Вычислительные технологии, 2024, том 29, №3, с. 81–91. © ФИЦ ИВТ, 2024
Computational Technologies, 2024, vol. 29, no. 3, pp. 81–91. © FRC ICT, 2024
ISSN 1560-7534
eISSN 2313-691X

DOI:10.25743/ICT.2024.29.3.007

Introducing TTMDY, a three-term modified DY conjugate gradient direction for large-scale unconstrained problems

K. Meansri[1,*], N. Benrabia[2], M. Ghiat[2], H. Guebbai[2], I. Hafaidia[2]
[1]University of Mohamed Cherif Messaadia Souk Ahras, 41000, Souk Ahras, Algeria
[2]University of 8 Mai 1945 Guelma, 24000, Guelma, Algeria
[*]Corresponding author: Khaoula Meansri, e-mail: k.meansri@univ-soukahras.dz
Received July 20, 2023, revised September 08, 2023, accepted October 11, 2023.

The main focus of our paper is a novel approach to enhance the MDY conjugate gradient direction. The key modification involves incorporating a third term, which plays a crucial role in determining the descent direction. By introducing this additional term, we transform the MDY conjugate gradient direction into a three-term conjugate gradient direction. This modification aims to improve the convergence properties of the algorithm and enhance its performance in solving optimization problems.

In comparison to traditional MDY conjugate gradient methods, our approach demonstrates improved convergence properties and achieves higher solution quality. Numerical results confirm the superiority of our proposed method in terms of optimization performance. This highlights the potential of our modified approach to effectively tackle a wide range of optimization problems in various domains. The results of our numerical experiments provide strong evidence of the efficacy of our modified three-term conjugate gradient direction.

Keywords: conjugate gradient direction, three-term conjugate gradient direction, descent condition, global convergence, Wolfe line search conditions, numerical tests.

Citation: Meansri K., Benrabia N., Ghiat M., Guebbai H., Hafaidia I. Introducing TTMDY, a three-term modified DY conjugate gradient direction for large-scale unconstrained problems. Computational Technologies. 2024; 29(3):81–91. DOI:10.25743/ICT.2024.29.3.007.

Introduction

Nonlinear conjugate gradient (NCG) algorithms have demonstrated their effectiveness in addressing large-scale unconstrained optimization problems [1–6]. These types of problems are commonly expressed in the following general format

$$\min\left\{f(x) : x \in \mathbb{R}^n\right\},$$

where, $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable function.

The NCG method generates a sequence $\{x_k\}_{k\in\mathbb{N}}$ for an initial starting point $x_0 \in \mathbb{R}^n$ using the iterative formula

$$x_{k+1} = x_k + \alpha_k d_k, k \geq 0,$$

where, x_k represents the current iterate point, and $\alpha_k > 0$ denotes the step length of the line search. The search direction, $d_k \in \mathbb{R}^n$ is defined by

$$d_k = \begin{cases} -g_k & \text{if } k = 0, \\ -g_k + \beta_k d_{k-1} & \text{if } k \geq 1, \end{cases}$$

where $g_k = \nabla f(x_k)$ represents the gradient of the function f at point x_k, and the parameter $\beta_k \in \mathbb{R}^*$ is known as the conjugate gradient coefficient. There are several classical formulas for β_k, including the Hestenes – Stiefel [7] (1952), Fletcher Reeves [8] (1964), Polak – Ribiére – Polyak [9, 10] (1969), Conjugate Descent [11] (1987), Liu – Storey [12] (1991), Dai – Yuan [13] (1999) formulas, which are respectively

$$\beta_k^{\mathrm{HS}} = \frac{g_k^t y_{k-1}}{d_{k-1}^t y_{k-1}}, \quad \beta_k^{\mathrm{FR}} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \quad \beta_k^{\mathrm{PRP}} = \frac{g_k^t y_{k-1}}{\|g_{k-1}\|^2},$$
$$\beta_k^{\mathrm{CD}} = \frac{\|g_k\|^2}{-g_{k-1}^t d_{k-1}}, \quad \beta_k^{\mathrm{LS}} = \frac{g_k^t y_{k-1}}{-g_{k-1}^t d_{k-1}}, \quad \beta_k^{\mathrm{DY}} = \frac{\|g_k\|^2}{d_{k-1}^t y_{k-1}},$$

where $y_{k-1} = g_k - g_{k-1}$ and $|\cdot|$ is the Euclidean norm.

