International Journal of Engineering & Technology, 7 (3.28) (2018) 307-311

International Journal of Engineering & Technology

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

Research paper

Modification on Spectral Conjugate Gradient Method for Unconstrained Optimization

Usman Abbas Yakubu[1,2], Mustafa Mamat[2]*, Mohamad Afendee Mohamed[2], Sukono[3], Mohd Rivaie[4]

[1]*Department of Mathematics, Northwest University, Kano. Nigeria*
[2]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Kuala Terengganu, Malaysia*
[3]*Fakultas MIPA, Universitas Padjadjaran, Bandung, Indonesia*
[4]*Department of Computer Science and Mathematics, Universiti Teknologi MARA (UiTM), Kuala Terengganu, Malaysia*
**Corresponding author E-mail: must@unisza.edu.my*

Abstract

The classical Newton's direction and spectral conjugate gradient direction are the prominent directions in solving large-scale unconstrained optimization problems. Using the standard secant equation, a modified spectral CG method (MSCG) is proposed, the scheme is a modification of Birgin and Martinez spectral CG method (SCG). Sufficient descent property as well as global convergence has been proved by strong Wolfe line search. Numerical outcome shows that the method is practically effective when compared with classical PRP, FR and spectral CG methods.

Keywords: *Global convergence; inexact line search; spectral CG; secant condition; sufficient descent property.*

1. Introduction

A modified spectral CG method was proposed using standard secant equation to outperform the prominent classical CG methods and recent spectral CG method. A spectral conjugate gradient method proposed originally by Barzilai and Borwein [7]. Raydan [1] introduce the spectral CG method (SGM) which was the combination of Barzila and Borwein (spectral) nonmonotone techniques with classical projected gradient ideas for solving large-scale or many variables problems. Recently, Birgin and Martinez [3] purposed a spectral CG method (SCG) which combined the spectral gradient and conjugate gradient ideas leading to efficient algorithm that performed excellently on advanced CG algorithms in many unconstrained optimization problems. Moreover, Perry [5] combining the classical CG search direction and the quasi-Newton direction. In our research, the new direction has been suggested by equating the spectral CG direction and Newton's direction which produce a new parameter named MSCG. The exhibited numerical results clearly indicated that the method was more efficient compared to classical and spectral CG methods for minimizing the general function. However, the new method has merged the advantages of spectral CG method and the classical method. A twenty eight standard test functions has been considered to test the practical effectiveness of the proposed method. Therefore, the nonlinear CG method is designed to solve this unconstrained optimization problem:

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

where $f: R^n \to R$ is continuous as well as differentiable function, g_k is a gradient vector of a function f and initial point $x_0 \in R^n$ is normally solved iteratively according to the recurrence expression

$$x_{k+1} = x_k + \gamma_k d_k, \qquad k = 0,1,2,3,4 \ldots \tag{2}$$

where x_k is a current iteration, $\gamma_k > 0$ is a step size obtained by line search procedure. However, d_k is a search direction defined as

$$d_0 = -g_0, \qquad d_{k+1} = -g_{k+1} + \beta_k d_k, \quad k = 0,1,2,3,4 \ldots \tag{3}$$

also $g_k = \nabla f(x)$, parameter $\beta_k \in R$ are the gradient vector and CG coefficient respectively. Some classical CG and spectral CG coefficients are given below, see also [6, 8, 11-12, 18-19], for further advance studies on properties of classical CG methods:

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \tag{4}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} \tag{5}$$

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

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

$$\beta_k^{CD} = -\frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}} \tag{8}$$

