Journal of Multidisciplinary Modeling and Optimization 3(1) (2020), 35-46.

A New Three-Term Conjugate Gradient Algorithm Based on the Dai-Liao and the Liu-Xu Conjugate Gradient Methods

Khalil K. ABBO[1] and Aynur J. NAMIK[2]

[1]Department of Mathematics,
Telafer University, Mosul, Iraq
kh_196538@yahoo.com

[2]Department of Mathematics
Mosul University, Mosul, Iraq
aynoor.Javdet@gmail.com

Received: 01.07.2020, Accepted: 31.08.2020, Published: 17.09.2020

Abstract — *Based on the Dai-Laio and Liu-Xu methods, we develop a new three-term conjugate gradient method for solving large-scale unconstrained optimization problem,. The suggested method satisfies both the descent condition and the conjugacy condition. For uniformly convex function, under standard assumption the global convergence of the algorithm is proved. Finally, some numerical results of the proposed method are given.*

Keywords: Unconstrained optimization, Descent methods, Conjugate Gradient methods.
Mathematics Subject Classification: 65K10, 90C26.

1 Introduction

We deal with the following unconstrained optimization problem:

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

where $f : R^n \to R$ is continuously differentiable and it‘s gradient $g = \nabla f$ is available. there are many different theories and algorithms that have been presented to solve problem (1), (see [1-4]). For solving problem (1), the iterative method is widely used and it‘s form is given by

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

where $x_k \in R^n$ is the $k-$th approximation to a solution of (1). $\alpha_k \in R$ is a step-length usually chosen to satisfy certain line search conditions [18]. and $d_k \in R^n$ is the search direction and defined by

Cite as: *K. K. Abbo and A. J. Namik, A new three-term conjugate gradient algorithm based on the Dai-Liao and the Liu-Xu conjugate gradient methods, Journal of Multidisciplinary Modeling and Optimization 3(1) (2020), 35-46.*

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad , \quad d_1 = -g_1$$

where $\beta_k \in R$ is a parameter which characterizes the conjugate gradient method. For general nonlinear functions, different choices of β_k lead to different conjugate gradient methods. Well-known formulas for β_k are called the Fletcher-Reeves (FR) [10], Hestenes -Stiefel (HS) [11], and Polak-Ribiere (PR) [15] are given by

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \qquad \beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \qquad \beta_k^{PR} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}$$

where $y_k = g_{k+1} - g_k$ and $\|.\|$denotes to $\ell_2$norm.

The line search in conjugate gradient algorithms is often based on the standared Wolfe Conditions (WC) [19]:

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

$$g_{k+1}^T d_k \ge \sigma g_k^T d_k , \tag{5}$$

Where d_k is a descent direction and $0 < \rho \le \sigma < 1$. However, for some conjugate gradient algorithms, a stronger version of the Wolfe line search conditions (SWC) given by (4) and

$$\left| g_{k+1}^T d_k \right| \le -\sigma g_k^T d_k , \tag{6}$$

is needed to ensure the convergence and to enhance the stability.

The form represents the pure conjugacy condition

$$d_{k+1}^T y_k = 0 \tag{7}$$

for nonlinear conjugate gradient methods. The extension of the conjugacy condition was studied by Perry [14]. He tried to accelerate the conjugate gradient method by incorporating the second-order information into it. Specifically, he used the secant condition

$$H_{k+1} y_k = s_k \tag{8}$$

of quasi-Newton methods, where a symmetric matrix H_{k+1} is an approximation to the inverse Hessian. For quasi-Newton methods, the search direction d_{k+1} can be calculated in the form

