Australian Journal of Basic and Applied Sciences, 4(9): 4223-4239, 2010
ISSN 1991-8178

New Hybrid Globally Convergent CG-Algorithms for Nonlinear Unconstrained Optimization

[1]Abbas Y. AL-Bayati and [2]Barah M. AL-Baro

[1,2]Department Mathematics, Mosul University, Mosul, IRAQ

Abstract: Problem Statement: Conjugate gradient methods, which we have investigated in this study, were widely used in optimization, especially for large scale optimization problems, because it does not need the storage of any matrix. The purpose of this construction is to find new CG-algorithms suitable for solving large scale optimization problems with obtaining better numerical results. Approach: In this study, we made three linear combinations for the proposed family of CG-methods, the first resulted from the linear combination of Andrei 2007 and Powell 1984 methods; the second yielded from Andrei 2007 and Dai-Liao 2001 methods while the third was a combination of Andrei 2007 and Yabe-Takano 2003 methods. **Results**: Numerical results, showed that the presented new CG-algorithms have been proved to be effective algorithms in solving large scale optimization problems and gave us a good numerical results. Conclusion: Our new proposed algorithms always produce descent search directions and were shown to be globally convergent under some assumptions.

Key words: Unconstrained optimization, conjugate gradient method, descent directions, global convergent methods.

INTRODUCTION

Our problem was to minimize a function of n variables

$$\text{minimize } f(x), \tag{1}$$

where *f*, smooth and the gradient $g(x) \equiv \nabla f(x)$ was available. Conjugate gradient methods for solving (1) were iterative methods of the form

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

where $\alpha_k > 0$ was a step length and d_k was a search direction. Let g_k denote g(x_k). The search direction at the first iteration was the steepest descent direction, i.e., $d_0 = -g_0$. The consequent search direction were defined by

$$d_{k+1} = -g_{k+1} + \beta_{k+1} d_k, \tag{3}$$

where β_{k+1} is a scalar. For a strictly convex quadratic function $f(x) = (1/2)x^T Gx - b^T x, G \in R^{n\times n}$ a symmetric positive definite matrix, and for the exact one-dimensional minimizer α_k given by $\alpha_k = \|g_k\|^2 / d_k^T G d_k$, then the method (2)-(3) was called the linear conjugate gradient method, where $\| . \|$ denotes the Euclidean norm.

The linear conjugate gradient method was originally proposed by (Hestenes and Stiefel, 1952) for solving the linear system of equations G*x*=*b* and several formulas of β_k were considered, which were equivalent for strictly convex quadratic objective function. Within the framework of linear conjugate gradient methods, the

Corresponding Author: Abbas Y. AL-Bayati, Department Mathematics, Mosul University, Mosul, IRAQ

conjugacy condition was defined by $d_i^T G d_j = 0, \quad i \neq j$ for search directions and this condition guarantees the finite termination of linear conjugate gradient methods.

On the other hand, (2)-(3) called the nonlinear conjugate gradient method for general unconstrained optimization problem. The nonlinear conjugate gradient method was first proposed by (Fletcher and Reeves, 1964). Within the framework of nonlinear conjugate gradient methods, the conjugacy condition replaced by

$$d_{k+1}^T y_k = 0, \tag{4}$$

where $y_k = g_{k+1} - g_k$, and

$$d_{k+1}^T G d_k = \frac{1}{\alpha_k} d_{k+1}^T G(x_{k+1} - x_k) = \frac{1}{\alpha_k} d_{k+1}^T (g_{k+1} - g_k) = \frac{1}{\alpha_k} d_{k+1}^T y_k$$

hold for the strictly convex quadratic objective function, or the mean value theorem stated

$$d_{k+1}^T y_k = \alpha_k d_{k+1}^T \nabla^2 f(x_k + \omega \alpha_k d_k) d_k$$

for some $\omega \in (0, 1)$. Thus condition (4) defined that the search directions d_{k+1} and d_k were mutually conjugate with respect to the Hessian matrix $\nabla^2 f(x)$ at some point. Well-known formulas for β_k were the Fletcher-Reeves (FR), Polak-Ribière-Polyak (PRP) and Hestenes-Stiefel (HS) formulas and they were given by

$$\beta_{k+1}^{FR} = \|g_{k+1}\|^2 / \|g_k\|^2, \tag{5}$$

$$\beta_{k+1}^{PRP} = g_{k+1}^T y_k / \|g_k\|^2, \tag{6}$$

$$\beta_{k+1}^{HS} = g_{k+1}^T y_k / d_k^T y_k, \tag{7}$$

The global convergence properties of the FR, PRP and HS methods without regular restarts have been studied by many researchers. To establish the convergence results of these methods, it was usually required that the step length α_k should satisfy the strong Wolfe conditions:

$$f(x_k) - f(x_k + \alpha_k d_k) \;\geq\; -\delta \alpha_k g_k^T d_k \tag{8}$$

$$\left| g(x_k + \alpha_k d_k)^T d_k \right| \;\leq\; \sigma \left| g_k^T d_k \right| \tag{9}$$

where $0 < \delta < \sigma < 1$. Some convergence analysis even require that the α_k be computed by the exact line search, i.e. $f(x_k + \alpha_k d_k) = \min_{\alpha > 0} f(x_k + \alpha\, d_k)$ On the other hand, many other numerical methods for unconstrained optimization were proved to be convergent under the Wolfe conditions:

$$f(x_k) - f(x_k + \alpha_k d_k) \;\geq\; -\delta \alpha_k g_k^T d_k \tag{10}$$

$$g(x_k + \alpha_k d_k)^T d_k \;\geq\; \sigma g_k^T d_k. \tag{11}$$

These line search strategies required the descent condition

$$g_k^T d_k < 0, \quad for\ all\ \ k \tag{12}$$

however most of conjugate gradient methods don't always generate a descent condition, so condition (12) was usually assumed in the analysis and implementations.

Now, the algorithm of general CG-algorithm was described as follows:

Outline of the Standard CG-Algorithm:

Step0. *Given* $x_0 \in R^n$, *set* $d_0 = g_0$ *and* $k := 0$ *. If* $g_0 = 0$ *, then stop.*

Step1. *Compute* $\alpha_k > 0$ *satisfying the Wolfe conditions* (10) *and* (11).

Step2. *Let* $x_{k+1} = x_k + \alpha_k d_k$ *. If* $g_{k+1} = 0$ *, then stop.*

Step3. *Compute* β_{k+1} *and generate* d_{k+1} *by* (3).