In unconstrained optimization problems, there exist various methods for selecting β_k, commonly referred to as classical conjugate gradient methods. Numerous researchers have introduced modifications to β_k. For instance, Yu.E. Nesterov, A.V. Gasnikov and their colleagues, who have presented a lot of work in this field, are among the references [14–17]. On the other hand, Hager and Zhang [18] proposed a modification of the HS method, which is known as the CG-DESCENT method, as follows

$$\beta_k^{N+} = \max\left\{\beta_k^N, \eta_k\right\},$$

where,

$$\beta_k^N = \beta_k^{\mathrm{HS}} - \frac{2\,\|y_{k-1}\|^2}{(y_{k-1}^t d_{k-1})^2} g_k^t d_{k-1}, \quad \eta_k = \frac{-1}{\|d_k\|^2 \min\left\{\|g_k\| .\eta\right\}}.$$

By introducing a constant $\eta > 0$, this modification demonstrates that the resulting descent vector is more efficient when combined with an inexact line search. On the other hand, H. Liu, S. Sun and X. Li [19] introduced a modification to the classical DY method in order to achieve β_k^{MDY} (modification for Dai – Yuan method), it is given by

$$\beta_k^{\mathrm{MDY}} = \frac{\|g_k\|^2}{\mu |d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}}, \quad \mu > 1. \tag{1}$$

Which satisfies the sufficient descent condition $g_k^t d_k \leqslant -(1 - 1/\mu)\|g_k\|^2$, and is globally convergent when combined the Wolfe line search [20, 21], as give by

$$f(x_k + \alpha_k d_k) - f(x_k) \leqslant \rho \alpha_k g_k^t d_k, \tag{2}$$
$$g_k^t d_{k-1} \geqslant \sigma g_{k-1}^t d_{k-1}, \tag{3}$$

where $0 < \rho < \sigma < 1$.

To enhance the efficiency of classical conjugate gradient directions in large-scale problems, researchers have recently proposed three-term conjugate gradient (TTCG) directions. Many of these directions are derived from the classical conjugate gradient direction. Beale [22] was the first to introduce a TTCG direction using β_k^{HS}, in the form

$$d_k = -g_k + \beta_k^{\mathrm{HS}} d_{k-1} + \gamma_k d_l,$$

where, $1 \leqslant l < k$ and $\gamma_k = \dfrac{g_k^t y_l}{d_l^t y_l}$. Nazareth [23] proposed the TTCG direction to obtain

$$d_k = -y_{k-1} + \left(\frac{y_{k-2}^t y_{k-2}}{y_{k-2}^t d_{k-2}}\right) d_{k-2} + \left(\frac{y_{k-1}^t y_{k-1}}{y_{k-1}^t d_{k-1}}\right) d_{k-1}.$$

Zhang et al [24, 25] introduced a three term PRP conjugate gradient method (TTPRP) and a three term FR conjugate gradient method (TTFR). These two modifications ensure that a descent direction is obtained, and when combined with Armijo line search, they guarantee global convergence. Building upon these ideas, Zhang et al. [26] proposed a three term HS conjugate gradient method (TTHS), which also provides a descent direction and global convergence with the standard Wolfe line search. For further exploration in this category, we can provide you with some references [27–33].

Our work focuses on the development of a new TTCG (three-term conjugate gradient) direction, which involves incorporating an additional term, $\omega_k s_{k-1}$, into the classical MDY conjugate gradient direction. The parameter ω_k is chosen to satisfy the condition $g_k^t d_k = -\|g_k\|^2$, ensuring the formulation of the desired equation. By introducing this new direction, we aim to guarantee a descent direction and establish the global convergence of our method by adhering to the Wolfe line search conditions (2), (3).

To assess the effectiveness of our proposed TTCG method, we conducted a series of numerical experiments. These experiments involved measuring various factors, including computation time, the number of iterations, and the number of gradient evaluations. These quantitative evaluations were carried out to validate and demonstrate the efficacy of our novel approach.

1. New three-term conjugate gradient direction

In this section, our study focuses on modifying the direction of the MDY conjugate gradient. Firstly, we start with the classical DY conjugate gradient direction, which can be expressed using the following formula

$$d_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k^{\mathrm{DY}} d_{k-1}, & k \geq 1, \end{cases}$$

where, $\beta_k^{\mathrm{DY}} = \dfrac{\|g_k\|^2}{d_{k-1}^t y_{k-1}}$.

Secondly, H. Liu, S. Sun and X. Li [19] introduced a modification to the classical DY method in order to obtain the conjugate gradient direction. This modification can be represented by the following formula

$$d_k = \begin{cases} -g_0, & k = 0, \\ -g_k + \beta_k^{\mathrm{MDY}} d_{k-1}, & k \geq 1, \end{cases}$$

where the parameter β_k^{MDY} is defined in equation (1).

In our new method, we have modified the MDY conjugate gradient direction to a three term conjugate gradient direction by introducing the additional term $\omega_k s_{k-1}$. We refer to this modified direction as TTMDY. Therefore, the new direction can be obtained by

$$d_k = \begin{cases} -g_0, & k = 0, \\ -g_k + \beta_k^{\text{MDY}} d_{k-1} - \omega_k s_{k-1}, & k \geq 1. \end{cases} \tag{4}$$

Where, the term ω_k determined by our method to verify the descent direction, and $s_{k-1} = x_k - x_{k-1} = \alpha_{k-1} d_{k-1}$.