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

By (8) and (9), the relation

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

holds. By taking this relation into account, Perry replaced the conjugacy condition (7) by the condition

$$d_{k+1}^T y_k = -g_{k+1}^T s_k. \tag{10}$$

Dai and Liao [8] generalized the condition (10) to the following

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k, \tag{11}$$

where $t \geq 0$ is a scalar. The case $t = 0$, (11) reduces to the usual conjugacy condition (7). On the other hand, the case $t = 1$, (11) becomes Perry‘s condition (10). To ensure that the search direction d_k satisfies condition (11), by substituting $d_{k+1} = -g_{k+1} + \beta_k d_k$ into (11), they had

$$-g_{k+1}^T y_k + \beta_{k+1} d_k^T y_k = -t g_{k+1}^T s_k.$$

This gives the Dai-Liao formula

$$\beta_k^{DL} = \frac{g_{k+1}^T (y_k - t s_k)}{d_k^T y_k}. \tag{12}$$

We note that the case $t = 1$ reduces to the Perry formula

$$\beta_k^{P} = \frac{g_{k+1}^T (y_k - s_k)}{d_k^T y_k}. \tag{13}$$

Furthermore, if $t = 0$, then β^{DL} reduces to the β^{HS}. The approach of Dai and Liao (DL) has been paid special attention to by many researches. In several efforts, modified secant equations have been applied to make modifications on the DL method. It is remarkable that numerical performance of the DL method is very dependent on the parameter t for which there is no any optimal choice [6].

This paper is organized as follows. In section 2 we briefly review the Three-term conjugate gradient methods. In section 3, the proposed algorithm is stated. The properties and convergent results of the new method are given in Section 4. Numerical results and one conclusion are presented in Section 5 and in Section 6, respectively.

2 Three-Term Conjugate Gradient (CG) methods

Recently many researchers have been studied three- term conjugate gradient methods. For example Narushima, Yab and Ford [13] have proposed a wider class of three term conjugate gradient methods (called 3TCG) which always satisfy the sufficient descent condition. Shanno in [17] used the well-known BFGS quasi-Newton method to obtain the following three-term CG method.

$$d_{k+1} = -g_{k+1} + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \left(1 + \frac{\|y_k\|^2}{s_k^T y_k} \right) \frac{g_{k+1}^T s_k}{s_k^T y_k} \right] s_k + \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k \tag{14}$$

Furthermore, Liu and Xu in [12] was generalized the Perry conjugate gradient algorithm (13), the search directions were formulated as follows

$$d_{k+1}^{LX} = -g_{k+1} + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \left(\tau_k + \frac{\|y_k\|^2}{s_k^T y_k} \right) \frac{g_{k+1}^T s_k}{s_k^T y_k} \right] s_k + \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k \tag{15}$$

where τ_k is parameter, which is Liu-Xu three-term conjugate gradient methods. When $\tau_k s_k^T y_k > 0$, the search directions defined by (15) satisfy the descent property

$$d_{k+1}^T g_{k+1} < 0$$

Or the sufficient descent property

$$d_{k+1}^T g_{k+1} \le -c_0 \|g_{k+1}\|^2, \qquad c_0 > 0 \tag{16}$$

Notice that if $\tau_k = 1$, then (15) reduces to the (14). It is remarkable that there is no any optimal choice for τ_k, However different values used for τ_k in [3], for example

$$\tau_k = 1, \quad \tau_k = c_1 \frac{y_k^T y_k}{s_k^T y_k}, \ldots$$

3 A New Three-Term Conjugate Gradient (CG) Method

The aim of this section is to derive a new three-term conjugate gradient method Aynur and Khalil (AK3 say) by using Liu-Xu (LX) method (15) and Dai - Liao (DL) CG method (3) and (12). consider the search direction given by DL

$$d_{k+1}^{DL} = -g_{k+1} + \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k - t \frac{s_k^T g_{k+1}}{s_k^T y_k} s_k \quad , \tag{17}$$

Letting $t = \frac{s_k^T y_k}{\|y_k\|^2}$ in equation (17) we get

$$d_{k+1} = -g_{k+1} + \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k - \frac{s_k^T g_{k+1}}{\|y_k\|^2} s_k \tag{18}$$

Now equating the equations (15) and (18) i.e

$$d_{k+1} = d_{k+1}^{LX} \ .$$

With simple algebra and with the change signal of the last term in d_{k+1}^{LX} we get

$$\tau_k = \frac{s_k^T y_k}{\|y_k\|^2} - \frac{\|y_k\|^2}{s_k^T y_k} \tag{19}$$

Substitute (19) in the equation (15) to obtain the new search direction

$$d_{k+1}^{AK3} = -g_{k+1} + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{\|y_k\|^2} \right] s_k - \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k \tag{20}$$

