Mathematics and Statistics 10(2): 293-300, 2022
DOI: 10.13189/ms.2022.100202 http://www.hrpub.org

A New Algorithm for Spectral Conjugate Gradient in Nonlinear Optimization

Ahmed Anwer Mustafa

Department of Mathematics, Faculty of Education, University of Zakho, Kurdistan Region, Iraq

Received November 18, 2021; Revised January 28, 2022; Accepted February 16, 2022

Cite This Paper in the following Citation Styles

(a): *[1] Ahmed Anwer Mustafa , "A New Algorithm for Spectral Conjugate Gradient in Nonlinear Optimization," Mathematics and Statistics, Vol. 10, No. 2, pp. 293 - 300, 2022. DOI: 10.13189/ms.2022.100202.*

(b): *Ahmed Anwer Mustafa (2022). A New Algorithm for Spectral Conjugate Gradient in Nonlinear Optimization. Mathematics and Statistics, 10(2), 293 - 300. DOI: 10.13189/ms.2022.100202.*

Abstract CJG is a nonlinear conjugation gradient. Algorithms have been used to solve large-scale unconstrained enhancement problems. Because of their minimal memory needs and global convergence qualities, they are widely used in a variety of fields. This approach has lately undergone many investigations and modifications to enhance it. In our daily lives, the conjugate gradient is incredibly significant. For example, whatever we do, we strive for the best outcomes, such as the highest profit, the lowest loss, the shortest road, or the shortest time, which are referred to as the minimum and maximum in mathematics, and one of these ways is the process of spectral gradient descent. For multidimensional unbounded objective function, the spectrum conjugated gradient (SCJG) approach is a strong tool. In this study, we describe a revolutionary SCG technique in which performance is quantified. Based on assumptions, we constructed the descent condition, sufficient descent theorem, conjugacy condition, and global convergence criteria using a robust Wolfe and Powell line search. Numerical data and graphs were constructed utilizing benchmark functions, which are often used in many classical functions, to demonstrate the efficacy of the recommended approach. According to numerical statistics, the suggested strategy is more efficient than some current techniques. In addition, we show how the unique method may be utilized to improve solutions and outcomes.

Keywords New Spectral Conjugated Gradient, Optimization with No Constraints, Analytical Convergence, Conjugacy Requirement, Sufficient Descent Inequality

1. Introduction

Gradient's procedures are among the most efficient algorithms easy implementation, convergence properties, and capacity to provide various unconstrained multi-objective optimization problems. The CJG approach is widely used for optimization because of its quick convergence rate, small memory footprint, and simple iterations [1]. Here you may find a simple definition of an unrestricted optimization strategy.

$$\text{Min.} f(x),\ x \in R^n \tag{1}$$

where R^n is an n-dimensional Euclidean space and $f : R^n \to R$ is a continuously differentiable function. The CJG technique creates a sequence of iterates [2]. There are several steps to the CJG technique, including iteration.

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

where x_k is the iteration point at the moment, $\alpha_k > 0$ is a step length and d_k is the direction of the search. The first direction of search is usually the gradient's negative value which is the steepest descent direction [3], i.e., $d_0 = -g_0$, A recursive definition follows the following directions:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad k \geq 1 \tag{3}$$

in which g_k=$\nabla f(x_k)$. Different β_k will result in various conjugate gradient algorithms. The following are some well-known β_k formulas.

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k},\ \beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k},\ \beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k},\ \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k},\ \ \beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k},\ \beta_k^{CD} = \frac{g_{k+1}^T d_k}{-d_k^T g_k}$$

$$\beta_{k+1}^{hybrid} = \frac{g_{k+1}^T(y_k - ts_k)}{\max\{y_k^T d_k, \|g_k\|^2\}}\ ,\ \beta_k^{AA3} = \frac{g_{k+1}^T y_k}{\|d_k\|^2}\left(1 - \eta\frac{g_{k+1}^T y_k}{\|d_k\|^2}\right),\ \beta_k^{AA1} = \beta_k^{HS} + (\lambda - 1)\frac{g_{k+1}^T d_k}{d_k^T d_k}$$

Where $y_k = g_{k+1} - g_k$. The above corresponding methods, HS is known as Hestenes and Steifel [4], FR is Fletcher and Reeves [5], PR is Polak and Ribiere [6], LS is Liu and Storey [7], DY is Dai and Yuan [8], Conjugate Descent [9], hybrid by Zhang, L. [10], Ahmed A. Mustafa [11], and lastly Ahmed A. Mustafa and Salah G. Shareef [12]. Many researchers have examined the convergence of the CJG method under various line searches, and some have used an exact line search to derive the step size (ELS). Others employ a line search known as the strong Wolfe line search condition (SWL), which is described as follows:

$$f(x_k + \alpha_k d_k) \le f(x_k) + c_1\alpha_k g_k^T d_k \tag{4}$$

$$|g(x_k + \alpha_k d_k)^T d_k| \le -c_2 g_k^T d_k \tag{5}$$

Where $0<c_1<c_2<1$

The spectral CJG-technique (SCJG), which was initially offered by Barzilai and Borwein [13], is another well-known method that may be utilized to address the problem (1). The following factors determine the direction d_{k+1}:

$$d_k = \begin{cases} -g_k, & k = 0 \\ -\theta_k g_{k+1} + \beta_k d_k, & k \ge 1 \end{cases} \tag{6}$$

where θ_k is the spectral gradient parameter. The SCJG-method surpasses the more powerful CJG method. Based on certain acceptable assumptions, many researchers have proposed a spectral conjugate gradient as Andrei, Jiang, and Raydan [14,15,16] and the benefits of these algorithms over many traditional methods.

