Journal of the Nigerian Mathematical Society
Vol. 39, Issue 2, pp. 223-237, 2020

A MODIFIED SPECTRAL CONJUGATE GRADIENT METHOD FOR SOLVING UNCONSTRAINED MINIMIZATION PROBLEMS

A. A. DANHAUSA[1], M. R. ODEKUNLE[2] AND A. S. ONANAYE[3]

ABSTRACT. The development a modified spectral conjugate gradient method for solving unconstrained minimization problems is considered in this paper. A new Conjugate (update) parameter isobtained by the idea of Dai-Kou's technique for generating conjugate parameters. A new spectral parameter is also presented based on quasi-Newton direction and quasi-Newton condition. Under the strong Wolfe line search, the proposed method (DOO) is proved to be globally convergent. Numerical results showed that the algorithm takes lesser number of iterations to obtain the minimum of a given function.

Keywords and phrases: Unconstrained minimization problem, spectral conjugate gradient method, global convergence
2010 Mathematical Subject Classification: 90C25, 90C30

1. INTRODUCTION

Let us consider the unconstrained optimization problem:

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

where $f : R^n \to R$ is continuously differentiable. The gradient vector of $f(x)$ is denoted by $g(x)$, that is , $g(x) = \nabla f(x)$. Conjugate Gradient Methods (CGMs) are very effective for solving large-scale unconstrained optimization problem . The basic approach of the traditional CGMs is to generate a sequence $\{x_k\}$of iterates according to

$$x_{k+1} = x_k + \alpha_k d_k \tag{2}$$

Received by the editors August 17, 2019; Revised January 08, 2020 ; Accepted: February 05, 2020

www.nigerianmathematicalsociety.org; Journal available online at https://ojs.ictp.it/jnms/

[1]Corresponding author

[2],

[3]Supported by ICTP, grant number T4538

where α_k is the step size minimizing f approximately along d_k from x_k by a suitable line search. And the next search direction is generated by

$$d_{k+1} = -g(x_{k+1}) + \beta_k d_k \tag{3}$$

with a suitable conjugate (update) parameter β_k , where $d_1 = -g(x_1)$

It is an established fact that different conjugate parameters correspond to different CGMs and their numerical performance and convergence also varies. Well-known formulas for β_k are the ones developed by Dai and Yuan[5], Fletcher conjugate descent[7], Fletcher and Reeves[8], Hestenes and Stiefel [9], Liu and Storey[12], and Polak-Rebie're- Polyak [13]. To attain good computational performance and to maintain the attrative feature of strong global convergence, researchers paid special attention to spectral CGMs. Barzilai and Borwein [3] considered a spectral conjugate gradient method for solving large scale unconstrained optimization problem (1) where they define their step-length as $x_{k+1} = x_k - \frac{1}{\alpha_k} g_k$. g_k is the gradient vector of f at x_k and the scalar α_k is given by $\alpha_k = \frac{s_{k-1}^T y_{k-1}}{s_{k-1}^T s_{k-1}}$ where $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1.}$ Every iteration of their method requires only O(n) floating point operations and a gradient evaluation. No matrix computations and no line searches are required during the process. More interesting from a theoretical point of view is that the method does not guarantee descent in the objective function. Andrei [1], proposed a spectral CGM, where the search direction is given by $d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_k^N s_{k,} \quad d_k = -g_1$ and θ_{k+1} is the spectral parameter. The directions yielded by d_{k+1}, β_k^N and θ_{k+1} possess descent property as $g_k^T d_{k+1 \leq -\left(\theta_{k+1} - \frac{1}{4}\right) \|g_{k+1}\|^2}$. This shows that the direction is descent only in case $\theta_{k+1} > \frac{1}{4}$. Therefore, to obtain descent in any case, Andrei reset $\theta_{k+1} = 1$ in case $\theta_{k+1} \leqslant \frac{1}{4}$. Dai and Kou [4] proposed a family of generating conjugate parameters as:

$$\beta_{k(\tau_k)} = \frac{y_k^T g_{k+1}}{d_k^T y_k} - \left(\tau_k + \frac{\|y_k\|^2}{s_k^T y_k} - \frac{s_k^T y_k}{\|s_k\|^2}\right) \frac{s_k^T g_{k+1}}{d_k^T y_k}$$

where τ_k is a pending parameter. Dai and Kou [4] emphasized that:

$\tau_k = \frac{s_k^T y_k}{\|s_k\|^2}$ where $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1.}$

Jinbao, et al. [10], proposed a new spectral conjugate gradient method for large-scale unconstrained optimization based on quasi-Newton direction and quasi-Newton condition. They introduced a new approach for generating spectral parameters by a new double-truncating technique, which can ensure both the sufficient descent property of the search directions and the bounded property of the sequence of spectral parameters. Despite the success recorded by this method, it does not guarantee descent direction in some instances in the objective function. This leads to large number of iterations to achieve convergence.