$$\beta_k^{RMIL} = \frac{g_k^T(g_k - g_{k-1})}{\|d_{k-1}\|^2} \tag{9}$$

$$\beta_k^{SCG} = \frac{(\theta_k y_k - s_k)^T g_{k+1}}{s_k^T y_k} \tag{10}$$

Therefore, in the above equations g_k, g_{k-1} are gradient vectors of function f at a point x_k, x_{k-1} respectively. Denoting $s_k = x_{k+1} - x_k = \gamma_k d_k$ and $y_k = g_{k+1} - g_k$, also this notation $\|.\|$ represent a Euclidian norm. Generally, consideration has been established on a global convergence of CG methods. Zoutendijk [4] proved that FR method on exact line search converges globally. Later, Powell [6] disproved the result and noted clearly that the method has demonstrated a poor practical behaviour and its convergence is not global. PRP method is more dependable and authentic CG method among others [9]. However, a spectral CG method is more effective in terms of a numerical execution than the other methods because it comprises the advantages of a spectral parameter to the CG method by constructing a new search direction in [1]. Nowadays, the CG methods has a very good reputations and advantages for solving relatively large-scale problems. Nevertheless, Perry [5] proposed a modified conjugate gradient algorithm, the result shows that their method is encouraging. See [3, 10, 12-17], for advance studies on CG methods. Consequently, in this research we prompted by [3] and determine to modify the direction by employing the standard secant condition and compare its performances with classical CG and spectral CG method with less iterations and CPU time in second respectively. This paper is coordinated as follows: Section 2 demonstrates our new method. Section 3 shows a convergence result. Numerical experiments are presented in Section 4. Finally, conclusions follows in Section 5.

2. Description of the New Method

A spectral CG method are proposed originally by Barzilai and Borwein [7], the direction is generated as $d_k = -\theta_k g_k + \beta_k s_{k-1}$, where $s_{k-1} = \gamma_{k-1} d_{k-1}$ and θ_k is a spectral parameter. The details analysis of asymptotic behaviour of [7] and related techniques is presented in Dai and Fletcher (2005). Birgin and Martinez [3] purposed a spectral CG method and they computed spectral parameter as $\theta_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}}$; which derived by approximating the secant condition given in [7], see Raydan [1] for more details. Moreover, in our new direction we equate the spectral CG direction and Newton's direction generated by

$$d_0 = -g_0,$$
$$d_{k+1} = -\theta_k g_{k+1} + (1-\theta_k)\beta_k d_k, \quad k = 0,1,2,3,4\ldots \quad (11)$$

Besides, since the parameter β_k is any scalar, this little change or modification of a spectral CG method does not alter the significance of the parameter Andrei [14].
Recall that from the prominent Newton's direction,

$$d_{k+1} = -H_k^{-1} g_{k+1} \quad (12)$$

Considering equation (11) then we have

$$(1-\theta_k)\beta_k d_k = \theta_k g_{k+1} + d_{k+1} \quad (13)$$

From (12) it implies $H_k d_{k+1} = -g_{k+1}$ and to have a better approximation, we multiply both sides by s_k^T

$$s_k^T H_k d_{k+1} = -s_k^T g_{k+1} \quad (14)$$

From the secant equation, we know that

$$H_k s_k = y_k \quad \equiv \quad s_k = H_k^{-1} y_k \quad (15)$$
$$s_k^T H_k^T = y_k^T \quad \equiv \quad s_k^T = (H_k^{-1})^T y_k^T \quad (16)$$

It is very important to note that H_k in this particular case is symmetric matrix for all value of k. Hence equation (16) can be written as

$$s_k^T H_k = y_k^T \quad (17)$$

Substituting equation (17) into (14) imply $y_k^T d_{k+1} = -s_k^T g_{k+1}$ where

$$d_{k+1} = -\frac{s_k^T}{y_k^T} g_{k+1} \quad (18)$$

Also substitute (18) in (13) this gives

$$(1-\theta_k)\beta_k d_k = \theta_k g_{k+1} - \frac{s_k^T}{y_k^T} g_{k+1}$$
$$\beta_k^{MSCG} = \frac{(\theta_k y_k - s_k)^T g_{k+1}}{(1-\theta_k) y_k^T d_k} \quad (19)$$

Conventionally, we assumed that $\theta_k \neq 1$. Having deduced the CG parameter β_k^{MSCG} in (19), we then present our new direction as

$$d_0 = -g_0,$$
$$d_{k+1} = -\theta_k g_{k+1} + (1-\theta_k)\beta_k^{MSCG} d_k, \quad k = 0,1,2,3,4\ldots \quad (20)$$

Algorithm 2.1 (MSCG Method)

