RAIRO Operations Research
Will be set by the publisher

A NEW FAMILY OF HYBRID THREE-TERM CONJUGATE GRADIENT METHOD FOR UNCONSTRAINED OPTIMIZATION AND ITS APPLICATION TO REGRESSION ANALYSIS

SABRINA BEN HANACHI[1], BADREDDINE SELLAMI[2] AND MOHAMMED BELLOUFI[3]

Abstract. We know a large variety of conjugate gradient algorithms (CG) for solving unconstrained optimization problems. In this paper, based on the three famous Liu-Storey (LS), Fletcher-Reeves (FR) and Polak-Ribiére-Polyak (PRP) conjugate gradient methods , a new hybrid CG projection method is proposed. Furthermore, the search direction satisfies the sufficient descent condition independent of the line search. Also, under the Wolfe line search we prove the global convergence of the new method. Numerical experiments are performed and reported, which show that the proposed method is efficient and promising. The application of the proposed method for solving regression models of COVID-19 is provided.

Keywords: Unconstrained optimization, hybrid conjugate gradient method, sufficient descent, convex combination, global convergence.

Mathematics Subject Classification. 65K05, 90C25, 90C26, 90C27, 90C30.

1. Introduction

Unconstrained optimization is a branch of optimization in which we minimize an objective function that depends on real variables with the total absence of restrictions on their values of those variables, we consider the general unconstrained optimization problems as follows :

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

where $f : \mathbb{R}^n \to \mathbb{R}$ is the continuously differentiable function and its first derivative is represented by $g(x) = \nabla f(x)$. Though, many optimization algorithms that are robust with rapid convergence are available to solve the above nonlinear optimization model, many researchers still refer to the conjugate gradient algorithm (CG) because it uses low memory and good convergence properties. This method was first established by Hestenes and Stiefel [15] and is used to solve unconstrained linear optimization problems. Then, in 1964, Fletcher and Reeves [12] extended the form of the conjugate gradient method to solving unconstrained nonlinear minimization problems. The results of the expansion inspired researchers to suggest a new conjugate gradient method which has good computational performance and at the same time good convergence properties

2022.

[1] Laboratory Informatics and Mathematics (LiM). Mohamed Cherif Messaadia University, Souk-Ahras, Algeria.
e-mail: `s.benhanachi@univ-soukahras.dz`
[2] Laboratory Informatics and Mathematics (LiM). Mohamed Cherif Messaadia University, Souk-Ahras, Algeria.
e-mail: `bsellami@univ-soukahras.dz`
[3] Laboratory Informatics and Mathematics (LiM). Mohamed Cherif Messaadia University, Souk-Ahras, Algeria.
e-mail: `m.belloufi@univ-soukahras.dz`

[5]. Generally, the iterates of the CG methods are usually determined through the following recursive computational scheme :

$$x_{k+1} = x_k + \alpha_k d_k. \quad k = 0, 1, \dots n, \tag{2}$$

and

$$d_{k+1} = \begin{cases} -g_{k+1}, & \text{if } k = 0, \\ -g_{k+1} + \beta_k d_k, & \text{if } k \geq 1, \end{cases} \tag{3}$$

where x_k is the current iteration, g_k is the gradient of f at the point x_k, d_k is the search direction, $\beta_k \in \mathbb{R}$ is the conjugate parameter which characterizes different versions of the CG methods and $\alpha_k > 0$ is the step size that can be obtained by many line search techniques. Among them, exact line search, weak Wolfe line search, or strong Wolfe line search but in our research we use strong Wolfe line search which is defined by the following conditions [22], [23] :

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

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

where scalars δ and σ satisfy $0 < \delta \leq \sigma < 1$.

As we said the CG algorithms differ by the choice of the coefficient β_k. The most common standard CG methods are are the Hestenes-Stiefel (HS) method [15], the Fletcher-Reeves (FR) method [12], the Conjugate-Descent (CD) method [13], the Dai-Yuan (DY) method [7], the Liu-Storey (LS) method [18] and the Polak-Ribiére-Polyak (PRP) method [19, 20], respectively and are defined as :

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

$$\beta_k^{DY} = \frac{||g_{k+1}||^2}{y_k^T s_k} \qquad \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k} \qquad \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{||g_k||^2},$$

where $y_k = g_{k+1} - g_k$, $s_k = \alpha_k d_k$ and $||.||$ stands for the Euclidean norm, if f is strictly convex quadratic function, all the above methods are equivalent, while behaves differently for general non quadratic functions.

One of the important classes of CG methods is the hybrid conjugate gradient algorithms. The hybrid computational schemes perform have better computational performances and strong convergence properties better than conventional CG methods because they take the advantages of the two parameters that were used to build it. Thus, for this reason many researchers carred about the hybrid or mixed conjugate gradient methods. Djordjevic' [9], proposed the following hybrid method: $\beta_k^{hyb} = \theta_k \beta_k^{FR} + (1 - \theta_k)\beta_k^{HS}$, Andrei [4], proposed the following hybrid method: $\beta_k^{c} = \theta_k \beta_k^{DY} + (1 - \theta_k)\beta_k^{HS}$, Li and Sun [16], proposed the following hybrid method: $\beta_k^{N} = \theta_k \beta_k^{MMWU} + (1 - \theta_k)\beta_k^{FR}$, Liu, J.K. and Li, Sij [17], proposed the following hybrid method: $\beta_k^{hyb} = \theta_k \beta_k^{DY} + (1 - \theta_k)\beta_k^{LS}$. In addition, Sabrina and al [14] propose a new hybrid CG method based on combination of FR, PRP and DY conjugate gradient algorithms in which