Note that, if line search is exact i.e $g_{k+1}^T s_k = 0$ then the search direction d^{AK3} reduces to the well-known Hestenes and Stiefel β^{HS}, furthermore if $g_{k+1}^T s_k = 0$ and successive gradients are orthogonal i.e $g_{k+1}^T g_k = 0$ then d^{AK3} reduces to the CD-Fletcher method defined by $\beta_k^{CD} = -\frac{g_{k+1}^T g_{k+1}}{s_k^T g_k}$.

In the following we summarize the our AK3 algorithm.

Algorithm (AK3)

Step (1): Select a starting point $x_1 \in dom\ f$ and $\varepsilon > 0$, compute $f_1 = f(x_1)$ and $g_1 = \nabla f(x_1)$. Select some positive values for ρ and σ. Set $d_1 = -g_1$ and $k = 1$.

Step (2): Test for convergence . If $\|g_k\|_\infty \le \varepsilon$, then stop x_k is optimal ; otherwise go to Step (3).

Step (3): Determine the step length α_k, by using the Wolfe line search conditions (4)-(5).

Step (4): Update the variables as: $x_{k+1} = x_k + \alpha_k d_k$. Compute f_{k+1}, g_{k+1}, $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$.

Step (5): Compute the search direction as: If $y_k^T s_k \ne 0$ then $d_{k+1} = d_{k+1}^{AK3}$ else $d_{k+1} = -g_{k+1}$.

Step (6): Set $k = k+1$ and go to Step 2.

4 Convergence Analysis

In this section. We investigate the global convergence property of the algorithm (AK3). For this purpose we make the following Assumptions:

1. The level set $S = \{x \in R^n : f(x) \le f(x_0)\}$ is bounded, i.e. there exists positive constant $B > 0$ such that, for all $x \in S$, $\|x\| \le B$.

2. In a neighborhood N of S the function f is continuously differentiable and its gradient is Lipschitz continuous, i.e. there exists a constant $L > 0$ such that $\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$, for all $x, y \in N$.

Under these assumptions on f, there exists a constant $\Gamma \ge 0$ such that $\|\nabla f(x)\| \le \Gamma$, for all $x \in S$. Observe that the assumption that the function f is bounded below is weaker than the usual assumption that the level set is bounded. Although the search directions generated by (20) are always descent directions, to ensure convergence of the algorithm we need to constrain the choice of the step length α_k. The following proposition shows that the Wolfe line search always gives a lower bound for the step length α_k. Based on the above assumptions we shall show that our method satisfies the conjugacy condition, the sufficient descent condition, and globally convergent with Wolfe line search conditions.

Theorem 1. Suppose that the step-size α_k satisfies the standard Wolfe conditions, consider the search directions d_k generated from (20) then the search directions d_{k+1} are conjugate for all k that is .

$$d_{k+1}^T y_k = -c_0 g_{k+1}^T s_k$$

Where c_0 positive constant.

Proof:

$$\begin{aligned} y_k^T d_{k+1}^{AK3} &= -y_k^T g_{k+1} + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{\|y_k\|^2}\right] y_k^T s_k - \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k^T y_k \\ &= -y_k^T g_{k+1} + y_k^T g_{k+1} - \frac{s_k^T y_k}{\|y_k\|^2} g_{k+1}^T s_k - \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k^T y_k \\ &= -\left(\frac{s_k^T y_k}{\|y_k\|^2} + \frac{\|y_k\|^2}{s_k^T y_k}\right) g_{k+1}^T s_k \,. \end{aligned}$$

By Wolfe condition $s_k^T y_k > 0$ we have

$$\therefore \frac{s_k^T y_k}{\|y_k\|^2} + \frac{\|y_k\|^2}{s_k^T y_k} = c_0 > 0 \,.$$

Therefore $d_{k+1}^T y_k = -c_0 g_{k+1}^T s_k$.