Consider the following steps below:
Step 1: Given a starting point $x_0 \in R^n$, compute $d_0 = -g_0$ and set $k = 0$
Step 2: Compute β_k as given in formula (19) above
Step 3: Compute $d_{k+1} = -\theta_k g_{k+1} + (1-\theta_k)\beta_k^{MSCG} d_k$, if $\|g_k\| = 0$, then stop.
Step 4: Compute $\gamma_k > 0$ by strong Wolfe line search

$$f(x_k + \gamma_k d_k) \leq f(x_k) + \delta \gamma_k g_k^T d_k \quad (21)$$
$$|g(x_k + \gamma_k d_k)^T d_k| \leq -\sigma g_k^T d_k \quad (22)$$

where $0 < \delta < \sigma < 1$
Step 5: Update the new point as given in the recurrence expression (2).
Step 6: If $f(x_{k+1}) < f(x_k), \|g_k\| < \varepsilon$ then stop, otherwise go to 1 above with $k = k+1$.

3. Global Convergence Analysis

3.1. Sufficient descent condition

In this section, we illustrates a sufficient descent property of MSCG method.

$$g_k^T d_k \leq -C\|g_k\|^2 \;\; for\; k \geq 0 \; and \; C > 0 \quad (23)$$

Theorem 3.1: *Suppose a CG method generated by* (20) *as a search direction and* β_k *is given as* (19), *and* γ_k *satisfies any line search then* (23) *holds* $\forall\, k \geq 0$.

Proof: We proceed by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (23) is satisfied as $k = 0$. Now, assume it is true for $k \geq 0$. Condition (23) as well hold, and then from equation (20) by multiplying both sides by g_{k+1}^T, we have

$$g_{k+1}^T d_{k+1} = -\theta_k \|g_{k+1}\|^2 + (1-\theta_k)\beta_k^{MSCG} g_{k+1}^T d_k$$
$$= -\theta_k \|g_{k+1}\|^2 (1-\theta_k) \frac{(\theta_k y_k - s_k)^T}{(1-\theta_k) y_k^T d_k} \|g_{k+1}\|^2 d_k$$
$$= -\theta_k \|g_{k+1}\|^2 + \frac{(\theta_k y_k^T - s_k^T)}{y_k^T} \|g_{k+1}\|^2$$
$$g_{k+1}^T d_{k+1} = -\frac{s_k^T}{y_k^T} \|g_{k+1}\|^2 \quad (24)$$

Therefore, the search direction satisfied the descent properties $g_{k+1}^T d_{k+1} < 0$. Hence, (23) is true for $k+1$.

3.2. Global convergence properties

To analyze global convergence properties of MSCG, we will show β_k^{MSCG} and θ_k are bounded.

$$|\theta_k| = \left|\frac{s_k^T s_k}{s_k^T y_k}\right| \leq \frac{\|s_k\|^2}{L\|s_k\|^2} = \frac{1}{L} \quad (25)$$
$$|\beta_k^{MSCG}| = \left|\frac{(\theta_k y_k - s_k)^T g_{k+1}}{(1-\theta_k) y_k^T d_k}\right| \leq \frac{(|\theta_k|\|y_k\| + \|s_k\|)\|g_{k+1}\|}{|1-\theta_k|\|y_k\|\|d_k\|} \quad (26)$$

Assumptions 3.1

i) A level set $\Omega = \{x \in R^n \mid f(x) \le f(x_0)\}$ is bounded, the function f is continuously differentiable in a neighborhood N of the level set Ω and x_0 is a starting point.
ii) $g(x)$ is globally Lipschitz continuous in N that is $\exists$ a constant $L > 0$ such that $\|g(x) - g(y)\| \le L\|x - y\|$ for any $x, y \in N$.

The assumptions on the function f imply that $\exists$ a positive constant $\gamma > 0$ such that

$$\|g(x)\| \le \gamma \text{ for all } x \in N \qquad (27)$$

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \qquad (28)$$

Proof of this lemma is in [4].

Lemma 3.2: *Suppose the assumptions 3.1 holds. Let the sequences $\{g_k\}$ and $\{d_k\}$ given by the MSCG method. If there exist a constant $\tau > 0$ such that*

$$\|g_k\| \ge \tau, \text{ for all } k > 0. \qquad (29)$$

Then, we have

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} < \infty \qquad (30)$$

Based on (28) and (29), we can prove easily.

