Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 1, October 2022, pp. 339~345
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i1.pp339-345

New memoryless self-scaling quasi Newton strategy on large scale unconstrained optimization problems

Aseel M. Qasim[1], Zinah F. Salih[2]
[1]Department of Mathematics, College of Education of Pure Sciences, University of Mosul, Mosul, Iraq
[2]Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq

Article Info

Article history:

Received Apr 13, 2022
Revised Jun 25, 2022
Accepted Jul 13, 2022

Keywords:

Conjugacy coefficient
Conjugate gradient
Descent direction
Globlal convergence
Memory less

ABSTRACT

In unconstrained optimization algorithms, we employ the memoryless quasi Newton procedure to construct a new conjugacy coefficient for the conjugate gradient approaches. This newer updating formula was adapted by scaling the well-known broyden fletcher glodfarb shanno (BFGS) formula by a self-scaling factor in order to reach to the new form of the conjugacy coefficient which makes a satisfactory result in the descent direction and satisfies the globally convergent features when compared the proposed method to HS standard conjugate gradient approach. The theorems are studied in detail and moreover the numerical results of this paper is depend on a Fortran programming which are extremely stable.

Corresponding Author:

Zinah F. Salih
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul
Mosul, Iraq
Email: zn_f2020@uomosul.edu.iq

1. INTRODUCTION

In this paper, the well-known large-scale unconstrained minimization method has been considered:

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

Where $f: R^n \to R$ is continuously differentiable and the matrix of the first partial derivative $g(x) = \nabla f(x)$ is available. $x \in R^n$, n is a dimensional of the vector x, the CG algorithm is among the most efficient optimization algorithms for getting the minimum of the function (1) especially for large-scale problem [1]. Nevertheless, the CG algorithm is one of the more excepted choices in the big scale problem solving, as this method does not need any matrices [2]. The behaviour of the unconstrained optimization problem (1) is onset with a starting guess $x_0 \in R^n$, then the CG algorithm would output a sequence of points $\{x_i\}_{i=0}^{\infty}$ using the repeated form which is denoted in the next equation:

$$x_{i+1} = x_i + \omega_i t_i \tag{2}$$

where ω_i the length of a step is calculated with a suitable line search method and t_i is the direction of search, that is getting as follows:

$$t_{i+1} = \begin{cases} -g_{i+1} & for\ i = 0 \\ -g_{i+1} + \beta_i t_i & for\ i \geq 1 \end{cases} \tag{3}$$

where g_i is the gradient vector of the function $f(x)$ and B_i is a small value used to correct the path of search at x_i . There are a number of well-known conjugation method formulas given by see [3]-[7]:

$$\beta_i^{FR} = \frac{\|g_{i+1}\|^2}{\|g_i\|^2} \tag{4}$$

$$\beta_i^{PRP} = \frac{g_{i+1}^T y_i}{\|g_i\|^2} \tag{5}$$

$$\beta_i^{HS} = \frac{g_{i+1}^T y_i}{t_i^T y_i} \tag{6}$$

$$\beta_i^{DY} = \frac{\|g_{i+1}\|^2}{t_i^T y_i} \tag{7}$$

where $y_i = g_{i+1} - g_i$ is the difference gradient of the function $f(x)$ at the points x_{i+1}, x_i respectively, and more details for the coefficient B_i can be seen in [8]-[12].

These aforementioned methods have been studied by many researches including [13], [14], most of these methods studied the features of the conjugate gradient approach, recently there are many attempts to discover a recent formula for conjugate gradient methods which have good numerical execution and satisfying a global property and that is the same aim of our research, to establish this convergence property it is required to compute the step $\omega_i > 0$ with some conditions such as week wolfe condition (WWc) [15]:

$$f(x_i + \omega_i t_i) - f(x_i) \leq p_1 \omega_i \, g_i^T t_i \tag{8}$$

$$ş \, g_i^T t_i \leq g(x_i + \omega_i t_i)^T t_i \leq -p_2 \, g_i^T t_i \tag{9}$$

or by using strong wolfe condition (SWc) which satisfy (8):

