A New Preconditioned Inexact Line-Search Technique for Unconstrained Optimization

Abbas Y. Al-Bayati
College of Computer Sciences and Mathematics
University of Mosul

Ivan S. Latif
College of Scientific Education
University of Salahaddin

Received on: 18/12/2006 **Accepted on: 16/04/2007**

الملخص

في هذا البحث تم دراسة التقارب الشامل لخوارزمية جديدة من خوارزميات التدرج المترافق المشروطة باستخدام دوال غير مقيدة غير خطية محدبة. الخوارزمية الجديدة تعتمد على إيجاد خط بحث جديد مشابه لخط بحث Armijo التي تستخدم في إيجاد خطوات بحث أكبر ويستخدم نفس الصيغة لإيجاد اتجاه البحث في الخوارزمية الجديدة التي تقلل من كفاءة خوارزمية التدرج المترافق المستخدمة. تم استحداث خوارزمية جديدة للتدرج المترافق المشروط باستخدام خوارزمية لأشباه نيوتن. النتائج العملية لـ (25) دالة وبأبعاد مختلفة توضح بأن خط البحث الجديد مع الاتجاه الجديد للخوارزمية المقترحة أكثر كفاءة في إيجاد حلول الدوال اللاخطية وغير المقيدة مقارنة بالخوارزميات المماثلة في مجالات عدة.

ABSTRACT

In this paper, we study the global convergence properties of the new class of preconditioned conjugate gradient descent algorithm, when applied to convex objective non-linear unconstrained optimization functions.

We assume that a new inexact line search rule which is similar to the Armijo line-search rule is used. It's an estimation formula to choose a large step-size at each iteration and use the same formula to find the direction search. A new preconditioned conjugate gradient direction search is used to replace the conjugate gradient descent direction of ZIR-algorithm. Numerical results on twenty five well-know test functions with various dimensions show that the new inexact line-search and the new preconditioned conjugate gradient search directions are efficient for solving unconstrained nonlinear optimization problem in many situations.

Keywords: Preconditioned CG, Unconstrained Optimization, Self-Scaling VM-update, inexact Line-Search.

1. Introduction

Some important global convergence result for various methods using line-search procedures have been given [1], [4] the above mentioned line search methods are monotone descent for unconstrained optimization [10], [11]. Non monotone line-searches have been investigated also by many authors see [6], [9]. The Barzilai-Borwein method [2], [8] is a non monotone descent method which is an efficient algorithm for solving some special problem, Zirilli [12] extend the Armijo line search rule ant analyze the global convergence of the corresponding method.

In this paper, we extend the Armijo line-search rule so that we can design a new inexact line search technique and we choose the search directions of AL-Bayati Self-Scaling [3] variable metric update which based on two parameter family of rank-two updating formulae. Numerical results show that the new algorithm which enables us to choose large step-size at each iteration and reduce the number of functions. The new algorithm is efficient for solving unconstrained optimization problems.

We consider the following unconstrained optimization problem of n variables,

$$Min\, f(x), \qquad x \in R^n, \qquad \ldots(1)$$

where $f(x)$ is twice continuously differentiable and its gradient g is exist available. We consider iterations of the form

$$x_{k+1} = x_k + \alpha_k d_k \qquad \ldots(2)$$

where d_k is a search direction and α_k is the step-length obtained by means of one-dimensional search. In conjugate gradient method when the function is quadratic and the line search is exact, another broad class of methods may be defined by the following search direction:

$$d_k = -H_k^{-1} g_k \qquad \ldots(3)$$

where H_k is a non singular symmetric matrix. Important special cases are given by

$H_k = I$ (Steepest descent direction)

$H_k = \nabla^2 f(x_k)$ (Newton's direction)

Variable Metric (VM) methods are also of the form (3) and in this case H_k is not only a function of x_k, but depends also on H_{k-1} and x_{k-1}.

All these methods are implemented so that d_k, is a descent direction, i.e.

$$d_k^T g_k < 0 \qquad \ldots(4)$$

which guarantees that the function can be decreased by taking a small step along d_k for the Newton type method (3). We can ensure that d_k is a descent direction by defining H_k to be positive definite.

For conjugate gradient method, obtaining descent direction is not easy and requires a careful choice properties of line search methods and it can be studied by measuring the goodness of the search direction and by considering the length of the step. The quality of the angle between the steepest descent direction $-g_k$ and the search direction. We can define:

$$\cos\langle -g_k, d_k\rangle = -g_k^T d_k \big/ (\|g_k\|.\|d_k\|) \geq \eta_0 \qquad \ldots(5)$$

The length of the step is determined by the line search iteration. A strategy that will play a central role in this paper is to set scalars s_k, β, L, $\sigma > 0$ with:

$$s_k = -g_k^T d_k \Big/ \left(L\|d_k\|^2\right), \quad \beta \in (0,\, 1); \quad \sigma \in (0, 1/2).$$

Let α_k be the larges α in $\{s_k,\ \beta s_k,\ \beta^2 s_k, \ldots\}$ such that

$$f_k - f(x_k + \alpha d_k) \geq -\sigma \alpha g_k^{\mathrm{T}} d_k \qquad \ldots(6)$$

The inequality ensures that the function is reduced sufficiently, we will call these relations as Armijo condition.