Step4. *Set* $k := k+1$ *, and go to Step1.*

MATERIALS AND METHODS

Neculai Andrei (2007) Algorithm:

Andrei (Andrei, 2007) presented a modification of Dai and Yuan (Dai and Yuan, 1999) computational scheme

$$\beta_{k+1}^{DY} = g_{k+1}^T g_{k+1} / y_k^T s_k \,, \tag{13}$$

in order to satisfy both the sufficient descent condition and the conjugacy condition in the frame of conjugate gradient methods as:

$$d_{k+1} = -\theta_{k+1} g_{k+1} + \beta_{k+1}^A s_k, \qquad d_0 = -g_{0,} \tag{14}$$

$$\theta_{k+1} = g_{k+1}^T g_{k+1} / y_k^T g_{k+1} \,, \tag{15}$$

$$\beta_{k+1}^A = \frac{1}{y_k^T s_k} \left(g_{k+1} - \delta_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} s_k \right)^T g_{k+1}, \tag{16}$$

$$\delta_k = y_k^T g_{k+1} / g_{k+1}^T g_{k+1}, \tag{17}$$

relations (14)–(17) referred to the family of scaled conjugate gradient methods introduced by (Birgin and Martinez, 2001). In Andrei's algorithm the parameter β_{k+1} has been selected in such a manner that the sufficient descent condition was satisfied every iteration. Besides, the parameters θ_{k+1} and δ_k were chosen so that the conjugacy condition $y_k^T d_{k+1} = 0$ always holds, independently of the line search procedure.

From (15) and (17) Andrei observed that θ_{k+1}-1/(4δ_k)=(3/4)θ_{k+1}. Therefore, if for all k, $\theta_{k+1} \geq 0$, i.e. if $g_{k+1}^T y_k > 0$, then for all k the search direction d_{k+1} given by (14)-(17) satisfy the sufficient descent condition

Andrei proved the super linear convergence of his algorithm.

Powell (1984) Algorithm:

This method proposed by Powell (Powell, 1984), and analyzed by Gilbert and Nocedal (Gilbert and Nocedal, 1992). Powell showed that the Polak-Ribière (PR) method with exact line searches can cycle infinitely, without approaching a solution point.

Powell assumed that the line search always finds the first stationary point, and showed that there was a twice continuously differentiable function such that the sequence of gradients generated by the Polak-Ribière method stays bounded away from zero. Since Powell's example requires that some consecutive search directions become almost contrary, and since this can only be achieved (in the case of exact line searches) when β_{k+1}<0, Powell suggests modifying the (PR) method

$$\beta_{k+1}^{PR} = g_{k+1}^{T} y_k \big/ \|g_k\|^2, \tag{19}$$

by setting

$$\beta_{k+1}^{PR+} = \max\left\{\beta_{k+1}^{PR},\ 0\right\}. \tag{20}$$

Thus if a negative value of β_{k+1}^{PR} occurs this strategy will restart the iteration along the steepest descent direction. Gilbert and Nocedal (Gilbert and Nocedal, 1992) show that this modification of the Polak-Ribière method is globally convergent both for exact and inexact line searches. If negative values of β_{k+1}^{PR} occurred infinitely often, global convergence would follow, because an infinite number of steepest descent steps would be taken.

Dai-Liao (2001) Algorithm:

The conjugacy condition may be represented by the form (4). For many unconstrained optimization methods, including Quasi-Newton (QN) methods, the memory less BFGS method, and the limited memory BFGS method,
the search direction d_k can be written in the form

$$d_{k+1} = -H_{k+1} g_{k+1}, \tag{21}$$

where H_k is some $n\times n$ symmetric and positive definite matrix satisfying the (QN) equation:

$$H_{k+1} y_k = s_k, \tag{22}$$

$s_k = \alpha_k d_k$ is the step. By (21) and (22), we have that

$$d_{k+1}^{T} y_k = -(H_{k+1} g_{k+1})^{T} y_k \quad = -g_{k+1}^{T}(H_{k+1} y_k) = -g_{k+1}^{T} s_k. \tag{23}$$

The above relation implies that (4) holds for exact line search since in this case $g_{k+1}^{T} s_k = 0$. However, practical numerical algorithms normally adopt inexact line searches instead of exact line searches. For this reason, it seems more reasonable to replace the conjugacy condition (4) with the condition

$$d_{k+1}^{T} y_k = -t g_{k+1}^{T} s_k, \tag{24}$$

where $t \geq 0$ is a scalar. To ensure the search direction d_k in (3) satisfies the conjugacy condition (24), multiply (3) with y_k and use (17), yielding

$$\beta_{k+1}^{DL} = g_{k+1}^{T}(y_k - t s_k) \big/ d_k^{T} y_k. \tag{25}$$

Dai and Liao (Dai and Liao, 2001) suggested a modification of (25) from a viewpoint of global convergence for general functions, which is restricted the first term to non-negative values,

$$\beta_{k+1}^{DL+} = \max\left\{g_{k+1}^{T} y_k \big/ d_k^{T} y_k,\ 0\right\} - t\left(g_{k+1}^{T} s_k \big/ d_k^{T} y_k\right). \tag{26}$$

Yabe-Takano (2004) Algorithm:

Yabe and Takano (Yabe and Takano, 2004) derived a new conjugacy condition following to Dai and Liao (Dai and Liao, 2001). For this purpose, they make use of the modified secant condition

$$\begin{aligned} &H_{k+1} \hat{y}_k = s_k \\ &\hat{y}_k = y_k + \left(\theta_k \big/ s_k^{T} u_k\right) u_k. \end{aligned} \tag{27}$$

instead of the ordinary one (22). Let Z_k be defined by

$$z_k = y_k + \rho\left[\left(\theta_k / s_k^T u_k\right) u_k\right],$$
$$\theta_k = 6(f_k - f_{k+1}) + 3(g_k + g_{k+1})^T s_k, \tag{28}$$

where ρ a nonnegative parameter and $u_k \in R^n$ any vector such that $s_k^T u_k \neq 0$. Yabe and Takano consider the modified secant condition with Z_k

$$H_{k+1} z_k = s_k, \tag{29}$$

For the case ρ =0 and ρ =1, this condition corresponds to the usual secant condition (22) and the modified secant condition (27), respectively. It follows from (21) and (29) that

$$d_{k+1}^T z_k = -(H_{k+1} g_{k+1})^T z_k = -g_{k+1}^T (H_{k+1} z_k) = -g_{k+1}^T s_k.$$

Taking this relation into consideration, they replace condition (4) by the new condition

$$d_{k+1}^T z_k = -t g_{k+1}^T s_k, \tag{30}$$

where $t > 0$ is a scalar. To ensure that the search direction d_k satisfies this condition, they substituted (3) into (30)

as $-g_{k+1}^T z_k + \beta_{k+1} d_k^T z_k = -t g_{k+1}^T s_k$ and then they obtained a new β_{k+1} in the form

$$\beta_{k+1}^{YT} = g_{k+1}^T (z_k - t s_k) / d_k^T z_k. \tag{31}$$

Following to Dai and Liao algorithm, they replaced (31) by

$$\beta_{k+1}^{YT+} = \max\left\{ g_{k+1}^T z_k / d_k^T z_k,\ 0 \right\} - t\left(g_{k+1}^T s_k / d_k^T z_k \right). \tag{32}$$

and proved that the conjugate gradient method with (32) was globally convergent.

