International Journal of Engineering & Technology, 7 (3.28) (2018) 348-353

International Journal of Engineering & Technology

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

Research paper

An Extension of Polak-Ribière-Polyak Method

Mahmoud Dawahdeh[1]*, Mustafa Mamat[1], Ahmad Alhawarat[2], Mohd Rivaie[3], Mohamad Afendee Mohamed[1]

[1]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia*
[2]*Department of Mathematics, College of Science, Isra University, Amman, Jordan*
[3]*Department of Computer Science and Mathematics, Universiti Technologi MARA (UiTM) Terengganu, Kuala Terengganu Campus, Malaysia*
**Corresponding author E-mail: mdawahdeh1976@yahoo.com*

Abstract

The conjugate gradient method has been widely used for finding solution for the large-scale unconstrained optimization. Fields such as computer science and engineering are the two most frequently engaged, because of its simplicity, the speed of getting the solution and the minimal storage requirement. This study presents an extended conjugate gradient method of Polak-Ribière-Polyak with the strong Wolfe-Powell (SWP) line search satisfying some properties such as sufficient descent and global convergence. For the purpose of experimentation, a set of 141 test problems have been used. The results showed that our proposed method has surpass the others in terms of efficiency and robustness.

Keywords: *Conjugate gradient method; global convergence; strong Wolfe-Powell; sufficient descent property; unconstrained optimization.*

1. Introduction

Conjugate gradient (CG) method is a line search algorithm mostly known for its wide application in solving unconstrained optimization problems. Its low memory requirements and global convergence properties makes it one of the most preferred method in real life application such as in economics, engineering and physics [1]. They are designed to solve problem of the form:

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

where $f: R^n \to R$ is continuously differentiable function. The $g(x) = \nabla f(x)$ is the gradient. For an initial point ($x_0 \in R^n$), the nonlinear CG method generates a sequence by using the recurrence relation

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

where x_k is the current iteration point, $\alpha_k > 0$ is a step length obtained by a line search and. d_k is the search direction which defined by:

$$d_k = \begin{cases} -g_k & , \ k = 0 \\ -g_k + \beta_k d_{k-1} & , \ k \geq 1 \end{cases} \tag{3}$$

where the scalar β_k is known as the conjugate gradient coefficient and $g_k = g(x_k)$. There are two categories of line search that are exact and inexact line search, which can be used to compute α_k. The following is the definition for an exact line search.

$$f(x_k + \alpha_k d_k) = \min f(x_k + \alpha d_k) \ , \ \alpha \geq 0 . \tag{4}$$

By and large, exact line search requires the step length to have an exact value and as such drives the rise in the computational power. To address this problem, researchers typically employ inexact line search. One that is commonly known inexact line search is the strong Wolfe-Powell (SWP).

The method relies heavily on the function reduction and it searches for α_k via narrowing the search area. This way, α_k becomes nearer to the local minimum.. SWP can be defined as follows:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{5}$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leq \sigma |g_k^T g_k| \tag{6}$$

where $0 < \delta < \sigma < 1$.

The SWP is a modified of a line search called weak Wolfe-Powell (WWP) which is given in (5) and

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T g_k \tag{7}$$

The scalar β_k comes with different formulas such that found in Hestenes-Stiefel (HS) [2], Fletcher-Reeves (FR) [3], Polak-Ribière-Polyak (PRP) [4], Gilbert and Nocedal (PRP+) [5], Conjugate Descent (CD) [6], Liu and Storey (LS) [7], Dai-Yuan (DY) [8], Wei et al. (WYL) [9], defined as follows:

$$\beta_K^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}, \tag{8}$$

$$\beta_K^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \tag{9}$$

$$\beta_K^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}, \tag{10}$$

$$\beta_K^{PRP+} = max\{\beta_K^{PRP}, 0\}, \tag{11}$$

$$\beta_K^{CD} = \frac{\|g_k\|^2}{d_{k-1}^T g_{k-1}}, \tag{12}$$

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

$$\beta_K^{LS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T g_{k-1}}, \quad (13)$$

where $y_{k-1} = g_k - g_{k-1}$.

$$\beta_K^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T (g_k - g_{k-1})}, \quad (14)$$

$$\beta_k^{WYL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2} \quad (15)$$

Initially, Hestenes-Stiefel formula [2] was proposed in 1952 to be used for solving the quadratic functions. This is followed by Fletcher and Reeves [3] which was introduced in 1964 specifically for nonlinear functions. Zoutendijk [10] initiated the study on the convergence properties of FR method, followed by another work from Al-Baali [11] and Guanghui et al. in [12] using the SWP line search with $\sigma < 1/2$.

