Control and Optimization in Applied Mathematics (COAM)
DOI. 10.30473/coam.2020.46435.1116
Vol. 3, No. 2, Autumn-Winter 2018(59-76), ©2016 Payame Noor University, Iran

Two Settings of the Dai-Liao Parameter Based on Modified Secant Equations

S. Nezhadhosein,[1], S. Mohammadkhan Sartip[2,*]

[1,2] Department of Mathematics, Payame Noor University,
P.O. Box. 19395-3697, Tehran, Iran

Received: May 28, 2019; **Accepted:** February 18, 2020.

Abstract. Following the setting of the Dai-Liao (DL) parameter in conjugate gradient (CG) methods, we introduce two new parameters based on the modified secant equation proposed by Li et al. (Comput. Optim. Appl. 202:523-539, 2007) with two approaches, which use an extended new conjugacy condition. The first is based on a modified descent three-term search direction, as the descent Hestenes-Stiefel CG method. The second is based on the quasi-Newton (QN) approach. Global convergence of the proposed methods for uniformly convex functions and general functions is proved. Numerical experiments are done on a set of test functions of the CUTEr collection and the results are compared with some well-known methods.

Keywords. Unconstrained optimization, Modified secant equations, Dai-Liao conjugate gradient method

MSC. 90C34; 90C40.

* Corresponding author
s_nejhadhosein@pnu.ac.ir, saharsartip@gmail.com
http://mathco.journals.pnu.ac.ir

1 Introduction

Conjugate gradient (CG) and quasi-Newton (QN) methods contain a class of unconstrained optimization algorithms, with some great properties such as low memory requirements and strong global convergence [34], which make them famous for engineers and mathematicians engaged in solving large-scale problems, as follows:

$$\begin{aligned} &\min f(x) \\ &x \in \mathbb{R}^n \end{aligned} \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth nonlinear function, and its gradient is available. The iterative formula of a CG method leads to a sequence of the approximate solutions, as $\{x_n\}$ with the following recursive formula:

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

where $x_0 \in \mathbb{R}^n$ is an initial solution and d_k is the search direction with following formula:

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

where $g_k = \nabla f(x_k)$ and β_k is a scalar called the CG (update) parameter. In Eqn (2) the α_k parameter is the step length at current iteration along d_k. Inexact line searches satisfy some certain line search conditions [22]. Among them, the so-called Wolfe conditions [22] have attracted particular attention in the convergence analyses and the implementation of CG methods, requiring that:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \tag{4}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \tag{5}$$

where $0 < \delta < \sigma < 1$. These conditions guarantee that $s_k^T y_k > 0$, where $y_k = g_{k+1} - g_k$, and s_k is defined in (2).

Different choices for the CG parameters lead to different CG methods. In early CG methods, the conjugate condition is based on the quadratic objective function and the exact line search, which is $d_k^T g_{k+1} = 0$. These methods lead to the classical linear CG methods such as Fletcher-Reeves (FR) [23], Hestenes-Stiefel (HS) [21], Polak-Ribie´re-Polyak (PRP)[9, 13] and Dai-Yuan (DY)[36]. Classic methods have same performance for linear CG methods, although they have different global convergence properties and numerical performance for general nonlinear objective functions or inexact line search (see [32]).

New nonlinear CG methods are presented with different approaches such as constructing descent or sufficient descent directions, new extended conjugacy conditions or a hybrid with QN methods. For example, Zhang et al. [18], construct some descent classic CG directions as three- terms CG, TTCG, methods. For instance in a special case, they proposed a three-term HS, TTHS, with the following search direction [18]:

$$d_{k+1}^{TTHS} = -g_{k+1} + \beta_k^{HS} d_k - \theta_{k+1} y_k, \tag{6}$$

where

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \quad \theta_{k+1} = \frac{g_{k+1}^T d_k}{d_k^T y_k} \tag{7}$$

It is also clear that if the exact line search is used, then $\theta_{k+1} = 0$, and the TTHS method is converted to the classic HS method. By replacing the HS method with other linear CG methods, some new descent methods, such as TTPR and TTFR can be achieved (see [18]). An attractive feature of these methods is that the direction has sufficient descent conditions, i.e. $d_k^T g_k = -\|g_k\|^2$ ($\|.\|$ is the Euclidean norm), which is independent of line search [18]. In addition, Babaie-Kafaki and Ghanbari [28] apply the idea of TTHS method, Eqns (6)-(7), using a modified BFGS, proposed by Li and Fukushima [6], and introduce a modified TTCG, named MTTHS, as follows:

$$d_{k+1}^{MTTHS} = -g_{k+1} + \beta_k^{MHS} d_k - \theta_k^M z_k, \tag{8}$$

where

$$\beta_k^{MHS} = \frac{g_{k+1}^T z_k}{d_k^T z_k}, \quad \theta_k^M = \frac{g_{k+1}^T d_k}{d_k^T z_k}, \tag{9}$$

and

$$z_k = g_{k+1} - g_k + c\|g_k\|^r s_k \triangleq y_k + c\|g_k\|^r s_k, \tag{10}$$

where $r \geq 0$ and $c > 0$ are some constants, in Eqn (10), z_k plays a vital role in the global convergence of the MBFGS method for nonconvex function [17]. Similarly, Sugiki et al. [15] proposed another modified TTCG method, using a TTCG method, proposed by Narushima et al. [37] and a general form of the modified secant conditions, which generate a search direction with sufficient descent conditions.

At first time, Perry [3] to find more efficient CG methods, incorporated the standard secant equation to conjugacy condition and proposed his method to approximate the directions of CG to QN direction, as in the following:

$$d_{k+1}^P = -g_{k+1} + \beta_k^P d_k = -Q_{k+1}^P g_{k+1}, \tag{11}$$

where Q_{k+1}^P is the direction matrix, as a nonsymmetric matrix which approximates the inverse Hessian of the objective function at current iteration, and β_k^P is the Perry CG parameter, which are defined as follows:

$$\beta_k^P = \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T s_k}{d_k^T y_k}, \quad Q_{k+1}^P = I - \frac{s_k y_k^T}{y_k^T s_k} + \frac{s_k s_k^T}{y_k^T s_k} \tag{12}$$

As mentioned, from Wolfe conditions in means (4)-(5), we have $s_k^T y_k > 0$, so the matrix in (12) is well-defined. In Perry approach, the direction matrix, Q_{k+1}^P, is not symmetric and also does not satisfy the secant equations [5]. To overcome these defects, Shanno [5] combined the Perry method and memoryless BFGS method to introduce a new CG direction as follows:

$$d_{k+1}^S = -Q_{k+1}^S g_{k+1} = -(I - \frac{s_k y_k^T + y_k s_k^T}{s_k^T y_k} + (1 + \frac{y_k^T y_k}{s_k^T y_k})\frac{s_k s_k^T}{s_k^T y_k})g_{k+1} \tag{13}$$

In 2001, Dai, and Liao [35] extended the Perry conjugate condition and introduced the new nonlinear CG method as follows:

$$d_{k+1}^{DL} = -g_{k+1} + \beta_k^{DL} d_k = -Q_{k+1}^{DL} g_{k+1}, \tag{14}$$

where

$$\beta_k^{DL} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - t\frac{g_{k+1}^T s_k}{d_k^T y_k}, \quad Q_{k+1}^{DL} = I + \frac{s_k^T y_k}{s_k^T y_k} - t\frac{s_k^T s_k}{s_k^T y_k}, \tag{15}$$

where t is a nonnegative DL parameter. Note that if t =0, then β_k^{DL} reduces β_k^{HS}, Eqn(7), if $t = 1$, then β_k^{DL} reduces to β_k^{P}, Eqn(12).