In the following section, we proposed and analyzed three new hybrid conjugate gradient algorithms as three convex combinations the first of Andrei and Powell algorithms, the second of Andrei and Dai-Liao algorithms and the third of Andrei and Yabe-Takano algorithms, and in our algorithms we used a hybrid line search procedure which optimized α_k which suggested by Al-Bayati and Rassam (Al-Bayati and Rassam, 2009).

NEW1 CG-Algorithm:

In this algorithm we combined two well-known β_i's as:

$$\begin{aligned}\beta_{k+1}^{NEW1} &= \phi_k \beta_{k+1}^{PR+} + (1-\phi_k)\alpha_k \beta_{k+1}^{A} \\ &= \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0 \right\} \right\} + (1-\phi_k)\alpha_k \left\{ \frac{1}{y_k^T s_k} \left(\|g_{k+1}\|^2 - \frac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right) \right\}\end{aligned} \tag{33}$$

where φ_k is a scalar param eter satisfying $0 \le \phi_k \le 1$. In order to keep $\beta_{k+1} \ge 0$, for the case $\phi_k = 0$, $\beta_{k+1}^{NEW1} \ge 0$ since $\beta_{k+1}^{A} \ge 0$, and for the case $\phi_k = 1$, $\beta_{k+1}^{NEW1} \ge 0$ since $\beta_{k+1}^{PR+} \ge 0$ too.

In addition, the search direction d_{k+1} is the convex combination of Andrei's search direction d_{k+1}^{A} and Powell's search direction d_{k+1}^{PR+} that is

$$d_{k+1}^{NEW1} = \phi_k d_{k+1}^{PR+} + (1-\phi_k) d_{k+1}^{A},$$

It is easy to see that

$$d_{k+1}^{NEW1} = -(\phi_k + (1-\phi_k)\theta_{k+1})g_{k+1} + (\phi_k \beta_{k+1}^{PR+} + (1-\phi_k)\alpha_k \beta_{k+1}^{A})d_k$$
$$= -\overline{\theta}_{k+1} g_{k+1} + \beta_{k+1}^{NEW1} d_k \tag{34}$$

To generate a descent search direction, Suppose that $\beta_{k+1} \geq 0$and the current search direction d_k was a descent direction, i.e., $g_k^T d_k < 0$. To find β_{k+1} let

$$g_{k+1}^T d_{k+1} = -\overline{\theta}_{k+1} \|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k < 0. \tag{35}$$

by using $g_k^T d_k < 0$ and $\beta_k \geq 0$, the right hand side may be rewritten as

$$\begin{aligned} -\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k &= -\overline{\theta}_{k+1} \|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k \ \text{-} \beta_{k+1} g_k^T d_k + \beta_{k+1} g_k^T d_k \\ &= -\overline{\theta}_{k+1} \|g_{k+1}\|^2 + \beta_{k+1} d_k^T y_k + \beta_{k+1} g_k^T d_k \\ &\leq -\overline{\theta}_{k+1} \|g_{k+1}\|^2 + \beta_{k+1} d_k^T y_k. \end{aligned} \tag{36}$$

therefore, the non-positivity of (36) was sufficient to investigate from condition (35). Condition (36) reduced to the condition

$$\|g_{k+1}\|^2 \geq -\theta_{k+1}^{-1} \beta_{k+1} d_k^T y_k \tag{37}$$

by summarizing these relations, we have the following lemma.

Lemma 1.

We assume $\beta_{k+1} \geq 0$. If β_{k+1} satisfies inequality (37)*, then d_{k+1} is a descent direction for the objective function.*

Now, we will construct a new β_{k+1} that satisfies the condition (37). For this purpose, we choose φ_k that satisfies the inequality (37).

Case (i). For exact line search $d_k^T g_{k+1} = s_k^T g_{k+1} = 0$

$$\begin{aligned} \|g_{k+1}\|^2 &\geq \overline{\theta}_{k+1}^{-1} \beta_{k+1} d_k^T y_k \\ &= \left\{ \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0 \right\} \right\} + (1-\phi_k)\alpha_k \left\{ \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right\} \right\} \overline{\theta}_{k+1}^{-1} d_k^T y_k \\ &= \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k + \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \overline{\theta}_{k+1}^{-1} d_k^T y_k. \end{aligned}$$

This yields,

$$\left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k} \right) \|g_{k+1}\|^2 \geq \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

For simplicity of notation, we define

$$\mu_k = \max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \qquad ,$$

then we have

$$\left(1-\alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k}\right)\|g_{k+1}\|^2 \geq \phi_k \mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \qquad .$$

Let φ_k be defined as follows:

$$\begin{cases} (case1) \quad \textit{For the case } \mu_k \leq 0, \\ \qquad \textit{we choose } \phi_k \textit{ such that } 0 \leq \phi_k \leq 1. \\ (case2) \quad \textit{For the case } \mu_k > 0, \\ \qquad \textit{we choose } \phi_k \textit{ such that } 0 \leq \phi_k \leq \overline{\phi}_k, \end{cases} \tag{38}$$

where

$$\overline{\phi}_k = \min\left\{\frac{1}{\mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k}\left(1-\alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k}\right)\|g_{k+1}\|^2,\ 1\right\}$$

Then β_{k+1}^{NEW1} satisfies condition (38).