Using an exact line search, Elijah and Ribiere [4] proved that PRP method is globally convergent. However, in later time, Powell [13] proved that there could be a non-convex function that could drive PRP method to not globally converge by using a counterexample. As a result, Powell introduced nonnegative PRP method (PRP+), which was later studied by Gilbert and Nocedal [5] proved to be globally convergence with $\beta_K^{PRP+} = max\{\beta_K^{PRP}, 0\}$, under complicated line searches. However, for a general nonlinear function, the convergence of PRP+ under the SWP line search cannot be guaranteed. Consequently, Wei et al. [9] proposed a positive conjugate gradient method that imitates the original PRP method. Further improvement from this idea can be found in [14–15].

$$\beta_k^{VHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T y_{k-1}} \quad (16)$$

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{w|g_k^T d_{k-1}| + \|g_{k-1}\|^2} \quad (17)$$

where $w \geq 1$.

Following to that, based on a modified version of β_k^{WYL}, Zhang [16] proposed the following

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2} \quad (18)$$

Many other studies CG formulas which focus on the robustness and the efficiency, see [17–20]. The structure of this paper is as follows. Section 2 describes our new CG formula and the algorithm. In Section 3, the convergence analysis presented. Numerical results are done in Section 4. Finally in section 5 we conclude our findings.

2. The Newly Proposed Algorithm

This section introduces β_k^{DMAR} an extended version to β_k^{WYL} and β_k^{NPRP} method. It is known as DMAR method. The idea of the newly proposed formula mainly comes from Zhang [16]. DMAR denotes Dawahdeh, Mustafa, Ahmad, and Rivaie, and its formula is given in (19).

$$\beta_k^{DMAR} = \begin{cases} \frac{\|g_k\|^2 - \mu_k |g_k{}^T g_{k-1}|}{\|g_{k-1}\|^2}, & \|g_k\|^2 \geq \mu_k |g_k{}^T g_{k-1}| \\ 0 & , \ otherwise \end{cases}. \quad (19)$$

where $\|.\|$ is the Euclidean norm. We define μ_k as

$$\mu_k = \frac{\|g_k\|}{\|y_{k-1}\|^2} = \frac{\|g_k\|}{\|g_k - g_{k-1}\|^2}.$$

It is worth noting that

$$0 \leq \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_k - g_{k-1}\|^2} |g_k{}^T g_{k-1}|}{\|g_{k-1}\|^2} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2} = \beta_K^{FR} \text{ and so,}$$

$$0 \leq \beta_K^{DMAR} < \beta_K^{FR}. \quad (20)$$

Hence, according to the argument presented in [5], β_k^{DMAR} will inherit all of the advantages and properties of β_K^{FR}.

Algorithm 1

S1: Select the initial point x_0. Set the initial search direction $d_0 = -g_0$ and let $k = 0$.
S2: Calculate d_k from (3) and (19).
S3: Calculate α_k from (5) and (6).
S4: Recalculate x_{k+1} from (2).
S5: If $\|g_k\| \leq 10^{-6}$ then exit; else let $k = k + 1$ and go to S2.

3. Global Convergence Analysis

This section is dedicated to the study of the convergence properties of β_k^{DMAR}. Keep in mind that the search direction must satisfy sufficient descent condition as a basis for proving global convergence. Consider the following

$$g_k^T d_k \leq 0$$

Then, we have $f(x_{k+1}) < f(x_k)$. If it is extended to the following,

$$g_k^T d_k \leq -C\|g_k\|^2, \text{ when } k \geq 0, C > 0. \quad (21)$$

We said that (21) satisfies sufficient descent condition. Now, consider the coefficient β_k^{DMAR} satisfies the following

$$0 \leq \beta_k^{DMAR} \leq \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$
$$0 \leq \beta_{k+1}^{DMAR} \leq \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad (22)$$

Lemma 1: Given g_k and d_k with $\sigma \leq 3/4$, then $\forall k \geq 0$, we have

$$\|g_k\| < 2\|d_k\|,$$
$$\|g_k\|^2 < 4\|d_k\|^2,$$
$$-\|d_k\| < -\frac{\|g_k\|}{2}. \quad (23)$$

Proof: We prove using mathematical induction. Obviously the statement is true for $k = 0$. Suppose that (23) also holds true for some $k > 0$., then

$$\|g_{k+1} + d_{k+1}\|^2 = (g_{k+1} + d_{k+1})^T (g_{k+1} + d_{k+1})$$
$$= \|g_{k+1}\|^2 + \|d_{k+1}\|^2 + 2g_{k+1}^T d_{k+1}.$$