$$|g(x_i + \omega_i t_i)^T t_i| \leq -p_2 \, g_i^T t_i \tag{10}$$

where $0 < p_1 < ş$, $p_2 \geq 0$. There are many other formulas that have been proposed by various scholars, for more details see [16]-[20]. The search direction is also important to determine the amount of the function that is ensuring the reduction of the search direction therefore we use quasi Newton method:

$$t_i = - G_i^{-1} g_i \tag{11}$$

where G_i is a matrix which is asymmetric and non-singular of the accession of the Hessian matrix which is denoted as a matrix of identity in the first step. The structure of this article is sequential as: first, a recent formula for the coefficient $\beta_i^{\wedge}$ is derived, while the sufficient descent property and global convergence is presented in the next section, after that, the numerical facts results are presented. Finally, the conclusion is presented in the last section.

2. NEW FORMULA OF $\beta_i^{\wedge}$

The self-scaling quasi-Newton will be utilized to scale the Hessian matrix G_i, [21], [22] scale some terms of broyden fletcher glodfarb shanno (BFGS). Our technical method is to scale all the terms of BFGS by multiplying G_i by a scalar ệ, then the direction becomes:

$$t_{i+1} = -ệ \, G_{i+1} g_{i+1} \tag{12}$$

where ệ is a self-scaling factor and there are several types of the scalar ệ such as [23], [24]:

$$ệ = \frac{y_i^T y_i}{y_i^T s_i} \qquad \text{(AlBayati)} \tag{13}$$

$$ệ = \frac{y_i^T s_i}{g^T H g_i} \qquad \text{(AlBayati \& Maha)} \tag{14}$$

$$ệ = \frac{6}{s_i^T y_i}\left[f(x_i) - f(x_{i+1}) + s_i^T g_{i+1}\right] - 2 \qquad \text{(Biggs)} \tag{15}$$

Now the (BFGS) formula can be written in the form:

$$G_{i+1} = G_i - \left(\frac{G_i y_i s_i^T + s_i y_i^T G_i}{s_i^T y_i}\right) + \frac{s_i s_i^T}{s_i^T y_i} + \frac{y_i^T G_i y_i}{s_i^T y_i} . \frac{s_i s_i^T}{s_i^T y_i} \tag{16}$$

To scale the Hessian matrix G we have to use the self-scaling quasi-Newton method, by multiplying (16) by ệ then:

$$G_{i+1} = ệ\left[G_i - \left(\frac{G_i y_i s_i^T + s_i y_i^T G_i}{s_i^T y_i}\right) + \frac{s_i s_i^T}{s_i^T y_i} + \frac{y_i^T G_i y_i}{s_i^T y_i} . \frac{s_i s_i^T}{s_i^T y_i}\right] \tag{17}$$

Now if we choose ệ from (15), and replace G_i by I then (17) will be refer to memory less (BFGS) and written as:

$$G_{i+1} = ệ\left[I - \left(\frac{y_i s_i^T + s_i y_i^T}{s_i^T y_i}\right) + \frac{s_i s_i^T}{s_i^T y_i} + \frac{y_i^T y_i}{s_i^T y_i} . \frac{s_i s_i^T}{s_i^T y_i}\right] \tag{18}$$

In order to compute the new parameter $\beta_i^{\wedge}$, both sides of (18) will be multiplied by g_{i+1}, from $t_{i+1} = -ệ\, G_{i+1} g_{i+1}$ and from (3):

$$-g_{i+1} + \beta_i^{\wedge} t_i = ệ\left(-g_{i+1} + \frac{s_i^T g_{i+1}}{s_i^T y_i} y_i + \frac{y_i^T g_{i+1}}{s_i^T y_i} s_i - \frac{s_i^T g_{i+1}}{s_i^T y_i} s_i - \frac{y_i^T y_i}{s_i^T y_i} \frac{s_i^T g_{i+1}}{s_i^T y_i} s_i\right) \tag{19}$$

now if we use ệ from biggs (15) and for simplicity, we assume the term:

$$F = f_i - f_{i+1} + s_i^T g_{i+1} \tag{20}$$

to simplify the steps of derivation, we have:

