first group that is related to PRP is better in performances than the second one which is related to FR and CD. The first group performance is less efficient than AMRO method; therefore, it solves only 85% of the problems, but it is better than the second group that solves about 80% of the problems. The AMRO method outperformed the other CGs.

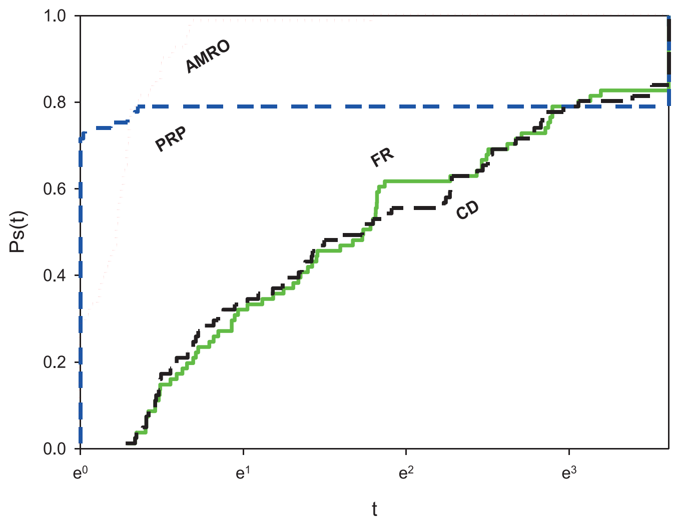

Figure 1: Performance profile based on the number of iterations

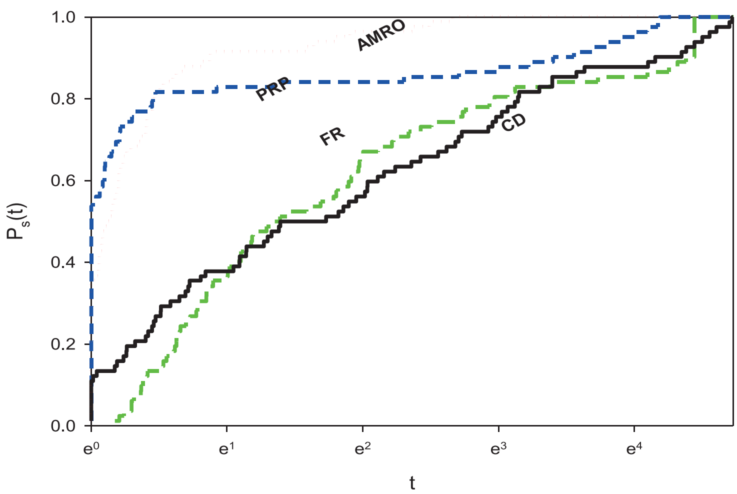

Figure 2**:** Performance profile based on the CPU time

Acknowledgement

The authors are very grateful to the editors and the referees for their valuable suggestions and advice.

References

Andrei N. 2008. An unconstrained optimization test functions collection. *Advanced Modeling and Optim*ization **10**:147-161

Dolan E. & More J.J. 2002. Benchmarking optimization software with performance profile. *Mathematical Programming* **91**(2): 201-213.

Dai Y.H. & Yuan. 1999. A nonlinear conjugate gradient method with a strong global convergent property. *SIAM J. Optim*. **10**: 177-182.

Dai Y.H. & Yuan. 2000. *Nonlinear Conjugate Gradient Methods*. Shanghai: Science Press of Shanghai.

Fletcher R. 1987. Practical Method of Optimization. 2nd ed. New York: Wiley.

Fletcher-Reeves C.1964. Function minimization by conjugate gradients. *Comput. J.* **7**: 149-154.

Gilbert J.C. & Nocedal J. 1992. Convergence properties of conjugate gradient methods for optimization. *SIAM. J. Optim*. **2**(1): 21-42.

Hestenes M. R & Stiefel E. 1952. Method of conjugate gradient for solving linear equations. *J. Res. Nat. Bur. Stand.* **49**: 409-436.

Hilstrom K.E. 1977. A simulation test approach to the evaluation of nonlinear optimization algorithms. *ACM. Trans. Math. Softw* **3:** 305-315.

Liu Y. & Storey C. 1992. Efficient generalized conjugate gradient algorithms, part 1: theory. *J. Optim. Theory Appl.***69:** 129–137.

Polak E. & Ribierre G. 1969. Note on the convergence of conjugate directions. *Rev. Francaise Informat Recherche Operationelle* **3:** 35-43.

Powell M.J.D. 1977. Restart procedure for the conjugate gradient method. *Math. Prog*. **2:** 241-254.

Powell M.J.D. 1984. Non-convex minimization calculation and the conjugate gradient method. *Lecture Notes in Mathematics* 1066. Berlin: Springer-Verlag.

Zoutendijk G. 1970. Nonlinear programming computational methods. In J. Abadie (Ed.): *Integer and Nonlinear Programming*, pp. 37-86. Amsterdam: North Holland.

Rivaie M., Mustafa M., Leong W. J. & Mohd. I. 2012. A new conjugate gradient coefficient for large scale nonlinear unconstrained optimization. *Int. Journal of Math Analysis* **6**(23): 1131-1146.

School of Informatics and Applied Mathematics
Universiti Malaysia Terengganu (UMT)
MALAYSIA
*E-mail:awad.abdalla26@yahoo.com**

Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin
Campus Tembila, 22200 Besut, MALAYSIA
E-mail: must@unisza.edu.my

Department of Computer Sciences and Mathematics
Universiti Teknologi MARA (UiTM) Terengganu
Campus Kuala Terengganu, MALAYSIA
E-mail: rivaie75@yahoo.com

School of Informatics and Applied Mathematics
Universiti Malaysia Terengganu (UMT)
MALAYSIA
E-mail: osman_om@hotmail.com

*Corresponding author