Theorem 2. Suppose that the step-size α_k satisfies the standard Wolfe conditions (WC), consider the search directions d_k generated from (20) then the search directions d_{k+1} satisfies the sufficient descent condition $d_k^T g_k \le -c\|g_k\|^2$, for all k.

Proof: The proof is by induction.

If $k = 1 \Rightarrow d_1 = -g_1$, $\therefore d_1^T g_1 = -\|g_1\|^2$

know let $s_k^T g_k < -c\|g_k\|$ to proof for $k+1$, multiply (20) by g_{k+1}^T to get

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{\|y_k\|^2}\right] s_k^T g_{k+1} - \frac{g_{k+1}^T s_k}{s_k^T y_k} y_k^T g_{k+1}$$

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

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

By Couchy-Shwartiz inequality and Lipschitz condition we get

$$\frac{\left(g_{k+1}^T s_k\right)^2}{\|g_{k+1}\|^2 \ \|y_k\|^2} \le \frac{\|g_{k+1}\|^2 \ \|s_k\|^2}{\|g_{k+1}\|^2 \ \|y_k\|^2} = \frac{\|s_k\|^2}{\|y_k\|^2} \le \frac{\|s_k\|^2}{L^2\|s_k\|^2} \le \frac{1}{L^2}$$

Therefore

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

where

$$c = 1 + \frac{1}{L^2} > 0 .$$

Proposition 1 ([15,16]). Suppose that d_k is a descent direction and that the gradient ∇f satisfies the Lipschitz condition for all x on the line segment connecting x_k and x_{k+1} , If the line search satisfies the Wolfe conditions (4) and (5), then

$$\alpha_k \ge \frac{(1-\sigma)\left|g_k^T d_k\right|}{L\|d_k\|^2}. \tag{21}$$

Proposition 2 ([12]). Suppose that assumptions (1) and (2) hold. Consider the algorithm (2) and (20), where d_k is a descent direction and α_k is computed by the general Wolfe line search (4) and (5). Then

$$\sum_{k=0}^{\infty} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty \ . \tag{22}$$

Proposition 3 ([16]). Suppose that assumptions (1) and (2) hold, and consider any conjugate gradient algorithm (2), where d_k is a descent direction and α_k is obtained by the Strong Wolfe line search (4) and (6).If

$$\sum_{k\ge1} \frac{1}{\|d_k\|^2} = \infty , \tag{23}$$

Then $\liminf_{k\to\infty} \|g_k\| = 0$.

For uniformly convex functions, we can prove that the norm of the direction d_{k+1} generated by (20) is bounded above. Therefore, by Proposition 3, we can prove the following result.

Theorem 3. Suppose that assumptions (1) and (2) hold, and consider the algorithm (2) and (20), where d_k is a descent direction and α_k is computed by the strong Wolfe line search (4) and (6). Suppose that f is a uniformly convex function on S, i.e. there exists a constant $\mu > 0$ such that $(\nabla f(x) - \nabla f(y))^T (x-y) \geq \mu \|x-y\|^2$ for all $x, y \in N$; then $\lim\inf_{k\to\infty} \|g_k\| = 0$.

Proof: The proof is obtained by Contradiction.

$$|\beta_k| = \left| \frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{g_{k+1}^T s_k}{\|y_k\|^2} \right| \leq \frac{|g_{k+1}^T y_k|}{|s_k^T y_k|} + \frac{|g_{k+1}^T s_k|}{\|y_k\|^2}$$

Since f is uniformly convex then $s_k^T y_k \geq \mu \|s_k\|^2$ where $\mu > 0$.

$$\therefore |\beta_k| \leq \frac{\|g_{k+1}\| \ \|y_k\|}{\mu \|s_k\|^2} + \frac{\|s_k\| \ \|g_{k+1}\|}{\|y_k\|^2}$$