For extending the global convergence properties of general objective functions, Dai, and Liao [35] considered a truncated form of the DL method, with an extended DL parameter, namely β_k^{DL+}, and the following direction:

$$d_{k+1}^{DL+} = -g_{k+1} + \beta_k^{DL+} d_k = -g_{k+1} + \left(\max\{\frac{g_{k+1}^T y_k}{d_k^T y_k}, 0\} - t\frac{g_{k+1}^T s_k}{d_k^T s_k}\right) d_k \tag{16}$$

As a famous descent CG method, independent from a type of line search, Hager and Zhang (HZ) [31] introduced the following CG parameter:

$$\beta_k^{HZ} = \frac{g_{k+1}^T y_k}{d_k^T y_k} - 2\frac{\|y_k\|^2}{d_k^T y_k}\frac{g_{k+1}^T d_k}{d_k^T y_k} \tag{17}$$

HZ method is an adaptive version of the DL parameter corresponding to $t = 2\frac{\|y_k\|^2}{s_k^T y_k}$ in Eqn (15). Another adaptive DL parameter is based on scaled memoryless BFGS, suggested by Dai and Kou (DK) [33], as follows:

$$\beta_k^{DK}(\tau_k) = \frac{g_{k+1}^T y_k}{d_k^T y_k} - (\tau_k + \frac{\|y_k\|^2}{s_k^T y_k} - \frac{s_k^T y_k}{\|s_k\|^2})\frac{g_{k+1}^T s_k}{d_k^T y_k} \tag{18}$$

In which τ_k is a parameter corresponding to the scaling factor in the scaled memoryless BFGS method.

Although the setting of the DL parameter is an open problem in CG methods [2], many efforts have been made by researchers to adjust it. As instance, in descent approach based on an eigenvalue study, the authors in [25] proposed a descent class of DL method, namely, DDL. An exciting feature of the proposed class is that the HZ and DK methods are individual cases of it, as efficient nonlinear CG methods. The DDL search direction is as follows [25]:

$$d_{k+1}^{DDL} = -g_{k+1} + \beta_k^{p,q} d_k = -(I + \frac{d_k^T y_k}{d_k^T y_k} - t_k^{p,q}\frac{d_k^T s_k}{d_k^T y_k})g_{k+1}, \tag{19}$$

where $t_k^{p,q}$ is DL parameter as follows:

$$t_k^{p,q} = p\frac{\|y_k\|^2}{s_k^T y_k} - q\frac{s_k^T y_k}{\|s_k\|^2}, \tag{20}$$

where p and q are nonnegative constants, which $p < \frac{1}{4}$ and $q \geq \frac{1}{4}$. For more information about setting the DL parameter, see [16, 24, 27, 29, 38].

Another conjugacy approach in CG methods is based on the different types of modified secant equations instead of standard secant equation in DL method. To review the different types of modified secant equations see the Introduction section of [25].

Here, motivated by $DL+$ approach, similar to [17, 13, 10, 11], we apply the modified secant equation proposed by Li et al. [10], named MSL, for a new extended conjugacy condition and then using two approaches, similar to $DL+$, we adjust its parameter. Therefore, the advantages of the new proposed nonlinear CG method are using the second-order information of the objective function, by a modified secant equation, and setting the $DL+$ parameter to improve in the search directions, simultaneously.

The remainder of this paper is organized as follows. In Section 2, we introduce a new extended conjugacy condition based on MSL [10]. Then we discuss two approaches to setting the parameter. In the first approach, we use the MTTHS descent method (8)-(9). In second approach, we try to match the direction matrix of the CG method to the Shanno quasi-Newton direction matrix, Q^S_{k+1}, Eqn (13). Then, we discuss their global convergence. In Section 3, we numerically compare our methods with the DL, HZ, and DK methods and report comparative testing results. Finally, we make conclusions in Section 4.

2 New Nonlinear Conjugate Gradient Methods

In this section, based on MSL [10], we first introduce a new extended, modified conjugate condition for CG methods, and then we describe two methods for calculating the parameter.

2.1 Conjugacy condition based on MSL

Using modifies secant equations are common in CG and QN methods for solving unconstrained optimization problems. For example Zhang et al. [13] and Zhang and Xu [14] proposed new QN methods based on a modified secant equation. Moreover, Yube, and Takano [11] applied this equation for a nonlinear CG with global convergence properties. New versions of this modified secant equation can be seen in [26, 20, 12]. Zhang and Zhou [17] applied a modified BFGS method for a nonlinear CG method, which is proposed by Li and Fukushima [6]. Li et al. [10] used with the modified secant equation in [39, 40]. Suugiki et al. [15], unify the above-modified secant equations as a general form and proposed a TTCG method with sufficient descent property.

As special case, here, we apply the conjugacy condition proposed by Li et al. [10], which further studied by [39, 40]. This condition is based on the modified secant equation, MSL, as follows [10]:

$$B_{k+1}s_k = \overline{y}_k, \quad \overline{y}_k = y_k + A_k u_k, \tag{21}$$

where B_{k+1} is an approximation of the Hessian matrix of the objective function, $u_k \in \mathbb{R}^n$ is a vector that satisfies $s_k^T u_k \neq 0$ and $A_k = \frac{\bar{\theta_k}}{s_k^T u_k}$ where $\bar{\theta_k} = max\{\theta_k, 0\}$ and θ_k is as follows:

$$\theta_k = 2(f_k - f_{k+1}) + (g_k + g_{k+1})^T s_k \tag{22}$$

The modified secant equation in Eqn (21), is based on a revised form of the modified secant equation proposed in [39, 40]. According to (3), similar to DL conjugate condition [35], the new extended conjugacy condition based on (21) is presented as follows:

$$d_{k+1}^T \overline{y}_k = -t^{\overline{DL+}} g_{k+1}^T s_k, \tag{23}$$

which is named $\overline{DL+}$ conjugate condition. Using CG direction in (3) and (23), we have the following CG parameter:

$$\beta_k^{\overline{DL+}} = \frac{g_{k+1}^T \overline{y}_k}{d_k^T \overline{y}_k} - t^{\overline{DL+}} \frac{g_{k+1}^T s_k}{d_k^T \overline{y}_k} \tag{24}$$

For $t^{\overline{DL+}} = 0$, the $\overline{DL+}$ method is converted to the MHS method in Eqn (9). By replacing the (24) in (3) and rearranging the vectors, we have the following new search direction:

$$d_{k+1}^{\overline{DL+}} = -Q_{k+1}^{\overline{DL+}} g_{k+1} = -(I + \frac{s_k^T \overline{y}_k}{s_k^T \overline{y}_k} - t^{\overline{DL+}} \frac{s_k^T s_k}{s_k^T \overline{y}_k}) g_{k+1} \tag{25}$$

Then the associate CG method is called $\overline{DL+}$ and its parameter, $t^{\overline{DL+}}$, is called $\overline{DL+}$ (update) parameter.
Now similar to $DL+$ parameter, the setting of the $\overline{DL+}$ parameter is an vital issue. In following, we use two approaches to set it.