2. Zirlli Inexact Line-Search Algorithm (Zir):

Inexact line-search rule was implemented the following assumptions [7], [11].

(H1) The function $f(x)$ has a lower bound on the level set
$L(x_0) = \{x \in R^n \mid f(x) - f(x_0)\}$ where x_0 is given

(H2) The gradient $g(x)$ of $f(x)$ is Lipschitz continuous in an open convex set B that contains L_0 the; i.e., there exists L such that

$$\|g(x) - g(y)\| \le L\|x - y\|, \quad \forall\, x, y \in B \qquad \ldots(7)$$

The modified Armijo line search rule as [1]:
Set scalars S_k, β, L_k, μ and σ with $s_k = -g_k^T d_k \big/ \left(L\|d_k\|^2\right)$, $\beta \in (0,\ 1), L_k > 0, \mu \in [0, 2)$ and $\sigma \in (0,\ 1/2)$.

Let α_k be the larges α in $\{s_k,\ \beta s_k,\ \beta^2 s_k, \ldots\}$ such that

$$f(x_k + \alpha d_k) - f_k \le \sigma\alpha\left[g_k^T d_k + \left(\frac{1}{2}\right)\alpha\mu\|d_k\|^2\right] \qquad \ldots(8)$$

2.1. Outlines of the Zir Algorithm:

The implementable inexact line search algorithm is stated as follows [12]:

Stepl: Given some parameters, $\sigma \in (0,\ 1/2)$, $x_0 \in R^n$, $\beta \in (0,\ 1)$, $\mu \in (0, 2)$, $L_0 = 1$ let and set $K = 0$, ε is a small parameter.

Step2: If $\|g_k\| \le \varepsilon$ then stop. Else go to step3.

Step3: Choose d_k, to satisfy the angle property (5) and set $d_k = -g_k$.

Step4: Set $x_{k+1} = x_k + \alpha_k d_k$, where α_k is defined by the modified Armijo line search rule (8).

Step5: Set $V_k = x_{k+1} - x_k$; $Y_k = g_{k+1} - g_k$ and L_{k+1} is determined by

$$L_{k+1} = \frac{\|y_k\|}{\|V_k\|} \qquad \ldots(9)$$

Step6: Set $k = k + 1$ and go to step 2.

2.2. Some Properties of the Zir Algorithm:

Theorem 2.2.1: Assume that (H1) and (H2) hold, the search direction d_k satisfies (4) and α_k is determined by the modified Armijo line-search rule. Zir Algorithm generates an infinite sequence $\{x_n\}$ with

$$0 < L_k < m_k L \qquad \ldots(10)$$

where m_k is appositive integer and $m_k \le M_0 \le \infty$ with M_0 being large positive constant then

$$\sum_{k=1}^{\infty}\left(\frac{g_k^T d_k}{\|d_k\|}\right)^2 < +\infty \qquad \ldots(11)$$

for the details of the proof see [12].

Corollary 2.2.1: If the condition in theorem 2.2.1 hold then

$$\lim_{k\to\infty}\left(\frac{g_k^T d_k}{\|d_k\|}\right)=0 \qquad \ldots(12)$$

In fact, Assumption (H2) can be replaced by the following weaker assumption. (H2') the gradient $g(x)$ of $f(x)$ is uniformly continuous on an open convex set B that contains L_0 see [9].

3. A New Proposed Preconditioned Inexact Line-Search Algorithm (New):

In this section we propose a new algorithm which implements the step-size α_k with inexact line search rule. This formula is implemented with AL-Bayati self-scaling [3] variable metric update.

3.1. Outlines of the New Algorithm:

The outlines of the new proposed Algorithm are stated as follows:

Step1: Given some parameters $\sigma \in (0, \frac{1}{2})$, $x_0 \in R^n$, $\beta \in (0,1)$, $M = 10^8$, H_0 is identity positive definite matrix and $L_0^* = 0.1$. Let, ε is a small parameter and set $K = 0$.

Step2: If $\|g_k\| \le \varepsilon$ then stop. Else go to step3.

Step3: Choose d_k to satisfy the angle property (5) and satisfy the new search direction.

$$d_k = \begin{cases} -H_k g_k, & if \;\; k = 1, \\ -H_k g_k + L_k^* d_k, & if \;\; k \ge 1, \end{cases} \qquad \ldots(13)$$

Step4: Set $x_{k+1} = x_k + \alpha_k d_k$ where α_k is defined later by a new modified line search rule (19), (20).