In this paper, the conjugate parameter of Polak, Ribie're and Polyak [13] will be combined with that of Dai and Kou [4] to obtain new β_k^p; quasi-Newton direction and condition will be used to obtain the required new search direction.

2.DERIVATION OF THE METHOD

2.1 Selection of the Modified Conjugate (Update) CG Parameter

Consider a spectral CGM of the form:

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_k d_{k,} \qquad d_1 = -g_1 \tag{4}$$

where β_k is the conjugate (update) parameter and θ_{k+1} is the spectral parameter. Combining the conjugate parameter of Polak, Ribie're and Polyak(β_k^{PRP}) with the conjugate parameters of Dai and Kou [4], new conjugate parameter is obtain as:

$$\beta_{k(\tau_k)} = \beta_k^{PRP} - \left(\tau_k + \frac{\|y_k\|^2}{s_k^T y_k} - \frac{s_k^T y_k}{\|s_k\|^2}\right) \frac{s_k^T g_{k+1}}{d_k^T y_k} \tag{5}$$

but

$$\tau_k = \frac{s_k^T y_k}{\|s_k\|^2} \quad \text{and} \quad \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2} = \frac{y_k^T g_{k+1}}{\|g_k\|^2} \tag{6}$$

Substitute (6) into (5) and simplify, to obtain β_k^P as

$$\beta_k^P = \frac{y_k^T g_{k+1}}{\|g_k\|^2} - \left(\frac{s_k^T y_k}{\|s_k\|^2} + \frac{\|y_k\|^2}{s_k^T y_k} - \frac{s_k^T y_k}{\|s_k\|^2}\right) \frac{s_k^T g_{k+1}}{d_k^T y_k}$$

$$\beta_k^P = \frac{y_k^T g_{k+1}}{\|g_k\|^2} - \frac{\|y_k\|^2 d_k^T g_{k+1}}{(d_k^T y_k)^2} \quad \forall \quad s_k^T = d_k^T \tag{7}$$

Substituting (7) into (4), the direction becomes

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_k^P d_{k,} \quad d_1 = -g_1 \tag{8}$$

2.2 Selection of the New Spectral Parameter (θ_{k+1}^P)

To obtain the new spectral parameter (θ_{k+1}^{P}) , the search direction d_{k+1} generated by (8) could be approximated by quasi-Newton direction as,

$$d_{k+1} \approxeq -H_{k+1}^{-1} g_{k+1} \tag{9}$$

where H_{k+1} is an approximation of the Hessian $\nabla^2 f(x_{k+1})$.
Equating (8) and (9) to have

$$\theta_{k+1} g_{k+1} = \frac{g_{k+1}}{H_{k+1}} + \beta_k^P d_{\,k}$$

Multiply through by H_{k+1} and S_k^T to have

$$\theta_{k+1} H_{k+1} g_{k+1} = g_{k+1} + \beta_k^P d_{\,k}\, H_{k+1} \tag{10}$$

From quasi-Newton condition

$$S_k^T H_{k+1} = y_k^T \tag{11}$$

Substitute (11) into (10), and then substitute for β_k^P to have

$$\theta_{k+1}^P = \frac{y_k^T d_k}{\|g_k\|^2} - \frac{1}{y_k^T g_{k+1}} \left(\frac{\|y_k\|^2 d_k^T g_{k+1}}{d_k^T y_k} - S_k^T g_{k+1} \right) \tag{12}$$

Thus, the new search direction in (8) becomes

$$d_{k+1} = -\theta_{k+1}^P g_{k+1} + \beta_k^P d_{k,} \quad d_1 = -g_1 \tag{13}$$

with β_k^P and θ_{k+1}^P defined by (7) and (12) respectively.