2.2 Setting $\overline{DL+}$ parameter

To set the $\overline{DL+}$ parameter, we apply two approaches. The first is based on the descent direction, and the second is based on the QN approach.

2.2.1 Descent approach

In linear search methods, the descent direction is vital to convergence analysis. Since the $\overline{DL+}$ direction may not satisfy the descent condition, similar to [25] for DL method, here we try to satisfy the descent condition of $\overline{DL+}$ method using the MTTHS direction in (8)-(9), [28]. For this purpose, consider the following subproblem:

$$\min \| d_{k+1}^{\overline{DL+}} - d_{k+1}^{MTTHS} \| \tag{26}$$

Using simple algebraic calculations, we get the $\overline{DL+}$ parameter as following:

$$t_{k_1}^{\overline{DL+}*} = \frac{1}{a_2}(a_1 - a_3 + \frac{a_4}{\|d_k\|^2} d_k^T z_k), \tag{27}$$

where

$$a_1 = \frac{g_{k+1}^T \overline{y}_k}{d_k^T \overline{y}_k}, \qquad a_2 = \frac{g_{k+1}^T s_k}{d_k^T \overline{y}_k}$$
$$a_3 = \frac{g_{k+1}^T z_k}{d_k^T z_k}, \qquad a_4 = \frac{g_{k+1}^T d_k}{d_k^T z_k},$$

where z_k and $\overline{y}_k$ are defined in Eqns (10) and (21), respectively. After some simplification, the Eqn(27) can be written as follows:

$$t_{k_1}^{\overline{DL+}*} = -\frac{\overline{y}_k^T g_{k+1}}{s_k^T g_{k+1}} \tag{28}$$

However, the parameter $t_{k_1}^{\overline{DL+}*}$ should be nonnegative. So, we use the following modified form of this parameter given:

$$t_{k_1}^{\overline{DL+}*} = max\{t_{k_1}^{\overline{DL+}*}, 0\} \tag{29}$$

So, by replacing (29) in (14), we get a new nonlinear $\overline{DL}$ direction as following:

$$d_{k+1}^{NDL-1} = -g_{k+1} + (\frac{g_{k+1}^T y_k}{d_k^T y_k} - t_{k_1}^{\overline{DL+}*} \frac{g_{k+1}^T s_k}{d_k^T y_k})d_k, \tag{30}$$

where $t_{k_1}^{\overline{DL+}*}$ is defined in Eqn (29). The CG method based on the search direction d_{k+1}^{NDL-1}, called ”NDL-1” method.

2.2.2 QN approach

Since QN methods apply the second derivative information in search directions, so they are useful in solving large scale unconstrained optimization problems. Therefore, to access the CG direction matrices to approximate the inverse Hessian matrix, similar to [3] in the QN method, we enhance the efficiency of CG method. For this reason, we approach the matrix direction of the $\overline{DL+}$ method, $Q_{k+1}^{\overline{DL+}}$, to the Shanno quasi-Newton direction matrix, Q_{k+1}^S, Eqn (13). Therefore, Consider the following subproblem:

$$t_{k_2}^{\overline{DL+}*} = argmin\|Q_{k+1}^{\overline{DL+}} - Q_{k+1}^S\|_F, \tag{31}$$

where $\|.\|_F$ is Frobenius norm. Using the property $tr(AA^T) = \|A\|_F^2$ and after some algebraic calculations, we have

$$t_{k_2}^{\overline{DL+}*} = 1 + \frac{\overline{y}_k^T \overline{y}_k}{s_k^T \overline{y}_k} - \frac{s_k^T \overline{y}_k}{\|s_k\|^2} \tag{32}$$

Now, similar to (29), we propose the following $\overline{DL}$ parameter:

$$t_{k_2}^{\overline{DL+}*} = \{t_{k_2}^{\overline{DL+}*}, 0\} \tag{33}$$

So, by replacing (33) in (14), we get another new $\overline{DL+}$ direction as following:

$$d_{k+1}^{NDL-2} = -g_{k+1} + (\frac{g_{k+1}^T \overline{y}_k}{d_k^T \overline{y}_k} - t_{k_2}^{\overline{DL+}*} \frac{g_{k+1}^T s_k}{d_k^T \overline{y}_k})d_k, \tag{34}$$

where $t_{k_2}^{\overline{DL+*}}$ is defined in Eqn (33). The CG method based on d_{k+1}^{NDL-2}, called "NDL-2" method. Now, we discuss the global convergence of the "NDL-1" and "NDL-2" methods. So, we need to make the following underlying assumptions on the objective function, commonly used in the convergence analysis of the CG methods [34].

Assumption (A):

Let the objective function f is strongly convex and ∇f is Lipschitz continuous on the level set

$$S = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\} \tag{35}$$

That is there exists constants $\mu > 0$ and L such that

$$(\nabla f(x) - \nabla f(y))^T(x-y) \geq \mu\|x-y\|^2, \quad \forall x, y \in S \tag{36}$$

and

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x-y\|, \quad \forall x, y \in S \tag{37}$$

From Eqns (36)-(37), there exists a positive constant Γ such that for all $x \in S$; $\|\nabla f(x)\| \leq \Gamma$.

Lemma 1. [30] Let the Assumption (A) holds. Consider any CG method in the form of (2)-(3) in which for all $k \geq 0$, the search direction d_k is a descent direction, and the step length α_k is determined to satisfy the Wolfe conditions, (4)-(5). If

$$\sum_{k\geq 0} \frac{1}{\|d_k\|^2} = \infty \tag{38}$$

then the method converges in the sense that

$$\liminf_{k\to\infty} \|g_k\| = 0 \tag{39}$$

Theorem 1. Let the Assumption (A) holds for the objective function f in (1). Consider a CG method in the form of (2)-(3) with the CG direction defined by (30), "NDL-1" method, in which the step length α_k is computed such that the Wolfe conditions (4)-(5) are satisfied. If the objective function f is uniformly convex on $\mathcal{S}$, then the method converges in the sense that (39) holds.