Since $g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \beta_{k+1}^{DMAR} g_{k+1}^T d_k$, we deduce that

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

which implies that

$$\|g_{k+1} + d_{k+1}\|^2 \leq \|d_{k+1}\|^2 - \|g_{k+1}\|^2 + 2|\beta_{k+1}^{DMAR}||g_{k+1}^T d_k|.$$

Applying the SWP conditions and $0 \leq \beta_{k+1}^{DMAR}$ from (22), we come to the relation

$$\|g_{k+1} + d_{k+1}\|^2 \leq \|d_{k+1}\|^2 - \|g_{k+1}\|^2 + 2\sigma\, \beta_{k+1}^{DMAR} |g_k^T d_k|.$$

Next we substitute (22) and apply Cauchy-Schwartz inequality to obtain

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

which implies

$$\|g_{k+1} + d_{k+1}\|^2 \leq \|d_{k+1}\|^2 - \|g_{k+1}\|^2 + 2\sigma \|g_{k+1}\|^2 \left(\frac{\|d_k\|}{\|g_k\|}\right)$$
$$= \|d_{k+1}\|^2 - \|g_{k+1}\|^2 - 2\sigma \|g_{k+1}\|^2 \left(\frac{-\|d_k\|}{\|g_k\|}\right)$$

By applying the induction hypothesis in (23), we have

$$\|g_{k+1} + d_{k+1}\|^2 \leq$$
$$\|d_{k+1}\|^2 - \|g_{k+1}\|^2 - 2\sigma \|g_{k+1}\|^2 \left(\frac{-\|g_k\|}{2\|g_k\|}\right)$$
$$= \|d_{k+1}\|^2 - \|g_{k+1}\|^2 - 2\sigma \|g_{k+1}\|^2 \left(\frac{-1}{2}\right)$$
$$= \|d_{k+1}\|^2 - \|g_{k+1}\|^2 + \sigma \|g_{k+1}\|^2,$$

which means that

$$\|g_{k+1} + d_{k+1}\|^2 \leq \|d_{k+1}\|^2 + (\sigma - 1)\|g_{k+1}\|^2.$$

Hence,

$$\|g_{k+1} + d_{k+1}\|^2 + (1-\sigma)\|g_{k+1}\|^2 \leq \|d_{k+1}\|^2.$$

Since $1 - \sigma > 0$, we get

$$(1-\sigma)\|g_{k+1}\|^2 \leq \|d_{k+1}\|^2,$$

which implies that

$$\|g_{k+1}\|^2 \leq \left(\frac{1}{1-\sigma}\right) \|d_{k+1}\|^2.$$

Therefore,

$\|g_{k+1}\|^2 \leq 4\|d_{k+1}\|^2$, whenever $\sigma \leq 3/4$, and then

$$\|g_{k+1}\| < 2\|d_{k+1}\|,$$

That guarantee (23) is true for $k+1$. As such, the proof is completed.
Based on Lemma 1, d_k and g_k, with $\sigma \leq 3/4$, is somehow related as

$$\frac{\|g_k\|^2}{4} < \frac{\|d_k\|^2}{1}, \text{ or } \frac{1}{\|d_k\|^2} < \frac{4}{\|g_k\|^2}, \text{ for all } k \geq 0. \quad (24)$$

Based on sufficient descent property found in (21), the coming theorem leads the way to global convergence.

Theorem 1: Given g_k and d_k with $\sigma \leq 3/4$, then for all $k \geq 0$, we have a relation

$$\frac{-1}{1-\sigma} < \frac{g_k^T d_k}{\|g_k\|^2} < \frac{2\sigma - 1}{1-\sigma} \quad (25)$$

Proof: Using mathematical induction. Obviously, the statement is true when $k = 0$. Let assume (25) is true, for some $k > 0$, it follows that

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

which implies

$$\frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} = -1 + \beta_{k+1}^{DMAR} \frac{g_{k+1}^T d_k}{\|g_{k+1}\|^2}$$
$$= -1 + \beta_{k+1}^{DMAR} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} \frac{g_{k+1}^T d_k}{\|g_k\|^2} \quad (26)$$

From the SWP condition, we have

$$\sigma g_k^T d_k \leq g_{k+1}^T d_k \leq -\sigma g_k^T d_k,$$

which together with that $0 \leq \beta_{k+1}^{DMAR}$, we get

$$\sigma\, \beta_{k+1}^{DMAR} g_k^T d_k \leq \beta_{k+1}^{DMAR} g_{k+1}^T d_k \leq -\sigma\, \beta_{k+1}^{DMAR} g_k^T d_k \quad (27)$$