2. Derivation of the New Method with Its Algorithm

2.1. Derivation of the New Formula

There are many proposed spectral conjugate gradient formulas for example

$d_k = -\theta_k g_{k+1} + \beta_k^{BZA} d_k$ where $\theta_k = \frac{t\alpha_k g_{k+1}^T d_k}{g_{k+1}^T d_k \|g_k\|^2} + \frac{d_k^T y_k}{d_k^T y_k + |g_{k+1}^T d_k|}$, $\mu = 0.01$ and $t = 0.1$ see [17,18].

$d_k = -\theta_{k+1} g_{k+1} + \beta_{k+1}^{JYJLL} d_k$ where $\theta_{k+1} = \theta_{k+1}^{JYJLL} = 1 + \frac{|g_{k+1}^T d_k|}{-g_k^T d_k}$ see [19,20].

We'll calculate a new spectral parameter in this section. The SCJG-search method's orientation is normally as follows:

$$d_{k+1} = -\theta_k g_{k+1} + \beta_k d_k,\ k \ge 0 \tag{7}$$

In 2015, Shuo-Wang proposed the following formula for calculating the value of beta: see [21].

$$\beta_k^{MCD} = \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right), \mu > \frac{1}{4} \tag{8}$$

Put (8) in (7), we have

$$d_{k+1} = -\theta_k g_{k+1} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k \tag{9}$$

Multiply both sides of the above equation by y_k^T, to obtain

$$d_{k+1}^T y_k = -\theta_k g_{k+1}^T y_k + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k^T y_k$$

The following conjugacy requirement is established by Dai and Liao $d_{k+1}^T y_k = -t\, g_{k+1}^T v_k, t > 0$ see [22].

Substitute $d_{k+1}^T y_k$, we have

$$-t\, g_{k+1}^T v_k = -\theta_k\, g_{k+1}^T y_k + (1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k})\, d_k^T y_k$$

$$\Rightarrow \theta_k g_{k+1}^T y_k = t\, g_{k+1}^T v_k + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k^T y_k$$

Dividing both sides of the above equation by $g_{k+1}^T y$, we get

$$\theta_k = \frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right)\frac{d_k^T y_k}{g_{k+1}^T y},$$

$$\text{where } t > 0 \text{ and } \mu > \frac{1}{4} \tag{10}$$

Put (10) in (9) we obtain the new search direction

$$d_{k+1} = -\left(\frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right)\frac{d_k^T y_k}{g_{k+1}^T y_k}\right) g_{k+1} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k \tag{11}$$

2.2. Outline of the New Method

Step(1): Select $x_0 \in R^n, \varepsilon = 10^{-5}, t > 0$ and $\mu > \frac{1}{4}$

Step(2): Set $k = 0$, Find $f(x_0)$, $g_0, d_k = -g_k$

Step(3): Compute $\alpha_k > 0$ satisfying the strong Wolfe condition

$$f(x_k + \alpha_k d_k) \le f(x_k) + c_1\alpha_k g_k^T d_k$$

$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \le c_2 |g_k^T d_k|$$

Where $0 < c_1 < c_2 < 1$

Step(4): Evaluate $x_{k+1} = x_k + \alpha_k d_k$, $g_{k+1} = \nabla f(g_{k+1})$, If $\|g_{k+1}\| < \varepsilon$ stop.

Step(5): Calculate $d_{k+1} = -\theta_k g_{k+1} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k$ where $\theta_k = \frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right)\frac{d_k^T y_k}{g_{k+1}^T y_k}$

Step(6): If $|g_{k+1}^T g_k| > 0.2\|g_{k+1}\|^2$ go to step(2) else $k = k + 1$ and go to step(3)

3. Convergence Analysis

The novel algorithms' descent constraint, enough descent quality, conjugacy requirement, and global convergence are all developed in this section. To this aim, we make some assumptions on the function of the objective as follows:

(1) At the start point x_0, f is limited below on the level set R^n continuous and differentiable in area N of the level set $S = \{x \in R^n : f(x) \leq f(x_0)\}$.

(2) In N, the gradient $g(x)$ is Lipschitz continuous, hence for any $x, y \in N$, there exists a constant $L > 0$ such that $\|g(x) - g(y)\| \leq L\|x - y\|$.

Theorem3.1: Consider a CJG method with the use direction of search given as (11), then, condition $d_{k+1}^T g_{k+1} \leq 0$ will hold for all $k \geq 0$ with exact and inexact line search.