We have

$$d_k = -g_k + \beta_k^{\text{MDY}} d_{k-1} - \omega_k s_{k-1}.$$

By multiplying it by g_k^t and utilizing equation (1), we obtain

$$g_k^t d_k = -\|g_k\|^2 + \frac{\|g_k\|^2}{\mu |d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}} g_k^t d_{k-1} - \omega_k g_k^t s_{k-1}.$$

So, for

$$g_k^t d_k = -\|g_k\|^2. \tag{5}$$

We find

$$\omega_k = \frac{\|g_k\|^2 g_k^t d_{k-1}}{(\mu |d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}) g_k^t s_{k-1}} \quad \forall k \in \mathbb{N}. \tag{6}$$

Proposition 1. ω_k is well defined

$$|\omega_k| = \left| \frac{\|g_k\|^2 g_k^t d_{k-1}}{(\mu |d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}) g_k^t s_{k-1}} \right| \leqslant \frac{\|g_k\|^2 |g_k^t d_{k-1}|}{(\mu |d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}) |g_k^t s_{k-1}|} \leqslant \frac{\|g_k\|^2 |g_k^t d_{k-1}|}{d_{k-1}^t y_{k-1} |g_k^t s_{k-1}|}.$$

By the Wolfe line search condition (3), and $s_{k-1} = \alpha_{k-1} d_{k-1}$, we have

$$|\omega_k| \leqslant \frac{\|g_k\|^2}{\alpha_{k-1}(1-\sigma) \left| d_{k-1}^t g_{k-1} \right|}. \tag{7}$$

The following theorem is required to prove the descent direction of proposed method.

Theorem 1. *If $\mu > 1$and the definition of ω_k given by equation (6), then d_k is a descent direction for all $k \in \mathbb{N}$. This condition must be satisfied*

$$g_k^t d_k = -\|g_k\|^2 \quad \forall k \geqslant 0. \tag{8}$$

Proof. For $k = 0$, we have $d_0 = -g_0$, then

$$g_0^t d_0 = -\|g_0\|^2.$$

For $k \geqslant 1$, from condition (5), we can deduce that the descent direction satisfies. ∎

Algorithm 1. We present our algorithm in the following steps
Step0: Choose a starting point $x_0 \in \mathbb{R}^n$ and the parameter $\mu > 1$, $\varepsilon > 0$.
Compute $f_0 = f(x_0)$ and $g_0 = \nabla f(x_0)$.
Set $d_0 = -g_0$ and $k = 0$.
Step1: If $\|g_k\| \leq \varepsilon$ stop.
Otherwise, go to step2.
Step2: Determine the step lenght α_k with Wolfe line search conditions (2), (3).

Step3: Compute $g_k = \nabla f(x_k)$, $y_k = g_k - g_{k-1}$ and $s_k = x_k - x_{k-1}$.
Step4: Calculate the direction $d_k = -g_k + \beta_k^{\text{MDY}} d_{k-1} - \omega_k s_{k-1}$. with β_k^{MDY} formula (1) with $\mu > 1$ and ω_k defined by (6).
Step5: Generate the next iterate by $x_{k+1} = x_k + \alpha_k d_k$.
Step6: Set $k = k + 1$, then return to Step1.

2. Global convergence

In this section, we focus on examining the global convergence of algorithm 1. Therefore, it is essential to consider the following two fundamental assumptions.

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

Assumption 2. f is a continuously differentiable function in a neighborhood $\aleph$ of Γ. Namely, there exists a constant $\chi > 0$, such that

$$\|x\| \leqslant \chi \qquad \forall x \in \aleph. \tag{9}$$

Its gradient $g(x)$ is Lipschitz continuous in $\aleph$, namely, there exists a constant $L > 0$, such that

$$\| g(x_1) - g(x_2) \| \leqslant L \| x_1 - x_2 \| \qquad \forall x_1,\, x_2 \in \aleph. \tag{10}$$

Remark 1. Applying assumptions 1 and 2, we can conclude that for all $x \in \aleph$ there exists a positive constant $\nu > 0$, satisfying the following condition

$$\| g(x) \| \leq \nu \quad \forall x \in \aleph. \tag{11}$$

In order to establish the global convergence of our method, we rely on the following two results.