Combining (26) with (27), we come to the relation

$$-1 + \sigma\, \beta_{k+1}^{DMAR} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} \frac{g_k^T d_k}{\|g_k\|^2} \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2}$$
$$\leq -1 - \sigma\, \beta_{k+1}^{DMAR} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} \frac{g_k^T d_k}{\|g_k\|^2}.$$

Since $0 \leq \beta_{k+1}^{DMAR}$, we apply induction hypothesis (25) to obtain

$$-1 - \beta_{k+1}^{DMAR} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} \left(\frac{\sigma}{1-\sigma}\right) \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2}$$
$$\leq -1 + \beta_{k+1}^{DMAR} \frac{\|g_k\|^2}{\|g_{k+1}\|^2} \left(\frac{\sigma}{1-\sigma}\right).$$
$$-1 - \left(\frac{\sigma}{1-\sigma}\right) \leq \frac{g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^2} \leq -1 + \left(\frac{\sigma}{1-\sigma}\right).$$

Hence,

$$\frac{-1}{1-\sigma} < \frac{g_k^T d_k}{\|g_k\|^2} < \frac{2\sigma - 1}{1-\sigma},$$

This shows that the statement holds for $k+1$, and thus concludes the proof. When working with global convergence analysis of the CG method with SWP line search, we frequently assume the following assumption, which we would consider hold throughout our studies unless otherwise stated:

i) The $\Omega = \{ x \in R^n : f(x) \leq f(x_0) \}$ is a bounded level set.
ii) The f is continuously differentiable function with Lipschitz gradient, also continuous within neighborhood N of Ω, that is, $\exists\, L > 0$ such that

$$\|g(x) - g(y)\| \leq L\|x - y\|, \forall\, x, y \in N.$$

The following lemma was proved by Zoutendijk in [10], it is used to analyze the global convergence of the CG method.

Lemma 2: For any CG method of the form (2) and (3), with d_k and α_k obtained via one-dimensional search direction, the following Zoutendijk condition holds

$$\sum_{K=0}^{\infty} \|g_k\|^2 \cos^2\theta_k < \infty, \quad (28)$$

where θ_k is the angle between $-g_k$ and d_k, given by

$$\cos\theta_k = \frac{-g_k^T d_k}{\|g_k\|\|d_k\|}. \quad (29)$$

See [10] for the proof of Lemma 2.

Lemma 3: Consider a sequence x_k, $\forall k \geq 0$ with $\sigma \leq 3/4$, we have

$$\sum_{K=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$$

Proof: Multiplying (25) by $\frac{\|g_k\|}{\|d_k\|}$ and by using (29), we get

$$c_2 \frac{\|g_k\|}{\|d_k\|} < \cos\theta_k < c_1 \frac{\|g_k\|}{\|d_k\|}, \text{ for all } k \geq 0, \quad (30)$$

where $c_1 = \frac{1}{1-\sigma}$ and $c_2 = \frac{2\sigma-1}{1-\sigma}$.
Combining (28) and (30) together, we have the relation

$$\sum_{K=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$$

Therefore, this proof is completed.

Theorem 2: Consider a sequence x_k with $\sigma \leq 3/4$. By assume that Assumption 1 holds, we have

$$\lim_{K\to\infty}\inf \|g_k\| = 0.$$

Proof: Suppose the opposite, hence $\exists\, \lambda > 0$ which is a constant and integer k_1 s.t.

$$\|g_k\| \geq \lambda, \forall\, k > k_1,$$

which means

$$\frac{1}{\|g_k\|^2} \leq \frac{1}{\lambda^2}, \text{ for all } k > k_1 \text{ and } \|g_k\| \neq 0. \quad (31)$$

Rewriting (3) as $d_k + g_k = \beta_k^{DMAR} d_{k-1}$, and squaring it we get

$$\|d_k\|^2 + \|g_k\|^2 + 2g_k^T d_k = (\beta_k^{DMAR})^2 \|d_{k-1}\|^2,$$

then,

$$\|d_k\|^2 = -\|g_k\|^2 - 2g_k^T d_k + (\beta_k^{DMAR})^2 \|d_{k-1}\|^2.$$

Applying Theorem 1, we have

$$\|d_k\|^2 < -\|g_k\|^2 + \left(\frac{2}{1-\sigma}\right) \|g_k\|^2 + (\beta_k^{DMAR})^2 \|d_{k-1}\|^2,$$

which leads to

$$\|d_k\|^2 < \left(\frac{1+\sigma}{1-\sigma}\right) \|g_k\|^2 + (\beta_k^{DMAR})^2 \|d_{k-1}\|^2. \quad (32)$$