$$\beta_k^{hyb} = \delta_k \beta_K^{FR} + \gamma_k \beta_K^{PRP} + (1 - \delta_k - \gamma_k)\beta_K^{DY},$$

where

$$\gamma_k = -\frac{g_k^T g_{k+1}||g_k||^2 + \delta_k(y_k^T s_k - ||g_k||^2)||g_{k+1}||^2}{(g_{k+1}^T y_k)(y_k^T s_k) - ||g_{k+1}||^2||g_k||^2}, \quad \text{where, } 0 < \delta_k < 1.$$

Inspired by this research, in this study we propose a new hybrid CG method based on combination of LS, FR and PRP conjugate gradient algorithms for solving unconstrained optimization problems. In addition, in this study, we also apply the new method for solving a model of COVID-19 outbreak around the globe in which the data is taken from January to September 2020.

The paper is structured as follows. In Section 2, we will describe the proposed method with its corresponding algorithm, and further established the descent condition and convergence under inexact line search. In Section 3, we present the numerical experiments to show the efficiency of our new method in Section 4. Finally, a breif conclusion is drawn in section 5.

2. Proposed method, algorithm

In this paper we propose another combination of LS, FR and PRP conjugate gradient algorithms. We use the following conjugate gradient parameter:

$$\beta_k^{New} = \delta_k \beta_k^{LS} + \gamma_k \beta_k^{FR} + (1 - \delta_k - \gamma_k)\beta_k^{PRP} \tag{6}$$

As a consequence, the direction d_k is given by :

$$d_{k+1} = -g_{k+1} + \beta_k^{New} d_k. \tag{7}$$

The parameters δ_k, γ_k in (6) satisfying $0 \leq \delta_k, \gamma_k \leq 1$ which will be determined in a particular way that will be described later. It should be noted that :

- If $\delta_k = 1$ and $\gamma_k = 0$, then $\beta_k^{New} = \beta_k^{LS}$.
- If $\delta_k = 0$ and $\gamma_k = 1$, then $\beta_k^{New} = \beta_k^{FR}$.
- If $\delta_k = 0$ and $\gamma_k = 0$, then $\beta_k^{New} = \beta_k^{PRP}$.
- If $\delta_k = 0$ and $0 < \gamma_k < 1$, then $\beta_k^{New} = \gamma_k \beta_k^{FR} + (1 - \gamma_k)\beta_k^{PRP}$ i.e β_k^{New} is a convex combination of β_k^{FR} and β_k^{PRP}. see. [8]
- If $\gamma_k = 0$ and $0 < \delta_k < 1$, then $\beta_k^{New} = \delta_k \beta_k^{LS} + (1 - \delta_k)\beta_k^{PRP}$ i.e β_k^{New} is a convex combination between β_K^{LS} and β_K^{PRP}. see [1].
- If $1 - \delta_k - \gamma_k = 0$, $0 < \delta_k, \gamma_k < 1$, then $\gamma_k = 1 - \delta_k$. Then $\beta_k^{New} = \delta_k \beta_k^{LS} + (1 - \delta_k)\beta_k^{FR}$ i.e β_K^{New} is a convex combination between β_k^{LS} and β_k^{FR}. see. [10]

Finally, if $\delta_k \in]0, 1[$, $\gamma_k \in]0, 1[$ and $0 < \delta_k + \gamma_k < 1$, then we have a **new hybrid CG method** as a convex combination of three methods "LS, FR and PRP". From (6) and (7), it is clear that:

$$d_{k+1} = \begin{cases} -g_{k+1}, & k = 1 \\ \\ -g_{k+1} + \delta_k \frac{g_{k+1}^T y_k}{-d_k^T g_k} d_k + \gamma_k \frac{||g_{k+1}||^2}{||g_k||^2} d_k + (1 - \delta_k - \gamma_k)\frac{g_{k+1}^T y_k}{||g_k||^2} d_k, & k > 1. \end{cases} \tag{8}$$

To select the parameters δ_k and γ_k we use the traditional conjugacy condition i.e. ($d_{k+1}^T y_k = 0$), so we have the following Lemma.

Lemma 1. *If the conjugacy condition $d_{k+1}^T y_k = 0$ is satisfied at every iteration, we get*

$$\delta_k = \frac{\left(g_{k+1}^T y_k(||g_k||^2 - d_k^T y_k) - \gamma_k (g_{k+1}^T g_k)(d_k^T y_k)\right)}{(||g_k||^2 + d_k^T g_k)(-g_{k+1}^T y_k)(d_k^T y_k)} d_k^T g_k, \qquad \textit{where} \;\; 0 < \gamma_k < 1 \tag{9}$$

Proof. multiplying (8) by y_k from the left and using the conjugacy condition, we obtain

$$0 = -g_{k+1}^T y_k + \delta_k \beta_k^{LS} d_k^T y_k + \gamma_k \beta_k^{FR} d_k^T y_k + (1 - \delta_k - \gamma_k)\beta_k^{PRP} d_k^T y_k,$$

$$g_{k+1}^T y_k = \delta_k \left(\frac{g_{k+1}^T y_k}{-g_k^T d_k} \right) d_k^T y_k + \gamma_k \left(\frac{||g_{k+1}||^2}{||g_k||^2} \right) d_k^T y_k + (1 - \delta_k - \gamma_k) \left(\frac{g_{k+1}^T y_k}{||g_k||^2} \right) d_k^T y_k.$$