Lemma 1. Assuming that assumptions 1 and 2 are satisfied, let's consider the sequence $\{x_k\}_{k\in\mathbb{N}}$ generated by algorithm 1. Additionally, let $d_k \in \mathbb{R}$ be a descent direction based on the condition (8), and α_k be obtained through Wolfe line search (2), (3). If we have the following condition

$$\sum_{k=0}^{\infty} \frac{1}{\| d_k \|^2} = \infty. \tag{12}$$

Then

$$\lim_{k\to\infty} \inf \| g_k \| = 0.$$

Lemma 2. [34] Suppose that assumptions 1 and 2 hold, and the sequence $\{x_k\}_{k\in\mathbb{N}}$ is generated by the algorithm 1. Additionally, let $d_k \in \mathbb{R}$ be a descent direction by the condition (8), and α_k satisfies the Wolfe line search (2), (3). Then, under these conditions, we can conclude that

$$\alpha_{k-1} \geqslant \frac{(1-\sigma) \mid g_{k-1}^t d_{k-1} \mid}{L \| d_{k-1} \|^2}. \tag{13}$$

Proof. By using the Wolfe conditions (2), (3), along with applying the Cauchy Schwarz inequality and the condition (10), we get

$$L\alpha_{k-1} \|d_{k-1}\|^2 \geqslant d_{k-1}^t(g_k - g_{k-1}) \geqslant (1-\sigma) \mid g_{k-1}^t d_{k-1} \mid .$$

Hence, we have demonstrated the validity of (13). ■

Remark 2. From lemma 1, it follows that the value of α_k obtained through algorithm 1 is not equal to zero. Consequently, there exists a positive constant $\zeta > 0$ such that

$$\alpha_k \geqslant \zeta \quad \forall k \geq 0. \tag{14}$$

To establish the global convergence of our modified algorithm 1, we introduce the following theorem:

Theorem 2. *Let $\{x_k\}_{k\in\mathbb{N}}$ be the sequence generated by algorithm 1, with d_k calculated using (4) such that it satisfies the condition for being a descent direction according to (8). Additionally, α_k is obtained through the Wolfe line search (2), (3). Assuming that assumptions 1 and 2 hold and the condition (12) is satisfied, then*

$$\lim_{k\to\infty} \inf \mid\mid g_k \mid\mid = 0. \tag{15}$$

Proof. We will prove by contradiction. Let's assume that (15) is not true, which means there exists $\varepsilon > 0$ such that for all $k \geqslant 0$, the following condition holds

$$\mid\mid g_k \mid\mid > \varepsilon \quad \forall k \geq 0. \tag{16}$$

By using the definition of β_k^{MDY} and the Wolfe line search condition (3), we get

$$\left|\beta_k^{\mathrm{MDY}}\right| = \left|\frac{\|g_k\|^2}{\mu|d_{k-1}^t g_k| + d_{k-1}^t y_{k-1}}\right| \leqslant \frac{\|g_k\|^2}{d_{k-1}^t y_{k-1}}.$$

By the Wolfe line search condition (3), we get

$$\left|\beta_k^{\mathrm{MDY}}\right| \leqslant \frac{\|g_k\|^2}{(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|}. \tag{17}$$

On the other hand, considering the definition of d_k as given in equation (4), we have

$$d_k = -g_k + \beta_k^{\mathrm{MDY}} d_{k-1} - \omega_k s_{k-1}.$$

This implies

$$\|d_k\| \leqslant \|g_k\| + \left|\beta_k^{\mathrm{MDY}}\right| \|d_{k-1}\| + |\omega_k| \|s_{k-1}\| .$$

From (17) and (7), we have

$$\|d_k\| \leqslant \|g_k\| + \frac{\|g_k\|^2}{(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|} \|d_{k-1}\| + \frac{\|g_k\|^2}{\alpha_{k-1}(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|} \|s_{k-1}\| .$$

As $s_{k_1} = \alpha_{k-1} d_{k-1}$, so

$$\|d_k\| \leqslant \|g_k\| + \frac{\|g_k\|^2}{(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|} \|d_{k-1}\| + \frac{\|g_k\|^2 \|d_{k-1}\|}{(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|}.$$

From $\alpha_{k-1}d_{k-1} = \|x_k - x_{k-1}\|$, (11), (9) and (14), we get

$$\|d_k\| \leqslant \nu + 2\frac{\nu^2\chi}{\zeta(1-\sigma)\left|d_{k-1}^t g_{k-1}\right|}.$$

By (8) and (16), we have

$$\|d_k\| \leqslant \Delta.$$

Where $\Delta = \nu + 2\dfrac{\nu^2\chi}{\zeta(1-\sigma)\varepsilon^2}$.