Case (ii). For inexact line search $d_{k+1}^T g_{k+1} > 0$, $s_k^T g_{k+1} > 0$,

Let $\eta_k = \dfrac{1}{y_k^T s_k}\left(\|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1}\right)$

$$\begin{aligned} \|g_{k+1}\|^2 &\geq \overline{\theta}_{k+1}^{-1} \beta_{k+1} d_k^T y_k \\ &= \left\{\phi_k \left\{\max\left\{\frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0\right\}\right\} + 1-\phi_k)\alpha_k \eta_k\right\} \overline{\theta}_{k+1}^{-1} d_k^T y_k \\ &= \phi_k \left(\max\left\{\frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0\right\} - \alpha_k \eta_k\right) \overline{\theta}_{k+1}^{-1} d_k^T y_k + \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k. \end{aligned}$$

This yields,

$$\begin{aligned} &\|g_{k+1}\|^2 - \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \\ &\quad \geq \phi_k \left(\max\left\{\frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0\right\} - \alpha_k \eta_k\right) \overline{\theta}_{k+1}^{-1} d_k^T y_k \end{aligned}$$

For simplicity of notation, we define

$$\mu_k = \max\left\{\frac{g_{k+1}^T y_k}{g_k^T g_k},\ 0\right\} - \alpha_k \eta_k \qquad ,$$

then we have

$$\|g_{k+1}\|^2 - \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \geq \varphi_k \mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

as (38) we have

$$\bar{\phi}_k = \min\left\{\frac{1}{\mu_k \bar{\theta}_{k+1}^{-1} d_k^T y_k}\left(\|g_{k+1}\|^2 - \alpha_k \eta_k \bar{\theta}_{k+1}^{-1} d_k^T y_k\right),\ 1\right\} \tag{39}$$

Then β_{k+1}^{NEW1} satisfies condition (37).

Outline of NEW1 CG-Algorithm:

For $x_0 \in R^n$, *set* $d_0 = -g_0$, *and* $k = 0$. *If* $g_0 = 0$, *then stop.*

Compute α_k *satisfying the new condition*

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta_1 \alpha_k g_k^T d_k - \delta_2 \alpha_k^2 \|d_k\|^2 \tag{40}$$

$$\left|g_{k+1}^T d_k\right| < \sigma \left|d_k^T g_k\right| \tag{41}$$

Set $x_{k+1} = x_k + \alpha_k d_k$. *If* $g_{k+1} = 0$, *then stop.*

Compute β_{k+1} *by* (33) *and* (38) *or* (39). *In addition, generate* d_{k+1} *by* (34).
Set k=k+1, *and go to Step2.*

NEW2 CG-Algorithm:

In this algorithm, we combined two well-known β_i's as:

$$\begin{aligned}\beta_{k+1}^{NEW2} &= \phi_k \beta_{k+1}^{DL+} + (1-\phi_k)\alpha_k \beta_{k+1}^{A} \\ &= \phi_k \left\{\max\left\{\frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0\right\} - t\,\frac{g_{k+1}^T s_k}{d_k^T y_k}\right\} + (1-\phi_k)\alpha_k \left\{\frac{1}{y_k^T s_k}\left(\|g_{k+1}\|^2 - \frac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1}\right)\right\},\end{aligned} \tag{42}$$

where φ_k is a scalar parameter satisfying $0 \le \phi_k \le 1$. In order to keep $\beta_{k+1} \ge 0$. we control the parameter t so that $\beta_k^{DL+} \ge 0$is satisfied. Concretely, for the case $\left(g_{k+1}^T s_k / d_k^T y_k\right) \le 0$, we can choose any $t \ge 0$, and for the case $\left(g_{k+1}^T s_k / d_k^T y_k\right) > 0$, we choose t such that

$$0 \le t \le \left(d_k^T y_k / g_{k+1}^T s_k\right) \max\left\{g_{k+1}^T y_k / d_k^T y_k\ ,\ 0\right\} \quad .$$

In addition, the search direction d_{k+1} is the convex combination of Andrei's search direction d_{k+1}^A and Dai and Laio's search direction d_{k+1}^{DL+} that is

$$d_{k+1}^{NEW2} = \phi_k d_{k+1}^{DL+} + (1-\phi_k) d_{k+1}^A \quad .$$

It is easy to see that

$$\begin{aligned} d_{k+1}^{NEW2} &= -(\phi_k + (1-\phi_k)\theta_{k+1}) g_{k+1} + (\phi_k \beta_{k+1}^{DL+} + (1-\phi_k)\alpha_k \beta_{k+1}^A) d_k \\ &= -\bar{\theta}_{k+1} g_{k+1} + \beta_{k+1}^{NEW2} d_k. \end{aligned} \tag{43}$$

where $\theta_{k+1} = \left(g_{k+1}^T g_{k+1} / y_k^T g_{k+1}\right)$, and $\bar{\theta}_{k+1} = \phi_k + (1-\phi_k)\theta_{k+1}$. Now, we will construct a new β_{k+1} that satisfies the condition (37). For this purpose, we choose φ_k that satisfies the following inequality (37).

Case (i). For exact line search $d_k^T g_k = s_k^T g_{k+1} = 0,$

$$\|g_{k+1}\|^2 \geq \overline{\theta}_{k+1}^{-1} \beta_{k+1} d_k^T y_k$$
$$= \left\{ \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} \right\} + \ (1-\phi_k)\alpha_k \left\{ \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right\} \right\} \overline{\theta}_{k+1}^{-1} d_k^T y_k$$
$$= \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k + \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \overline{\theta}_{k+1}^{-1} d_k^T y_k.$$

This yields,

$$\left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k}\right) \|g_{k+1}\|^2 \geq \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

For simplicity of notation, we define

$$\mu_k = \max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k},$$

then we have

$$\left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k}\right) \|g_{k+1}\|^2 \geq \phi_k \mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k.$$

as (38) we have

$$\overline{\phi}_k = \min\left\{ \frac{1}{\mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k} \quad \left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k}\right) \|g_{k+1}\|^2,\ 1 \right\} \tag{44}$$

Then β_{k+1}^{NEW2} satisfies condition (37).

Case (ii). For inexact line search $d_{k+1}^T g_{k+1} > 0$, $s_k^T g_{k+1} > 0$,

Let $\eta_k = \dfrac{1}{y_k^T s_k} \left(\|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right)$

$$\|g_{k+1}\|^2 \geq \overline{\theta}_{k+1}^{-1} \beta_{k+1} d_k^T y_k$$
$$= \left\{ \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T y_k} \right\} + \quad (1-\phi_k)\alpha_k \eta_k \right\} \overline{\theta}_{k+1}^{-1} d_k^T y_k$$
$$= \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T y_k} - \alpha_k \eta_k \right) \ \overline{\theta}_{k+1}^{-1} d_k^T y_k + \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k.$$

This yields,

$$\|g_{k+1}\|^2 - \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \ge \phi_k \left(\max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - \ t \frac{g_{k+1}^T s_k}{d_k^T y_k} - \alpha_k \eta_k \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

For simplicity of notation, we define

$$\mu_k = \max\left\{ \frac{g_{k+1}^T y_k}{d_k^T y_k},\ 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T y_k} - \alpha_k \eta_k \quad ,$$

then we have

$$\|g_{k+1}\|^2 - \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \ge \phi_k \mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

as (38) we have

$$\overline{\phi}_k = \min\left\{ \frac{1}{\mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k} \left(\|g_{k+1}\|^2 - \alpha_k \eta_k \overline{\theta}_{k+1}^{-1} d_k^T y_k \right),\ 1 \right\} \tag{45}$$

Then β_{k+1}^{NEW2} satisfies condition (37).