Proof: If $k = 0$, then we will have $d_1^T g_1 \leq -\|g_1\|^2$. Hence the condition of descent is hold. Assume that $d_k^T g_k \leq 0, \forall k$. Now, we prove the search direction (11) is the descent direction at $(k+1)$. Multiply all ends of the equation (11) by g_{k+1}^T to obtain

$$d_{k+1}^T g_{k+1} = -\left(\frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k} + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right)\frac{d_k^T y_k}{g_{k+1}^T y_k}\right)\|g_{k+1}\|^2 + \left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right) d_k^T g_{k+1} \Rightarrow d_{k+1}^T g_{k+1}$$
$$= -\frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k}\|g_{k+1}\|^2 - \frac{d_k^T y_k}{g_{k+1}^T y_k}\|g_{k+1}\|^2 + \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\frac{d_k^T y_k}{g_{k+1}^T y_k}\|g_{k+1}\|^2 + d_k^T g_{k+1} - \frac{\mu\left(g_{k+1}^T d_k\right)^2}{d_k^T g_k} \Rightarrow d_{k+1}^T g_{k+1}$$
$$= -\frac{t\, \alpha_k g_{k+1}^T d_k\left(g_{k+1}^T y_k\right)}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 - \frac{d_k^T y_k\left(g_{k+1}^T y_k\right)}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 - \frac{\mu g_{k+1}^T d_k}{\|g_k\|^2}\frac{\left(g_{k+1}^T y_k\right) d_k^T y_k}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 + d_k^T g_{k+1} + \frac{\mu\left(g_{k+1}^T d_k\right)^2}{\|g_k\|^2}$$

Put $g_{k+1}^T y_k = \|g_{k+1}\|^2 + g_{k+1}^T d_k$ and we multiply and divide the last two terms by $\frac{d_k^T g_{k+1}}{d_k^T g_{k+1}}$ and $\frac{d_k^T y_k}{d_k^T y_k}$ respectively. This implies that

$$d_{k+1}^T g_{k+1} = \frac{t\, \alpha_k \|g_{k+1}\|^2(-g_{k+1}^T d_k \|g_{k+1}\|^2 - (g_{k+1}^T d_k)^2)}{(g_{k+1}^T y_k)^2} + \frac{d_k^T y_k(-\|g_{k+1}\|^2 - g_{k+1}^T d_k)}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 + \frac{\mu}{\|g_k\|^2}\frac{(-\|g_{k+1}\|^2 g_{k+1}^T d_k - (g_{k+1}^T d_k)^2) d_k^T y_k}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 + \frac{(d_k^T g_{k+1})^2}{d_k^T g_{k+1}} + \frac{\mu (g_{k+1}^T d_k)^2}{\|g_k\|^2}\frac{d_k^T y_k}{d_k^T y_k}$$

We know that $g_{k+1}^T d_k \leq d_k^T y_k$ and by Wolfe condition $g_{k+1}^T d_k \geq c_2 {g_k}^T d_k \Rightarrow -g_{k+1}^T d_k \leq c_2\|g_k\|^2 \Rightarrow c_2 {g_k}^T d_k \leq d_k^T y_k \Rightarrow -c_2 {g_k}^T d_k \geq -d_k^T y_k$. This implies that $\|g_k\|^2 \geq \frac{-1}{c_2} d_k^T y_k \Rightarrow -c_2\|g_k\|^2 \leq d_k^T y_k$, we use the above relations we have

$$d_{k+1}^T g_{k+1} \leq \frac{t\, \alpha_k \|g_{k+1}\|^2(c_2\|g_k\|^2\|g_{k+1}\|^2 - (g_{k+1}^T d_k)^2)}{(g_{k+1}^T y_k)^2} + \frac{d_k^T y_k(-\|g_{k+1}\|^2 + c_2\|g_k\|^2)}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 + \frac{\mu}{\|g_k\|^2}\frac{(\|g_{k+1}\|^2 c_2\|g_k\|^2 - (g_{k+1}^T d_k)^2) d_k^T y_k}{(g_{k+1}^T y_k)^2}\|g_{k+1}\|^2 - \frac{(d_k^T g_{k+1})^2}{d_k^T y_k} - \frac{\mu (g_{k+1}^T d_k)^2}{\|g_k\|^2}\frac{d_k^T y_k}{c_2\|g_k\|^2}$$

Therefore $\mu, c_2, \|g_k\|^2$ $d_k^T y_k$ and $(g_{k+1}^T d_k)^2$ are greater than zero, then

$$d_{k+1}^T g_{k+1} \leq \frac{t\, \alpha_k \|g_{k+1}\|^2\left(c_2\|g_k\|^2\|g_{k+1}\|^2\frac{d_k^T y_k}{d_k^T y_k} - \left(g_{k+1}^T d_k\right)^2\right)}{\left(g_{k+1}^T y_k\right)^2} + \frac{d_k^T y_k\left(-\|g_{k+1}\|^2 + c_2\|g_k\|^2\frac{d_k^T y_k}{d_k^T y_k}\right)}{\left(g_{k+1}^T y_k\right)^2}\|g_{k+1}\|^2 + \frac{\mu}{\|g_k\|^2}\frac{\left(\|g_{k+1}\|^2 c_2\|g_k\|^2\frac{d_k^T y_k}{d_k^T y_k} - \left(g_{k+1}^T d_k\right)^2\right) d_k^T y_k\|g_{k+1}\|^2}{\left(g_{k+1}^T y_k\right)^2}$$