Step5: Set $V_k = x_{k+1} - x_k$, $Y_k = g_{k+1} - g_k$ and L_{k+1} is determined by

$$L_{k+1}^* = \min\left\{ L_k^*, \frac{g_k^T H_k Y_k}{\|V_k\|^2}, \frac{Y_k^T H_k Y_k}{\|V_k\|} \right\}, \qquad \ldots(14)$$

Step6: Update H_k by H_{k+1}, see [3]

$$H_{k+1} = \left(H_k - \frac{H_k Y_k Y_k^T H_k}{Y_k H_k^T Y_k} + W_k W_k^T \right)\mu_k + \frac{V_k V_k^T}{V_k^T Y_k} \qquad \ldots(15)$$

$$W_k = \left(Y_k^T H_k Y_k\right)^{1/2}\left[\frac{V_k}{V_k^T Y_k} - \frac{H_k Y_k}{Y_k^T H_k Y_k} \right] \qquad \ldots(16)$$

$$\mu_k = \frac{Y_k^T H_k Y_k}{V_k Y_k} \qquad \ldots(17)$$

Step7: If available storage is exceeded then employ a restart option either with $k = n$ or $g_{k+1}^T g_{k+1} > g_{k+1}^T g_k$ i.e. orthogonality condition is not satisfy see [7].

Steps: Set $k = k + 1$ and go to step2.

3.2. Some Theoretical Properties of the New Algorithm:

We analyze the global convergence of the proposed new inexact line-search algorithm. For the proof of convergence we adopt the assumptions (H1), (H2') on the function *f* which is commonly used and we suppose that $\{H_k\}$ is a sequence of positive definite matrices. Assume also that there exist parameters $v_{\min} > 0$ and $v_{\max} > 0$ such that $\forall\, d \in R^n$

$$v_{\min} d^T d \le d^T H_k d \le v_{\max} d^T d \quad ...(18)$$

this condition would be satisfied for instance, if $H_k \equiv H$ and H is positive definite as in Al-Bayati VM-update [3]. We analyze the conjugate gradient algorithm that use the following modified line-search formula: Set scalars $S_k,\ \beta,\ L_k,\ \mu$ and σ with

$$s_k = -g_k^T d_k \Big/ \left(L \|d_k\|_{H_k}^2\right) \qquad \ldots(19)$$

where, $\beta \in (0, 1)$, $L_k^* > 0$ is a new parameter, $\mu \in [0, 2)$ and $\sigma \in (0, v_{\min}/\mu)$. Note that the specification of σ ensures $\dfrac{\rho\mu}{v_{\min}} < 1$.

Let α_k be the larges α in $\{s_k,\ \beta s_k,\ \beta^2 s_k, ...\}$ such that

$$f(x_k + \alpha d_k) - f_k \le \sigma\alpha\left[g_k^T d_k + \left(\frac{1}{2}\right)\alpha\mu L_k \|d_k\|^2 \right] \qquad \ldots(20)$$

where $\|d_k\|_{H_k} = \sqrt{d_k^T H_k d_k}$

Lemma 3.2.1: Suppose that x_k is given by the new proposed algorithm defined by {(2), (13), (14) and (19)} then

$$g_{k+1}^T d_k = \rho_k g_k^T d_k \qquad \ldots(21)$$

holds for all k ,where

$$\rho_k = 1 - \frac{\phi_k g_k^T d_k}{L_k^* \|d_k\|_{H_k}^2} \qquad \ldots(22)$$

and L_k^* is known as a new scalar defined in (14). Let

$$\phi_k = \begin{cases} 0 & for\ \alpha_k = 0 \\ \dfrac{y_k^T V_k}{\|V_k\|^2} & for\ \alpha_k \ne 0 \end{cases} \qquad \ldots(23)$$

Proof:

The case of $\alpha_k = 0$ implies that $\rho_k = 1$ and $g_{k+1} = g_k$ hence (21) is valid, we now prove for the case of $\alpha_k \ne 0$ from (2) and new modified inexact line search α_k we have

$$g_{k+1}^T d_k = g_k^T d_k + \left(g_{k+1} - g_k\right)^T d_k$$

$$= g_k^T d_k + \alpha_k^{-1}\left(g_{k+1} - g_k\right)^T \left(x_{k+1} - x_k\right) \text{ from (2) we have } d_k = \alpha_k^{-1}\left(x_{k+1} - x_k\right)$$

$$= g_k^T d_k + \alpha_k^{-1}\phi_k \|x_{k+1} - x_k\|^2 \text{ from (23)}$$

$$= g_k^T d_k + \alpha_k^{-1}\phi_k \|d_k\|^2$$

$$= g_k^T d_k - \left(\frac{g_k^T d_k}{L_k \|d_k\|_{H_k}^2} \right) \phi_k \|d_k\|^2$$