By assumption (2) and Lipschitz continuity, we have $\|y_k\| \leq L\|s_k\|$ .we get

$$|\beta_k| \leq \frac{\Gamma L}{\mu \|s_k\|} + \frac{\Gamma}{L^2 \|s_k\|} = \Gamma\left(\frac{L}{\mu} + \frac{1}{L^2}\right)\frac{1}{\|s_k\|}$$

$$|\eta_k| = \left|\frac{g_{k+1}^T s_k}{s_k^T y_k}\right| = \frac{|g_{k+1}^T s_k|}{|s_k^T y_k|} \leq \frac{\|g_{k+1}\| \ \|s_k\|}{\mu \|s_k\|^2} \leq \frac{\Gamma}{\mu \|s_k\|}$$

$$\begin{aligned}
\therefore \|d_{k+1}\| &\leq \|g_{k+1}\| + |\beta_k| \ \|s_k\| + |\eta_k| \ \|y_k\| \\
&\leq \Gamma + \Gamma\left(\frac{L}{\mu} + \frac{1}{L^2}\right)\frac{1}{\|s_k\|} \ \|s_k\| + \left(\frac{\Gamma}{\mu\|s_k\|}\right) \ L\|s_k\| \\
&\leq \Gamma + \Gamma\left(\frac{L}{\mu} + \frac{1}{L^2}\right) + \frac{\Gamma L}{\mu} \\
&\leq \Gamma\left(1 + \frac{2L}{\mu} + \frac{1}{L^2}\right)
\end{aligned}$$

$$\|d_{k+1}\| \leq \Gamma b$$

where $b = \left(1 + \frac{2L}{\mu} + \frac{1}{L^2}\right)$

$$\therefore \frac{1}{\left\|d_{k+1}\right\|} \geq \frac{1}{\Gamma b}$$

Taking the sum for both sides and considering $\left\|d_1\right\| = \left\|g_1\right\|^2 \geq \Gamma$

$$\sum_{k=0}^{\infty} \frac{1}{\left\|d_{k+1}\right\|^2} = \Gamma + \sum_{k=0}^{\infty} \frac{1}{\Gamma b} = \Gamma + \frac{1}{\Gamma b} \sum_{k=0}^{\infty} 1 = \infty$$

Contradiction we have $\lim \inf_{k \to \infty} \left\|g_k\right\| = 0$.

5 Numerical Results and Comparison

In this section, we report some numerical results obtained with an implementation of the AK3 algorithm. The code of the AK3 Algorithm is written in Fortran and compiled with f77 (default compiler settings), taken from N. Andrei web page. We selected 71 Large-scale unconstrained optimization test functions in the generalized or extended form presented in [1]. For each test function, we undertook ten numerical experiments with the number of variables increasing as n=100, 200,…, 1000. The algorithm implements the Wolfe line search conditions with $\rho = 0.0001, \ \sigma = 0.9$ and the same stopping criterion $\left\|g_k\right\|_2 \leq 10^{-6}$. In all algorithms we considered in this numerical study the maximum number of iterations is limited to 1000. The comparisons of algorithms are given in the following context. Let f_i^{ALG1} and f_i^{ALG2} be the optimal values found by ALG1 and ALG2, for problem i = 1. . . 710, respectively. We say that, in the particular problem i, the performance of ALG1 was better than the performance of ALG2 if

$$\left|f_i^{ALG1} - f_i^{ALG2}\right| < 10^{-3}$$

and the number of iterations (iter), or the number of function-gradient evaluations (fg) or the CPU time of ALG1 was less than the number of iterations, or the number of function-gradient evaluations, or the CPU time corresponding to ALG2 respectively. Figures (1), (2) and (3) shows the Dolan and Moré [5] (iterations (iter), function-gradient evaluations (fg) and CPU time) performance profile of AK3 versus DL, LX, CD and HS conjugate gradient algorithms. In a performance profile plot, the top curve corresponds to the method that solved the most problems in a(iter) or (fg) or CPU time that was within a given factor of the best((iter) or (fg) or CPU time). The percentage of the test problems for which a method is the fastest is given on the left axis of the plot. The right side of the plot gives the percentage of the test problems that were successfully solved by these algorithms, respectively. The right is a measure of the robustness of an algorithm. When comparing AK3 with the DL and LX subject (iter, fg, CPU) as in figures(1), (2) and (3) we see that AK3 is the top performer.