2.3 Descent Property of the Search Direction

The direction defined by (13) has the following descent property.

Theorem 1

The direction d_{k+1} generated by β_k^p satisfies $g_{k+1}^T d_{k+1} \leq -(\theta_{k+1}^p - \frac{1}{4}) \|g_{k+1}\|^2$, $\quad k = 0, 1, 2, ...,$ where $\theta_1^p = 1$. particularly, d_{k+1} is a descent direction when $\theta_{k+1} > \frac{1}{4}$.

Proof

For $k = 0$, from (13) the relation is true.

$$g_1^T d_1 = -\|g_1\|^2 \leq -(1 - \frac{1}{4}) \|g_1\|^2 .$$

For $k \geq 1$

Multiply (13) by g_{k+1}^T and Substitute for β_k^p to have

$$g_{k+1}^T d_{k+1} = -\theta_{k+1}^P \|g_{k+1}\|^2 + \frac{g_{k+1}^T y_k g_{k+1}^T d_k}{\|g_k\|^2} - \frac{\|y_k\|^2 (d_k^T g_{k+1})^2}{(d_k^T y_k)^2} \tag{14}$$

Taking into acccount the inequality $U^TV = \frac{1}{2}\left(\|U\|^2 + \|V\|^2\right)\ \forall U, V \in \mathbb{R}^n$, letting $U = \frac{1}{\sqrt{2}}g_{k+1}$ and $V = \sqrt{2}\frac{y_k g_{k+1}^T d_k}{d_k^T y_k}$, we have

$$\frac{g_{k+1}^T y_k g_{k+1}^T d_k}{d_k^T y_k} \leq \frac{1}{4}\|g_{k+1}\|^2 + \frac{\|y_k\|^2 (g_{k+1}^T d_k)^2}{(d_k^T y_k)^2} \tag{15}$$

Substitute (15) into (14) to have

$$g_{k+1}^T d_{k+1} = -\theta_{k+1}^P \|g_{k+1}\|^2 + \frac{1}{4}\|g_{k+1}\|^2 + \frac{\|y_k\|^2 (g_{k+1}^T d_k)^2}{(d_k^T y_k)^2} - \frac{\|y_k\|^2 (d_k^T g_{k+1})^2}{(d_k^T y_k)^2}$$

$$g_{k+1}^T d_{k+1} = -\left(\theta_{k+1}^P - \frac{1}{4}\right)\|g_{k+1}\|^2.$$

Therefore, the desired result holds ■

2.4 Wolfe Line Search

The Wolfe-type line search can bring good convergence of the associated CGMs. The two Wolfe-type line searches that will be used to generate the step size α_k are:

1. Standard Wolfe line Search , where $0 < \delta \leq \sigma < 1$

$$f(x_k + \alpha_K d_k) - f(x_k) \leq \delta \alpha_K g_k^T d_k \tag{16}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k$$

2. Strong Wolfe line search , where $0 < \delta \leq \sigma < 1$

$$f(x_k + \alpha_K d_k) - f(x_k) \leq \delta \alpha_K g_k^T d_k$$
$$\left|g(x_k + \alpha_k d)^T d_k\right| \leq -\sigma g_k^T d_k \tag{17}$$

Algorithm DOO (Danhausa, Odekunle and Onanaye)

The steps of our algorithm can be given as:

Step 1. Given an initial point $x_o \in R^n$, accuracy tolerance $\epsilon > 0$, positive parameters δ and σ such that $\delta < \sigma$. Let $d_1 = -g_{1,}$ set $k := 1$.

Step 2. Generate a step size by (16) and (17)

Step 3. Evaluate x_{k+1} by (2) and compute $g_{k+1} = \nabla f(x_{k+1})$. If $|g_{k+1}| < \epsilon$, stop. Otherwise, go to Step 4.

Step 4. Compute the conjugate (update) parameter β_k^P and then generate $\theta_{k+1.}^P$ by (7) and (12)

Step 5. Generate the next direction by (13). Set $k := k + 1$, go to Step 2.

3. CONVERGENCE ANALYSIS

In this section, the analysis of the convergence of Algorithm DOO is considered. For the proof of the global convergence, the following basic assumptions are needed.

Assupmtion 1

1. The level set $S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$ is bounded.

2. In a neighbourhood U of S, the function f is differentiable and it is Lipschitz continuous, i.e, thre exists a constant $L > 0$, such that $\|g(x) - g(y)\| \leq L\|x - y\|$ $\forall x, y \in U$.