$$= \left(1 - \frac{1}{L_k} \phi_k \frac{\|d_k\|^2}{\|d_k\|_{H_k}^2} \right) g_k^T d_k \text{ from (23)}$$

$$= \rho_k g_k^T d_k$$

The proof is complete. #

Theorem 3.2.1: If (H1) and (H2') hold, then the new algorithm generates an infinite number of sequences $\{\alpha_k\}$ and satisfy

$$0 < L_k^* < mL < M \qquad \ldots(24)$$

where *m* is a positive integer, *M* is a large positive constant then

$$\lim_{k \to \infty} \left(\frac{-g_k^T d_k}{\|d_k\|} \right) = 0$$

Proof:

Let $K_1 = \{k \mid \alpha_k = s_k\},\ K_2 = \{k \mid \alpha_k < s_k\}$

Case (l):

If $k \in K_1$ then

$$f(x_k + \alpha d_k) - f_k \le \sigma\alpha \left[g_k^T d_k + \left(\frac{1}{2}\right) \alpha_k L_k^* \|d_k\|^2 \right],\ L^* \text{ is a new parameter defined by (41)}$$

$$= -\sigma\left[g_k^T d_k \big/ L_k^* \|d_k\|^2 \right] \left[g_k^T d_k - \left(\frac{1}{2}\right) \mu g_k^T d_k \right]$$

$$= -\left[\sigma \left(1 - \left(\frac{1}{2}\right)\mu \right) \Big/ L_k^* \right] \left(g_k^T d_k \right)^2 \Big/ \|d_k\|^2$$

Thus

$$f(x_k + \alpha d_k) - f_k \le -\left[\sigma \left(1 - \left(\frac{1}{2}\right)\mu \right) \Big/ L_k^* \right] \left(g_k^T d_k \right)^2 \Big/ \|d_k\|^2, \quad k \in K_1 \qquad \ldots(27)$$

Let

$$\eta_k = -\sigma \left(1 - \left(\frac{1}{2}\right)\mu \right) \Big/ L_k^*, \quad k \in K_1$$

By (24) we have

$$\eta_k = -\sigma \left(1 - \left(\frac{1}{2}\right)\mu \right) \Big/ L_k^*$$

$$\leq -\sigma\left(1-\left(\frac{1}{2}\right)\mu\right)\Big/ mL$$

$$\leq -\sigma\left(1-\left(\frac{1}{2}\right)\mu\right)\Big/ ML$$

$$< 0$$

Let

$$\eta' \leq -\sigma\left(1-\left(\frac{1}{2}\right)\mu\right)\Big/ ML$$

This and (27) imply that $\eta_k \leq \eta'$ and

$$f_{k+1} - f_k \leq \eta'\left(g_k^T d_k \big/ \|d_k\|^2\right)^2, \quad k \in K_1 \qquad \ldots(28)$$

Thus if $k \in K_1$, from (28) we can prove that

$$\lim_{k\in K_1, k\to\infty}\left(\frac{-g_k^T d_k}{\|d_k\|}\right) = 0. \quad \#$$

Case (2):

If $k \in K_2$ then $\alpha_k < s_k$ this show that s_k, can not satisfy the new suggested line search and thus $\alpha_k < \beta s_k$ we show that $\alpha = \alpha_k/\beta$ were α_k be the larges α in $\{s_k,\ \beta s_k,\ \beta^2 s_k, \ldots\}$ can not satisfy (14) and thus

$$f(x_k + \frac{\alpha_k d_k}{\beta}) - f_k \leq \sigma\alpha_k\beta\left[g_k^T d_k + \frac{\left(\frac{1}{2}\right)\alpha_k \mu L_k^* \|d_k\|^2}{\beta}\right]$$

using the mean-value theorem on the left hand side of the above inequality, we see that there exists $\theta_k \in [0,1]$ such that

$$g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right) > \sigma\alpha_k/\beta\left[g_k^T d_k + \frac{\left(\frac{1}{2}\right)\alpha_k \mu L_k^* \|d_k\|^2}{\beta}\right]$$

Therefore

$$g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right)^T d_k > \rho\left[g_k^T d_k + \frac{\left(\frac{1}{2}\right)\alpha_k \mu L_k^* \|d_k\|^2}{\beta}\right] \qquad \ldots(30)$$

in this case of $k \in K_2$, by (19) and (20) we have

$$f(x_k + \alpha d_k) - f_k \leq \sigma\alpha\left[g_k^T d_k + \left(\frac{1}{2}\right)\alpha_k L_k^* \|d_k\|^2\right]$$

$$\le \sigma\alpha\left[g_k^T d_k + \left(\frac{1}{2}\right)s_k L_k^* \|d_k\|^2\right]$$

$$\le \sigma\alpha_k\left[1+\left(\frac{1}{2}\right)\mu\right]g_k^T d_k$$

By (H1) we have

$$\lim_{k\in K_2, k\to\infty}\left(\frac{-g_k^T d_k}{\|d_k\|}\right)=0 \qquad \text{...(31)}$$