Since $-c_2\|g_k\|^2 \leq d_k^T y_k$ then

$$d_{k+1}^T g_{k+1} \leq \frac{t\, \alpha_k \|g_{k+1}\|^2\left(-c_2\|g_k\|^2\|g_{k+1}\|^2\frac{d_k^T y_k}{c_2\|g_k\|^2} - \left(g_{k+1}^T d_k\right)^2\right)}{\left(g_{k+1}^T y_k\right)^2} + \frac{d_k^T y_k\left(-\|g_{k+1}\|^2 - c_2\|g_k\|^2\frac{d_k^T y_k}{c_2\|g_k\|^2}\right)}{\left(g_{k+1}^T y_k\right)^2}\|g_{k+1}\|^2 + \frac{\mu}{\|g_k\|^2}\frac{\left(-\|g_{k+1}\|^2 c_2\|g_k\|^2\frac{d_k^T y_k}{c_2\|g_k\|^2} - \left(g_{k+1}^T d_k\right)^2\right) d_k^T y_k}{\left(g_{k+1}^T y_k\right)^2}\|g_{k+1}\|^2 \Rightarrow$$

$$d_{k+1}^T g_{k+1} \leq -\frac{t\, \alpha_k \|g_{k+1}\|^2\left(\|g_{k+1}\|^2 d_k^T y_k + \left(g_{k+1}^T d_k\right)^2\right)}{\left(g_{k+1}^T y_k\right)^2} - \frac{\mu}{\|g_k\|^2}\frac{\left(\|g_{k+1}\|^2 d_k^T y_k + \left(g_{k+1}^T d_k\right)^2\right) d_k^T y_k}{\left(g_{k+1}^T y_k\right)^2}\|g_{k+1}\|^2 \qquad (12)$$

Therefore

t, μ, c_2 $\alpha_k, \|g_{k+1}\|^2$ $d_k^T y_k, \|g_k\|^2, (g_{k+1}^T y_k)^2$ and $(g_{k+1}^T d_k)^2$ are greater than zero then $d_{k+1}^T g_{k+1} \leq 0$

Theorem3.2: Suppose that the direction of search is given by (11). We assume that the step size satisfies strong Wolfe conditions (4) and (5). Then, the following result:

$$g_{k+1}^T d_{k+1} \leq -C\|g_{k+1}\|^2 \text{ holds for any } k \geq 0.$$

Proof: From (12) we have $d_{k+1}^T g_{k+1} \leq -\frac{d_k^T y_k(\|g_{k+1}\|^2 + d_k^T y_k)}{\left(g_{k+1}^T y_k\right)^2}\|g_{k+1}\|^2$

Let $C = \frac{d_k^T y_k(\|g_{k+1}\|^2 + d_k^T y_k)}{(g_{k+1}^T y_k)^2}$ which is positive, then $g_{k+1}^T d_{k+1} \leq -C\|g_{k+1}\|^2$

Theorem3.3: Assume that the sequence $\{x_k\}$ is generated by (2), then the direction of search (11) satisfies the conjugacy condition that is $d_{k+1}^T G_{k+1} d_k = d_{k+1}^T y_k = -t\, g_{k+1}^T v_k = 0$. See [19]

Proof: Multiply both sides of (11) by y_k^T, we have

$$d_{k+1}^T y_k = -\frac{t\, g_{k+1}^T v_k}{g_{k+1}^T y_k} g_{k+1}^T y_k - \frac{d_k^T y_k}{g_{k+1}^T y_k} g_{k+1}^T y_k + \frac{\mu g_{k+1}^T d_k}{d_k^T g_k} g_{k+1}^T y_k + d_k^T y_k - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k} d_k^T y_k$$

Since $g_{k+1}^T y_k$ is scalar then

$$d_{k+1}^T y_k = -\, t g_{k+1}^T v_k - d_k^T y_k + \frac{\mu g_{k+1}^T d_k}{d_k^T g_k} d_k^T y_k + d_k^T y_k - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k} d_k^T y_k$$

Therefore μ $d_k^T y_k$ $d_k^T g_k$, $g_{k+1}^T d_k$ and $g_{k+1}^T y_k$ are scalars. Implies that

$$d_{k+1}^T y_k = -t g_{k+1}^T v_k = 0$$

Theorem3.4: Suppose that assumption (i) holds. Consider any conjugate gradient of the form (11) where d_k is a descent search direction and we take α_k obtained by strong Wolfe conditions (4) and (5). Then, Zoutendijk condition holds, i.e.

$$\sum_{k=0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

For proof see [23]. From the previous information, we can obtain the following convergence theorem of the conjugate gradient methods.