Lemma 1

Consider a CGM of the form (2) for solving the unconstrained optimazation problem where d_k satisfies the descent condition $g_k^T d_k < 0$ and the step size satisfies the Standard Wolfe line searches (16) If Assumption 1 holds then $\sum\limits_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$

Proof

From (16), we have

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

$$(g_{k+1} - g_k)^T d_k \geq (\sigma - 1) g_k^T d_k \tag{18}$$

Noting that $\{x_k\} \subset S$ and by Lipschitz condition we have,

$$(g_{k+1} - g_k)^T d_k \leq L\|\alpha_k d_k\| \|d_k\| \leq L|\alpha_k| \|d_k\|^2 \leq L\alpha_k \|d_k\|^2 \tag{19}$$

Equating (18) and (19), we have

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

Substituting (20) into (16), to have

$$f_k - f_{k+1} \leq \delta \frac{(1 - \sigma)}{L} \frac{(g_k^T d_k)^2}{\|d_k\|^2}$$

Summing this relation over k and noting the boundedness of $\{f_k\}$, we can see that the sequence for the iteration converges ■

Lemma 2

Consider a spectral CGM, where the sequence $\{x_k\}$ are generated by (2) and (13); the spectral parameter $\theta_{k+1}^p > \frac{1}{4}$ and the step size

α_k is generated by the strong Wolfe line search (24). If Assumption 1 holds then

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty$$

Proof

Since $\theta_{k+1}^p > \frac{1}{4}$, it is clear that $g_k^T d_k < 0$ by theorem 1.From (13), make $\beta_k^P d_k$ the subject and square both sides

$$(\beta_k^p)^2 \|d_k\|^2 = \|d_{k+1}\|^2 + 2\theta_{k+1}^p g_{k+1}^T d_{k+1} + (\theta_{k+1}^p)^2 \|g_{k+1}\|^2$$

Note that,

$$\frac{\|d_{k+1}\|^2}{\|d_k\|^2} - (\beta_k^p)^2 \geq -(\theta_{k+1}^p)^2 \frac{\|g_{k+1}\|^2}{\|d_k\|^2} \tag{21}$$

Multiply (13) by g_{k+1}

$$g_{k+1}^T d_{k+1} - \beta_k^p d_k^T g_{k+1} = -\theta_{k+1}^p \|g_{k+1}\|^2$$

Therefore,

$$|\beta_k^p| \left|d_k^T g_{k+1}\right| + |g_{k+1} d_{k+1}| \geq \theta_{k+1}^p \|g_{k+1}\|^2 \tag{22}$$

Let $\sigma_{\max} = \{\sigma, \sigma_{1,}\sigma_2\} > 0$ where $0 \leq \delta \leq \sigma \leq \sigma_1 \leq \sigma_2$. Then, the step size α_k satifies (22) plus

$$\left|g(x_k + \alpha_k d)^T d_k\right| \leq -\sigma_{\max} g_k^T d_k = \sigma_{\max} \left|g_k^T d_k\right| \tag{23}$$

Equation (23), together with (22), further shows that

$$\sigma |\beta_k^p| \left|d_k^T g_k\right| + \left|g_{k+1}^T d_{k+1}\right| \geq \theta_{k+1}^p \|g_{k+1}\|^2 \tag{24}$$

From the L.H.S of (24), in view of the fact that inequality $(a^2 + \sigma b^2) \leq (1 + \sigma^2)(a^2 + b^2)$ always holds for any nonnegative numbers a, b and σ and letting $a = |g_{k+1} d_{k+1}|$, $b = \left|\beta_k^p d_k^T g_k\right|$

$$(g_{k+1}^T d_{k+1} + \sigma \beta_k^p d_k^T g_k)^2 \leq (1 + \sigma^2) \left[(g_{k+1}^T d_{k+1})^2 + (\beta_k^p)^2 (d_k^T g_k)^2\right] \tag{25}$$