If there exist $\varepsilon > 0$ and an infinite subset $K_3 \subseteq K_2$ such that

$$\frac{-g_k^T d_k}{\|d_k\|} \ge \varepsilon, \quad \forall k \in K_3 \qquad \text{...(32)}$$

then by (31), (32) we have

$$\lim_{k\in K_3, k\to\infty}\alpha_k\|d_k\|=0 \qquad \text{...(33)}$$

by (30) we have

$$g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right)^T d_k \ge \rho g_k^T d_k, \quad k \in K_3 \qquad \text{...(34)}$$

where $\theta_k \in [0,1]$ is defined in the proof. By the Cauchy Schwarz inequality and (34) we have

$$\left\|g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right) - g_k\right\| = \left\|g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right) - g_k\right\|\frac{\|d_k\|^2}{\|d_k\|^2}$$

$$\ge \frac{\left[g\left(x_k + \frac{\theta_k\alpha_k d_k}{\beta}\right) - g_k\right]^T d_k}{\|d_k\|^2}$$

$$\ge \frac{-(1-\rho)g_k^T d_k}{\|d_k\|^2}, \qquad k \in K_3$$

by (H2') and (31) we obtain

$$\lim_{k\in K_3, k\to\infty}\left(\frac{-g_k^T d_k}{\|d_k\|}\right)=0$$

which contradicts (32) this show that

$$\lim_{k\in K_2, k\to\infty}\left(\frac{-g_k^T d_k}{\|d_k\|}\right)=0 \qquad \text{...(35)}$$

by (29), (35) and noting that $K_1 \cup K_2 = \{1,2,...\}$ we show that (25) holds. #

Lemma 3.2.2.: suppose that (H1), (H2') holds and x_k is given by the new proposed algorithm defined by {(2), (13), (14) and (19)} then

$$\sum_{d_k \neq 0} \frac{-g_k^T d_k}{\|d_k\|} < \infty \qquad \ldots(36)$$

Proof:

By the mean value theorem we have

$$f(x_{k+1}) - f(x_k) = \bar{g}^T (x_{k+1} - x_k)$$

from (19) we have

$$f(x_{k+1}) - f(x_k) \leq -\sigma \left[1 - \left(\frac{1}{2}\right)\mu L_k\right] \frac{(g_k^T d_k)^2}{\|d_k\|^2} \qquad \ldots(37)$$

which implies that $f(x_{k+1}) \leq f(x_k)$. It follows by assumption (Hl), (H2') that $\lim_{k\to\infty} f(x_k)$ exists thus from (18) and (37) we have

$$\frac{(g_k^T d_k)^2}{\|d_k\|^2} \leq v_{\max} \frac{(g_k^T d_k)^2}{\|d_k\|_{H_k}^2} \leq \frac{v_{\max}}{\sigma\left[1 - \left(\frac{1}{2}\right)\mu L_k\right]} [f(x_{k+1}) - f(x_k)]$$

this finishes our proof. #

4. Numerical results:

In this section, we compare the numerical behavior of the new algorithm with the Zir algorithm for different dimensions of test functions. Comparative test were performed with (25) (specified in the Appendices 1 and 2) well-Known test function see [5]. All the results are obtained with newly-programmed FORTRAN routines which employ double precautions. We solve each of these test function by the:

1- Zirlli algorithm (Zir).

2- The new algorithm (New).

and for each algorithm we used the following stopping criterion $\|g_{k+1}\| < 1 \times 10^{-5}$.

All the numerical results are summarized in Table (l), Table (2) and Table (3). They present the numbers of iterations (NOI) versus the numbers of function evaluations (NOF) that are need to obtain the condition $\|g_{k+1}\| < 1 \times 10^{-5}$ while Table (3) gives the percentage performance of the new algorithm based on both NOI and NOF against the original Zit algorithm.

The important thing is that the new algorithm solves each particular problem measured by NOI and NOF respectively, while the other algorithm may fail in some cases. Moreover, the new proposed algorithm always performs more stably and efficiently.

Namely there are about (50-52)% on NOI for all dimensions also there are (63-78)% improvements on NOF for all test functions.