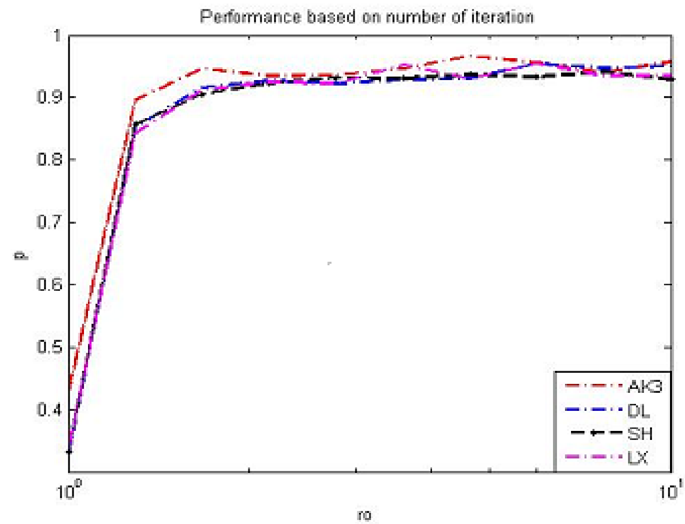

Figure 1: Performance based on number of iteration

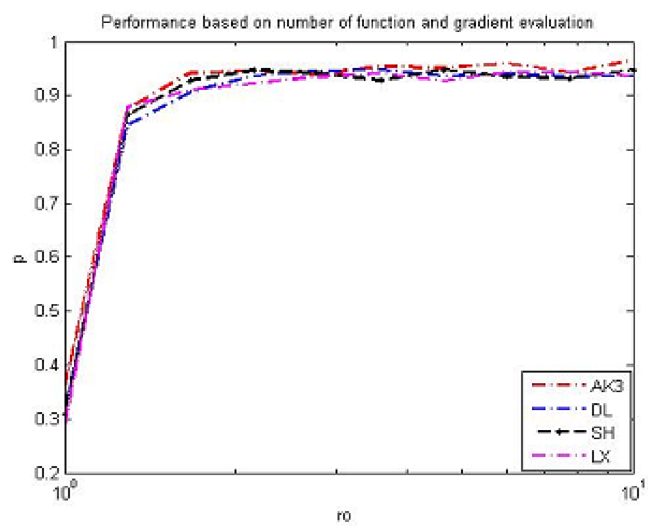

Figure 2: Performance based on number of function-gradient evaluation

6 Conclusion

In this paper, a new three –term conjugate gradient algorithm, as a modification of the DL and PS methods which generates sufficient descent and conjugate directions. Under suitable assumptions our method has been shown to converge globally. In numerical experiments, we have confirmed the effectiveness of the proposed method by using performance profile.

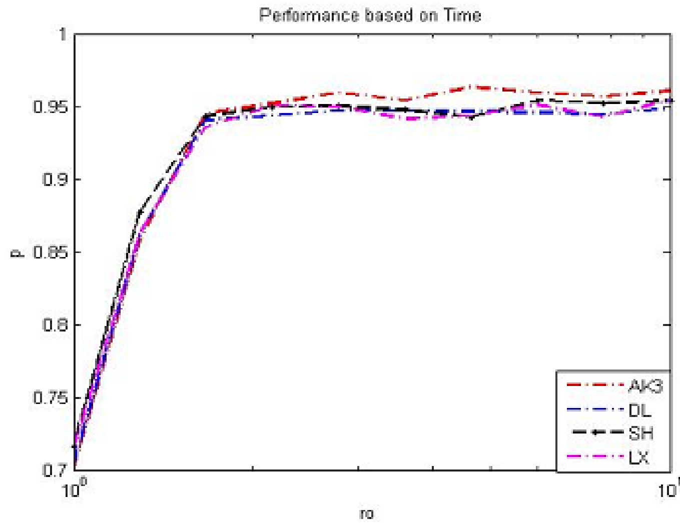

Figure 3: Performance based on time

Acknowledgment

The authors declare that this work is completely private effort and it has not been supported by any governmental, financial, educational or otherwise establishment.