Proof. For any uniform convex differentiable function f, there exists a positive constant μ such that (see Theorem 1.3.16 of [30])

$$y_k^T s_k \geq \mu\|s_k\|^2 \tag{40}$$

Also similar inequality can be proved by replacing y_k with $\overline{y}_k$. For this purpose we have

$$\overline{y}_k^T s_k = (y_k + \frac{\bar{\theta}_k}{s_k^T u_k} u_k)^T s_k = s_k^T y_k + max\{\theta_k, 0\} \geq s_k^T y_k \geq \mu\|s_k\|^2 \tag{41}$$

Note that, from the second equation of the Wolf conditions, Eqn (4), we have:

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k \tag{42}$$

On the other hand, from (21) we have:

$$\|\overline{yk}_k\| \leq \|y_k\| + \|A_k u_k\| = \|y_k\| + \|w_k\| \tag{43}$$

where $w_k = A_k u_k$ and A_k is defined in (21). Now we show that $\|\overline{y}_k\| \leq L_1\|s_k\|$. For this purpose, first of all, using Taylor expansion of θ_k in Eqn (22), we have:

$$|\theta_k| < M\|s_k\|^2 \tag{44}$$

Then, considering the Eqn(21), we have two cases for w_k, [10]: $w_k = \frac{\theta_k s_k}{\|s_k\|^2}$ or $w_k = \frac{\theta_k y_k}{s_k^T y_k}$, which θ_k is defined in (21). In the first case, from Eqns (37), (43) and (44),we get:

$$\|\overline{y}_k\| \leq \|y_k\| + \frac{|\theta_k|\|s_k\|}{\|s_k\|^2} = (L+M)\|s_k\| = M_1\|s_k\|, \tag{45}$$

where $M_1 = L + M$. In the second case, from (37), (40) and (44) we have:

$$\|\overline{y}_k\| \leq \|y_k\| + \frac{ML\|s_k\|^3}{\mu\|s_k\|^2} \leq L(1+\frac{M}{\mu})\|s_k\| = M_2\|s_k\|, \tag{46}$$

where $M_2 = L(1+\frac{M}{\mu})$. Now, let $L_1 = max\,\{M_1, M_2\}$, then we have:

$$\|\overline{y}_k\| \leq L_1\|s_k\| \tag{47}$$

Next we can show that $\|z_k\| \leq L_2\|s_k\|$, where z_k is defined in (10). From the eqns (10) and (37), we have:

$$\begin{aligned} \|z_k\| = \|y_k + c\|g_k\|^r s_k\| \leq \|y_k\| + c\|g_k\|^r\|s_k\| \leq L\|s_k\| + c\|g_k\|^r\|s_k\| \\ \leq (L + c\Gamma^r)\|S_k\| = L_2\|s_k\|, \end{aligned} \tag{48}$$

where $L_2 = L + c\Gamma^r$. Moreover, from (40) and (10) we have:

$$\begin{aligned} s_k^T z_k = s_k^T(y_k + c\|g_k\|^r s_k) = s_k^T y_k + c\|g_k\|^r\|s_k\|^2 \\ \geq (\mu + c\|g_k\|^r)\|s_k\|^2 \geq \mu\|s_k\|^2, \end{aligned} \tag{49}$$

which implies that $s_k^T z_k \geq \mu\|s_k\|^2$. Hence from this inequality and Eqns (41), (47), (48), (49), (5) and Cauchy-Shwartz inequality we have:

$$\begin{aligned} |t_{k_1}^{\overline{DL*}}| &= \frac{d_k^T\overline{y}_k}{g_{k+1}^T s_k}\left(\frac{g_{k+1}^T\overline{y}_k}{d_k^T\overline{y}_k} + \frac{g_{k+1}^T z_k}{s_k^T z_k} + \frac{g_{k+1}^T d_k}{\|d_k\|^2}\right) \\ &\leq \frac{\|s_k\|\|\overline{y}_k\|}{\sigma\|g_{k+1}\|\|s_k\|}\left(\frac{\|g_{k+1}\|\|\overline{y}_k\|}{\overline{\mu}\|s_k\|^2} + \frac{\|g_{k+1}\|\|z_k\|}{\mu\|s_k\|^2} + \frac{\|g_{k+1}\|\|s_k\|}{\|s_k\|^2}\right) \\ &\leq \frac{L_1}{\sigma}\left(\frac{L_1}{\overline{\mu}} + \frac{L_2}{\mu} + 1\right) \end{aligned} \tag{50}$$

That is $t_{k_1}^{\overline{DL+*}}$ is bounded for uniformly convex objective function. So, if we use the Wolfe conditions, (4)-(5), similar to Theorem (2.1) in [25], the search directions are bounded away, which with Lemma 1 complete the proof. □

Theorem 2. Let Assumption (A) holds for the objective function f in (1). Consider a CG method in the form of (2)-(3) with the CG direction defined by (34), "NDL-2" method, in which the step length α_k is computed such that the Wolfe conditions (4) and (5) are satisfied. If the objective function f is uniformly convex on $\mathcal{S}$, then the method converges in the sense that (39) holds.

Proof. Considering the Assumption (A) and the assumptions of Theorem 1, from eqns (36), (41), (45), (47) and definition of $t_{k_2}^{\overline{DL+}*}$, Eqn (33), we have:

$$|t_{k_2}^{\overline{DL+}*}| = |1 + \frac{\overline{y}_k^T \overline{y}_k}{s_k^T \overline{y}_k} - \frac{s_k^T \overline{y}_k}{\|s_k\|^2}| \leq 1 + \frac{\|\overline{y}_k\|^2}{|s_k^T \overline{y}_k|} + \frac{|s_k^T \overline{y}_k|}{\|s_k\|^2}$$
$$\leq 1 + \frac{L_1^2 \|s_k\|^2}{\|s_k\|^2} + \frac{\|s_k\| \|\overline{y}_k\|}{\|s_k\|^2} = 1 + L_1 + \frac{L_1^2}{\mu} \qquad (51)$$

So, similar to Theorem 1, the search directions are bounded away, and the proof is complete. □

In order to ensure the global convergence of the proposed CG methods, "NDL-1" and "NDL-2" methods, for general functions, we modify the CG parameter in Eqn (24), similar to [35, **?**], as follows:

$$\beta_{k_i}^{\overline{DL}+} = max\{\frac{g_{k+1}^T \overline{y}_k}{d_k^T \overline{y}_k}, 0\} - t_{k_i}^{\overline{DL+}*} \frac{g_{k+1}^T s_k}{d_k^T \overline{y}_k}, \ i = 1, 2 \qquad (52)$$

where $t_{k_i}^{\overline{DL+}*}$, $i = 1, 2$ is defined in (29) and (33), respectively. Theorem 3.6 of [35] ensures the global convergence of the methods, which are named $\overline{DL}+$, for general functions, if the search directions satisfy the sufficient descent condition.

3 Numerical Experiments

In this section, we present some numerical experiments, obtained by applying a MATLAB 8.8.0.1 (R2013a) implementation of the proposed nonlinear CG methods, "NDL-1" and "NDL-2". The numerical results are compared with the $DL+$ [35] with parameter $t = 0.1$ and DK [33] with parameter $\tau_k = \frac{\|y_k\|^2}{s_k^T y_k}$. We perform the implementations on a computer, Intel(R) Core (TM) A10-8700P CPU 3.20 Gigahertz 64-bit desktop with 8 Gigabyte RAM. Our experiments have been done on a set of test problems of unconstrained optimization problems of CUTEr collection [1]. Although the descent property may not always hold for the proposed method, the upward search direction seldom occurred in our experiments; when encountering, we restarted the algorithm with Powell Restart [30], which is $|g_k^T g_{k+1}| < 0.2\|g_{k+1}\|$.