Squaring the R.H.S of (24), together with (25), we have

$$(1 + \sigma^2) \left[(g_{k+1}^T d_{k+1})^2 + (\beta_k^p)^2 (d_k^T g_k)^2\right] \geq (\theta_{k+1}^p)^2 \|g_{k+1}\|^4$$

$$(g_{k+1}^T d_{k+1})^2 \geq c_k \|g_{k+1}\|^4 - (\beta_k^p)^2 (d_k^T g_k)^2 \tag{26}$$

where

$$c_k = \frac{(\theta^p_{k+1})^2}{1+\sigma^2_{\max}} > 0$$

Considering,

$$\frac{(g^T_{k+1}d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(d^T_k g_k)^2}{\|d_k\|^2} = \frac{1}{\|d_{k+1}\|^2}\left[(g^T_{k+1}d_{k+1})^2 + \frac{\|d_{k+1}\|^2}{\|d_k\|^2}(d^T_k g_k)^2\right] \tag{27}$$

Substitute (26) into (27), we have

$$\frac{(g^T_{k+1}d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(d^T_k g_k)^2}{\|d_k\|^2} \geq \frac{1}{\|d_{k+1}\|^2}\left[c_k\|g_{k+1}\|^4 + \left(\frac{\|d_{k+1}\|^2}{\|d_k\|^2} - (\beta^p_k)^2\right)(d^T_k g_k)^2\right] \tag{28}$$

Substitute (21) into (28), we have

$$\frac{(g^T_{k+1}d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(d^T_k g_k)^2}{\|d_k\|^2} \geq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left[c_k - (\theta^p_{k+1})^2\frac{(d^T_k g_k)^2}{\left\|\|d_k\|^2\right\|\left\|\|g_{k+1}\|^2\right\|}\right]$$

Substituting for c_k

$$\frac{(g^T_{k+1}d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(d^T_k g_k)^2}{\|d_k\|^2} \geq (\theta^p_{k+1})^2\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left[\frac{1}{(1+\sigma^2_{\max})} - \frac{(d^T_k g_k)^2}{\left\|\|d_k\|^2\right\|\left\|\|g_{k+1}\|^2\right\|}\right]$$

From Lemma 1, $\lim\limits_{k\to\infty}\frac{(g^T_k d_k)^2}{\|d_k\|^2} = 0$

In view of $\lim\limits_{k\to\infty}\frac{(g^T_k d_k)^2}{\|d_k\|^2} = 0$, from (29), we have

$$\lim_{k\to\infty}\frac{(g^T_k d_k)^2}{\|d_k\|^2}\frac{1}{\|g_{k+1}\|^2} = 0$$

Thus, for any sufficiently large k, taking into account (29) and $\theta^p_{k+1} > \frac{1}{4}$, we have

$$\frac{(g^T_{k+1}d_{k+1})^2}{\|d_{k+1}\|^2} + \frac{(d^T_k g_k)^2}{\|d_k\|^2} \geq \frac{1}{32(1+\sigma^2_{\max})}\frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$$

This together with Lemma 1 shows that conclusion (29) holds true.■

Theorem 2

Consider the sequence $\{x_k\}$ given by Algorithm DOO. If Assumption 1 holds, f is uniformly convex, then $\lim_{k\to\infty} \inf = 0, i.e$, the algorithm is globally convergent.

Assumption 2

For k large enough the inequalities

$$0 < g_{k+1}^T g_k \leq 2g_{k+1}^T g_{k+1}$$

Proof

Suppose that there exists a positive constant $\epsilon > 0$, such that

$$\|g_k\| \geq \epsilon \tag{30}$$

Square both side of (13)

$$\|d_{k+1}\|^2 = (\theta_{k+1}^p)^2 \|g_{k+1}\|^2 - 2\theta_{k+1}^p g_{k+1}^T d_{k+1} - 2(\theta_{k+1}^p)^2 \|g_{k+1}\|^2 + (\beta_k^p)^2 \|d_k\|^2 \tag{31}$$

From (13), make β_k^p the subject and substitute for β_k^p in the second term on the R.H.S of (31)

$$\|d_{k+1}\|^2 = (\beta_k^p)^2 \|d_k\|^2 - 2\theta_{k+1}^p g_{k+1}^T d_{k+1} - (\theta_{k+1}^p)^2 \|g_{k+1}\|^2$$