Finally, after some algebra we have:

$$\delta_k = \frac{\left(g_{k+1}^T y_k(||g_k||^2 - d_k^T y_k) - \gamma_k(g_{k+1}^T g_k)(d_k^T y_k)\right)}{(||g_k||^2 + d_k^T g_k)(-g_{k+1}^T y_k)(d_k^T y_k)} d_k^T g_k, \qquad \text{where} \ \ 0 < \gamma_k < 1.$$

□

Next, we give the algorithm of our proposed method below

Algorithm (SCH)

Step1.: *Initialization.* Given $x_0 \in \mathbb{R}^n$ and the parameters $0 < \delta \leq \sigma < 1$. Set $k = 0$. Compute $f(x_0)$, $g_0 = \nabla f(x_0)$. Consider $d_0 = -g_0$, set the initial guess: $\alpha_0 = 0$ and $\gamma_k = 0.5$.

Step 2.: *Test a criterion for stopping iterations,* if $||g_k|| < 10^{-6}$, then stop. Else continue with **Step 3**.

Step 3.: *Line search.* Compute $\alpha_k > 0$ by the strong Wolfe line search, i.e., α_k satisfies (4), (5).

Step 4.: *Generate.* $x_{k+1} = x_k + \alpha_k d_k$. Compute $f(x_{k+1})$, $g_{k+1} = \nabla f(x_{k+1})$ and $y_k = g_{k+1} - g_k$.

Step 5.: *Compute* δ_k as in E.q (9).

Step 6.: Calculate β_k^{New} by E.q (6).

Step 7.: *Computation of the search direction.* Compute $d = -g_{k+1} + \beta_k^{New} d_k$. If the restart criterion of Powell

$$|g_{k+1}^T g_k| \geq 0.2||g_{k+1}||^2,$$

is satisfied, then $d_{k+1} = -g_{k+1}$. Otherwise define $d_{k+1} = d$.

Step 8.: Put $k = k + 1$ and continue with **Step 2**.

2.1. The Sufficient Descent Condition

In this study, we will etablish the sufficient descent of our new method which plays an important role in the global convergence analysis. Thus, we need the following assumptions,

Assumption 2. *The level set $S = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$ is bounded, i.e. there exists a constant $B > 0$, such that*

$$||x|| \leq B, \ \ \textit{for all} \ \ x \in S. \tag{10}$$

Assumption 3. *In a neighborhood $\mathcal{N}$ of S the function f is continuously differentiable and its gradient $\nabla f(x)$ is Lipschitz continuous, i.e. there exists a constant $0 < L < \infty$ such that*

$$||\nabla f(x) - \nabla f(y)|| \leq L||x - y||, \quad \textit{for all} \ \ x, y \in \mathcal{N}. \tag{11}$$

Under Assumptions (2) and (3) on f, there exists a constant $\Gamma \geq 0$, such that

$$||\nabla f(x)|| \leq \Gamma, \tag{12}$$

for all $x \in S$ [3].

The following Theorem proves that the search direction obtained by the new method satisfies the sufficient descent condition.

Theorem 4. *Let generated sequences* $\{g_k\}$ *and* $\{d_k\}$ *by a* ***SCH*** *algorithm and let* $(g_{k+1}^T y_k)(g_{k+1}^T d_k) < 0$*. Also let* $\{||s_k||\}$ *tend to zero, and let there exist some constants* $\delta_1, \delta_2 > 0$*, such that*

$$||g_k||^2 \geq \delta_1 ||d_k||^2, \tag{13}$$

$$||g_{k+1}||^2 \leq \delta_2 ||d_k||, \tag{14}$$

and if $0 < \delta_k, \gamma_k < 1$*, then* d_k *satisfies the sufficient descent condition for all* k*.*

Proof. We will show that d_k satisfies the sufficient descent condition for $k = 0$, the proof is trivial one, i.e. $d_0 = -g_0$, it holds $g_0^T d_0 = -||g_0||^2$. Next, we need to show that for all $k \geq 1$, d_k satisfies the sufficient descent condition, we have

$$d_{k+1} = -g_{k+1} + \delta_k \frac{g_{k+1}^T y_k}{-d_k^T g_k} d_k + \gamma_k \frac{||g_{k+1}||^2}{||g_k||^2} d_k + (1 - \delta_k - \gamma_k) \frac{g_{k+1}^T y_k}{||g_k||^2} d_k. \tag{15}$$

Multiplying the above equation by g_{k+1}^T from the left, we get

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

since $0 < \delta_k, \gamma_k < 1$ and $(g_{k+1}^T y_k)(g_{k+1}^T d_k) < 0$, we obtain

$$g_{k+1}^T d_{k+1} \leq -||g_{k+1}||^2 + \frac{||g_{k+1}||^2}{||g_k||^2} g_{k+1}^T d_k.$$