By substituting (22) into (32), we obtain

$$\|d_k\|^2 < \left(\frac{1+\sigma}{1-\sigma}\right) \|g_k\|^2 + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2. \quad (33)$$

By multiplying both sides of (33) by $\left(\frac{1}{\|g_k\|^4}\right)$, we get

$$\begin{aligned}\frac{\|d_k\|^2}{\|g_k\|^4} &< \left(\frac{1+\sigma}{1-\sigma}\right)\frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} \\ &= \left(\frac{1+\sigma}{1-\sigma}\right)\frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^2}\left(\frac{1}{\|g_{k-1}\|^2}\right) \\ &= \left(\frac{1+\sigma}{1-\sigma}\right)\frac{1}{\|g_k\|^2} - \left(-\frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^2}\right)\left(\frac{1}{\|g_{k-1}\|^2}\right). \quad (34)\end{aligned}$$

From (23), we have $\left(-\frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^2}\right) < -\frac{1}{4}$, so (34) becomes

$$\frac{\|d_k\|^2}{\|g_k\|^4} < \left(\frac{1+\sigma}{1-\sigma}\right)\frac{1}{\|g_k\|^2} - \left(-\frac{1}{4}\right)\left(\frac{1}{\|g_{k-1}\|^2}\right)$$

$$= \left(\frac{1+\sigma}{1-\sigma}\right)\frac{1}{\|g_k\|^2} + \left(\frac{1}{4}\right)\left(\frac{1}{\|g_{k-1}\|^2}\right). \quad (35)$$

Combining (31) and (35) together, we have

$$\frac{\|d_k\|^2}{\|g_k\|^4} < \left(\frac{1+\sigma}{1-\sigma} + \frac{1}{4}\right)\frac{1}{\lambda^2}, \text{ for all } k \geq k_1 + 1,$$

which means that

$$\frac{\|g_k\|^4}{\|d_k\|^2} > \left(\frac{4-4\sigma}{5+3\sigma}\right)\lambda^2, \text{ for all } k \geq k_1 + 1. \quad (36)$$

Since (36) is true for all $k \geq k_1 + 1$, then

$$\sum_{K=k_1+1}^{n} \frac{\|g_k\|^4}{\|d_k\|^2} > \sum_{K=k_1+1}^{n} \left(\frac{4-4\sigma}{5+3\sigma}\right)\lambda^2 = \left(\frac{4-4\sigma}{5+3\sigma}\right)\lambda^2 (n - k_1). \quad (37)$$

From (37), we get

$$\sum_{K=0}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} > \sum_{K=k_1+1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = \lim_{n\to\infty} \sum_{K=k_1+1}^{n} \frac{\|g_k\|^4}{\|d_k\|^2} > \lim_{n\to\infty} \sum_{K=k_1+1}^{n} \left(\frac{4-4\sigma}{5+3\sigma}\right)\lambda^2 (n - k_1) = \infty.$$

However, this is contradictory to Lemma 3. Hence, the proof is concluded.

4. Results and Discussion

This section aims at evaluating our methods against other existing methods such as the WYL and NPRP, using some standard test function as shown in Table 1 [21-23]. There are four different parameters considered namely the time taken by the CPU, the number of iterations, the number of function evaluations and the number of gradient evaluations.

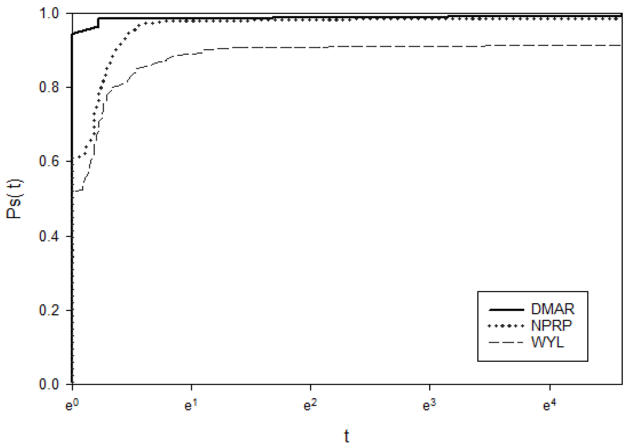

Fig. 1: The number of iterations

Let $\delta = 0.00$, $\sigma = 0.1$., and ε is chosen to be 10^{-6}. Here, $\|g_k\| \leq 10^{-6}$ is set to hold the stopping criteria. Any success is recorded for having the iterations number of not more than 1000. Our experimentation is done with Matlab R2017a subroutine running on a PC powered by an Intel R Core TM, i5-2410 processor having 3 GB of memory.