Moreover, we used the active approximate Wolfe conditions described in (4)-(5) with parameters $\sigma = 0.9$ and $\rho = 10^{-4}$. The same stop condition is considered for all methods, which are $\|g_k\|_\infty \leq 10^{-6}$ and the maximum number of iterations is limited to 1000. Table 1, shows our comparing data contains the test problems, dimensions (n), the total number of function evaluations (f_n) and the total number of gradient evaluations (g_n), respectively.

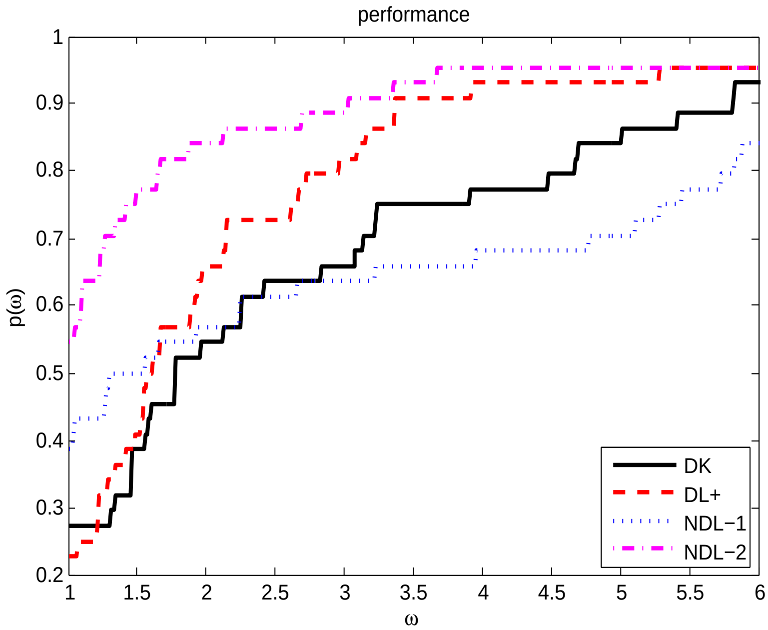

Figure 1: Performance profiles based on the number of iterations for ”NDL-1”, ”NDL-2”, DL+ and DK methods.

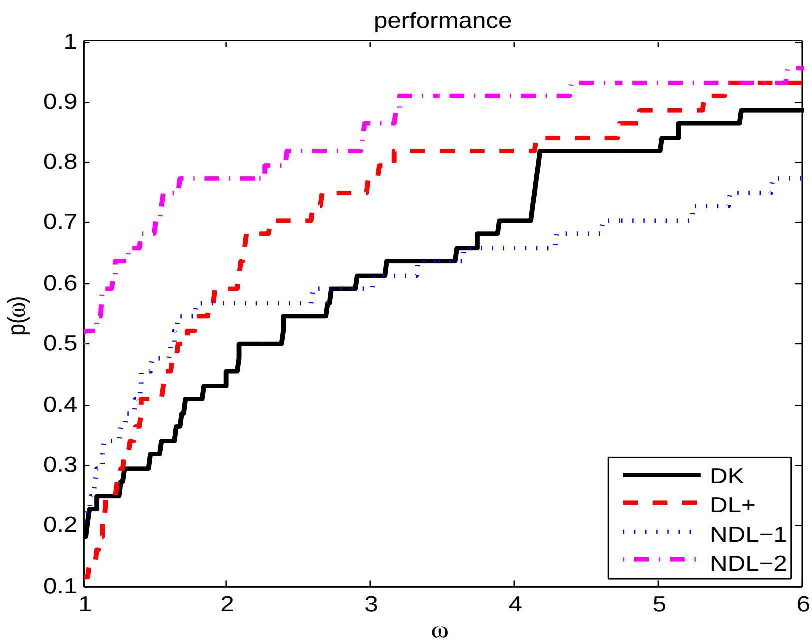

Figure 2: Performance profiles based on CPU time for ”NDL-1”, ”NDL-2”, DL+ and DK methods.

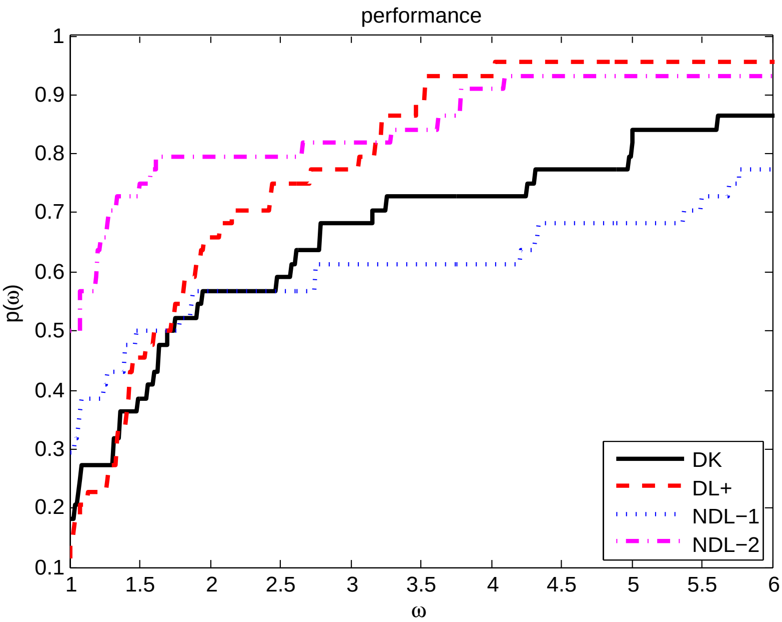

Figure 3: Performance profiles based on $n_f + 3n_g$ for "NDL-1", "NDL-2", DL+ and DK methods.

For more comparison on our numerical results, we apply the performance profile introduced by Dolan and More´ [8].