Observe that, since $g_{k+1}^T d_k = y_k^T d_k + g_k^T d_k$ and since $g_k^T d_k < 0$, then $g_{k+1}^T d_k < y_k^T d_k$, so the last relation becomes

$$\begin{aligned} g_{k+1}^T d_{k+1} &\leq -||g_{k+1}||^2 + \frac{||g_{k+1}||^2}{||g_k||^2} y_k^T d_k. \\ &\leq -||g_{k+1}||^2 + \frac{||g_{k+1}||^2}{||g_k||^2} ||y_k|| \; ||d_k||, \end{aligned}$$

in the other hand, according to (11), we get

$$||y_k|| = ||g_{k+1} - g_k|| \leq L||s_k||,$$

by using the above equation, (13) and (14), we get

$$\begin{aligned} g_{k+1}^T d_{k+1} &\leq -||g_{k+1}||^2 + \frac{\delta_2}{\delta_1} ||d_k||, \\ &\leq -||g_{k+1}||^2 + \frac{\delta_2}{\delta_1} \alpha ||s_k||. \end{aligned}$$

Due to the assumption that $||s_k|| \to 0$, so $\frac{\delta_2}{\delta_1}\alpha||s_k|| \to 0$, so there exists ω such that $0 < \omega \leq 1$. Therefore

$$\frac{\delta_2}{\delta_1} \alpha ||s_k|| \leq \omega ||g_{k+1}||^2.$$

This gives,

$$g_{k+1}^T d_{k+1} \leq -||g_{k+1}||^2 + \omega ||g_{k+1}||^2,$$

i.e.

$$g_{k+1}^T d_{k+1} \leq -(1 - \omega) \; ||g_{k+1}||^2. \tag{16}$$

So, it is proved that d_{k+1} satisfied the sufficient descent condition. □

2.2. Convergence analysis

In the analysis below, we will establish the global convergence properties of the proposed method. First, we need the following Proposition and Zoutendijk conditions.

Proposition 5. *[17] Suppose that Assumptions (2) and (3) hold, if d_k is a descent direction and α_k satisfies*

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k, \ \ 0 < \sigma < 1. \tag{17}$$

Then,

$$\alpha_k \geq \frac{(1-\sigma)}{L} \, \frac{|d_k^T g_k|}{||d_k||^2}. \tag{18}$$

Proof. It follows (17), the Lipschitz condition, the Cauchy-Bunyakovsky-Schwartz inequality, it holds that

$$-(1-\sigma) d_k^T g_k \leq \sigma d_k^T g_k - d_k^T g_k \leq d_k^T (g_{k+1} - g_k) \leq d_k^T L \alpha_k d_k \leq L \alpha_k ||d_k||^2. \tag{19}$$

Since d_k is a descent direction and $\sigma < 1$, formula (18) holds immediately. □

According to the Proposition (5), Assumptions (2) and (3), the strong Wolfe conditions and (16), we conclude that α_k which obtained in our new method is not equal to zero i.e there exists a constant $\lambda > 0$ such that

$$\alpha_k \geq \lambda, \text{ for all } k \geq 0. \tag{20}$$

The Zoutendijk condition [27] is often utilized to prove the global convergence of the CG method. The following Lemma shows that Zoutendijk condition holds for the proposed method under the strong Wolfe conditions of formulas (4) and (5).

Lemma 6. *Suppose that Assumptions (2) and (3) hold. Consider common iterate (2), where d_k is a descent direction and α_k is determined by the Wolfe line search (4) and (5). Then the Zoutendijk condition*

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{||d_k||^2} < \infty. \tag{21}$$

The following Theorem gives the global convergence of **SCH** method.

Theorem 7. *Suppose that Assumption (2) and (3) hold, let $\{x_k\}$ be generated by **SCH** Algrithm. Then*

$$\lim_{k \to \infty} \inf ||g_k|| = 0. \tag{22}$$

Proof. We prove this theorem by contradiction. Suppose that formula (22) is not true. Then there exists a constant $c > 0$ in which

$$||g_k|| \geq c, \ \ \forall k \geq 1. \tag{23}$$

From theorem (4) it follows that

$$g_k^T d_k \leq -K ||g_k||^2, \text{ for all } \ K, \tag{24}$$

in the other hand, according to (11), we get

$$||y_k|| = ||g_{k+1} - g_k|| \leq L ||s_k|| \leq LD, \tag{25}$$

where $D = \max\{||x-y||, x, y \in \mathcal{N}\}$ is the diameter of $\mathcal{N}$ and $s_k = x_{k+1} - x_k$.

We have

$$d_{k+1} = -g_{k+1} + \beta_K^{New} d_k.$$
$$||d_{k+1}|| \leq ||g_{k+1}|| + |\beta_K^{New}|||d_k||. \quad (26)$$