Divide both side by $\|g_{k+1}\|^4$, we have

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} = \frac{(\beta_k^p)^2 \|d_k\|^2}{\|g_{k+1}\|^4} - \frac{2\theta_{k+1}^p g_{k+1}^T d_{k+1}}{\|g_{k+1}\|^4} - \frac{(\theta_{k+1}^p)^2 \|g_{k+1}\|^2}{\|g_{k+1}\|^4} \tag{32}$$

$$g_k^T d_k = -\|g_k\|^2 \quad \Rightarrow g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 \tag{33}$$

Substitute (33) into (32), we have

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} = \frac{(\beta_k^p)^2 \|d_k\|^2}{\|g_{k+1}\|^4} - \frac{(\theta_{k+1}^p - 1)^2}{\|g_{k+1}\|^2} + \frac{1}{\|g_{k+1}\|^2} \tag{34}$$

Substitute (7) into (34) (Also recall that $y_k = g_{k+1} - g_k$) :

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} = \frac{\|d_k\|^2}{\|g_{k+1}\|^4} \left[\frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|^2} - \frac{\|y_k\|^2 d_k^T g_{k+1}}{d_k^T (g_{k+1} - g_k)} \right]^2 - \frac{(\theta_{k+1}^p - 1)^2}{\|g_{k+1}\|^2} + \frac{1}{\|g_{k+1}\|^2}$$

From Assumption 2, $\quad 0 < g_k g_{k+1} \leq 2g_{k+1}^T g_{k+1}$

$$\frac{\|d_{k+1}\|^2}{\|g_{k+1}\|^4} \leq \frac{\|d_k\|^2}{\|g_k\|^4} + \frac{1}{\|g_{k+1}\|^2}$$

Let $k = k - 1$, then

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2}$$

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \sum_{i=0}^{k-1} \frac{1}{\|g_i\|^2}$$

From(30), $\|g_k\| \geq \epsilon$
Therefore,

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \frac{k}{\epsilon^2}$$

Taking the inverse of both sides and applying summation

$$\sum_{k\geq 1} \frac{\|g_k\|^4}{\|d_k\|^2} \geq \epsilon^2 \sum_{k\geq 1} \frac{1}{k} = +\infty$$

which contradicts Lemma 2, therefore $\lim_{k\to\infty} \inf \|g_k\| = 0$.
Hence, the proof of the desired result has been completed ■

4. NUMERICAL EXPERIMENT

Here, we intend to check the numerical performance of the proposed Algorithm and compare it with the methods [2, 10, 11], with a large number of numerical experiments on large-scale instances.

4.1 Preparation

The testing problems consist of 96 instances from ten to ten thousand variables. For the sake of analyzing and comparing the numerical results, for all tested methods, the parameters and stop criterion will be uniformly chosen as $\sigma = 0.9$, $\delta = 0.00001$, stop criterion : (i) $\|g_k\| < 10^{-6}$ or (ii) Itr ¿ 2000, where ITR denotes the number of iterations, and the step size is always generated by (17). All codes are written in MatLab 8.5 and run on a Samsung PC with 1.67 GHz, 2 GB RAM memory and Windows 7.0 operating system. When observing the numerical performance of a given optimal method, the number of iterations (ITR), function evaluations (NF), gradient evaluations (NG), Central Processing Unit time (CT) and so on are important factors. Particularly, CPU time is the critical observation point. So, to show the performance difference between the tested methods more clearly, we adopt the performance profiles [6] to summarize the numerical performances of the proposed

Algorithm and compare it with methods [2, 10, 11], with a large number of numerical experiments on large-scale instances. These performance profiles allow us to compare objectively the different methods with respect to robustness and efficiency. We say that a given solver is robust for solving a given optimization problem if it succeeds in finding an optimal solution, and it is efficient if it requires fewer CT, NF, NG, ITR and so on. Efficiency and robustness rates are readable on the left and right vertical axes of the associated performance profiles, respectively.

4.2 Numerical Experiment

Experiment I—numerical performance of our method (N-M) compare with methods [2, 10, 11] with respect to number of iterations.

Here, we test our method (N-M) on 96 middle-large-scale instances using 48 problems from ten to ten thousand variables, and compare it with Loannis and Panagiotis's spectral CGM [11] (LP), Birgin and Martinez's spectral CGM [2] (BM), and Jinbao, Qian, Xianzhen, Youfang and Jianghua's spectral CGM [10] (JQXYJ). All the tests are carried out under ceteris paribus conditions.
In Figure 1, N-M slightly outperforms the LP, BM, and JQXYJ methods respectively, as the former is always on the top curve of the other three methods. The N-M solves 30% of the test problems with the least number of iteration while the LP, the BM and JQXYJ approximately solve about 15%, 27% and 22% of the test problems respectively. Therefore, N-M outperforms the other three method with respect to iterations numbers. Similar result is obtained when we considered the total numbers of function evaluations and the respective gradient evaluations.