Therefore, by applying Lemma 1, we can conclude that equation (15) is true. This contradicts equation (16), leading us to the conclusion that (15) holds. Thus, we have proven (15). ■

3. Numerical results

In this section, we provide numerical test results that compare the performance of our TTCG algorithm 1, which implements the Wolfe line search conditions (2), (3) with $\rho = 0.0001$ and $\sigma = 0.1$, using the parameter $\mu = 1.1$. We compare it with three conjugate gradient methods MDY (1) with the parameter $\mu = 1.1$, TTFR [25] with Armijio line search and CG-DESCENT as presented in [18]. For that we selected 50 unconstrained optimization test problems from [35] and each problem was tested with varying numbers of variables: 2, 50, 100, 200, 500, 1000, 2000, 3000, $\ldots$, 3500. In all the algorithms the same stopping criterion $|g_k|^2 \leq 10^{-7}$ and we considered in this numerical study the maximum number of iterations is limited to 50 000. All code implementations were compiled in MATLAB 2013, using the compiler settings on a PC machine with an Intel Core i3-2348M CPU @ 2.30 GHz and 4.00 GB RAM. To compare the performance of the algorithms, we utilized performance profiles, as provided by Dolan and Moré [36]. This allowed us to assess and compare the performance of each algorithm objectively.

Figures 1–3 display the performance profiles of TTMDY versus MDY, TTFR and CG-DESCENT based on CPU time, the number of iterations, and the number of gradient evaluations, respectively. These profiles provide a visual representation of the performance