From (6), we obtain

$$\begin{aligned}
|\beta_K^{New}| &\leq |\beta_K^{LS}| + |\beta_K^{FR}| + |\beta_K^{PRP}|, \\
&= \frac{|g_{k+1}^T y_k|}{|d_k^T g_k|} + \frac{||g_{k+1}||^2}{||g_k||^2} + \frac{|g_{k+1}^T y_k|}{||g_k||^2}. \\
&\leq \frac{||g_{k+1}||\,||y_k||}{K||g_k||^2} + \frac{||g_{k+1}||^2}{||g_k||^2} + \frac{||g_{k+1}||\,||y_k||}{||g_k||^2}. \\
&\leq \frac{\Gamma LD}{Kc^2} + \frac{\Gamma^2}{c^2} + \frac{\Gamma LD}{c^2} = M,
\end{aligned}$$

where the first inequality follows from $0 < \delta_k, \gamma_k < 1$ and $1 - \delta_k - \gamma_k < 1$, the second inequality applies the Cauchy Schwarz inequality and (24), the last inequality uses (12), (23) and (25).
Thus, it follows from (3) and (20) that

$$||d_{k+1}|| \leq ||g_{k+1}|| + |\beta_K^{New}|\; ||d_k|| \leq ||g_{k+1}|| + \frac{|\beta_K^{New}|||s_k||}{\alpha_k} \leq \Gamma + \frac{MD}{\lambda} = W,$$

which implies that

$$\begin{aligned}
||d_{k+1}|| \leq W \implies & \sum_{k\geq 1} \frac{1}{||d_k||^2} = +\infty \\
\implies & \sum_{k\geq 1} \frac{(g_k^T d_k)^2}{||d_k||^2} = +\infty.
\end{aligned}$$

Which contradicts Lemma (6), hence, (23) does not hold, and the claim (22) is proved. □

3. Numericals Analysis

This section is devoted to test the implementation of the new method. Basing on this, we compare the computational performance of the prorposed method with some known algorithms such as the LS, FR and PRP. For this comparisons, we consider 400 unconstrained optimization test problems from CUTE library [6] along with other large-scale optimization problems presented in [2]. We selected 30 large-scale unconstrained optimization problems in extended or generalized form. What is more, each problem is tested for a number of variables: $n = 2, 4, \ldots, 25000$. The analysis was based on the number of iterations and central processing unit CPU time. For the numerical tests, the iterations are terminated when $||g_k||_\infty < 10^{-6}$, where $||.||_\infty$ is the maximum absolute component of a vector, the parameters in the strong Wolfe line searches are chosen to be $\delta = 10^{-3}$ and $\sigma = 10^{-4}$ and the hybridization parameter $\gamma_k = 0.5$. All programs are written in Matlab and compiler settings on the PC machine with Intel(R) Core(TM) i3- 4030U CPU @ 1.90 GHz 1.90 GHz processor and 4GB RAM memory and windows 7 professional system.

The comparisons of these methods are given in the following two sides. On the first side, for the *ith* problem, let f_i^{M1} and f_i^{M2} be the optimal value found by $M1$ method and $M2$ method, respectively. We say that, for the praticular problem *ith*, the performance of $M1$ method was better than the performance of $M2$ method if

$$|f_i^{M1} - f_i^{M2}| < 10^{-3}, \quad (27)$$

and number of iterations, or CPU time of $M1$ method is less than those of $M2$ method, respectively. In the other side, in order to obtain complete comparisons in CPU time, we used the profile of Dolan and Moré [11] to evaluate and compare the performance of the set of methods S on a test set P. Assume that S consists of n_s methods, P consists of n_p problems. For each problem $p \in P$ and method $s \in S$, denote $t_{p,s}$ be the computing time required to solve problem p by method s, and the comparison between different methods is based on the performance ratio deffined by $r_{p,s} := t_{p,s} | \min_{s \in S} t_{p,s}$. Then the performance profile is given by

$$\rho_s(\tau) = \frac{1}{n_p} size\{p \in P : \log_2 r_{p,s} \leq \tau\}, \ \ \forall \tau \in \mathbb{R}^+,$$

where $\rho : \mathbb{R} \to [0, 1]$ and $1 \leq s \leq n_s$. The function ρ_s is the distribution function for the performance ratio. Moreover, ρ_s for a method is a nondecreasing, piecewise constant function, continuous from the right at each breakpoint. Note that $\rho_s(\tau)$ is the probability for method $s \in S$ that $\log_2 r_{p,s}$ is within a factor $\tau \in \mathbb{R}^+$ of the best possible ratio. Obviously, when τ takes certain value, a method with high value of $\rho_s(\tau)$ is preferable or represent the best method.

Figures 1 and 2 represent the performance profiles of the new method versus LS, FR and PRP based on the CPU time and number of iterations ,respectively. From the two figures, we can see that the new method is superior to the other conjugate gradient methods on the testing problems.

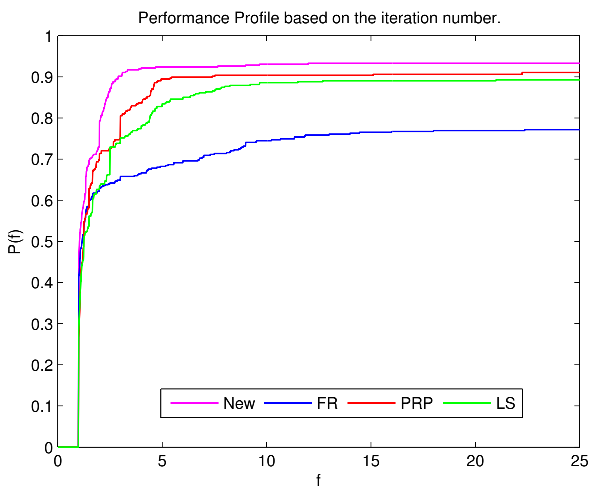

Figure 1. Performance Profile based on the iteration number.