Outline of NEW2 CG-Algorithm:

Is same as in NEW1 CG-algorithm except that β_{k+1} *computed by* (42) *and* (44) *or* (45). *In addition, generate* d_{k+1} *by* (43).

NEW3 CG-Algorithm:

In this algorithm we combined two well-known $\beta_i's$ as:

$$\begin{aligned} \beta_{k+1}^{NEW3} &= \phi_k \beta_{k+1}^{YT+} + (1-\phi_k)\alpha_k \beta_{k+1}^{A} \\ &= \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0 \right\} - t \frac{g_{k+1}^T s_k}{d_k^T z_k} \right\} + (1-\phi_k)\alpha_k \left\{ \frac{1}{y_k^T s_k} \left(\|g_{k+1}\|^2 - \frac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right) \right\}, \end{aligned} \tag{46}$$

$$z_k = y_k + \rho\left(\frac{\theta_k}{s_k^T u_k} u_k \right) \rho\ \ge 0,$$

$$\theta_k = 6\left(f_k - f_{k+1}\right) + 3\left(g_k + g_{k+1}\right)^T s, u_k \ \textit{is any vecter}\ \ni\ s_k^T u_k \ne 0.$$

where φ_k is a scalar parameter satisfying $0 \le \phi_k \le 1$. In order to keep $\beta_{k+1} \ge 0$ we control the parameter t so that $\beta_k^{YT+} \ge 0$is satisfied. Concretely, for the case $\left(g_{k+1}^T s_k / d_k^T z_k\right) \le 0$, we can choose any $t \ge 0$, and for the case $\left(g_{k+1}^T s_k / d_k^T z_k\right) > 0$, we choose t such that

$$0 \le t \le \left(d_k^T z_k / g_{k+1}^T s_k\right) \max\left\{ g_{k+1}^T z_k / d_k^T z_k\ ,\ 0 \right\}.$$

In addition, the search direction d_{k+1} is the convex combination of Andrei's search direction d_{k+1}^A and Yabe and Takano's search direction d_{k+1}^{YT+}, that is

$$d_{k+1}^{NEW3} = \phi_k d_{k+1}^{YT+} + (1-\phi_k) d_{k+1}^A \quad .$$

It is easy to see that

$$\begin{aligned} d_{k+1}^{NEW3} &= -(\phi_k + (1-\phi_k)\theta_{k+1})g_{k+1} + \ (\phi_k \beta_{k+1}^{YT+} + (1-\phi_k)\alpha_k \beta_{k+1}^{A})d_k \\ &= -\overline{\theta}_{k+1} g_{k+1} + \beta_{k+1}^{NEW3} d_k . \end{aligned} \tag{47}$$

where $\theta_{k+1} = \left(g_{k+1}^T g_{k+1} / y_k^T g_{k+1} \right)$, and $\overline{\theta}_{k+1} = \phi_k + (1-\phi_k)\theta_{k+1}$. Now, we will construct a new β_{k+1} that satisfies the condition (37). For this purpose, we choose φ_k that satisfies the inequality (37).

Case (i). For exact line search $d_k^T g_k = s_k^T g_{k+1} = 0,$

$$\begin{aligned} \|g_{k+1}\|^2 &\geq \overline{\theta}_{k+1}^{-1} \beta_{k+1} d_k^T y_k \\ &= \left\{ \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0 \right\} \right\} + (1-\phi_k)\alpha_k \left\{ \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right\} \right\} \overline{\theta}_{k+1}^{-1} d_k^T y_k \\ &= \phi_k \left(\max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k + \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \overline{\theta}_{k+1}^{-1} d_k^T y_k . \end{aligned}$$

This yields,

$$\left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k} \right) \|g_{k+1}\|^2 \geq \phi_k \left(\max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k} \right) \overline{\theta}_{k+1}^{-1} d_k^T y_k$$

For simplicity of notation, we define

$$\mu_k = \max\left\{ \frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0 \right\} - \alpha_k \frac{\|g_{k+1}\|^2}{y_k^T s_k},$$

then we have

$$\left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k} \right) \|g_{k+1}\|^2 \geq \phi_k \mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k .$$

as (38) we have

$$\overline{\phi}_k = \min\left\{ \frac{1}{\mu_k \overline{\theta}_{k+1}^{-1} d_k^T y_k} \ \left(1 - \alpha_k \overline{\theta}_{k+1}^{-1} \frac{d_k^T y_k}{y_k^T s_k} \right) \|g_{k+1}\|^2,\ 1 \right\} \tag{48}$$

Then β_{k+1}^{NEW3} satisfies condition (37).

Case (ii). For inexact line search $d_{k+1}^T g_{k+1} > 0$, $s_k^T g_{k+1} > 0$,

Let $\eta_k = \dfrac{1}{y_k^T s_k} \left(\|g_{k+1}\|^2 - \dfrac{y_k^T g_{k+1}}{y_k^T s_k} s_k^T g_{k+1} \right)$

$$\begin{aligned}\|g_{k+1}\|^2 &\geq \bar{\theta}_{k+1}^{-1}\beta_{k+1}d_k^T y_k \\ &= \left\{\phi_k \left\{\max\left\{\frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0\right\} - t\,\frac{g_{k+1}^T s_k}{d_k^T z_k}\right\} + \quad (1-\phi_k)\alpha_k\eta_k\right\}\bar{\theta}_{k+1}^{-1}d_k^T y_k \\ &= \phi_k\left(\max\left\{\frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0\right\} - t\,\frac{g_{k+1}^T s_k}{d_k^T z_k} - \alpha_k\eta_k\right)\bar{\theta}_{k+1}^{-1}d_k^T y_k + \alpha_k\eta_k\bar{\theta}_{k+1}^{-1}d_k^T y_k.\end{aligned}$$