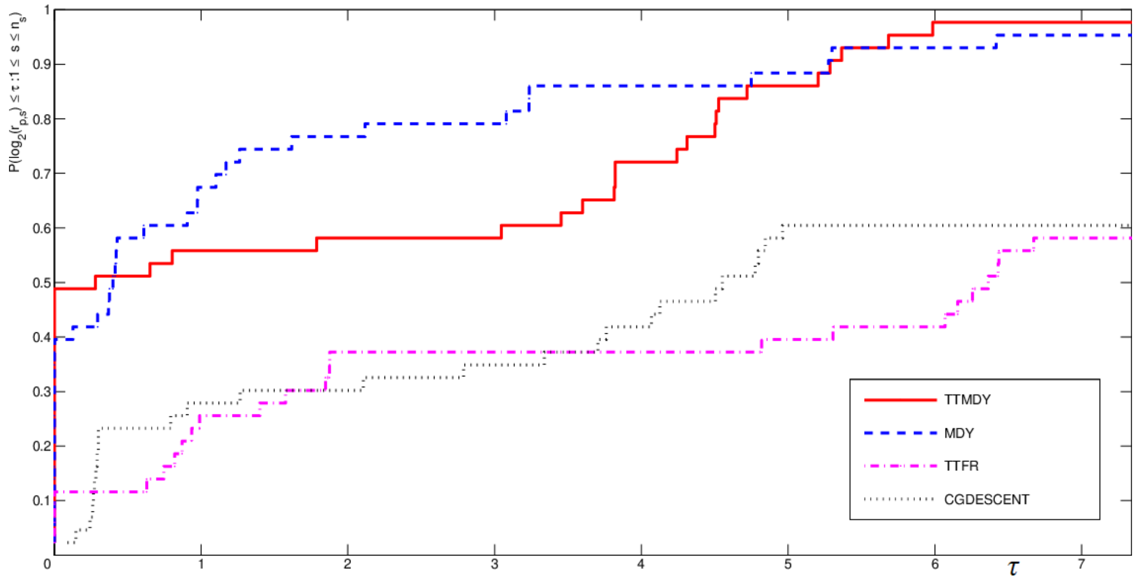

Fig. 1. Performance profile for CPU time

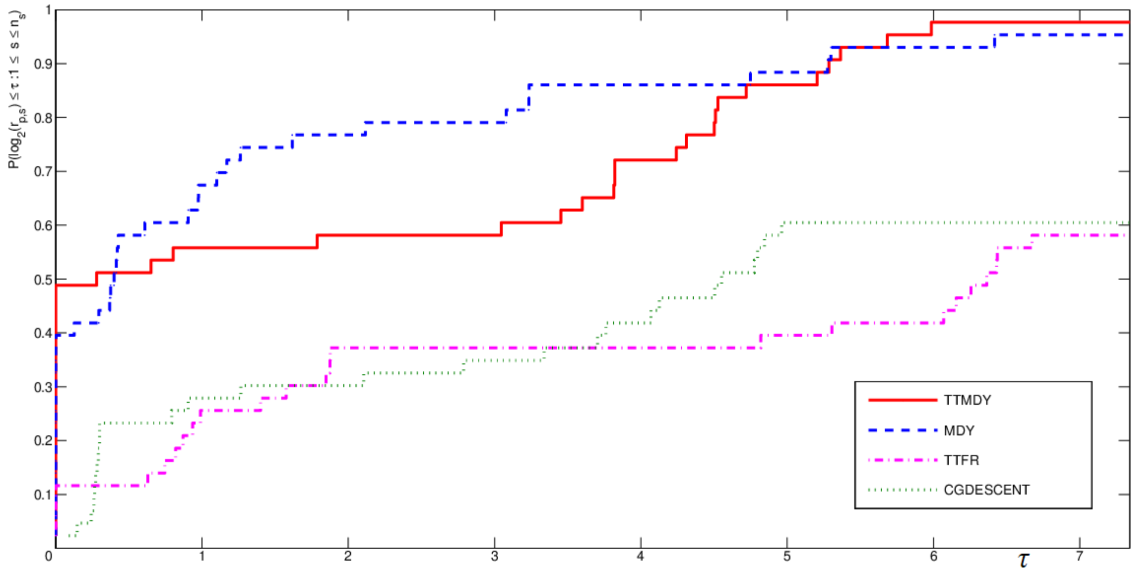

Fig. 2. Performance profile for the number of iterations

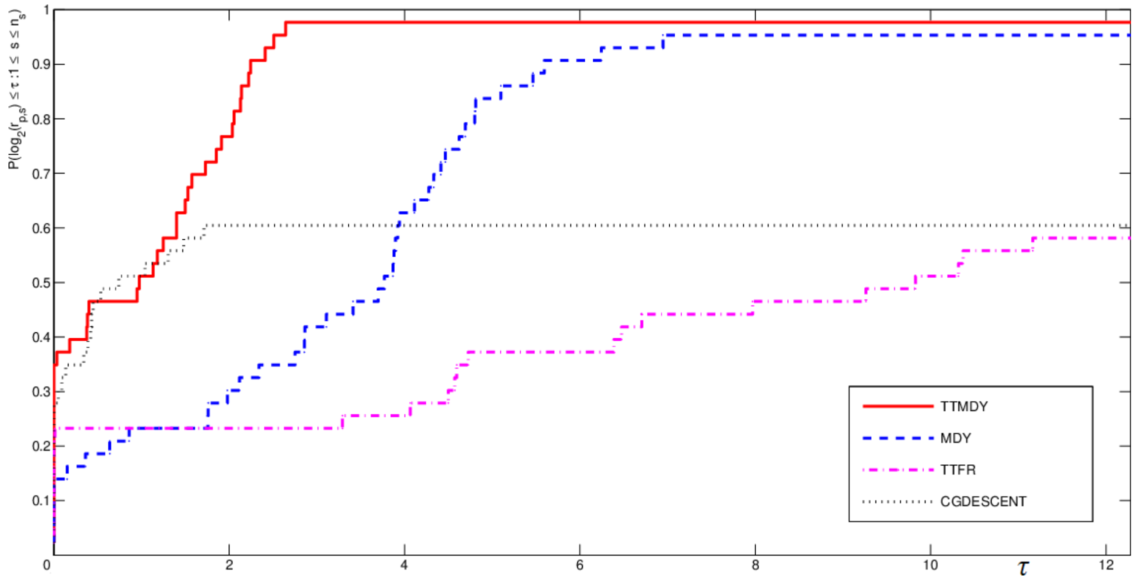

Fig. 3. Performance profile for the number of gradient evaluations

of the algorithms in terms of these metrics, which were evaluation using the profiles of Dolan and Moré.

Based on the analysis of Fig. 1–3, it is evident that our TTCG algorithm 1 outperforms the other algorithms in terms of efficiency in terms of time, number of iterations, and error. This suggests that our algorithm provides better results and requires fewer resources to converge to the desired solution.

Conclusion

The three-term conjugate gradient method has emerged as a vital tool for tackling large-scale unconstrained optimization problems. In this paper, we have presented a new TTCG direction, known as TTMDY, which is derived from the classical MDY conjugate gradient direction. The TTMDY direction satisfies the descent condition, ensuring effective optimization, and we have demonstrated its global convergence by employing the Wolfe line search. Through extensive numerical experiments, we have obtained compelling results, including measurements of time, number of iterations, and number of gradient evaluations. These results unequivocally demonstrate that our proposed TTCG algorithm surpasses other methods in terms of both speed and efficiency. This clear advantage positions our chosen method as superior to alternative approaches.

Looking ahead, our work opens up several promising perspectives for further research. Firstly, investigating the application of the TTMDY direction to constrained optimization problems could provide valuable insights into its adaptability and performance in more complex scenarios. Additionally, exploring variations or extensions of the TTCG method, such as incorporating different line search conditions or considering different step size selection strategies, could lead to even more efficient and accurate optimization algorithms.

References

[1] **Jiang X., Yang H., Yin J., Liao W.** A three-term conjugate gradient algorithm with restart procedure to solve image restoration problems. Journal of Computational and Applied Mathematics. 2023; (424):115020.

[2] **Khoshsimaye-Bargard M., Ashrafi A.** A family of the modified three-term Hestenes–Stiefel conjugate gradient method with sufficient descent and conjugacy conditions. Journal of Applied Mathematics and Computing. 2023; (69):2331–2360.