Theorem3.5: Suppose that assumption (i) is true. Consider conjugate gradient method of the form (11) and α_k is obtained by strong Wolfe conditions (4) and (5) and d_k is a descent search direction than either

$$\lim_{k\to\infty} \|g_k\| = 0 \quad \text{Or} \quad \sum_{k=0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty$$

Proof: To prove Theorem 2.3, we use contradiction. If Theorem 2.3 is not true, then there exists a constant $\rho > 0$, such that

$$\|g_i\| \geq \rho,\ \forall i \geq 0 \tag{13}$$

Rewrite (7) and (8), we get

$$d_{k+1} + \theta_k g_{k+1} = \beta_k^{MCD} d_k \tag{14}$$

Squaring the above equation, we have

$$\|d_{k+1}\|^2 = (\beta_k^{MCD})^2 \|d_k\|^2 - 2\theta_k g_{k+1}^T d_{k+1} - \theta_k^2 \|g_{k+1}\|^2 \tag{15}$$

Dividing both sides of equation (15) by $(g_{k+1}^T d_{k+1})^2$, therefore we end up with

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} = \frac{(\beta_k^{MCD})^2 \|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{2\theta_k}{g_{k+1}^T d_{k+1}} - \frac{\theta_k^2 \|g_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2}$$

$$= \frac{(\beta_k^{MCD})^2 \|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \left(\frac{1}{\|g_{k+1}\|} + \frac{\theta_k \|g_{k+1}\|}{g_{k+1}^T d_{k+1}}\right)^2 + \frac{1}{\|g_{k+1}\|^2} \leq \frac{(\beta_k^{MCD})^2 \|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2}$$

Substitute β_k^{MCD} we obtain

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{\left(1 - \frac{\mu g_{k+1}^T d_k}{d_k^T g_k}\right)^2 \|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2}$$
$$\leq \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{2\mu g_{k+1}^T d_k}{d_k^T g_k} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{\mu^2 (g_{k+1}^T d_k)^2}{(d_k^T g_k)^2} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2}$$
$$\leq \frac{\|d_k\|^2 d_k^T y_k}{(g_{k+1}^T d_{k+1})^2 d_k^T y_k} - \frac{2\mu g_{k+1}^T d_k}{d_k^T g_k} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{\mu^2 (g_{k+1}^T d_k)^2 d_k^T y_k}{(d_k^T g_k)^2 d_k^T y_k} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2}$$

We know that $g_{k+1}^T d_k \leq d_k^T y_k$ and by Wolfe condition $g_{k+1}^T d_k \geq c_2 {g_k}^T d_k \Rightarrow c_2 {g_k}^T d_k \leq d_k^T y_k \Rightarrow -c_2 {g_k}^T d_k \geq -d_k^T y_k$. This implies that $\|g_k\|^2 \geq \frac{-1}{c_2} d_k^T y_k \Rightarrow -c_2 \|g_k\|^2 \leq d_k^T y_k$ then

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{-\|d_k\|^2 d_k^T y_k}{(g_{k+1}^T d_{k+1})^2 c_2 \|g_k\|^2} - \frac{2\mu c_2 g_k^T d_k}{d_k^T g_k} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} - \frac{\mu^2 (g_{k+1}^T d_k)^2 d_k^T y_k}{(d_k^T g_k)^2 c_2 \|g_k\|^2} \frac{\|d_k\|^2}{(g_{k+1}^T d_{k+1})^2} + \frac{1}{\|g_{k+1}\|^2}$$

Since $c_2, \mu, \|g_k\|^2$ $d_k^T y_k, \|d_k\|^2, (d_k^T g_k)^2$ and $(g_{k+1}^T d_{k+1})^2$ are greater than zero, and $d_k^T g_k$ is a scalar, then

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \leq \frac{1}{\|g_{k+1}\|^2}$$

Hence $k=0$ the above inequality yield $\frac{\|d_1\|^2}{(g_1^T d_1)^2} \leq \frac{1}{\|g_1\|^2}$. Hence for all k, we conclude that $\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{1}{\|g_k\|^2}$

Therefore $\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=0}^{k} \frac{1}{\|g_i\|^2}$ So, by (13) $\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=0}^{k} \frac{1}{\rho^2} \Rightarrow \frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{1}{\rho^2} \sum_{i=0}^{k} 1 \Rightarrow \frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \frac{k}{\rho^2} \Rightarrow \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \frac{\rho^2}{k}$

When we add up both sides, then get $\sum_{k=0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \rho^2 \sum_{k=0} \frac{1}{k} = \infty \ \Rightarrow \sum_{k=0} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \geq \infty$

Which contradicts the Zountendijk condition in Theorem 3.4 The proof is then complete.

4. Numerical Results

This portion is devoted to testing the implementation of a new direction of search. We compare the fresh search direction of the conjugate gradient algorithm with Conjugate Gradient by (CD) and (MCD). The comparative tests involve well-known nonlinear problems (classical test function) with different functions $100 \leq N \leq 5000$. In all cases the stopping condition $\|g_{k+1}\| \leq 1 \times 10^{-5}$. The cubic interpolation procedure used function and gradient values in the line search routine. The results given in table 1 specifically quote the number of iteration NOIS and the number of function NOFS. Experimental results in table 1, confirm that the fresh direction of search of the conjugate gradient algorithm is superior to the standard algorithm (CD) and (MDC) concerning the number of iterations NOIS and the number of functions NOFS.