Theorem 3.2: *Suppose the assumptions 3.1 holds. Consider a CG method (2) and (20), α_k satisfies (21), (22) and β_k is given by (19). Then*

$$\lim_{k\to\infty} \|g_k\| = 0 \qquad (31)$$

Proof: From the search direction (20), we have $d_{k+1} + \theta_k g_{k+1} = (1-\theta_k)\beta_k^{MSCG} d_k$, square both sides of the equation, we have

$$(d_{k+1} + \theta_k g_{k+1})^2 = \left((1-\theta_k)\beta_k^{MSCG} d_k\right)^2$$

$$\|d_{k+1}\|^2 = \left((1-\theta_k)\beta_k^{MSCG}\right)^2 \|d_k\|^2 - 2\theta_k g_{k+1}^T d_{k+1} - \theta_k^2 \|g_{k+1}\|^2 \qquad (32)$$

Substituting (24) and (26) into (32), we obtain the following

$$\|d_{k+1}\|^2 \le \left(|1-\theta_k| \frac{(|\theta_k|\|y_k\| + \|s_k\|)}{|1-\theta_k|\|y_k\|\|d_k\|} \|g_{k+1}\|\right)^2$$

$$\|d_k\|^2 + 2\theta_k \frac{s_k^T}{y_k^T} \|g_{k+1}\|^2 - \theta_k^2 \|g_{k+1}\|^2$$

$$\|d_{k+1}\|^2 \le \left(\frac{4|\theta_k|\|s_k\|}{\|y_k\|} + \frac{\|s_k\|^2}{\|y_k\|^2}\right) \|g_{k+1}\|^2$$

Recall that from (25) above $|\theta_k| \le \frac{1}{L}$ which θ_k is bounded for some constant $L > 0$. Therefore,

$$\|d_{k+1}\|^2 \le \left(\frac{4\|s_k\|}{L^2\|s_k\|} + \frac{\|s_k\|^2}{L^2\|s_k\|^2}\right) \|g_{k+1}\|^2$$

$$\|d_{k+1}\|^2 \le \frac{5}{L^2} \|g_{k+1}\|^2 = M\|g_{k+1}\|^2 \qquad (33)$$

Multiply both side of (33) with $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$, then we have

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \qquad (34)$$

From the Lemma 3.1 above, $\lim_{k\to\infty} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} < 0$. It implies that Theorem 3.2 does not hold true, then the $\lim_{k\to\infty} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} = \infty$ and from equation (34) this is true $\infty \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So Theorem 3.2 is true for a sufficient large k.
An alternative proof of Zoutendijk condition of the MSCG method.

Theorem 3.3: *Suppose the assumptions 3.1 holds. Consider a CG method (2), (20) and assume d_k is a descent direction α_k satisfies (21), (22) and β_k is given by (19). If*

Lemm

$$\sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} = \infty, \qquad (35)$$

then we have

$$\liminf_{k\to\infty} \|g_k\| = 0 \qquad (36)$$

Proof: Suppose by contradiction (36) does not hold, that is $\exists$ $\tau > 0$ s.t $\|g_k\| \ge \tau$ for all $k > 0$. Also, by strong Wolfe condition using Lipschitz continuity and uniform convexity, we have from equation (20) by multiplying both sides with g_{k+1}^T

$$|g_{k+1}^T d_{k+1}| \le -|\theta_k|\|g_{k+1}\|^2 + \sigma|1-\theta_k||\beta_k^{MSCG}||g_{k+1}^T d_k|$$

$$\|g_{k+1}\|\|d_{k+1}\| \le -|\theta_k|\|g_{k+1}\|^2 + \sigma|1-\theta_k| \frac{(|\theta_k|\|y_k\| + \|s_k\|)}{|1-\theta_k|\|y_k\|\|d_k\|} \|g_{k+1}\|^2\|d_k\|$$

$$= -|\theta_k|\|g_{k+1}\|^2 + \sigma\left(|\theta_k| + \frac{\|s_k\|}{\|y_k\|}\right)\|g_{k+1}\|^2$$

$$= -|\theta_k|\|g_{k+1}\|^2 + \sigma|\theta_k|\|g_{k+1}\|^2 + \frac{\sigma}{L}\|g_{k+1}\|^2$$

$$= \left(-\frac{1}{L} + \frac{\sigma}{L} + \frac{\sigma}{L}\right)\|g_{k+1}\|^2 = \frac{2\sigma - 1}{L}\|g_{k+1}\|^2 = W\|g_{k+1}\|^2$$

$$\|d_{k+1}\| \le W\|g_{k+1}\| = W\tau = M \qquad (37)$$

which means that $\|d_{k+1}\|$ is bounded above, since $\|g_k\| \ge \tau$ for all $k > 0$. it follows that (37) contradicts with (35) implies that $\liminf_{k\to\infty}\|g_k\| = 0$.