Table 1: Experiments results of the proposed methods about the total function evaluations (f_n) and gradient evaluations (g_n)

		$DL+$	DK	$NDL-1$	$NDL-2$
$Problem$	n	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$
$AKIVA$	2	2 \ 2	2 \ 2	2 \ 2	2 \ 2
$ALLINITU$	4	451 \ 313	626 \ 408	622 \ 404	421 \ 326
$ARGLINA$	200	17 \ 17	18 \ 18	7 \ 7	11 \ 11
$ARGLINB$	200	45267 \ 2003	33853 \ 766	33853 \ 766	41262 \ 1258
$ARGLINC$	200	76960 \ 3407	153903 \ 3479	153903 \ 3479	35670 \ 1081
$ARWHEAD$	5000	8136 \ 1077	36169 \ 3122	71908 \ 6014	668 \ 7225 \ 669
$BARD$	3	6018 \ 2672	23350 \ 8756	22325 \ 8326	4163 \ 1749
$BDQRTIC$	5000	19057 \ 1820	142535 \ 10001	143643 \ 10001	22755 \ 1717
$BEALE$	2	3490 \ 1413	2753 \ 949	2719 \ 946	990 \ 421
$BIGGS6$	6	4271 \ 3784	1670 \ 1319	7928 \ 7719	530 \ 429
BOX	10000	11453 \ 1123	115065 \ 10001	119587 \ 10001	6361 \ 587
$BOX3$	3	60 \ 59	29 \ 28	54 \ 53	1016 \ 998
$BRKMCC$	2	455 \ 179	3087 \ 965	3103 \ 970	1496 \ 599
$BROWNAL$	200	85271 \ 10001	122892 \ 8870	139672 \ 10001	24583 \ 1892
$BROWNDEN$	4	23116 \ 1816	70244 \ 5015	29466 \ 2129	16501 \ 1286

Continued on next page

Table 1 – *Continued from previous page*

		$DL+$	DK	$NDL-1$	$NDL-2$
$Problem$	n	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$
$BROYDN7D$	5000	34827 \ 7796	49687 \ 10001	48910 \ 10001	26146 \ 6441
$BRYBND$	5000	4336 \ 1063	1845 \ 495	1779 \ 511	2659 \ 702
$CHAINWOO$	4000	21411 \ 2385	54444 \ 6153	91905 \ 10001	16734 \ 1961
$CHNROSNB4$	50	13102 \ 1757	67274 \ 8503	66035 \ 8374	10379 \ 1464
$CHNRSNBM$	50	15446 \ 20934	54485 \ 7030	47455 \ 6218	10953 \ 1571
$CLIFF$	2	32270 \ 10001	10037 \ 10001	13987 \ 10001	37944 \ 10001
$CUBE$	2	3891 \ 739	81088 \ 100014	81353 \ 10001	2893 \ 575
$CURLY10$	10000	104174 \ 10001	108942 \ 10001	108925 \ 10001	100641 \ 10001
$CURLY20$	10000	122180 \ 10001	127249 \ 10001	127199 \ 10001	121138 \ 10001
$CURLY30$	10000	132988 \ 100014	138686 \ 10001	138608 \ 10001	130225 \ 10001
$DECONVU$	63	18656 \ 5170	6881 \ 1898	8575 \ 2421	11110 \ 3554
$DENSCHNA$	2	25 \ 25	33 \ 33	26 \ 26	27 \ 27
$DENSCHNB$	2	16 \ 16	17 \ 17	10 \ 10	12 \ 12
$DENSCHNC$	2	1642 \ 651	2146 \ 1417	3643 \ 1164	867 \ 439
$DENSCHND$	3	2354 \ 283	741 \ 96	100410 \ 3841	3038 \ 288
$DENSCHNE$	3	21 \ 18	21 \ 18	12 \ 9	16 \ 13
$DENSCHNF$	2	6528 \ 1158	2136 \ 375	12319 \ 2203	5669 \ 1006
$DIXMAANC$	3000	20 \ 184	19 \ 17	14 \ 12	15 \ 13
$DIXMAANA$	3000	17 \ 16	18 \ 17	11 \ 10	13 \ 12
$DIXMAANB$	3000	19 \ 18	18 \ 17	11 \ 10	14 \ 13
$DIXMAANC$	3000	20 \ 18	19 \ 17	14 \ 12	15 \ 13
$DIXMAAND$	3000	609 \ 76	21 \ 17	16 \ 12	3717 \ 289
$DIXMAANE$	3000	279 \ 278	1313 \ 1312	1093 \ 1092	143 \ 142
$DIXMAANF$	3000	760 \ 759	625 \ 624	393 \ 392	496 \ 495
$DIXMAANG$	3000	219 \ 217	425 \ 423	276 \ 274	794 \ 792
$DIXMAANH$	3000	9281 \ 863	13364 \ 1332	157256 \ 10001	5123 \ 553
$DIXMAANI$	3000	302 \ 287	117 \ 116	75 \ 74	546 \ 156
$DIXMAANJ$	3000	1191 \ 1188	580 \ 579	367 \ 366	630 \ 629
$DIXMAANK$	3000	322 \ 320	414 \ 412	254 \ 252	626 \ 624
$DIXMAANL$	3000	107360 \ 10001	6648 \ 616	136154 \ 10001	147794 \ 10001
$DIXMAANM$	15	231 \ 231	172 \ 172	810 \ 810	199 \ 199
$DIXMAANN$	15	207 \ 206	1299 \ 1298	754 \ 753	174 \ 173
$DIXMAANO$	15	203 \ 201	1309 \ 1307	747 \ 745	171 \ 169
$DIXMAANP$	15	194 \ 191	1309 \ 1306	742 \ 739	176 \ 173
$DIXON3DQ$	10000	208 \ 208	286 \ 281	1218 \ 1218	1003 \ 1003
$DJTL$	2	13213 \ 606	36052 \ 1461	30600 \ 1246	11650 \ 532
$DQDRTIC$	5000	11776 \ 2069	27788 \ 4556	27788 \ 4556	8663 \ 1591
$DQRTIC$	5000	13757 \ 843	8647 \ 693	5849 \ 398	14899 \ 993
$EDENSCH$	2000	3449 \ 797	3215 \ 1299	3762 \ 1361	3509 \ 1182