Table 1. Comparative Performance of Three Algorithms Standard CD, MCD and New Algorithm

No. of Test	Test Function	N	Standard Formula (CD)		Standard Formula (MCD)		New Formula	
			NOIS	NOFS	NOIS	NOFS	NOIS	NOFS
1	OSP	100	52	189	68	230	49	167
		2000	322	1027	359	1309	198	601
		5000	446	1444	1192	3975	256	781
2	Miele	100	68	246	209	493	39	108
		2000	70	261	1480	3551	45	127
		5000	74	284	2495	6003	45	127
3	G-Central	100	18	142	86	523	22	111
		2000	27	263	111	797	23	125
		5000	28	278	94	613	23	125
4	Beal	100	12	30	85	182	12	27
		2000	12	30	96	204	12	27
		5000	12	30	119	249	12	27
5	Sum	100	14	85	39	192	14	73
		2000	29	136	44	243	32	154
		5000	38	198	60	300	30	145
6	Cubic	100	14	40	1681	3387	13	34
		2000	15	44	1042	2120	13	34
		5000	15	44	891	1808	13	34
7	Fred	100	9	25	337	709	8	23
		2000	9	25	358	751	8	23
		5000	9	25	222	479	8	23
8	Rosen	100	30	85	472	972	30	82
		2000	30	85	853	1732	30	82
		5000	30	85	732	1490	32	87
9	TRI	100	76	153	244	489	75	151
		2000	417	835	1801	3603	417	835
		5000	674	1349	2405	4811	674	1349
10	Resip	100	5	16	52	133	5	15
		2000	5	16	52	133	5	15
		5000	6	18	52	133	6	17
Totals			**2566**	**7488**	**17731**	**41614**	**2149**	**5529**

Table 2. Comparing the rate of improvement between the new algorithm and the standard algorithm (CD) and (MCD)

Tools	CD-Method	New Method	MCD-Method	New Method
NOIS	100%	83.7490%	100%	12.1200%
NOFS	100%	73.8381%	100%	13.2854%

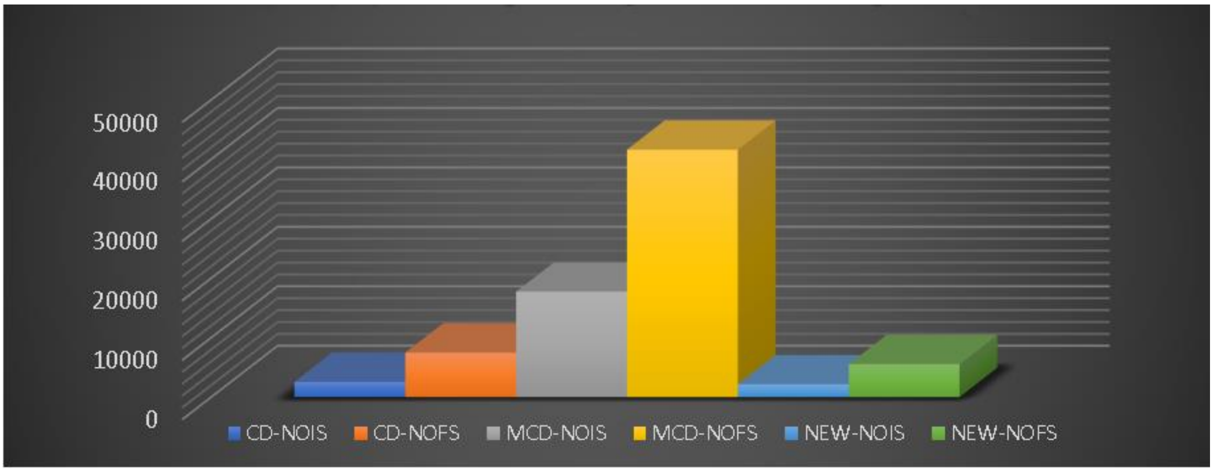

Figure 1. Percentage of comparison between algorithms

The above diagram is an explanation of Table 2.

Table 2 and figure 1 show the rate of improvement in the new algorithm with the standard algorithms (CD) and (MCD). The numerical results of the new method are better than the standard algorithm. As we notice that (NOIS), (NOFS) of the standard algorithm (CD) are about 100% that means the fresh algorithm has improved on the standard algorithm (CD) prorate (16.251%) in (NOIS) and prorate (26.1619%) in (NOFS) and standard algorithm (MCD) is about 100% that means the novel algorithm has improved on the classical algorithm (MCD) prorate (87.88%) in (NOIS) and prorate (86.2854%) in (NOFS) in general, the new (54.1446%) compared with classical algorithms (CD) and (MCD).

Figure 2 shows the comparison between the new algorithm and the classical algorithms (CD) and (MCD) according to the total number of iterations (NOIS) and the total number of functions (NOFS).

5. Conclusion

In this paper, we proposed a new spectral conjugate gradient method that has some properties of global convergence. Numerical results have shown that this new algorithm performs better than (CD) and (MCD). In the future, we can, and in some way, we proposed many new spectral conjugate gradients of unconstrained optimization.