[3] **Prolubnikov A.V.** Approaches to discrete optimization problems with interval objective function. Computational Technologies. 2021; 26(6):82–109. DOI:10.25743/ICT.2021.26.6.007. (In Russ.)

[4] **Ibrahim A., Muangchoo K., Mohamed N.S., Abubakar A.B.** Derivative-free SMR conjugate gradient method for constraint nonlinear equations. Journal of Mathematics and Computer Science. 2022; 24(2):147–164. DOI:10.22436/jmcs.024.02.06.

[5] **Dong X.** A modified nonlinear Polak–Ribiere–Polyak conjugate gradient method with sufficient descent property. Calcolo. 2020; (57):30.

[6] **Sulaiman I.M., Mamat M., Owoyemi A.E., Ghazali P.L., Rivaie M., Malik M.** The convergence properties of some descent conjugate gradient algorithms for optimization models. Journal of Mathematics and Computer Science. 2020; 22(3):204–215.

[7] **Hestenes M.R., Stiefel E.** Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards. 1952; (49):409–436. DOI:10.6028/JRES.049.044.

[8] **Fletcher R., Reeves C.M.** Function minimization by conjugate gradients. Computer Science, Mathematics. (1964):149–154. DOI:10.1093/COMJNL/7.2.149.

[9] **Polak E., Ribiere G.** Note sur la convergence de directions conjugées. Rev. Francaise Informat Recherche Operatinelle, 3e Année 16. (1969):35–43.

[10] **Polyak B.T.** The conjugate gradient method in extreme problems. USSR Computational Mathematics and Mathematical Physics. 1969(9):94–112. DOI:10.1016/0041-5553(69)90035-4.

[11] **Fletcher R.** Practical method of optimization. Chichester, N.Y.: Wiley; 1987: 436.

[12] **Liu Y., Storey C.** Effcient generalized conjugate gradient algorithms, part 1: theory. Journal of Optimization Theory and Applications. 1991; (69):129–137.

[13] **Dai Y.-H., Yuan Y.-X.** A nonlinear conjugate gradient method with a strong global convergence property. SIAM Journal on Optimization. 1999; 10(1). DOI:10.1137/S1052623497318992.

[14] **Guminov S.V., Nesterov Yu.E., Dvurechensky P.E., Gasnikov A.V.** Accelerated primal-dual gradient descent with linesearch for convex, nonconvex, and nonsmooth optimization problems. Doklady Mathematics. 2019; (99):125–128.

[15] **Gasnikov A.V., Nesterov Yu.E.** Universal method for stochastic composite optimization problems. Computational Mathematics and Mathematical Physics. 2018; 58(1):48–64. DOI:10.1134/S0965542518010050.

[16] **Gasnikov A.V., Gasnikova E.V., Nesterov Yu.E., Chernov A.V.** Efficient numerical methods for entropy-linear programming problems. Computational Mathematics and Mathematical Physics. 2016; 56(4):514–524.

[17] **Gasnikov A.V., Nesterov Yu.E., Spokoiny V.G.** On the efficiency of a randomized mirror descent algorithm in online optimization problems. Computational Mathematics and Mathematical Physics. 2015; 55(4):580–596.

[18] **Hager W., Zhang H.** A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM Journal on Optimization. 2005; DOI:10.1137/030601880. Available at: https://www.semanticscholar.org/paper/A-New-Conjugate-Gradient-Method-with-Guaranteed-and-Hager-Zhang/45efe38ea906de376db50a86b8a11ef566821e42.

[19] **Hailin L., Cheng S.S., Li X.** A conjugate gradient method with sufficient descent and global convergence for unconstrained nonlinear optimization. Applied Mathematics E-Notes. 2011; (11):139–147.

[20] **Wolfe P.** Convergence conditions for ascent methods. SIAM Review. 1969; 11(2). DOI:10.1137/1011036. Available at: https://msulaiman.org/onewebmedia/Wolfe%20conditions.pdf.

[21] **Wolfe P.** Convergence conditions for ascent methods. II: some corrections. SIAM Review. 1971; 13(2). DOI:10.1137/1013035.

[22] **Beale E.M.** A derivation of conjugate-gradients. Numerical methods for non-linear optimization. N.Y.: Academic Press; 1972: 39–43.

[23] **Nazareth L.** A conjugate direction algorithm without line searches. Journal of Optimization Theory and Applications. 1977; (23):373–387.

[24] **Zhang L., Zhou W., Li D.** A descent modified Polak – Ribiére – Polyak conjugate gradient method and its global convergence. IMA Journal of Numerical Analysis. 2006; (26):629–640.

[25] **Zhang L., Zhou W., Li D.** Global convergence of a modified Fletcher – Reeves conjugate gradient method with Armijo-type line search. Numerische Mathematik. 2006; (104):561–572.