4. Results and Discussion

We proceed with numerical experiments to test Algorithm 2.1 and make a comparison of the performance of MSCG with PRP, FR and SCG methods. For both algorithms, codes were written in *MatlabR2012* subroutine programming using Intel® Core™ i5-3317U, 1.7GHz with 4 GB RAM memory. We consider $\varepsilon = 10^{-6}$ as a stopping criteria as suggested by Hillstrom for each of the problems and $\|g_k\| < \varepsilon$. We represents failure due to; (i) Memory requirement (ii) Number of iterations exceed 1000 (iii) If $\|g_k\|$ is not a number (NaN). The methods were tested on 28 test problems functions by employing test problems in [10] as shown in Table 1 with 4 different initial values. The results in Figure 1 and 2 were obtained using the performance profile acquainted by [2]. Performances are established on CPU time as well as number of iterations respectively. The highest value of $P_s(t)$ and the method reached the top foremost will be regarded as a best performing method. We can boldly say that our modified spectral CG method has performed very good since it solved all the test problems functions, the method also has absolute potentials when compared with PRP, FR and SCG methods. Figure 1 and 2 show that MSCG, SCG, and PRP methods successfully reach the solution points. Note that by referring the "successful" in Table 2 and 3 means that MSCG method has less number of iterations and less CPU time when compared to SCG, FR and PRP while denoting of "unsuccessful" means that MSCG produce more iterations and CPU time. If MSCG provide same iterations or

CPU time with other methods, then its equivalent. For both Table 2 and 3, the results presented are in percentage.

Table 1: Test Problems functions

Functions	Dimensions	Initial Points
Trecanni	2	(2,2), (11,11), (17,17), (-17,-17)
Booth	2	(4,4), (-4,-4), (21,21), (109,109)
Tree hump Camel	2	(3,3), (-3,-3), (7,7), (51,51)
Six hump Camel	2	(11,11), (41,41), (51,51), (61,61)
Zettl	2	(-3,-3), (5,5), (100,100), (200,200)
Leon	2	(2,2), (4,4), (25,25), (-25,-25)
Quartic	4	(10,.,10), (25,.,25),(100,.,100),(-100,.,-100)
Colville	4	(-2,.,-2), (4,.,4), (7,.,7), (9,.,9)
Wood	4	(7,.,7), (9,.,9), (-9,.,-9), (-16,.,-16)
Gen. Tridiagonal 1	10	(7,.,7), (-7,.,-7), (14,.,14), (-14,.,-14)
Gen. Tridiagonal 2	10	(3,.,3), (5,.,5), (7,.,7),), (14,.,14),
Fletcher	10	(3,.,3), (-10,.,-10), (13,.,13), (-13,.,-13)
Hager	50	(2,.,2), (5,.,5), (8,.,8), (10,.,10)
Quadratic Penalty QP1	100	(50,.,50),(100,.,100),(200,.,200),(300,.,300)
Dixon and Price	100	(49,.,49), (-49,.,-49), (81,.,81), (-81,.,-81)
Quadratic QP1	100	(40,.,40), (-40,.,-40),49,.,49), (-49,.,-49)
Quadratic QP2	100	(8,.,8), (-8,.,-8), (11,.,11), (-11,.,-11)
Raydan 1	2,4,10,100	(8,.,8), (14,.,14), (20,.,20), (30,.,30)
Ext. Tridiagonal 1	100,1000	(10,., (-10,.,-10), (100,.,100), (-100,.,-100)
Quadratic QF2	100,1000	(100,100),(-100,-100),(200,200),(-200,200)
Extended Penalty	2,4,10,100,1000	(50,.,50),(100,.,100),(200.,,200),(500,.,500)
Ext. Maratos	2,4,10,100,1000	(2,.,2), (10,.,10) (25,.,25), (49,.,49)
Freud. and Roth	100,1000,10000, 100000	(2,.,2), (3,.,3), (-5,.,-5), (-10,.,-10)
Himmelblau	100,1000,10000, 100000	(8,.,8), (-14,.,-14),(-50,.,-50), (55,.,55)
White and Holst	100,1000,10000, 100000	(2,…,2), (-2,..,-2), (4,..,4), (13,..,13)
shallow	100,1000,10000, 100000	(11,.,11), (-11,.,-11), (25,.,25), (49,.,49)
Gen. Quartic	100,1000,10000, 100000	(2,…,2), (-2,..,-2), (4,...,4), (-4,..,-4)
Rosenbrock	100,1000,10000, 100000, 1000000	(2,..,2), (5,..,5), (10,..,10), (25,..,25)