This yields,

$$\|g_{k+1}\|^2 - \alpha_k\eta_k\bar{\theta}_{k+1}^{-1}d_k^T y_k \geq \phi_k\left(\max\left\{\frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0\right\} - t\,\frac{g_{k+1}^T s_k}{d_k^T z_k} - \alpha_k\eta_k\right)\bar{\theta}_{k+1}^{-1}d_k^T y_k$$

For simplicity of notation, we define

$$\mu_k = \max\left\{\frac{g_{k+1}^T z_k}{d_k^T z_k},\ 0\right\} - t\,\frac{g_{k+1}^T s_k}{d_k^T z_k} - \alpha_k\eta_k,$$

then we have

$$\|g_{k+1}\|^2 - \alpha_k\eta_k\bar{\theta}_{k+1}^{-1}d_k^T y_k \geq \phi_k\mu_k\bar{\theta}_{k+1}^{-1}d_k^T y_k$$

as (38) we have

$$\bar{\phi}_k = \min\left\{\frac{1}{\mu_k\bar{\theta}_{k+1}^{-1}d_k^T y_k}\left(\|g_{k+1}\|^2 - \alpha_k\eta_k\bar{\theta}_{k+1}^{-1}d_k^T y_k\right),\ 1\right\} \tag{49}$$

Then β_{k+1}^{NEW3} satisfies condition (37).

Outline of NEW3 CG-Algorithm:

Is same as in NEW1 CG-algorithm except that β_{k+1} computed by (46) *and* (48) *or* (49). *In addition, generate d_{k+1} by* (47).

Convergence Results:

Assumption 1.

The level set $\Gamma := \{x : f(x) \leq f(x_0)\}$ at initial point was bounded.

In some neighborhood N of Γ, the objective function f was continuously differentiable, and its gradient was Lipschitz condition, there exists a constant L>0 such that

$$\|g(x) - g(\tilde{x})\| \leq L\|x - \tilde{x}\|, \quad \forall \quad x, \tilde{x} \in N \tag{50}$$

Now, we deal with the sufficient descent condition, namely, for some constant c≥0,

$$g_k^T d_k \leq -c\|g_k\|^2, \quad \textit{for all} \quad k \geq 0 \tag{51}$$

We will show that our algorithm guarantees (51).

Theorem 1.

Suppose that for a starting point x_0, for which Assumption (1) *holds. Let the sequence $\{x_k\}$ be generated By any one of the new Algorithms. If we choose β_{k+1} such that $\|g_{k+1}\|^2 \geq \theta_{k+1}^{-1}\beta_{k+1}d_k^T y_k, \quad \beta_{k+1} \geq 0$ and the step size α_k satisfied the conditions* (40)-(41)*, suppose that for all $k \geq 0$ there exist a positive constant $\overline{c}$ such that $0 \leq \theta_{k+1} \leq \overline{c}$ then the conjugate gradient method will satisfy the sufficient descent condition* (51) *with $c = \dfrac{\overline{c}}{1+\sigma}$.*

Proof.

The proof was done by induction. By noting that σ>0 implies $-1 < -\dfrac{1}{1+\sigma}$, the result clearly holds for k=1 since

$$g_0^T d_0 = -\|g_0\|^2 \leq -\frac{1}{1+\sigma}\|g_0\|^2 = -c\|g_0\|^2 ,$$

where $c = \dfrac{1}{1+\sigma}$.

Assume that (51) holds for some $k \geq 0$. It follows from (41) that

$$l_k \equiv \frac{g_{k+1}^T d_k}{g_k^T d_k} \in [-\sigma,\ \sigma] \quad and \quad d_k^T y_k > 0 \quad .$$

Then we have

$$l_k - 1 = \frac{g_{k+1}^T d_k}{g_k^T d_k} - 1 = \frac{g_{k+1}^T d_k - g_k^T d_k}{g_k^T d_k} = \frac{d_k^T y_k}{g_k^T d_k} \neq 0 \quad . \tag{52}$$

By (34), we have

$$\begin{aligned} g_{k+1}^T d_{k+1} &= g_{k+1}^T(-\theta_{k+1}g_{k+1} + \beta_{k+1}d_k) \\ &= -\theta_{k+1}\|g_{k+1}\|^2 + \beta_{k+1}g_{k+1}^T d_k. \end{aligned}$$

Based on the sign of $g_{k+1}^T d_k$, we consider the following two cases.

(i) The case $g_{k+1}^T d_k \leq 0$: We immediately have that

$$\begin{aligned} g_{k+1}^T d_{k+1} &= -\theta_{k+1}\|g_{k+1}\|^2 + \beta_{k+1}g_{k+1}^T d_k \\ &\leq -\theta_{k+1}\|g_{k+1}\|^2 \\ &\leq -\frac{1}{1+\sigma}\theta_{k+1}\|g_{k+1}\|^2 \\ &\leq \frac{c}{1+\sigma}\|g_{k+1}\|^2 \end{aligned}$$

(ii) The case $g_{k+1}^T d_k > 0$: The conditions on β_{k+1} and (52) yield

$$g_{k+1}^T d_k = -\theta_{k+1}\|g_{k+1}\|^2 + \beta_{k+1} g_{k+1}^T d_k$$
$$\le -\theta_{k+1}\|g_{k+1}\|^2 + \frac{\theta_{k+1}\|g_{k+1}\|^2}{d_{k+1}^T d_k}$$
$$= \left(-1 + \frac{g_{k+1}^T d_k}{d_k^T y_k}\right)\theta_{k+1}\|g_{k+1}\|^2$$
$$\le -\frac{1}{1+\sigma}\theta_{k+1}\|g_{k+1}\|^2$$
$$\le \frac{c}{1+\sigma}\|g_{k+1}\|^2.$$

By noting $\frac{1}{1+\sigma} \le \frac{1}{1+l_k} \le \frac{1}{1-\sigma}$, the inequality above implies that

$$g_{k+1}^T d_{k+1} \le \frac{\overline{c}}{l_k - 1}\|g_{k+1}\|^2$$
$$\le -\frac{\overline{c}}{l_k - 1}\|g_{k+1}\|^2$$

By summarizing the cases (i) and (ii), the sufficient descent condition (51) holds with $c\frac{\overline{c}}{1+\sigma}$ at the k+1 iteration, Therefore the proof was completed. ■

Theorem 2.

Let ∇f(x) be uniformly continuous on the level set $\Gamma = \left\{x \in R^n \,\middle|\, f(x) \le f(x_0)\right\}$ Let also the angle θ_k between -∇f(x) and the direction d_k was uniformly bounded away from 90°, i.e., satisfies

$$\theta_k \le \frac{\pi}{2} - \mu \quad \textit{fot some } \mu>0. \tag{53}$$

Then $\nabla f(x_k)$ for some k ; or $f(x) \to \infty; or\, \nabla f(x_k) \to 0$

Proof. See (Sun and Yuan, 2006)

Theorem 3.