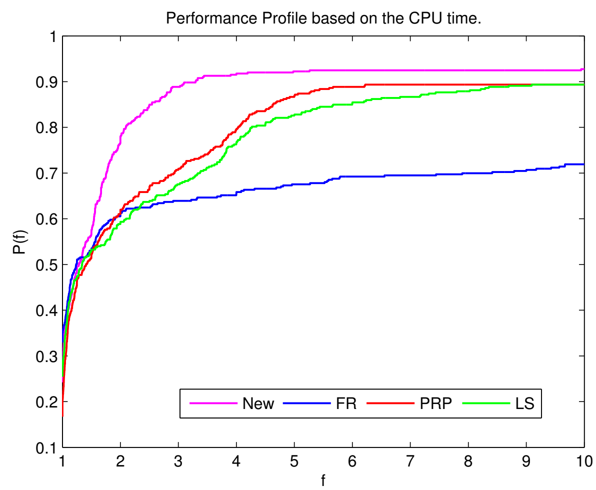

Figure 2. Performance Profile based on the CPU time.

4. Application of the New CG method to Regression Analysis

Novel coronavirus-19 (COVID-19) is a new chain of corona group viruses that had not been recognized in human history earlier than December 2019. It was first discovered in Wuhan, China [25] and has spread to various urban areas in China as well as approximately 196 different countries of the world. It has since been declared an outbreak by World Health Organization (WHO). It is difficult to take a single point of view on this virus's origin. It can be due to a seafood market exchange, or the people's migration from one location to another, or the transmission from animals to humans. Most people infected by the virus will develop mild to moderate symptoms, such as mild fever, cold, difficulty in breathing, and recover without special treatment. According to data reported by the WHO, on 20 October 2020, the laboratory declared that the number of confirmed cases is over 40 million with more than one million deaths recorded in 215 regions and countries around the world since the disease was first reported in Wuhan.

Mathematical modeling plays an important role in describing the epidemic of infectious diseases and thus overcome it at an early stage. Recently, numerous studies modeled various aspects of the coronavirus outbreak, and application of numerical methods on some COVID-19 models was also studied [21, 26]. This paper aims to investigate the performance of the proposed method on a parameterized COVID-19 regression model. For deriving the COVID-19 regression model, the study will consider the total confirmed cases of the infection from january 2020 until september 2020. The obtained data would be transformed into an unconstrained optimization problem which would later be solved using the proposed method.

Regression analysis is one of the most effective statistical modeling tools for modeling problems in the applied sciences, physical sciences, management, and many others. Based on the previous description, we can describe regression analysis as a statistical technique used to estimate the relationship between a dependent variable and one or more independent variables. The function of regression analysis is defined as follows :

$$y = h(x_1, x_2, \dots, x_p + \epsilon), \tag{28}$$

where $x_i, i = 1, 2, \dots, p, p > 0$ is the predictor, y is the response variable, and ϵ is the error. For any problem related to regression analysis the linear regression function can be derived by computing y such that:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_p x_p + \epsilon. \tag{29}$$

with $a_0, \dots, a_p$ representing the regression parameters, these parameters are estimated to minimize the error ϵ value. This scheme is often used when the relationship between x and y is approximated by a straight line. However, this cases rarely occur because most problems are often nonlinear in nature. Therefore, the nonlinear regression scheme is frequently used. In this paper, we considered the nonlinear regression one.

To derive the approximate function, we consider the data from the global confirmed cases of COVID-19 from January to September, 2020. Table 1 shows the description of the process that is considered from the statistics obtained from the World Health Organization [24]. We have data for nine months (Jan–Sept), the months of data collection would be denoted by x- variable and the confirmed cases corresponding to these months would be denoted by the y-variable. However, the data of eight months (Jan to Aug) would be considered for fitting the data, while the data for September 2020 would be reserved for error analysis.

Table 1. Statistics of confirmed cases of COVID-19, Jan–Sept, 2020

Monthly data (Jan–Sept) (x)	Data of confirmed COVID-19 cases (y)	Statistics of COVID-19 in %
1	2010	0.16
2	1852	0.14
3	58,863	4.7
4	74,019	6.0
5	115,577	9.3
6	172,158	13.9
7	293,238	23.6
8	269,338	21.7
9	254,423	20.5

From the above data, the approximate function for the nonlinear least square method is defined by

$$f(x) = -25932 + 14512x + 3294.5x^2. \tag{30}$$

The above function (30) will be utilized when approximating the y data values based on x data values from Jan - Aug. Let x_j denotes number of months and y_j be the the confirmed cases for that month. Based on this information, the above least squares method (30) is transformed into the following unconstrained minimization problems

$$\min_{x\in\mathbb{R}^n} f(x) = \sum_{j=1}^{n}((u_0 + u_1x_j + u_2x_j^2) - y_j)^2. \tag{31}$$

The data of the first eight months from the table 1 are utilized to formulate the nonlinear quadratic model for the least square method, which is further used to derive the unconstrained optimization model. On the basis of the above discussion, it is obvious that there exist some parabolic relations between the data x_j and the value of y_j with the regression function defined by (30) and the regression parameters u_0, u_1 and u_2

$$\min_{x\in\mathbb{R}^2} \sum_{j=1}^{n} E_j^2 = \sum_{j=1}^{n}((u_0 + u_1x_j + u_2x_j^2) - y_j)^2. \tag{32}$$

Next, using the data of Table 1, we transform (32) to obtain our nonlinear quadratic unconstrained minimization model as follows:

$$8u_1^2+72u_1u_2+408u_1u_3-1974110u_1+204u_2^2+2592u_2u_3-12593164u_2+8772u_3^2-84833792u_3+210479037915. \tag{33}$$

The above nonlinear quadratic model was constructed using data from Jan– Aug. While the data for Sept is reserved for relative error analysis of the predicted data. Now, we can apply the SCH, LS, FR and PRP methods for solving the model (33) under the strong Wolfe line search conditions (4, 5), we obtain the performance results based on iteration numbers and CPU time presented in Table 2.