Abbreviations

CJG conjugation gradient, SCJG spectral conjugated gradient, Min minimum, $g_k=\nabla f(x_k)$ gradient, β_k A parameter has different formulas, ELS exact line search, MCD modified of conjugate gradient, SWL strong Wolfe line search, NOIS number of iterations, NOFS number of functions.

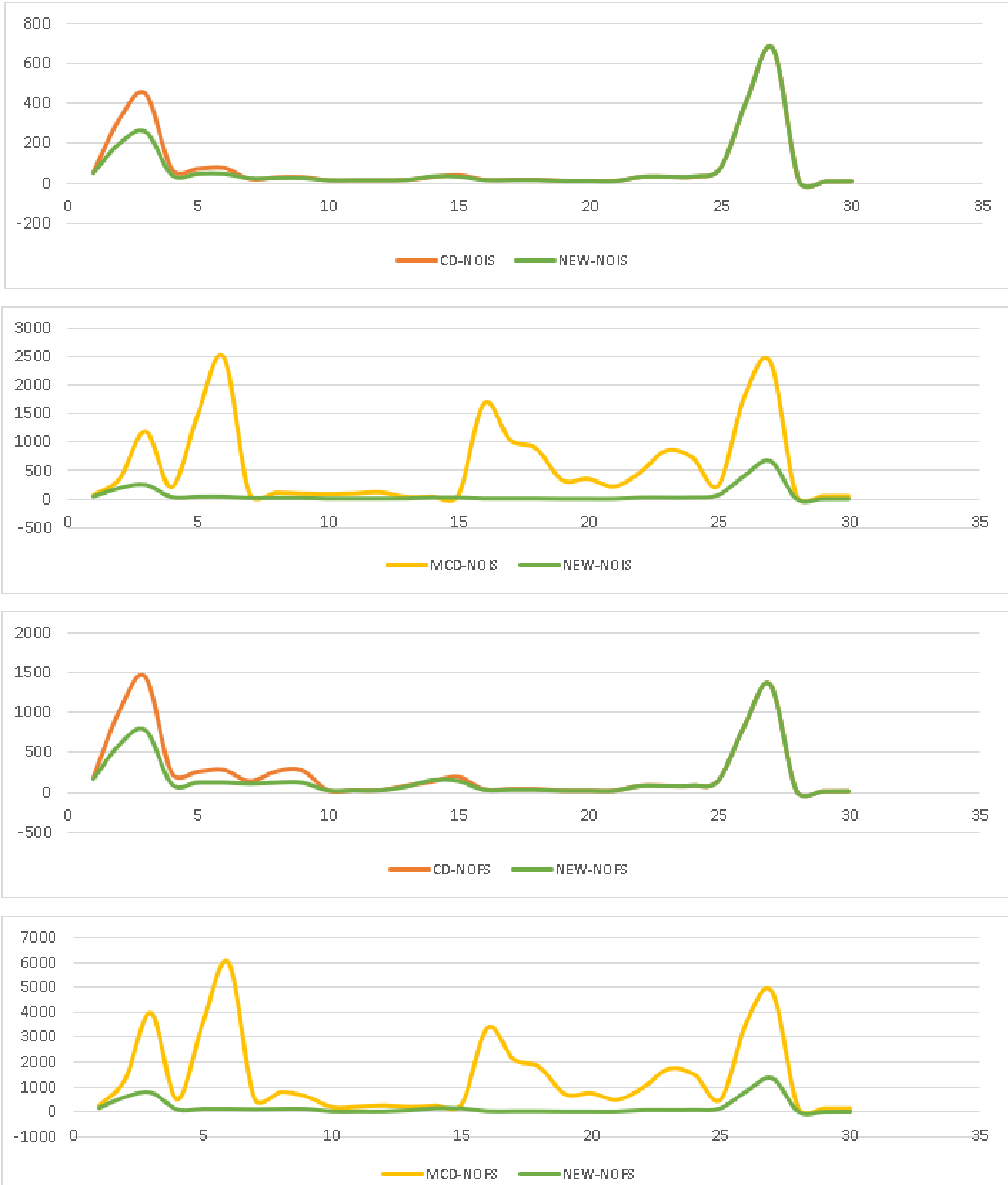

Figure 2. Preforming algorithms for the NOIS and the NOFS

REFERENCES

[1] F. N. Jardow and G.M. Al-Naemi. A new hybrid conjugates gradient algorithm as a convex combination of MMWU and RMIL nonlinear problems, Journal of Interdisciplinary Mathematics, vol. 24, no. 3 pp. 637-655, 2021. DOI: 10.1080/09720502.2020.1815346

[2] R. Mohd, A. Abdalrahman, M. Mustafa and M., Ismail. The convergence properties of a new type of conjugate gradient methods, Appl. Math. Sci., vol. 8, no. 1, pp. 33-44, 2014. http://dx.doi.org/10.12988/ams.2014.310578

[3] Mohammed S., M. Mamat, K. Kamfa, M. Danlami. A Descent Modification of Conjugate Gradient Method for Optimization Models, Iraqi Journal of Science, vol. 61, no. 7, pp. 1745-1750, 2020. DOI: 10.24996/ijs.2020.61.7.23