$$-g_{i+1} + \beta_i^{\wedge} t_i = \frac{-6 g_{i+1}}{s_i^T y_i}[F] + 2 g_{i+1} + \frac{6}{s_i^T y_i}[F] \frac{s_i^T g_{i+1} y_i + y_i^T g_{i+1} s_i}{s_i^T y_i} - 2\, \frac{s_i^T g_{i+1} y_i + y_i^T g_{i+1} s_i}{s_i^T y_i} - \frac{6}{s_i^T y_i}[F] \frac{s_i^T g_{i+1} s_i}{s_i^T y_i} + 2 \frac{s_i^T g_{i+1} s_i}{s_i^T y_i} - \frac{6}{s_i^T y_i}[F] \frac{y_i^T y_i}{s_i^T y_i} . \frac{s_i^T g_{i+1} s_i}{s_i^T y_i} + 2 \frac{y_i^T y_i}{s_i^T y_i} . \frac{s_i^T g_{i+1} s_i}{s_i^T y_i} \tag{21}$$

now by multiplying (21) by y_i^T and divide both side of (21) by $y_i^T t_i$:

$$\beta_i^{\wedge} = \frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{6 y_i^T g_{i+1}}{s_i^T y_i\, y_i^T t_i}[F] + 2 \frac{y_i^T g_{i+1}}{y_i^T t_i} + \frac{6}{s_i^T y_i}[F] . \frac{s_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . y_i^T s_i}{y_i^T t_i . s_i^T y_i} - 2 \frac{s_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . y_i^T s_i}{s_i^T y_i . y_i^T t_i} - \frac{6}{s_i^T y_i}[F] . \frac{s_i^T g_{i+1} . y_i^T s_i}{y_i^T t_i . s_i^T y_i} + 2 \frac{s_i^T g_{i+1} . y_i^T s_i}{s_i^T y_i . y_i^T t_i} - \frac{6}{s_i^T y_i}[F] \frac{y_i^T y_i}{s_i^T y_i} . \frac{s_i^T g_{i+1} . y_i^T s_i}{s_i^T y_i . y_i^T t_i} + 2 \frac{y_i^T y_i}{s_i^T y_i . y_i^T t_i} . \frac{s_i^T g_{i+1} . y_i^T s_i}{s_i^T y_i} \tag{22}$$

Since $s_i^T = \omega t_i^T$

$$\beta_i = \frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{6 y_i^T g_{i+1}}{\omega t_i^T y_i . y_i^T t_i}[F] + 2 \frac{y_i^T g_{i+1}}{y_i^T t_i} + \frac{6[F]}{\omega t_i^T y_i} . \frac{\omega t_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . \omega y_i^T t_i}{\omega y_i^T t_i . y_i^T t_i} - 2 \frac{\omega t_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . \omega y_i^T t_i}{\omega t_i^T y_i . y_i^T t_i} - \frac{6}{\omega t_i^T y_i}[F] . \frac{\omega t_i^T g_{i+1} . \omega t_i^T y_i}{\omega y_i^T t_i . t_i^T y_i} + 2 \frac{\omega t_i^T g_{i+1} . \omega t_i^T y_i}{\omega t_i^T y_i . y_i^T t_i} - \frac{6}{\omega t_i^T y_i}[F] . \frac{y_i^T y_i}{\omega t_i^T y_i} . \frac{\omega t_i^T g_{i+1} . \omega t_i^T y_i}{\omega t_i^T y_i . y_i^T t_i} + 2 \frac{y_i^T y_i . \omega t_i^T g_{i+1}}{\omega t_i^T y_i . y_i^T t_i} \tag{23}$$

$$= \frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{6[F]}{\omega t_i^T y_i}\left(\frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{t_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . y_i^T t_i}{y_i^T t_i . y_i^T t_i} + \frac{\omega t_i^T g_{i+1}}{y_i^T t_i} + \frac{y_i^T y_i}{t_i^T y_i} . \frac{t_i^T g_{i+1}}{y_i^T t_i}\right) + 2\left(\frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{t_i^T g_{i+1} . y_i^T y_i + y_i^T g_{i+1} . y_i^T t_i}{t_i^T y_i . y_i^T t_i} + \frac{\omega t_i^T g_{i+1}}{y_i^T t_i} + \frac{y_i^T y_i . t_i^T g_{i+1}}{t_i^T y_i . y_i^T t_i}\right) \tag{24}$$