Table 2. Test results for optimization of quadratic model for SCH, LS, FR and PRP.

Initial points	SCH		LS		FR		PRP	
	NOI	CPU	NOI	CPU	NOI	CPU	NOI	CPU
$(2, 2, 2)$	244	2.0670	408	2.3450	F	F	F	F
$(3, 3, 3)$	471	2.9430	519	3.1970	F	F	252	1.4530
$(10, 10, 10)$	410	2.7200	429	2.5140	F	F	F	F
$(13, 13, 13)$	632	4.4500	691	4.2880	F	F	F	F
$(30, 30, 30)$	344	2.1830	457	2.7360	F	F	F	F

To overcome the difficulty of computing the values of u_0, u_1, u_2 using matrix inverse, we implement the previously mentioned methods using different initial points. We terminate the computation if:

- The defined stopping criteria is satisfied. This is based on value defined for each function.
- The method is unable to solve the model.

4.1. Trend line method

In this subsection, we tend to estimate the confirmed cases of COVID-19 for a period of eight (8) months, using the proposed SCH, some known CG and least square methods. From the actual data obtained from Table 2, we use Microsoft Excel software to plot the trend line as shown in Figure (3). Furthermore, to show the efficiency of the proposed method, we compare the approximation functions of the SCH method, with the functions of the LS, FR, PRP and trend line methods. Based on the results presented in Table 2, it is obvious that the suggested SCH method is faster and more efficient compared to the used methods. On the other hand, from the plot, it is clear that the trend line equation obtained is in the form of a nonlinear quadratic equation. The ideal purpose of the regression analysis is to estimate $a_0, a_1, \ldots, a_p$ where the error ϵ is minimized. From the above discussion, we can conclude that the SCH method can be used as an alternative to the trend line method and the least squares method, which implies that the method is applicable to real-world situations.

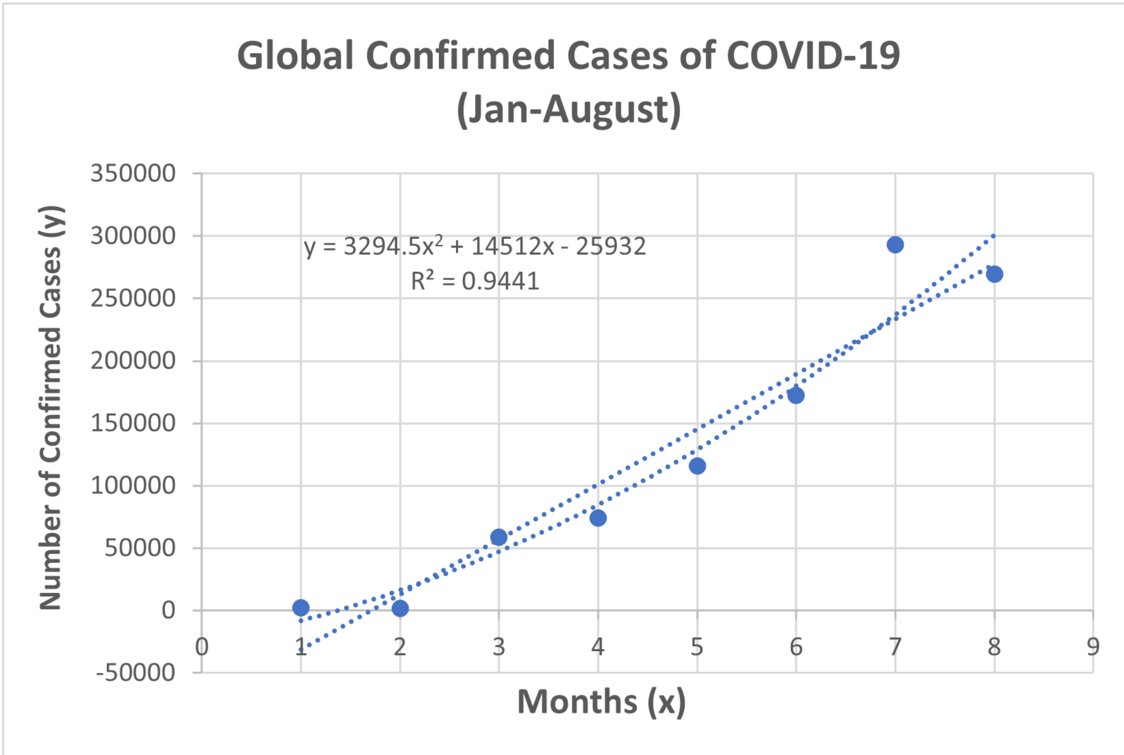

Figure 3. Nonlinear quadratic trend line for confirmed cases of COVID-19.