[4] M. R. Hestenes and E. Steifel. Method of the conjugate

gradient for solving linear equations, J. Res. Nat. Bur. Stand, vol. 6, no. 49, pp. 409-436, 1952. https://nvlpubs.nist.gov/nistpubs/jres/049/jresv49n6p409_A1b.pdf

[5] R. Fletcher, and C. Reeves. Function minimization by conjugate gradients, Comput. J., 7, pp. 149-154, 1964. https://academic.oup.com/comjnl/article/7/2/149/335311

[6] E. Polak and G. Ribiere. Note sur la convergence de directionsconjugees, Rev. Francaise Informat Recherche Operationelle. 3E Annee, vol. 3, no.16, pp. 35-43, 1969. http://www.numdam.org/item?id=M2AN_1969__3_1_35_0

[7] Y. Liu and C. Storey. Efficient generalized conjugate gradient algorithm's part 1: Theory, J. Comput. Appl. Math., vol. 69, no. 1, pp. 129-137, 1991. https://link.springer.com/article/10.1007/BF00940464

[8] Y.H. Dai & Y. Yuan. A nonlinear conjugate gradient with a strong global convergence property, SIAM J. Optim., vol. 10, no. 1 pp. 177–182, 1999. https://doi.org/10.1137/S1052623497318992

[9] R., Fletcher. "Conjugate Direction Methods" Practical Methods of Optimization Unconstrained Optimization (1). John Wiley & Sons, New York, USA, 2nd edition, 1987. https://www.wiley.com/en-ba

[10] Zhang, L. A derivative-free conjugate residual method using secant condition for general large-scale nonlinear equations. Numer. Algorithms, 83, pp. 1277–1293, 2020. https://doi.org/10.1007/s11075-019-00725-7

[11] M. A. Anwer. Enhance the Efficiency of RMIL's Formula for Minimum Problem, Journal of the University of Duhok, (Pure and Eng. Sciences), University of Duhok, vol. 24, no. 2, pp. 62-70, 2021. DOI: https://doi.org/10.26682/sjuod.2021.24.2.7

[12] M. A. Anwer and Sh. S. Gazi. Global convergence of new three terms conjugate gradient for unconstrained optimization, General Letters in Mathematics (GLM), Refaad in Jordan, vol. 11, no. 1, pp. 1-9, 2021. doi: 10.31559/glm2021.11.1.1

[13] B. Jonathan and B. J. M. Jonathan. Two-point step size gradient methods, IMA Journal of Numerical Analysis, vol. 8, no. 1, pp.141-148, 1988. https://aip.scitation.org/doi/abs/10.1063/1.4952138

[14] Andrei. Scaled conjugate gradient algorithm for unconstrained optimization. Computational Optimization and Application, vol. 38, no. 3, pp. 401-416, 2007. https://doi.org/10.1007/s10589-007-9055-7

[15] J. Huabin, D. Songhai, Z. Xiaodong, and W. Zhong. Global convergence of a modified spectral conjugate gradient method. Journal of Applied mathematics, pp. 1-13, 2012. https://doi.org/10.1155/2012/641276

[16] R. Marcos. The Barzilai and Borwein gradient method for the large-scale unconstrained minimization problem, SIAM Journal on Optimization, vol. 7, no. 1, pp. 26-33, 1997. https://doi.org/10.1137/S1052623494266365

[17] H. Sheekoo, Gh. M. Al-Naemi. Good Characteristics of The New Spectral Conjugate Gradient Method for Unconstrained Optimization, 2nd International Conference on Physics and Applied Sciences, Baghdad, Iraq, 26-27, May 2021, pp. 1-10, doi:10.1088/1742-6596/1963/1/012079

[18] Bakhtawar, S. Zabidin, and A. Ahmad. A New Modified Three-Term Hestenes–Stiefel Conjugate Gradient Method with Sufficient Descent Property and Its Global Convergence, Journal of Optimization, pp. 1-13, 2018. https://doi.org/10.1155/2018/5057096.

[19] J. Jinbao, H. Lin, J. Xianzhen. A hybrid conjugate gradient method with descent property for unconstrained optimization. J. Comput. Appl. Math., vol. 39, no. 3-4, pp. 1281-1290, 2015. https://doi.org/10.1016/j.apm.2014.08.008

[20] J. Jinbao, Y. Lin, J. Xianzhen, L. Pengjie, and L. Meixing. A Spectral Conjugate Gradient Method with Descent Property, Mathematics vol. 280 no. 8, pp 1-13, 2020. https://doi.org/10.3390/math8020280

[21] Ma Sh., and Anping W. Declining species full CD conjugate gradient method and its convergence, J Chongqing Technol Business Univ. (Nat Sci Ed), vol. 32, no. 1, pp. 1-4, 2015.

[22] Y. Dai and. L. Liao. New conjugacy conditions and related nonlinear conjugate gradient methods. Applied Mathematics and Optimization, vol. 43, pp. 87-101, 2001. https://link.springer.com/article/10.1007/s002450010019

[23] Zoutendijk G. Nonlinear Programming Computational Methods, Integer and Nonlinear Programming, J. Ababie (Ed), North-Ho land. Amsterd, 143, pp. 37-86, 1970. ttps://www.oalib.com/references/13789575