Continued on next page

Table 1 – *Continued from previous page*

		$DL+$	DK	$NDL-1$	$NDL-2$
Problem	n	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$
$EG2$	1000	1592 \ 358	8142 \ 866	6019 \ 799	1891 \ 363
$ENGVAL1$	5000	2434 \ 1092	1308 \ 574	641 \ 241	1535 \ 677
$ENGVAL2$	3	9163 \ 1069	104609 \ 10001	104593 \ 10001	6510 \ 807
$ERRINROS$	50	83940 \ 10001	71123 \ 10001	70231 \ 10001	74704 \ 10001
$ERRINRSM$	50	81101 \ 10001	82952 \ 10001	77688 \ 10001	83878 \ 10001
$EXPFIT$	2	4474 \ 1021	10474 \ 1730	11086 \ 1838	2818 \ 633
$EXTROSNB$	1000	76398 \ 10001	19469 \ 2277	49393 \ 5073	71250 \ 10001
$FLETBV3M$	5000	1062 \ 1062	328 \ 328	1149 \ 1149	136 \ 136
$FLETCBV3$	5000	10001 \ 10001	10041 \ 10001	10001 \ 10001	10017 \ 10001
$FLETCHBV$	5000	10001 \ 10001	10004 \ 10001	10001 \ 10001	10001 \ 10001
$FLETCHCR$	1000	74350 \ 9375	83444 \ 10001	83284 \ 10001	75752 \ 10001
$FMINSRF2$	5625	620 \ 620	402 \ 402	3670 \ 3670	399 \ 399
$FMINSURF$	5625	622 \ 622	485 \ 485	3678 \ 3678	443 \ 443
$FREUROTH$	5000	16915 \ 1829	77020 \ 9110	97310 \ 10001	11725 \ 1307
$GENHUMPS$	5000	70178 \ 10001	73098 \ 10001	71091 \ 10001	67735 \ 10001
$GENROSE$	500	26425 \ 3337	84050 \ 10001	83962 \ 10001	32775 \ 4311
$GULF$	3	36993 \ 10001	57851 \ 10001	59339 \ 10001	55008 \ 10001
$HAIRY$	2	12791 \ 1469	8686 \ 971	9463 \ 1074	10611 \ 1287
$HATFLDD$	3	15526 \ 10001	14806 \ 10001	11764 \ 7673	23089 \ 10001
$HATFLDE$	3	130 \ 124	1469 \ 819	1800 \ 1206	33097 \ 10001
$HATFLDFL$	3	769 \ 290	1460 \ 487	1370 \ 457	716 \ 180
$HEART6LS$	6	133006 \ 9038	113591 \ 10001	118906 \ 10001	114234 \ 7332
$HEART8LS$	8	10175 \ 1443	84531 \ 10001	82921 \ 10001	12482 \ 1764
$HELIX$	3	8600 \ 1268	14178 \ 2131	11806 \ 1781	7906 \ 1207
$HIELOW$	3	2 \ 2	2 \ 2	2 \ 2	2 \ 2
$HILBERTA$	2	159 \ 159	315 \ 315	181 \ 181	67 \ 67
$HILBERTB$	10	126 \ 110	291 \ 249	291 \ 249	289 \ 263
$HIMMELBB$	2	41 \ 27	107 \ 93	74 \ 60	32 \ 18
$JENSMP$	2	70 \ 10	117 \ 13	13111 \ 857	4212 \ 497
$KOWOSB$	4	81 \ 81	99 \ 80	141 \ 141	27 \ 25
$LIARWHD$	5000	4561 \ 573	126596 \ 10001	129301 \ 10001	31873 \ 2659
$LOGHAIRY$	2	331 \ 331	1748 \ 1748	10001 \ 10001	10001 \ 10001
$MANCINO$	100	35018 \ 1796	196345 \ 10001	196351 \ 10001	35270 \ 1811
$MARATOSB$	2	4201 \ 375	11849 \ 579	12329 \ 602	9751 \ 807
$MEXHAT$	2	18659 \ 1080	62163 \ 2941	77494 \ 3640	12668 \ 738
$NONCVXU2$	5000	18932 \ 10001	27496 \ 10001	27503 \ 10001	15803 \ 10001
$NONCVXUN$	5000	20774 \ 10001	32200 \ 10001	32296 \ 10001	19306 \ 10001
$NONDQUAR$	5000	1341 \ 309	1556 \ 692	2318 \ 1049	1556 \ 378
$OSBORNEA$	5	408 \ 45	135339 \ 10001	133752 \ 10001	24 \ 4

Continued on next page

Table 1 – *Continued from previous page*

		$DL+$	DK	$NDL-1$	$NDL-2$
$Problem$	n	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$	$f_n \setminus g_n$
$PALMER1C$	8	209967 \ 10001	259169 \ 10001	259585 \ 10001	240573 \ 10001
$PALMER2C$	8	173444 \ 10001	223136 \ 10001	223115 \ 10001	202956 \ 10001
$PALMER3C$	8	156749 \ 10001	205456 \ 10001	205364 \ 10001	185655 \ 10001
$PALMER4C$	8	156752 \ 10001	205456 \ 10001	205364 \ 10001	185582 \ 10001
$PALMER5C$	6	2687 \ 1250	1079 \ 450	1091 \ 455	1656 \ 859
$PALMER6C$	8	126013 \ 10001	169991 \ 10001	170001 \ 10001	152097 \ 10001
$PALMER8C$	8	128521 \ 10001	173751 \ 10001	173697 \ 10001	154526 \ 10001
$HIMMELBG$	2	10 \ 10	13 \ 7	7 \ 7	17 \ 12
$HIMMELBH$	2	16 \ 16	16 \ 16	11 \ 11	23 \ 23
$POWELLSG$	5000	4253 \ 1084	39270 \ 7596	34586 \ 6692	3560 \ 870
$POWER$	10000	35680 \ 1839	71636 \ 7038	104337 \ 10001	30141 \ 1621
$QUARTC$	5000	13757 \ 843	8647 \ 693	5849 \ 398	14899 \ 993
$ROSENBR$	2	4068 \ 875	83000 \ 10001	83733 \ 10001	2244 \ 491

Figure 1, to the number of iteration, and Figure 2, to the running time, shows that the "NDL-2" method slightly outperforms the "NDL-1", $DL+$ and the DK methods. In addition, Figure 3 shows that to the $n_f + 3n_g$, the "NDL-2" method is competitive with the $DL+$ method.

4 Conclusion

Here, using DL approach, we provide a new conjugacy condition by a modified secant equation proposed in [10]. To set the parameter of the new conjugacy condition, $\overline{DL+}$ parameter, two approaches are used. The convergence analysis is presented for uniformly convex and general nonlinear functions. The comparison of the new nonlinear CG methods with some well-known methods, shows that "NDL-2" method is better in the iteration criteria and in CPU time, although to the $n_f + 3n_g$ is comparative with $DL+$ method.

References

[1] Andrei N. (2008). "An unconstrained optimization test functions collection", Adv. Model. Optim, 10, 147-161.

[2] Andrei N. (2011). "Open problem in conjugate gradient algorithms for uncostraind optimization", Bull Malays Math Sci Soc, 34, 319-330.

[3] Perry A. (1978). "A modified conjugate gradient algorithm", Operations Research, 26, 1073-1078.

[4] Polyak B. T. (1969). "The conjugate gradient method in extremal problems", USSR Computational Mathematics and Mathematical Physics, 9, 94-112.

[5] Shanno D. F. (1978). "Conjugate gradient methods with inexact searches", Mathematics of operations research, 3, 244-256.

[6] Li D. H., Fukushima M. (2001). "A modified BFGS method and its global convergence in nonconvex minimization", Journal of Computational and Applied Mathematics, 129, 15-35.

[7] Li D. H., Fukushima M. (2001). "On the global convergence of the BFGS method for nonconvex unconstrained optimization problems", SIAM Journal on Optimization, 11, 1054-1064.

[8] Dolan E. D., Moré J. J. (2002). "Benchmarking optimization software with performance profiles", Mathematical programming, 91, 201-213.

[9] Polak E., Ribiere G. (1969). "Note sur la convergence de méthodes de directions conjuguées", ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 3, 35-43.

[10] Li G., Tang C., Wei Z. (2007). "New conjugacy condition and related new conjugate gradient methods for unconstrained optimization", Journal of Computational and Applied Mathematics, 202, 523-539.

[11] Yabe H., Takano M. (2004). "Global convergence properties of nonlinear conjugate gradient methods with modified secant condition", Computational Optimization and Applications, 28, 203-225.

[12] Livieris I. E., Pintelas P. (2013). "A new class of spectral conjugate gradient methods based on a modified secant equation for unconstrained optimization", Journal of Computational and Applied Mathematics, 239, 396-405.