We use the idea of performance profile introduced by Dolan and Mor´ [24], the results are shown in Fig. 1 through Fig. 4. From the figure, we can determine the best method as the one with high values of $p_s(t)$ or observably located in the upper right corner of the plot. The $p_s(t)$ tells the rate of successfully solving the test problem. Clearly from Figures 1, 3 and 4, DMAR method significantly outclassed the other two methods in terms of the number of iterations, the number of gradient evaluations and the number of function evaluations. Whereas, in Figure 2, somewhere after it begins, WYL has slightly outperformed DMAR for some values of input, whereas at all other times, DMAR performed the best.

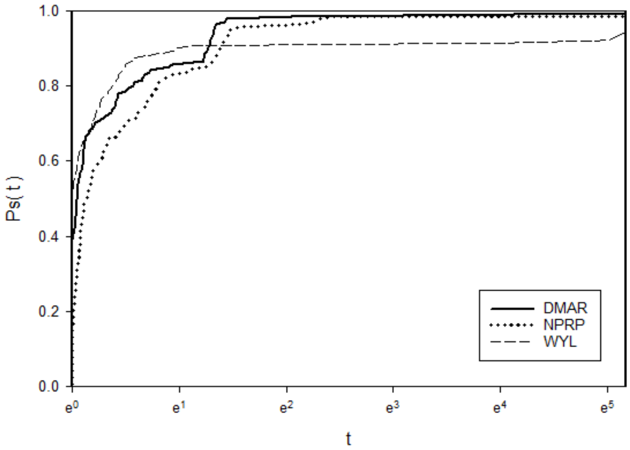

Fig. 2: The CPU time

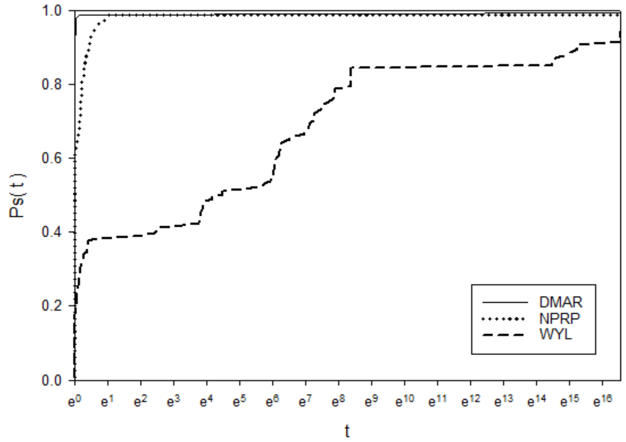

Fig. 3: The gradient evaluation

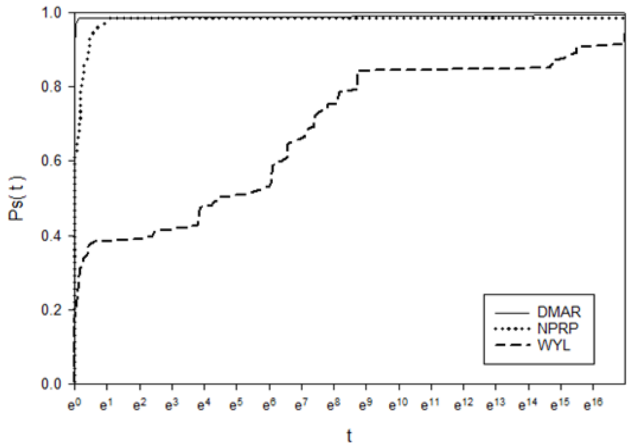

Fig. 4: The function evaluation

Table 1: Test functions

N Function	Dimension	Initial Points
1 GENROSEN2	2	(2,2)
2 QUARTC	2,500,1000, 5000,10000	(2, 2, . . ., 2), (5, 5, . .., 5) (10, 10, . ., 10) (15, 15,..)
3 Extended Block Diagonal	2,500, 1000,5000	(0.1, 0.1, . . ., 0.1)
4 Sphere	2,500,1000, 5000,10000	(1, 1, . . ., 1), (5, 5, . . ., 5)
5 Generalized Quartic GQ1	2,500,1000, 5000,10000	(1, 1, . . ., 1)
6 Diagonal 4	2,500,1000, 5000,10000	(1, 1, . . ., 1)
7 EDENSCH	2,500,1000, 5000,10000	(0, 0, . . ., 0)
8 DENSCHNB	2,500,1000, 5000,10000	(1, 1, . . ., 1)
9 DENSCHNC	2,500, 1000,10000	(2,2, . . ., 2)
10 Extended DENSCHNB	2,500,1000, 5000,10000	(1, 1, . . ., 1), (5, 5, . . ., 5)
11 Generlized Tridiagonal 2	2	(10, 10)
12 EXTROSNB	2	(1.5, 1.5)