Conflict of Interest Declaration

The authors declare that there is no conflict of interest statement.

Ethics Committee Approval and Informed Consent

The authors declare that declare that that there is no ethics committee approval and/or informed consent statement.

References

[1] N. Yilmaz and A. Sahiner, New global optimization method for non-smooth unconstrained continuous optimization. In: AIP Conference Proceedings. AIP Publishing LLC, 2017. p. 250002.

[2] A. Sahiner and S. A. Ibrahem, A new global optimization technique by auxiliary function method in a directional search, Optim. Lett., 13(2) 2019, 309-323.

[3] A. Sahiner, I. A. M. Abdulhamid and S. A. Ibrahem, A new filled function method with two parameters in a directional search, Journal of Multidisciplinary Modeling and Optimization, 2(1) 2019, 34-42.

[4] A. Sahiner, S. A. Ibrahem and N. Yilmaz, Increasing the Effects of Auxiliary Function by Multiple Extrema in Global Optimization. In: Numerical Solutions of Realistic Nonlinear Phenomena. Springer, Cham, 2020, 125-143.

[5] N. Andrei, An unconstrained optimization test function collection, Adv. Model. Optimization, (10) 2008, 147-161.

[6] N. Andrei, Open Problems in Nonlinear Conjugate Gradient Algorithms for Unconstrained Optimization, Bulletin of the Malaysian Mathematical Sciences Society, (34) 2011, 319–330.

[7] Y. Dai and C. Kou, A nonlinear conjugate gradient algorithm with an optimal property and an improved Wolfe line search, SIAM J. Optim., 23(1) 2013, 296-320.

[8] Y. H. Dai and L.Z. Liao, New Conjugacy Conditions and Related Nonlinear Conjugate Gradient Methods, Applied Mathematics and Optimization, Springer-Verlag, New York, 2001.

[9] E.D. Dolan and J.J. Mor´e, Benchmarking optimization software with performance profiles, Math. Programming, (91) 2002, 201-213.

[10] R. Fletcher and C.M. Reeves, Function Minimization by Conjugate Gradients. Computer Journal, (7) 1964, 149-154.

[11] M.R. Hestenes and E. Stiefel, Methods of Conjugate Gradients for Solving Linear Systems, Journal of Research of the National Bureau of Standards, 5 (49) 1952.

[12] D. Liu and G. Xu, Symmetric Perry conjugate gradient method. Comput. Optim Appl., (56) 2013.

[13] Y. Narushima, H. Yabe and J.A. Ford, A three-term conjugate gradient method with sufficient descent property for unconstrained optimization, SIAM Journal on Optimization, (21) 2013, 212-230.

[14] A. Perry, (1978). A Modified Conjugate Gradient Algorithms. Operations Research, 26, PP. 1073–1078.

[15] E. Polak and G. Ribiére, Note sur la convergence de directions conjuguée, Revue Francaise Information, (16) 1969, 35-43.

[16] M. J. D. Powell, Nonconvex Minimization Calculations and the Conjugate Gradient Method. in : Numerical Analysis (Dundee,1983), In Lecture Notes in mathematics, Springer-Verlag, Berlin, 1066, 122-141, 1984.

[17] D. F. Shanno, Conjugate gradient methods with inexact searches, Mathematics of Operations Research, (3) 1996, 244-256.

[18] W. Sun and Y. Yuan, Optimization Theory and Methods, Nonlinear programming, Springer Science, Business Media, LLC., New York, 2006.

[19] M. Wolfe, Numerical Methods For Unconstrained Optimization An Introduction, New York: Van Nostraned Reinhold, 1978.

[20] G. Zoutendijk, Nonlinear Programming, Computational Methods. Integer and Nonlinear Programming (J. Abadie ED.), North-Holland, Amsterdam, 1970.

Khalil K. Abbo, ORCID: https://orcid.org/0000-0001-5858-625X
Aynur J. Namik, ORCID: https://orcid.org/0000-0001-9100-3319