Table (1). Comparison between the New and Zri algorithms using different values of 12<N<5000 for 1st test functions

N. OF Test	TEST FUNCTION	Zir NOF(NOI)						New NOF(NOI)					
		N= 12	N= 36	N= 360	N= 1080	N= 4320	N= 5000	N= 12	N= 36	N= 360	N= 1080	N= 4320	N= 5000
1	EX-beal	804 644	855 684	956 764	1004 802	1073 855	1086 872	137 115	142 128	153 141	158 146	165 150	168 152
2	GEN-edger	53 22	55 24	59 26	61 27	63 28	65 29	24 21	25 23	27 25	28 26	30 28	31 29
3	Full Hession	85 19	116 22	183 25	265 32	154 16	154 16	25 21	30 27	38 35	42 39	47 44	48 45
4	GEN-Q2	164 162	164 162	160 157	159 156	159 156	159 156	160 137	160 137	160 137	160 137	160 137	160 137
5	Digonal4	91 17	97 18	103 19	109 20	115 21	118 22	22 15	22 15	23 16	23 16	23 16	23 16
6	GEN-quadratic	243 169	241 235	241 235	241 235	241 235	241 235	104 94	107 103	118 116	124 122	126 123	128 124
7	Digonal6	20 17	21 19	24 22	25 23	26 24	26 24	15 12	16 14	18 16	19 17	20 18	20 18
8	GEN-Wolf	207 166	277 254	301 289	364 324	382 328	394 332	274 250	296 286	325 323	360 332	380 361	389 364
9	GEN-Shallow	763 422	817 449	931 506	985 533	992 554	998 556	155 153	160 158	171 169	177 175	177 175	177 175
10	Quadratic	106 30	531 94	3043 373	3140 433	3280 482	3328 506	19 15	35 32	238 204	269 224	291 243	298 247
General TOTAL of 7 functions		2536 1668	3174 1961	6001 2416	6353 2585	6485 2699	6569 2748	935 833	993 923	1271 1185	1360 1234	1419 1295	1442 1307

Table (2). Comparison between the New and Zir algorithms using different values of 12< N<5000 for 2nd test functions

N. OF Test	TEST FUNCTION	Zir NOF(NOI)						New NOf(NOI)					
		12	36	360	1080	4320	5000	12	36	360	1080	4320	5000
1	GEN-Helical	F	F	F	F	F	F	70 53	71 57	73 59	74 60	76 62	76 62
2	Fred	F	F	F	F	F	F	139 125	142 136	147 144	149 146	153 150	153 150
3	liarwhid	F	F	F	F	F	F	107 95	694 424	240 227	214 200	214 200	214 200
4	starcase	F	F	F	F	F	F	22 17	52 44	446 426	1310 1279	1360 1292	1392 1301
5	TDP	F	F	F	F	F	F	144 130	211 183	1179 1087	2467 2163	2616 2282	2616 2282
6	Biggsb	F	F	F	F	F	F	16 10	32 23	232 214	668 214	684 656	702 674
7	Miele	F	F	F	F	F	F	184 146	194 160	269 187	268 195	268 195	268 195
8	GEN-Powell	F	F	F	F	F	F	203 131	209 141	211 145	216 145	217 246	218 247
9	EX-Fredent& Roth	F	F	F	F	F	F	196 177	199 191	204 201	206 203	208 204	208 204
10	TR1	F	F	F	F	F	F	74 61	260 230	1311 1001	1770 1196	1820 1204	1846 1242
11	Almost Peturbed quadratic	F	F	F	F	F	F	20 16	38 32	247 201	470 354	482 402	488 404
12	QDP	F	F	F	F	F	F	23 13	43 25	112 57	171 88	192 91	198 93
13	Gen-Centar	F	F	F	F	F	F	91 69	92 75	95 78	96 79	96 79	96 79
14	sinquadratic	F	F	F	F	F	F	146 75	224 124	385 206	190 161	190 161	190 161
15	OSP	F	F	F	F	F	F	833 483	2674 1918	2720 1964	2842 1981	2989 2018	2996 2068

Table (3). Percentage performance of the New algorithm against Zri algorithm for 100% in both NOI and NOF

N	Costs	NEW
12	NOF NOI	63.13 50.06
36	NOF NOI	68.72 52.93
360	NOF NOI	78.82 50.95
1080	NOF NOI	78.59 52.26
4320	NOF NOI	78.12 52.02
5000	NOF NOI	78.05 52.44

5. Conclusions:

In this paper, a new PCG-algorithm with a self-scaling VM-update and a new search direction formula is proposed. A modified formula of an inexact line search is implemented to solve a large-scale unconstrained optimization test functions. Our numerical results supports our claim and also indicate that the new algorithm sufficiently decrease the function values and iterations and it needs an extra line search conditions satisfied near the stationary point of the proposed line search procedure.