Let f(x) be continuously differentiable, and Assumption (1) *holds. Suppose that for all k≥0 there exists the positive constants γ,δ and M>m>0 such that* $0 < \gamma \le \overline{\theta}_k \le \delta$ *and*

$$m\|y\|^2 \le y^T \nabla^2 f(x) y \le M\|y\|^2, \forall y \in R^n, x \in \Gamma \tag{53}$$

Then the sequence generate by NEW1 CG-algorithm with exact line search converges to the minimizer x of f.*

Proof.

From theorem (2), we know that it is enough to prove that (53) holds, that is, there exists a constant λ>0 such that

$$-g_k^T d_k \ge \lambda\|g_k\|\|d_k\|, \tag{55}$$

which means $\cos\theta_k \ge \lambda > 0.$

Then, from Theorem (2), we have $g_k \to 0$ and $(x^*) = 0$ From (47), it follows that $\{x_k\} \to x^*$ which is a unique minimizer.

By using $g_k^T d_k = 0$ and (34). We have

$$g_k^T d_{k+1} = -\bar{\theta}_{k+1} \|g_{k+1}\|^2$$

Then (55) is equivalent to

$$\frac{\|g_{k+1}\|}{\|d_{k+1}\|} \geq \bar{\lambda}. \tag{56}$$

where $\bar{\lambda} = \lambda / \delta > 0$ by using $\alpha_k = -(g_k^T d_k / d_k^T G_k d_k)$, and (34), it follows that

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T G_k d_k} = \frac{\bar{\theta}_{k+1} \|g_k\|^2}{d_k^T G_k d_k}, \tag{57}$$

where

$$G_k = \int_0^1 G(x_k + t\alpha_k d_k) dt \tag{58}$$

By (58), the integral form of the mean-value theorem is

$$g_{k+1} - g_k = g(x_k + \alpha_k d_k) - g(x_k) = \alpha_k G_k d_k \tag{59}$$

Then, by (58) and (57), (33) becomes

$$\beta_{k+1}^{NEW1} = \phi_k \left\{ \max\left\{ \frac{g_{k+1}^T y_k}{g_k^T g_k}, 0 \right\} \right\} + (1-\phi)\alpha_k \left\{ \frac{1}{y_k^T S_k} \left[\|g_{k+1}\|^2 - \frac{g_k^T g_{k+1}}{y_k^T S_k} g_{k+1} \right] \right\}$$

where $0 \leq \phi_k \leq 1$

$$|\beta_{k+1}| \leq \frac{\delta \|g_{k+1}\| \|G_k d_k\|}{m\|d_k\|^2} + \frac{1}{m\|d_k\|^2} \frac{Mm\|d_k\|^2}{\|s\| \|G_k d_k\|^2} \left(\frac{\|g_{k+1}\| \|G_k d_k\|}{\|s\| \|G_k d_k\|} \|s_k\| \|g_{k+1}\| \right)$$

$$\leq \frac{M}{m} \frac{\delta \|g_{k+1}\|}{\|d_k\|^2} + \frac{M\|g_{k+1}\|}{m^2 \|s_k\| \|d_k\|} \left(1 + \frac{M}{m} \right)$$

$$= \frac{M}{m} \frac{\|g_{k+1}\|}{\|d_k\|} \left[\delta + \frac{\|g_{k+1}\|}{m\|s_k\|} \left(1 + \frac{M}{m} \right) \right]$$

$$= C_k \frac{M}{m} \frac{\|g_{k+1}\|}{\|d_k\|}$$

where $C_k = \delta + \dfrac{\|g_{k+1}\|}{m\|d_k\|} \left(1 + \dfrac{M}{m} \right)$. Therefore

$$\|d_{k+1}\| \le |\overline{\theta}_{k+1}|\|g_{k+1}\| + |\beta_{k+1}|\|d_k\|$$

$$\le \delta\|g_{k+1}\| + C_k \frac{M}{m}\|g_{k+1}\|$$

$$= \left(\delta + C_k \frac{M}{m}\right)\|g_{k+1}\|$$

$$\le \left(\delta + C \frac{M}{m}\right)\|g_{k+1}\|$$

$$= \xi\|g_{k+1}\|$$

where $C = \max_{\forall k}\{C_k\}$ and $\xi = \delta + C\dfrac{M}{m}$ Which gives .

The above inequality show (56) holds. We complete the proof of the NEW1 CG-algorithm.■ Similarly we can prove the convergence property of NEW2 and NEW3 CG-algorithms.

Numerical Results

In this section we present the numerical results of the new CG-algorithms. We have modified a readily program written by the Nocedal (1998) to be suitable for the new CG-algorithms NEW1, NEW2 and NEW3 and we have taken different values for the existent constants in the parameters $Ø_k$, t, and ρ. We implemented the new CG-algorithms by using a set of unconstrained optimization test problems (Andrie, 2009).

Table 1: Numerical results of NEW1 algorithm

Test Fun.	n	NEW1 CG-Algorithm			(PR+) Method	(A) Method
		$Ø_k$=0.001	$Ø_k$=0.5	$Ø_k$=0.999		
Extended EP1	100	7.931e+02	7.931e+02	7.931e+02	*failed*	*failed*
	1000	7.931e+03	7.931e+03	7.931e+03	7.931e+03	7.931e+03
ARWHEAD	800	*failed*	*failed*	8.870e-14	1.774e-12	*failed*
	900	*failed*	*failed*	0.000e+00	7.763e-13	*failed*
Extended Powell	100	*failed*	*failed*	4.888e-08	5.822e-07	1.310e-04
	700	*failed*	*failed*	1.452e-07	1.466e-06	1.211e-03
Miele	200	*failed*	3.243e-08	9.912e-07	1.240e-06	5.535e-06
	300	*failed*	2.103e-09	5.071e-07	4.890e-07	8.770e-06
Diagonal3	400	*failed*	-7.83e+04	-7.83e+04	-7.83e+04	-7.83e+04
	1000	*failed*	-4.95e+05	-4.95e+05	-4.95e+05	-4.95e+05
Extended Hiebert	700	8.750e+11	8.750e+11	8.750e+11	8.750e+11	8.750e+11
	900	1.125e+12	1.125e+12	1.125e+12	1.125e+12	1.125e+12
Extended Himmelblau	300	3.840e-11	8.748e-11	1.077e-14	2.645e-16	6.864e-11
	900	*failed*	2.931e-11	2.159e-13	4.660e-13	4.898e-10
Extended Trigonometric	200	*failed*	*failed*	*failed*	*failed*	*failed*
	500	*failed*	*failed*	*failed*	*failed*	*failed*