[26] **Li Z., Zhou W., Li D.** Some descent three-term conjugate gradient methods and their global convergence. Optimisation Methods and Software. 2007; 22(4):697–711.

[27] **Al-Bayati A.Y.** A new three-term non-linear conjugate gradient method for unconstrained optimization. Canadian Journal on Science and Engineering Mathematics. 2010; 1(5):108–124. Available at: https://scholar.google.com/citations?view_op=view_citation&hl=ru&user=a9VjHfoAAAAJ&citation_for_view=a9VjHfoAAAAJ:cFHS6HbyZ2cC.

[28] **Yasushi N., Yabe H., Ford J.** A three-term conjugate gradient method with sufficient descent property for unconstrained optimization. SIAM Journal on Optimization. 2011; 21(1):212–230. DOI:10.1137/080743573.

[29] **Neculai A.** A simple three-term conjugate gradient algorithm for unconstrained optimization. Journal of Computational and Applied Mathematics. 2013; (241):19–29.

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

[31] **Liu J.K., Feng Y.M., Zou L.M.** Some three-term conjugate gradient methods with the inexact line search condition. Calcolo. 2018; (55):1–16.

[32] **Mahdi M.M., Shiker M.A.K.** Three terms of derivative free projection technique for solving nonlinear monotone equations. Journal of Physics: Conference Series. 2020; 1591(1):012031.

[33] **Ibrahim S.M., Mamat M., Owoyemi A.E., Ghazali P.L., Rivaie M., Malik M.** The convergence properties of some descent conjugate gradient algorithms for optimization models. Journal of Mathematics and Computer Science. 2020; 22(3):204–215. DOI:10.22436/jmcs.022.03.02.

[34] **Liu J.K., Li S.J.** New hybrid conjugate gradient method for unconstrained optimization. Applied Mathematics and Computation. 2014; (245):36–43.

[35] **Neculai A.** An unconstrained optimization test functions collection. Advanced Modeling and Optimization. 2008; 10(1):147–161.

[36] **Dolan E.D., Moré J.J.** Benchmarking optimization software with performance profiles. Mathematical Programming. 2002; (91):201–213. Available at: `https://link.springer.com/article/10.1007/s101070100263`.

Вычислительные технологии, 2024, том 29, № 3, с. 81–91. © ФИЦ ИВТ, 2024 ISSN 1560-7534
Computational Technologies, 2024, vol. 29, no. 3, pp. 81–91. © FRC ICT, 2024 eISSN 2313-691X

ВЫЧИСЛИТЕЛЬНЫЕ ТЕХНОЛОГИИ

DOI:10.25743/ICT.2024.29.3.007

Представление TTMDY — трехчленный модифицированный метод сопряженных градиентов DY для крупномасштабных задач без ограничений

К. Минсри[1,*], Н. Бенрабиа[2], М. Гиат[2], Х. Гуэббай[2], И. Хафаидия[2]

[1]Университет Мохамеда Шерифа Мессаадия Сук Ахрас, 41000, Сук Ахрас, Алжир
[2]Университет 8 мая 1945 г., 24000, Гельма, Алжир
*Контактный автор: Каула Минсри, e-mail: k.meansri@univ-soukahras.dz

Поступила 20 июля 2023 г., доработана 08 сентября 2023 г., принята в печать 11 октября 2023 г.

Аннотация

В этой статье основное внимание уделяется представлению нового подхода к улучшению метода сопряженного градиента MDY. Существенная модификация предполагает включение третьего члена, который играет решающую роль в определении направления спуска. Вводя этот дополнительный член, мы преобразуем направление сопряженного градиента MDY в трехчленное направление сопряженного градиента. Целью данной модификации является улучшение свойств сходимости алгоритма и повышение его производительности при решении оптимизационных задач.

По сравнению с традиционными методами сопряженных градиентов MDY наш подход демонстрирует улучшенные свойства сходимости и обеспечивает более высокое качество решения. Численные результаты подтверждают превосходство предложенного метода с точки зрения производительности оптимизации. Это подчеркивает потенциал предложенного модифицированного подхода для эффективного решения широкого спектра задач оптимизации в различных областях. Результаты численных экспериментов убедительно свидетельствуют об эффективности метода модифицированного направления трехчленного сопряженного градиента.

Ключевые слова: направление сопряженного градиента, направление трехчленного сопряженного градиента, условие спуска, глобальная сходимость, условия поиска линии Вульфа, численные тесты.

Цитирование: Минсри К., Бенрабиа Н., Гиат М., Гуэббай Х., Хафаидия И. Представление TTMDY — трехчленный модифицированный метод сопряженных градиентов DY для крупномасштабных задач без ограничений. Вычислительные технологии. 2024; 29(3):81–91. DOI:10.25743/ICT.2024.29.3.007. (на английском)