Appendix 1:

All the test functions used in Table (l) for this paper arc from general literature. See [5]:

1. Generalized Beale Function:

$$f(x) = \sum_{i=1}^{n/2} \left[1.5 - x_{2i} + (1 - x_{2i})\right]^2 + \left[2.25 - x_{2i-1}(1 - x_{2i}^2)\right]^2 + \left[2.625 - x_{2i-1}(1 - x_{2i}^2\right]^2,$$

$$x_0 = [-1.,-1.,\ldots,-1.,-1.].$$

2. Generalized Edger Function:

$$f(x) = \sum_{i=1}^{n/2} (x_{2i-1} - 2)^4 + (x_{2i-1} - 2)^2 x_{2i}^2 + (x_{2i} + 1)^2,$$

$$x_0 = [1.,0.,\ldots,1.,0.].$$

3. Full Hessian Function:

$$f(x) = \left(\sum_{i=1}^{n} x_i\right)^2 + \sum_{i=1}^{n} (x_i \exp(x_i) - 2x_i - x_i^2),$$

$$x_0 = [1.,1.,\ldots,1.].$$

4. Generalized quadratic Function GQ2:

$$f(x)=(x_1^2-1)^2+\sum_{i=2}^{n}(x_i^2-x_{i-1}-2)^2,$$
$$x_0=[1.,1.,\ldots,1.].$$

5. Diagonal 4 Function:

$$f(x)=\sum_{i=1}^{n/2}\tfrac{1}{2}\left(x_{2i-1}^2+cx_{2i}^2\right),$$
$$x_0=[1,1,\ldots,1]\ ,\ c=100.$$

6. Generalized quadratic Function GQ1

$$f(x)=\sum_{i=1}^{n-1}x_i^2+(x_{i+1}+x_i^2)^2\ ,$$
$$x_0=[1.,1.,\ldots,1.].$$

7. Diagonal 6 Function:

$$f(x)=\sum_{i=1}^{n}(\exp(x_i)-(1+x_i)),$$
$$x_0=[1.,1.,\ldots,1.,1.].$$

8. Generalized Wolfe Function:

$$f(x)=(-x_1(3-x_1/2)+2x_2-1)^2+\sum_{i=1}^{n-1}\left(x_{i-1}-x_i(3-x_i/2+2x_{i+1}-1)\right)^2+\left(x_{n-1}-x_n(3-x_n/2)-1\right)^2,$$
$$x_0=[-1.,\ldots,-1.].$$

9. Generalized Shallow function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2-x_{2i})^2+(1-x_{2i-1})^2\ ,$$
$$x_0=[-2.,-2.,\ldots,-2.,-2.].$$

10. Quadratic Function QF2:

$$f(x)=\tfrac{1}{2}\sum_{i=1}^{n}i(x_i^2-1)^2-x_n\ ,$$
$$x_0=[0.5,0.5,\ldots,0.5].$$

Appendix 2:

All the test Functions used in Table (2) far this paper are from general literature .See [5]:

1. General Helical Function:

$$f(x)=\sum_{i=1}^{n/3}\left(100x_{3i}-10*H_i\right)^2+100(R_i-1)^2+x_{3i}^2,$$

$$\text{where } R_i = sqrt(x_{3i-2}^2 + x_{3i-1}^2), H_i = \begin{cases} (2\pi)^{-1}\tan^{-1}\dfrac{x_{3i-1}}{x_{3i-2}} & if\ x_{3i-2} > 0 \\ 0.5 + (2\pi)^{-1}\tan^{-1}\dfrac{x_{3i-1}}{x_{3i-2}} & if\ x_{3i-2} < 0 \end{cases}$$

$x_0 = [-1.,0.,0....,-1.,0.],0..$

2. Extended Fred Function:

$$f(x) = \sum_{i=1}^{n/2}(-13 + x_{2i-1} + (5 - x_{2i}) + (x_{2i} - 2)(x_{2i}))^2 + \sum_{j=1}^{n/2}(-29 + x_{2i-1} + (1 - x_{2i}) + (x_{2i} - 14)(x_{2i}))^2,$$

$x_0 = [1.,2.,...,n]$

3. Liarwhd Function (cut):

$$f(x) = \sum_{i=1}^{n} 4(-x_1 + x_i^2)^2 + \sum_{i=1}^{n}(x_i - 1)^2,$$

$x_0 = [4.,4.,...,4.]$.

4. Staircase2 Function:

$$f(x) = \sum_{i=1}^{n}\left[\left(\sum_{j=1}^{i} x_j\right) - i\right]^2,$$

$x_0 = [0.,0.,..., 0.]$.

5. Tridiagonal Perturbed Quadratic Function:

$$f(x) = x_i^2 + \sum_{i=2}^{n-1} ix_i^2 + (x_{i-1} + x_i + x_{i+1})^2,$$

$x_0 = [0.5,0.5.,...,0.5]$.

6. Biggsbl Function (CUTE):

$$f(x) = (x_i - 1)^2 + \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2 + (1 - x_n)^2,$$

$x_0 = [1.,1.,...,1.]$.