5. Conclusion

The hybrid CG methods are usually obtained based on the classical CG methods by integrating their advantages. In this paper we proposed a new hybrid conjugate gradient algorithms in which the famous parameter β_k is computed as a convex combination of $\beta_k^{LS}, \beta_k^{FR}$ and β_k^{PRP} algorithms. Based on some conditions, we show that The proposed algorithm enjoys the sufficient descent condition and converge globally under strong Wolfe line search. A numerical experiments are considered to illustrate the performance of the proposed method. The obtained results show that the new method is effective with better convergence rate compare to the methods of LS, FR and PRP. Moreover, our proposed method can solve the COVID-19 case model.

Acknowledgment

We would to thank the professors of Laboratory Informatics and Mathematics (LIM) (University of Souk Ahras, Algeria) who have always been helping and giving us advice.

References

[1] A.B Abubakar and P. Kumam and M. Malik and A.I Hassan. A hybrid conjugate gradient based approach for solving unconstrained optimization and motion control problems. *Mathematics and Computers in Simulation.* **201.** (2022). pp.640–657.
[2] N.Andrei, An unconstrained optimization test functions collection. *Adv. Model. Optim.* **10** (2008) pp. 147–161.
[3] N. Andrei. New hybrid conjugate gradient algorithms for unconstrained optimization. *Encyclopedia of Optimization* (2009) pp.2560-–2571.
[4] N. Andrei, Another nonlinear conjugate gradient algorithm for unconstrained optimization, *Optimization methods & software.* **24**, (2009) pp. 89–104.
[5] N. Andrei. Nonlinear Conjugate Gradient Methods for Unconstrained Optimization.*Springer* (2020).

[6] I. Bongartz, A. R. Conn, N. Gould and P. L. Toint, CUTE: constrained and unconstrained testing environments, *ACM Transactions on Mathematical Software (TOMS).* **21** (1995) pp. 123–160 .

[7] Y. H. Dai and Y. Yuan, A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property, *SIAM Journal on Optimization.* **10** (1999) pp.177–182 .

[8] S. S. Djordjevic', New Hybrid Conjugate GradientMethod as a Convex Combination of FR and PRP Methods, *Filomat.* **30** (2016) pp. 3083–3100.

[9] S.S. Djordjevic, New Hybrid Conjugate Gradient Method as a Convex Combination of HS and FR Conjugate Gradient Methods. *Journal of Applied Mathematics and Computation*, **2**, (2018) pp.366–378.

[10] S.S. Djordjević. New Hybrid Conjugate Gradient Method as a Convex Combination of LS and FR Conjugate Gradient Methods. *Acta Mathematica Scientia* , **39**, (2019) PP.214–228.

[11] E.D. Dolan and J.J. Moré, Benchmarking optimization software with performance profiles,*Mathematical programming.* **91** (2002) pp.201–213.

[12] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, *The computer journal.* **7** (1964) pp.149–154.

[13] R.Fletcher, Practical Methods Of Optimization, Unconstrained Optimization, *John Wiley and Son, New York.* **1** (1980).

[14] S. B. Hanachi. S. Badreddine. M. Belloufi, New iterative conjugate gradient method for nonlinear unconstrained optimization. *RAIRO-Operations Research*, **56**, (2022) pp. 2315–2327.

[15] M. R. Hestenes and E.L. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Standards.* **49** (1952) pp. 409–436.

[16] S. Li, and Z.B.Sun. A New Hybrid Conjugate Gradient Method and Its Global Convergence for Unconstrained Optimization. *International Journal of Pure and Applied Mathematics.* **63**, (2010) pp.84–93.

[17] A.J. Liu, and S. Li. New Hybrid Conjugate Gradient Method for Unconstrained Optimization. *Applied Mathematics and Computation*, **245**, (2014) pp. 36–43.

[18] D.Liu and C.Story, Efficient Generalized Conjugate Gradient Algorithms, part 1: Theory. *Journal of Optimization Theory and Applications.* **69** pp.129–137.

[19] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, *Revue Française D'informatique et De Recherche Opérationnelle, Série Rouge.* **3** (1969), pp.35–43.

[20] B. T. Polyak, The conjugate gradient method in extremal problems, *USSR Computational Mathematics and Mathematical Physics.* **9** (1969) pp.94–112.

[21] IM. Sulaiman, NA. Bakar, M. Mamat, BA. Hassan, M. Malik, AM. Ahmed. A new hybrid conjugate gradient algorithm for optimization models and its application to regression analysis. *Indones. J. Electr. Eng. Comput. Sci* **23** (2021). pp.1100–1109.

[22] P. Wolfe, Convergence Conditions for Ascent Methods, *SIAM Review.* **11** (1969) pp.226–235.

[23] P. Wolfe, Convergence Conditions for Ascent Methods. II: Some Corrections, *SIAM Review.* **13** (1971) pp.185-188.

[24] World heath organization: Report on coronavirus (COVID-19) (2020).

[25] World Health Organization (WHO): Naming the coronavirus disease (COVID-19) and the virus that causes it. Archived from the original on 28 February 2020. Retrieved, 28 February; 2020

[26] Wu, T.Joseph and Leung, Kathy and Leung, M.Gabriel. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *The Lancet* , **395**. (2020). pp.689–697 .

[27] G. Zoutendijk. Nonlinear programming, computational methods. *Integer and nonlinear programming*, (1970). pp.37–86.