[13] Zhang J. Z., Deng N. Y., Chen L. H. (1999). "New quasi-Newton equation and related methods for unconstrained optimization", Journal of Optimization Theory and Applications, 102, 147-167.

[14] Zhang J., Xu, C. (2001). "Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations", Journal of Computational and Applied Mathematics, 137, 269-278.

[15] Sugiki K., Narushima Y., Yabe H. (2012). "Globally convergent three-term conjugate gradient methods that use secant conditions and generate descent search directions for unconstrained optimization", Journal of Optimization Theory and Applications, 153, 733-757.

[16] Zhang K., Liu H., Liu Z. (2019). "A New Dai-Liao Conjugate Gradient Method with Optimal Parameter Choice", Numerical Functional Analysis and Optimization, 40, 194-215.

[17] Zhou W., Zhang L. (2006). "A nonlinear conjugate gradient method based on the MBFGS secant condition", Optimisation Methods and Software, 21, 707-714.

[18] Zhang L., Zhou W., Li D. (2007). ”Some descent three-term conjugate gradient methods and their global convergence”, Optimisation Methods and Software, 22, 697-711.

[19] Powell M. J. (1984). ”Nonconvex minimization calculations and the conjugate gradient method”. In Numerical analysis, Springer, Berlin, Heidelberg, 122-141.

[20] Peyghami M. R., Ahmadzadeh H., Fazli A. (2015). ”A new class of efficient and globally convergent conjugate gradient methods in the Dai–Liao family”, Optimization Methods and Software, 30, 843-863.

[21] Hestenes M. R., Stiefel E. (1952). ”Methods of conjugate gradients for solving linear systems”, Journal of research of the National Bureau of Standards, 49, 409-436.

[22] Wolfe P. (1969). ”Convergence conditions for ascent methods”, SIAM review, 11, 226-235.

[23] Fletcher R., Reeves C. M. (1964). ”Function minimization by conjugate gradients”. The computer journal, 7, 149-154.

[24] Babaie-Kafaki S. (2016). ”On optimality of two adaptive choices for the parameter of Dai–Liao method”, Optimization Letters, 10, 1789-1797.

[25] Babaie-Kafaki S., Ghanbari R. (2014). ”A descent family of Dai–Liao conjugate gradient methods”, Optimization Methods and Software, 29, 583-591.

[26] Babaie-Kafaki S., Ghanbari R., Mahdavi-Amiri N. (2010). ”Two new conjugate gradient methods based on modified secant equations”, Journal of computational and applied Mathematics, 234, 1374-1386. 1374-1386.

[27] Babaie-Kafaki S., Ghanbari R. (2014). ”The Dai–Liao nonlinear conjugate gradient method with optimal parameter choices”, European Journal of Operational Research, 234, 625-630.

[28] Babaie-Kafaki S., Ghanbari R. (2014). ”Two modified three-term conjugate gradient methods with sufficient descent property”, Optimization Letters, 8, 2285-2297.

[29] Babaie-Kafaki S., Ghanbari R. (2015). ”Two optimal Dai–Liao conjugate gradient methods”, Optimization, 64, 2277-2287.

[30] Sun W., Yuan Y. X. (2006). ”Optimization theory and methods: nonlinear programming”, Springer Science & Business Media.

[31] Hager W. W., Zhang H. (2005). ”A new conjugate gradient method with guaranteed descent and an efficient line search”, SIAM Journal on optimization, 16, 170-192.

[32] Hager W. W., Zhang H. (2006). ”A survey of nonlinear conjugate gradient methods”, Pacific journal of Optimization, 2, 35-58.

[33] Dai Y. H., Kou C. X. (2013). ”A nonlinear conjugate gradient algorithm with an optimal property and an improved Wolfe line search”, SIAM Journal on Optimization, 23, 296-320.

[34] Dai Y., Han J., Liu G., Sun D., Yin H., Yuan Y. X. (2000). ”Convergence properties of nonlinear conjugate gradient methods”, SIAM Journal on Optimization, 10, 345-358.

[35] Dai Y. H., Liao L. Z. (2001). ”New conjugacy conditions and related nonlinear conjugate gradient methods”, Applied Mathematics and Optimization, 43, 87-101.

[36] Dai Y. H., Yuan Y. (1999). ”A nonlinear conjugate gradient method with a strong global convergence property”, SIAM Journal on optimization, 10, 177-182.

[37] Narushima Y., Yabe H., Ford J. A. (2011). ”A three-term conjugate gradient method with sufficient descent property for unconstrained optimization”, SIAM Journal on Optimization, 21, 212-230.

[38] Aminifard Z., Babaie-Kafaki S. (2019). ”An optimal parameter choice for the Dai–Liao family of conjugate gradient methods by avoiding a direction of the maximum magnification by the search direction matrix”. 4OR, 17, 317-330.

[39] Wei Z., Li G., Qi L. (2006). ”New quasi-Newton methods for unconstrained optimization problems”, Applied Mathematics and Computation, 175, 1156-1188.

[40] Wei Z., Yu G., Yuan G., Lian Z. (2004). ”The superlinear convergence of a modified BFGS-type method for unconstrained optimization”, Computational optimization and applications, 29, 315-332.

دو روش تنظیم پارامتر دای-لیاو بر اساس معادلات سکانت اصلاح شده

نژادحسین، س.
استادیار، گروه ریاضی کاربردی،
ایران، تهران، دانشگاه پیام‌نور، گروه ریاضی، صندوق پستی ۱۹۳۹-۵۳۶۹۷
s_nejhadhosein@pnu.ac.ir

محمدخان سرتیپ، س.
کارشناس ارشد، گروه ریاضی کاربردی-نویسنده مسئول،
ایران، تهران، دانشگاه پیام‌نور، گروه ریاضی، صندوق پستی ۱۹۳۹-۵۳۶۹۷
saharsartip@gmail.com

تاریخ دریافت: ۲۷ خرداد ۱۳۹۸ تاریخ پذیرش: ۲۹ بهمن ۱۳۹۸

چکیده

در ادامه کار تنظیم پارامتر دای-لیاو در روش‌های گرادیان مزدوج، دو پارامتر جدید براساس معادلات سکانت اصلاح‌شده معرفی شده توسط لی و فوکوشیما، با دو رویکرد متفاوت که از یک شرط مزدوجی جدید استفاده می‌کند، ارایه کرده‌ایم. اولین پارامتر براساس روش ارایه شده توسط ژنگ و همکارانش به عنوان یک روش گرادیان مزدوج هستینس-استیفل است. دومین پارامتر براساس رویکرد شبه نیوتن است. همگرایی سراسری روش‌های پیشنهادی برای توابع محدب یکنواخت و توابع عمومی ثابت شده است. نتایج عددی با استفاده از مجموعه‌ای از مسایل CUTEr و مقایسه روش‌های پیشنهادی با تعدادی از روش‌های مشهور، به دست آمده است.

کلمات کلیدی

بهینه‌سازی نامقید، معادلات مرزی اصلاح‌شده، روش گرادیان مزدوج دای-لیاو.