7. Mill and Cornwell function:

$$f(x) = \sum_{i=1}^{n/4}[\exp(x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2(x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4],$$

$x_0 = [1.,2.,2.,2.,..,1.,2.,2.,2.]$.

8. Generalized Powell function:

$$f(x) = \sum_{i=1}^{n/3}\{3 - [\tfrac{1}{1+(x_i - x_{2i})^2}] - \sin(\tfrac{\pi x_{2i}x_{3i}}{2}) - \exp[-(\tfrac{x_i + x_{3i}}{x_{2i}} - 2)^2]\},$$

$x_0 = [0.,1.,2.,...,0.,1.,2.]$.

9. Extended Freudenstein & Roth Function:

$$f(x)=\sum_{i=1}^{n/2}\left(-13+x_{2i-1}+((5-x_{2i})x_{2i}-2)x_{2i}\right)^2+\left(-29+x_{2i-1}+((x_{2i}+1)x_{2i}-14)x_{2i}\right)^2 ,$$
$$x_0=[0.5.,-2.,0.5,-2.,...,0.5,-2.].$$

10.Extended Tridigonal-1 Function:

$$f(X)=\sum_{i=1}^{n/2}(x_{2i-1}+x_{2i}-3)^2+(x_{2i-1}-x_{2i}+1)^4 ,$$
$$x_0=[2,2,...,2].$$

11. Almost Perturbed Quadratic Function:

$$f(x)=\sum_{i=1}^{n}ix_i^2+\tfrac{1}{100}(x_1+x_n)^2 ,$$
$$x_0=[0.5,0.5,...,0.5].$$

12. Quadratic Diagonal Perturbed Function:

$$f(x)=\left(\sum_{i=1}^{n}x_i\right)^2+\sum_{i=1}^{n}\tfrac{i}{100}x_i^2 ,$$
$$x_0=[0.5,0.5,...,0.5].$$

13. Generalized Cant real Function:

$$f(x)=\sum_{i=1}^{n/4}\left[(\exp(x_{4i}-3)-x_{4i-2})^4+100(x_{4i-2}-x_{4i-1})^6+(\arctan(x_{4i-1}-x_{4i}))^4+x_{4i-3}\right],$$
$$x_0=[1.,2.,2.,2.,...,1.,2.,2.,2.].$$

14. Sinquad Function (CUTE):

$$f(x)=(x_i-1)^4+\sum_{i=1}^{n/2}\left(\sin(x_i-x_n)-x_1^2+x_i^2\right)^2+(x_n^2-x_1^2)^2 ,$$
$$x_0=[0.1,0.1,...,0.1].$$

15. Generalized OSP (Oren and Spedicato) Function:

$$f(x)=\left[\sum_{i=1}^{n}ix_i^2\right]^2,$$
$$x_0=[1.,...,1.].$$

REFERENCES

[1] L. Armijo, "Minimization of function having Lipschitz continuous first partial derivatives", pacific. J. Math., Vol.(16), pp. 1-13, (1966).

[2] Barzilai and J. M. Borwein, "Tow-point size gradient methods", IMA J. Numer. Anal., Vol.(8), PP. 141-148, (1988).

[3] A.Y. Al-Bayati, "A new family of self-scaling VM-algorithms", Journal of Education and Science, Mosul University, Mosul, Iraq (1991).

[4] A.I. Cohen, "Stepsize analysis for descent Methods", J. Optim. theo. & Appli., Vol. (33), PP. 187-205, (1981).

[5] J. Nocedal, "Unconstrained Optimization Test Function Research Institute for Informatics", Center for advanced Modeling and Optimization, Bucharestl, Romania. April 28, (2005).

[6] J. Nocedal, "Theory of Algorithms for Unconstrained Optimization", Acta. Numer., Vol. (l), pp. l99-242, (1992).

[7] J. Nocedal and J. S. Wright, "Numerical Optimization", Spring-Verlag, New York, (1999).

[8] M. Raydan, "On the Barazilai Borwein Gradient Choice of Steplength for the Gradient Method", IAM J. Numer. Anal., Vol.(13), pp.321-326, (1993).

[9] W.Y. Sun; J.Y. Han and J. Sun, "Global Convergence of Nonmonotone Descent Methods for Unconstrained Optimization Problems", J. Copmut. &c Appli. Math., Vol.(14), pp. 89-98, (2002).

[10] M.N. Vrahatis; G.S. Androulakis, and G.E. Manoussakis, "A New Unconstrained Optimization Method for Imprecise Function and Gradient Values", J Math. Anal tic Appl., Vol.(197), PP. 586-607, (1996).

[11] Y. Yuan, "Numerical Methods for Nonlinear Programming", Shanghai Scientific Technical Publishers, (1993).

[12] F. Zirilli, "New Inexact Line Search Method for Unconstrained Optimization", J. Optrm. Theo. & Appl., Vol. (127), PP. 425-446, (2005).