13 Raydan 2	2,500,1000, 5000,10000	(1, 1, . . ., 1), (10, 10, . . .,10)
14 HIMMELBC	2,500,1000, 5000,10000	(1, 1, . . ., 1)
15 DIXMAANA	6000,9000, 12000	(2, 2, . . ., 2)
16 DIXMAANB	300,6000, 9000,12000	(2, 2, . . ., 2)
17 BIGGSB1	2	(0.5, 0.5)
18 EG2	2,1000, 5000,10000	(1, 1, . . ., 1), (2, 2, . . ., 2)
19 DENSCHNF	2,500,1000, 5000,10000	(1.5, 1.5, . . ., 1.5)
20 HIMMELBH	2	(1.5, 1.5)
21 LIARWHD	2,1000, 5000,10000	(4, 4, . . ., 4)
22 TRIDIA	2	(1, 1)
23 Six Hump	2	(1, 1), (10, 10), (15, 15)
24 EG3	2	(1, 1)
25 A Quadratic QF2	2	(0.5, 0.5), (1, 1)
26 FLETCHCR		
27 Diagonal 1	2	(0.5, 0.5), (5, 5)
	2	(0, 0),(1, 1),
28 Hager		(2, 2) ,(3, 3)
29 Booth	2	(1, 1)
30 Zettl	2	(1, 1), (3, 3)
31 Tridiagonal Double Borded	2	(0, 0)
	2	(-1, -1)
32 STAIRCASES1		
33 Raydan 1	2	(1, 1)
34 Extended Trigonometric	2	(1, 1)
	2	
35 ENGVAL8		(0.2, 0.2)
	2	(2, 2)
36 DESCHANA	2,500,1000, 5000,10000	(1, 1, . . ., 1)

5. Conclusion

Many areas of computer science and engineering make use of conjugate gradient methods [25–29]. This study introduced a new breed of coefficient for a CG algorithm. The experimental results revealed the superiority over existing methods. Moreover, we proved the globally converged properties of DMAR method with the strong Wolf Powell (SWP) line search.

Acknowledgement

This project is funded by the Malaysian Fundamental Research Grant Scheme (FRGS/1/2017/STG06/Unisza/01/1) and Research Management, Innovation & Commercialization Center, Universiti Sultan Zainal Abidin.

References

[1] Cai J., Li Q., Li L., Peng H., & Yang Y., "A fuzzy adaptive chaoticant swarm optimization for economic dispatch",