Experiment II—numerical performance of our method (N-M) compare with methods [2, 10, 11] with respect to CPU time.

Here, we test our method (N-M) on 96 middle-large-scale instances from ten to ten thousand variables, and compare it with Loannis and Panagiotis's spectral CGM [11] (LP), Birgin and Martinez's spectral CGM [2] (BM), and Jinbao, Qian, Xianzhen, Youfang and Jianghua's spectral CGM [10] (JQXYJ) with respect to CPU time.
In Figure 2, the performance profile was compared with respect to CPU time of N-M with the LP, BM and JQXYJ methods. The

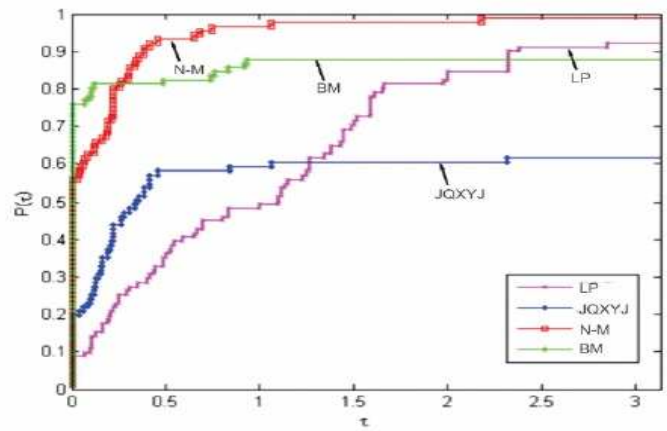

Fig. 1. Performance profile with respect to number of iteration of LP, JQXYJ, N-M and BM

Figure shows that the LP solve 25% of the test problems with the least CPU time, followed by the N-M method, BM method and JQXYJ method respectively. Since all methods are implemented with the same condition of line search, by taking the three factors all in all, we conclude that, N-M to generate the more efficient search direction, on the whole. This is further supported by the fact that run-based profiling can be very sensitive to hardware configuration and the possibility of other internal or external interruptions to the computer during computation

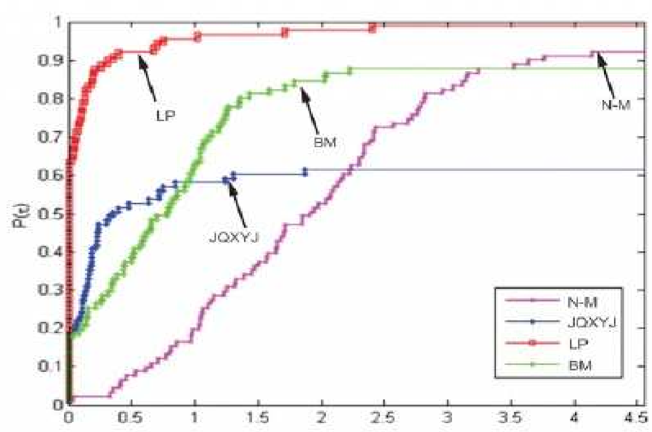

Fig. 2. Performance profile with respect to CPU time of N-M, JQXYJ, LP and BM

5. CONCLUDING REMARKS

In conclusion, a Modified Spectral Conjugate gradient Method was developed for solving unconstrained minimization problem, where the step length is generated by the Strong Wolfe line Search. When the objective function is uniformly convex, the proposed algorithm is proved to be globally convergent. A large number of numerical examples on large-scale instances are reported, and this shows that the proposed method is promising.

ACKNOWLEDGEMENTS

The author would like to thank the anonymous referee whose comments improved the original version of this manuscript.

REFERENCES

[1] N. Andrei, New Accelerated Conjugate Gradient Algorithms as a Mo Modification of Dai-Yuan's Computational scheme for Unconstrained Opimization, J. Comput. Appl. Math., 234, 3397-3410, 2010.

[2] E. G Birgin and J. M. Martinez, A Spectral Conjugate Gradient Method for Unconstrained Optimization, Applied Mathematics and Optimization, 43,117-128, 1999.