Table 2: Numerical results of NEW2 algorithm

Test Fun.		NEW2 CG-Algorithm			(DL+) Method	(A) Method
		$Ø_k$=0.5	$Ø_k$=0.999			
	n	t=0.9	t=0.5	t=0.9		
GDixon	200	7.931e+02	7.931e+02	7.931e+02	*failed*	*failed*
	1000	7.931e+03	7.931e+03	7.931e+03	7.931e+03	7.931e+03
DIXMAANA	100	*failed*	*failed*	8.870e-14	1.774e-12	*failed*
	500	*failed*	*failed*	0.000e+00	7.763e-13	*failed*
ARWHEAD	100	*failed*	*failed*	4.888e-08	5.822e-07	1.310e-04
	200	*failed*	*failed*	1.452e-07	1.466e-06	1.211e-03
Extended Himmelblau	600	*failed*	3.243e-08	9.912e-07	1.240e-06	5.535e-06
	700	*failed*	2.103e-09	5.071e-07	4.890e-07	8.770e-06
Extended EP1	100	*failed*	-7.83e+04	-7.83e+04	-7.83e+04	-7.83e+04
	200	*failed*	-4.95e+05	-4.95e+05	-4.95e+05	-4.95e+05
Extended Powell	200	8.750e+11	8.750e+11	8.750e+11	8.750e+11	8.750e+11
	300	1.125e+12	1.125e+12	1.125e+12	1.125e+12	1.125e+12
Miele	100	3.840e-11	8.748e-11	1.077e-14	2.645e-16	6.864e-11
	400	*failed*	2.931e-11	2.159e-13	4.660e-13	4.898e-10
Extended Trigonometric	500	*failed*	*failed*	*failed*	*failed*	*failed*
	600	*failed*	*failed*	*failed*	*failed*	*failed*

Table 3: Numerical results of NEW3 algorithm

Test Fun.	n	t	ρ	NEW3 Algorithm		(YT+) Method	(A) Method
				$\O_k$=0.5	$\O_k$=0.999		
ARGILNB with u_k=y_k	1000	0.0	1.0	4.634e+00	4.634e+00	4.634e+00	4.634e+00
		0.9	0.2	4.634e+00	4.634e+00	4.634e+00	4.634e+00
ARGILNB with u_k=s_k	1000	0.4	0.7	4.634e+00	4.634e+00	4.634e+00	4.634e+00
		0.5	0.6	4.634e+00	4.634e+00	4.634e+00	4.634e+00
DIXMAANB with u_k=y_k	800	0.0	1.0	*failed*	1.000e+00	1.000e+00	*failed*
		0.8	0.3	*failed*	1.000e+00	*failed*	*failed*
DIXMAANB with u_k=s_k	800	0.6	0.5	*failed*	*failed*	1.000e+00	*failed*
		0.7	0.4	*failed*	1.000e+00	1.000e+00	*failed*
Diagonal5 with u_k=y_k	1000	0.3	0.8	6.931e+02	6.931e+02	6.931e+02	6.931e+02
		1.0	0.0	6.931e+02	6.931e+02	6.931e+02	6.931e+02
Diagonal5 with u_k=s_k	1000	0.5	0.6	6.931e+02	6.931e+02	6.931e+02	6.931e+02
		0.6	0.5	6.931e+02	6.931e+02	6.931e+02	6.931e+02
DIXMAANJ with u_k=y_k	500	0.3	0.8	*failed*	*failed*	*failed*	*failed*
		0.5	0.6	*failed*	*failed*	*failed*	*failed*
DIXMAANJ with u_k=s_k	500	0.4	0.7	*failed*	*failed*	*failed*	*failed*
		0.9	0.2	*failed*	*failed*	*failed*	*failed*

Conclutions:

We have concluded that more large values of $\O_k$ gave us better results for NEW1 CG-algorithm and we have got a robust results when we took more large value of $\O_k$ with more small value of t for NEW2 CG-algorithm while for NEW3 CG-algorithm we have taken more large values of $\O_k$, and we have chosen two choices for the arbitrary vector u_k,(u_k=y_k and u_k=s_k).

The suggested algorithms like the original algorithm for some test functions don't converge within given accuracy in some test problems.

We have observed by taking large $\O_k$ that the suggested algorithms arrive at the limit point while the original algorithm failed.

REFERENCES

Al-Bayati, A.Y. and S.Y. Rassam, 2009, A modified globally convergence CG-algorithm with a hybrid line search procedure, Australian Journal of Basic and Applied Sciences, INSInet Publication, 3(3): 2188-2206.

Andrei, N., 2009. Another nonlinear conjugate gradient algorithm for unconstrained optimization," Optimization Methods & Software, 24(1): 89-104.

Andrei, N., 2007. Dai-Yuan conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization. Applied Mathematics Letters, 21: 165-171.

Birigin, F.G. and J.M. Martinez, 2001. A spectral conjugate gradient method for unconstrained optimization. Applied Mathematics and Optimization, 43: 117-128.

Dai, Y.H. and L.Z. Liao, 2001. New conjugacy conditions and related nonlinear conjugate gradient methods. Applied Mathematics and Optimization, Springer-Verlag, New York, 43: 87-101.

Dai, Y.H. and Y. Yuan, 1999. A nonlinear conjugate gradient method with a strong global convergence property. SIAM Journal on Optimization, 10(1): 177-182.

Fletcher, R. and C.M. Reeves, 1964. Function minimization by conjugate gradients. Computer Journal, 7: 149-154.

Gilbert, J.C. and J. Nocedal, 1992. Global convergence properties of conjugate gradient methods for optimization. SIAM Journal on Optimization, 2: 21-42.

Hestenes, M.R. and E. Stiefel, 1952. Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards, 5(49): 409-436.

Powell, M.J.D., 1984. Nonconvex minimization calculations and the conjugate gradient method. in Lecture Notes in mathematics 1066, Springer-Verlag, Berlin, 1066: 122-141.

Sun, W. and Y. Yuan, 2006. Optimization theory and methods, Nonlinear programming. Springer Science, Business Media, LLC.

Yabe, H. and M. Takano, 2004. Global convergence properties of nonlinear conjugate gradient methods with modified secant condition. Computational Optimization and Applications.