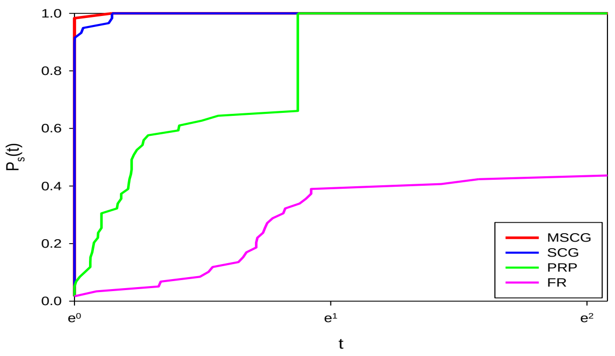

Fig. 1: MSCG vs SCG, PRP and FR methods on iterations

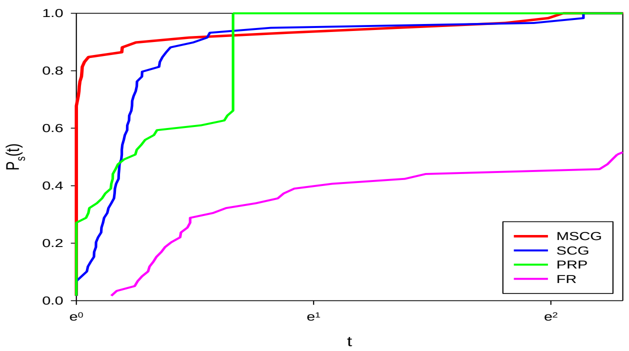

Fig. 2: MSCG vs SCG, PRP and FR methods on CPU time

We can boldly say that our modified spectral CG method have per-formed very good, since it solved all the test problems func-tions 100%, the method also has absolute potentials when com-pared with PRP, FR and SCG methods. Figure 1 and 2 show that MSCG, SCG, and PRP methods successfully reach the solution points.

Note that by referring the 'successful' in Table 2 and 3 means that MSCG method has less number of iterations and less CPU time when compared to SCG, FR and PRP while denoting of 'unsuc-cessful' means that MSCG produce results with more iterations and CPU time. If MSCG provide same iterations or CPU time with other methods, then its equivalent. For both Table 2 and 3, the results presented are in percentage.

Table 2: Percentage performance of MSCG vs SCG, PRP and FR on number of iterations

Method		SCG	PRP	FR
	Successful	8.47%	94.91%	98.30%
MSCG	Equivalent	91.53%	3.39%	1.70%
	Unsuccessful	0.0%	1.70%	0.0%

By the comparison above we observed that MSCG method almost closely similar as SCG method in Table 2, we see that 91.53% are the same pertaining number of iterations, nevertheless in Table 3 below it has 0.0% equivalence in terms of CPU time. Therefore, MSCG solves the test problems with highest percentage of suc-cess and equivalence of 100% compared with SCG, 100% over FR and 98.31% above the famous PRP in terms of number of iteration.

Table 3: Percentage performance of MSCG vs SCG, PRP and FR on CPU time

Method		SCG	PRP	FR
	Successful	86.44%	77.97%	93.20%
MSCGG	Equivalent	0.0%	0.0%	0.0%
	Unsuccessful	13.56%	22.03%	6.80%

However, percentage of success and equivalence of MSCG based on CPU time is 86.44% as compared with SCG, 93.20% over FR and 77.97% above PRP. Indeed the MSCG is promising.