[3] J. Barzilai and J. M. Borwien, Two Point Step Gradient Methods, IMA J. Numer. Anal., 8, 141-148, 1988.

[4] Y. H. Dai and C. X. Kou, A Nonlinear Conjugate Gradient Algorithm with an Optimal Property and an improved Wolfe line Search, SIAM J. Optim., 23, 296-320, 2013.

[5] Y. H. Dai and Y. Yuan, A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property, SIAM Journal on Optimization, 10, 177-182, 1999.

[6] E. D. Dolan and J. J. Mor'e, Benchmarking Optimization Software with Performance profiles, Math. Program. 91, 201-213, 2002.

[7] R. Fletcher, Low Storage Methods for Unconstrained Optimization, Lectures in Applied Mathematics, American Mathematical Society, Providence, RI, 26, 165-179, 1990.

[8] R. Fletcher and C. M. Reeves, Functional Minimization by Conjugate Gradients, Comput. J., 7, 149-154, 1964.

[9] M. R. Hestenes and E. Stiefel, Method of Conjugate Gradient for solving Linear System, J. Res. Nat. Bur. Standards, 49, 409-436, 1952.

[10] J. Jinbao, C. Qian, J. Xianzhen, Z. Youfang and Y. Jianghua, A new Spectral Conjugate Gradient Method for Large-scale Unconstrained Optimization, Optimization Methods and Software, DOI:10:1080/10556788.2016.1225213, 2016.

[11] E. L. Loannis and P Panagiotis, A new Spectral Conjugate Gradient Methods based om a Modified Secant Equation for Unconstrained Optimization, Elsevier, 396-405, 2012.

[12] Y. Liu and C. Storey, Efficient Generalized Conjugate Gradient Algorithms, part : Theory, J. Optim. Theory Apppl. 69, 129-137, 1991.

[13] E. Polak and G. Ribie're, Note Sur la convergence de directions conjuge'es, Rev. Fr. Inform. Rech. Oper. 16, 35-43, 1969.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, FEDERAL UNIVERSITY OF KASSHERE, GOMBE STATE, NIGERIA
E-mail address: danhausaa@gmail.com

DEPARTMENT OF MATHEMATICS, MODIBBO ADAMA UNIVERSITY OF TECHNOLOGY, YOLA, NIGERIA
E-mail address: odekunle@mautech.edu.ng

DEPARTMENT OF MATHEMATICAL SCIENCES, REDEEMER'S UNIVERSITY, EDE, OSUN STATE, NIGERIA
E-mail address: onanayea@run.edu.ng

APPENDIX

The test problems used in these analysis are the unconstrained problems in the CUTE test problem library.

1. Extended White and Holst function
2. Extended Rosenbrock Function
3. Extended Freudenstein and Roth Function
4. Extended Beale Function
5. Extended Wood Function
6. Perturbed Quadratic Function
7. Raydan 1 Function
8. Extended Tridiagonal 1 Function
9. Diagonal 4 Function
10. Extended Himmelblau Function
11. Extended Powell Function
12. FLETCHCR Function (CUTE)
13. NONSCOMP Function (CUTE)
14. Extended DENSCHNB Function (CUTE)
15. Extended Quadratic Penalty QP1 Function
16. Extended Penalty Function
17. Hager Function
18. BIGGSB1 Function (CUTE)
19. Extended Maratos Function
20. Six-Hump Camel Function
21. Three-Hump Camel Function
22. Booth Function
23. Trecanni Function
24. Zettl Function
25. Shallow Function
26. Generalized Quartic Function
27. Quadratic QF2 Function
28. Leon Function
29. Generalized Tridiagonal 1 Function
30. Generalized Tridiagonal 2 Function
31. POWER Function (CUTE)
32. Quadratic QF1 Function
33. CUBE Function (CUTE)
34. Extended Quadratic Penalty QP2 Function
35. Extended Quadratic Penalty QP1 Function
36. Quartic Function
37. Matyas Function
38. Colville Function
39. Dixon and Price Function
40. Sphere Function
41. Sum Squares Function
42. Powell Singular Function
43. Extended Wood

44. Test Function
45. Raydan 2 Function
46. Diagonal 3 Function
47. Diagonal 5 Function
48. Quadratic QF2