International Journal of Electrical Power and Energy Systems, Vol. 34, No. 1, (2012), pp. 154–160.
[2] Hestenes M. R. & Steifel E., “Method of conjugate gradient for solv-ing linear equations”, J. Res. Nat. Bur. Stand., 49, (1952), 409-436.
[3] Fletcher R. & Reeves C., “Function minimization by conjugate gradients”, Computer Journal, Vol. 7, No. 2, (1964), pp. 149-154.
[4] Polak E. & Ribiere G., Note sur la convergence de directions con-jugees. Rev. Francaise Informat Recherche Operationelle, 3E Annee 16, (1969), 35-43.
[5] Gilbert J. C. & Nocedal J., “Global convergence properties of conjugate gradient methods for optimization”, SIAM Journal on optimization, Vol. 2, No. 1, (1992), pp. 21-42.
[6] Fletcher R., Practical methods of optimization, John Wiley and Sons, 2001.
[7] Liu Y. & Storey C., “Efficient generalized conjugate gradient al-gorithms, part 1: theory”, Journal of Optimization Theory and Ap-plications, Vol. 69, No. 1, (1991), pp. 129-137.
[8] Dai Y. H. & Yuan Y., Nonlinear conjugate gradient methods, Shanghai Science and Technology Publisher, 2000.
[9] Wei Z., Yao S., & Liu L., “The convergence properties of some new conjugate gradient methods”, Applied Mathematics and Computation, Vol. 183, No. 2, (2006), pp. 1341–1350.
[10] Zoutendjik G., Nonlinear programming computational methods, in J. Abadie (Ed.), Interger and Nonlinear Programming, Amsterdam: North-Holland, (1970), pp. 37-86.
[11] Al-Baali M., “Descent property and global convergence of the Fletcher–Reeves method with inexact line search”, IMA Journal of Numerical Analysis, Vol. 5, No. 1, (1985), pp. 121-124.
[12] Guanghui L., Jiye H., & Hongxia Y., “Global convergence of the fletcher-reeves algorithm with inexact line search”, Applied Mathematics-A Journal of Chinese Universities, Vol. 10, No. 1, (1995), pp. 75–82.
[13] Powell M. J. D., “Nonconvex minimization calculations and the conjugate gradient method”, Lecture Notes in Mathematics, 1066, (1984), pp. 122-141.
[14] Shengwei Y., Wei Z., & Huang H., “A note about WYL’s conjugate gradient method and its applications”, Applied Mathematics and Computation, Vol. 191, No. 2, (2007), pp. 381–388.
[15] Dai Z. & Wen F., “Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property”, Applied Mathematics and Computation, Vol. 218, No. 14, (2012), pp. 7421–7430.
[16] Zhang L., “An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation”, Applied Mathematics and Computation, Vol. 215, No. 6, (2009), pp. 2269–2274.
[17] Rivaie M., Mamat M., June L. W., & Mohd I., “A new class of nonlinear conjugate gradient coefficients with global convergence properties”, Applied Mathematics and Computation, Vol. 218, No. 22, (2012), pp. 11323–11332.
[18] Alhawarat A., Mamat M., Rivaie M., & Mohd I., “A new modification of nonlinear conjugate gradient coefficients with global convergence properties”, International Journal of Mathematical, Computational, Statistical, Natural and Physical Engineering, Vol. 8, No. 1, (2014), pp. 54-60.
[19] Rivaie M., Mamat M., & Abashar A., “A new class of nonlinear conjugate gradient coefficients with exact and inexact line searches”, Applied Mathematics and Computation, Vol. 268, (2015), pp. 1152-1163.
[20] Abashar A., Mamat M., Rivaie M., Mohd I., & Omer O., “The proof of sufficient descent condition for a new type of conjugate gradient methods”, AIP Conference Proceedings, Vol. 1602, (2014), pp. 296-303.
[21] Bongartz I., Conn A. R., Gould N., & Toint P. L., “CUTE: constrained and unconstrained testing environment”, ACM Transactions on Mathematical Software, Vol. 21, No. 1, (1995), pp. 123–160.
[22] Andrei N., “An unconstrained optimization test functions collec-tion”, Adv. Model. Optim, Vol. 10, No. 1, (2008), pp. 147-161.
[23] Adorio E. P. & Diliman U. P., “Mvf-multivariate test functions library in C for unconstrained global optimization”, 2005, http://www.geocities.ws/eadorio/mvf.pdf.
[24] Dolan E. D. & Jorge J. M., “Benchmarking optimization software with performance profiles”, Mathematical Programming, Vol. 91, No. 2, (2002), pp. 201-213.
[25] Abidin N. Z., Mamat M., Dangerfield B., Zulkepli J. H., Baten M. A., & Wibowo A., “Combating obesity through healthy eating behavior: a call for system dynamics optimization”, Plos One, Vol. 9, No. 12, (2014), pp. 1-17.
[26] Mamat M., Subiyanto, & Kartono A., “Mathematical model of cancer treatments using immunotherapy, chemotherapy and biochemotherapy”, Appl. Math. Sci, Vol. 7, No. (5-8), (2013), 247-261.
[27] Mamat M., Deraman S. K., Noor N. M. M., & Rokhayati Y., “Diet problem and nutrient requirement using fuzzy linear programming approach”, Asian Journal of Applied Sciences, Vol. 5, No. 1, (2012), pp. 52-59.
[28] Rivaie M., Mamat M., Mohd I., & Fauzi M., “Comparative study of conjugate gradient coefficient for unconstrained optimization”, Aus. J. Bas. Appl. Sci, Vol. 5, No. 9, (2011), pp. 947-951.
[29] Mamat M., Rokhayati Y., Noor N. M. M., & Mohd I., “Optimizing human diet problem with fuzzy price using fuzzy linear programming approach”, Pakistan Journal of Nutrition, Vol. 10, No. 6, (2011), pp. 594-598.