5. Conclusion

Lastly, the modified spectral CG method satisfies a sufficient descent conditions and converges globally. Numerical outcomes by employing a set of large-scale test problems indicated that MSCG is highly efficient compared to the classical and spectral CG methods.

Acknowledgement

The authors thank the referees for their contributions and also grateful to the Malaysian government and Universiti Sultan Zainal Abidin for funding this research under the Fundamental Research Grant Scheme (FRGS/1/2017/STG06/Unisza/01/1).

References

[1] M. Raydan (1997). The Barzilai and J.M. Borwein gradient methods for the large scale unconstrained minimization in extreme problems. SIAM. J. Optim., 7 (1), 26-33.
[2] E. Dolan and J. J. More (2002). Benchmarking optimization software with performance profile. Math. Prog., 91, 201-213.
[3] E.G. Birgin and J. M. Martinez (2011). A spectral conjugate gradient method for unconstrained optimization. Appl. Math. Optim., 43 (2), 117-128.
[4] G. Zoutendijk (1970). Nonlinear programming, computational methods. Integer and Nonlinear Programming, 1970, 37-86.
[5] A. Perry (1978). A modified conjugate gradient algorithm. Operational Research, 26, 1073-1078.
[6] M. J. D. Powell (1984). Non-convex minimization calculations and the conjugate gradient method. Lecture Notes in Mathematics, 1066, 122-241.
[7] J. Barzilai and J. M. Borwein (1988). Two-point step size gradient methods, IMA J Numer Anal, 8, 141-148.
[8] N. Zull, M. Rivaie, M. Mamat, Z Salleh and Z. Amani (2015). Global convergence of a spectral conjugate gradient by using strong Wolfe line search. Appl. Math. Sci., 63, 3105-3117.
[9] W. W. Hager and H. Zhang (2006). A survey of nonlinear conjugate gradient methods. Pacific Journal of Optimization, 2(1), 35-58.
[10] N. Andrei (2008). An unconstrained optimization test functions collection. Adv. Modell. Optim., 10, 147-161.
[11] R. B. Yunus, M. Mamat, A. Abashar M. Rivaie, Z. Salleh and Z. Amani (2015). The convergence properties of a new kind of conjugate gradient method for unconstrained optimization. Appl. Math. Sci., 38, 1845-1856.
[12] M. Rivaie, M. Mamat and A. Albashar (2015). A new class of nonlinear conjugate gradient coefficients with exact and inexact line search. Appl. Math. and Comp., 268, 1152-1163.
[13] A. Y. Usman, M. Mamat, M. Rivaie, A. M. Mohamad and B. Y. Rabi'u (2018). Secant free condition of a spectral WYL and its global convergence properties. Far East Journal of Mathematical Science, 103(12), 1889-1902.
[14] A. Y. Usman, M. Mamat, M. Rivaie, A. M. Mohamad and J. Sabi'u (2018). A recent modification on Dai-Liao conjugate gradient method for solving symmetric nonlinear equations. Far East Journal of Mathematical Science, 103(12), 1961-1974.
[15] K. U. Kamfa, M. Mamat, A. Abashar, M. Rivaie, P. L. B. Ghazali and Z. Salleh (2015). Another modified conjugate gradient coefficient with global convergence properties. Applied Mathematical Sciences, 9, 1833-1844.
[16] N. Z. Abidin, M. Mamat, B. Dangerfield, J. H. Zulkepli, M. A. Baten and A. Wibowo (2014). Combating obesity through healthy eating behavior: A call for system dynamics optimization. Plos One, 9(12), 1-17.
[17] M. Mamat, Y. Rokhayati, N. M. M. Noor and I. Mohd (2011). Optimizing human diet problem with puzzy price using fuzzy linear programming approach. Pakistan Journal of Nutrition, 10(6), 594-598.
[18] A. Abashar, M. Mamat, M. Rivaie and I. Mohd (2014). Global convergence properties of a new class of conjugate gradient method for unconstrained optimization. Applied Mathematical Sciences Issue, 65-68, 3307-3319.
[19] A. Abashar, M. Mamat, M. Rivaie, I. Mohd and O. Omer (2014). The proof of sufficient descent condition for a new type of conjugate gradient methods. AIP Conference Proceedings, 1602, 296-303.