$$= \frac{y_i^T g_{i+1}}{y_i^T t_i} - \frac{6[F]}{\omega t_i^T y_i}\left(\frac{y_i^T g_{i+1} . y_i^T t_i - t_i^T g_{i+1} . y_i y_i - y_i^T g_{i+1} . y_i^T t_i + \omega t_i^T g_{i+1} . y_i^T t_i + y_i y_i . t_i^T g_{i+1}}{y_i^T t_i . y_i^T t_i} + 2 \frac{y_i^T g_{i+1} . y_i^T t_i - t_i^T g_{i+1} . y_i^T y_i - y_i^T g_{i+1} . y_i^T t_i + \omega t_i^T g_{i+1} t_i^T y_i + y_i^T y_i . t_i^T g_{i+1}}{t_i^T y_i . y_i^T t_i}\right)$$

$$= \frac{y_i^T g_{i+1}}{t_i^T y_i} - \frac{6[F]\omega}{\omega t_i^T y_i} . \frac{t_i^T g_{i+1}}{y_i^T t_i} + 2\omega \frac{t_i^T g_{i+1}}{y_i^T t_i} \tag{25}$$

Now subs (F) of (20) in (25) we get:

$$= \frac{y_i^T g_{i+1}}{t_i^T y_i} - \frac{t_i^T g_{i+1}}{t_i^T y_i . t_i^T y_i}[6(f_i - f_{i+1} + s_i^T g_{i+1})] + 2\omega \frac{t_i^T g_{i+1}}{y_i^T t_i}$$

$$\beta_i^{\wedge} = \frac{y_i^T g_{i+1}}{t_i^T y_i} - \frac{t_i^T g_{i+1}}{t_i^T y_i} . \left(6 \frac{(f_i - f_{i+1} + s_i^T g_{i+1})}{t_i^T y_i} - 2\omega\right) \tag{26}$$

And that is the final form of our new coefficient. If the direction is exact line search, then $B_i^{\wedge}$ will be reduced to Hesten Stefile method. However, if we used inexact line search with Wolfe type line search then our algorithm of a new method is as the following:

```
Algorithm (1)
Given x_0 ϵ R^n and ω_i > 0, set i=1.
Step1: set t_i = -g_i = -∇f(x_i) , if ||g_i|| < ϵ , then go to end.
Step2: determine ω_i > 0 satisfying Wolfe type line search in (8) and (9).
Step3: calculate a new iteration (x_{i+1}) by (2) and g_{i+1} .if ||g_{i+1}|| < ϵ then go to end,
else go to step (2).
```

3. THE SUFFICIENT DESCENT AND GLOBAL CONVERGENCE PROPERTY

3.1. Acceptance (1)

(a) Let the set $\psi = \{x_0 \epsilon R^n : f(\mathrm{x}) \le f(x_0)\}$ is bounded.

(b) Suppose ψ is a neighbourhood of ζ then f is continuously differentiable and the Lipchitz condition of the gradient is continuous of ψ. This means, there is $k > 0$ such that $\forall\, x$.

$$\|\mathrm{g}(\mathrm{x}) - g(\hat{x})\| \le k\|\mathrm{x} - \hat{x}\|, \hat{x} \in \psi$$

From acceptance (a) and (b) we can design the sequence $\{x_i\}\epsilon\zeta$, because f is decreasing. From acceptance (a) and (b), we can profit that $\forall\, x \in \zeta \exists c_1, c_2 > 0$ for which $\|\mathrm{x}\| \le c_1$, $\|\hat{x}\| \le c_2$ and the sequence $\{x_i\}\epsilon\zeta$ because $\{f(x_i)\}$ is decreasing, henceforward we will assume that assumption (a), (b) are hold and the objective function is bounded below.

Theorem (1): let that acceptance (1) it satisfies, and ω holds the Wolfe type line search (8) and (9) and $\beta_i^{\wedge}$ is given in (26) then (3) holds the property of descent.

Proof: for (i=1) we get $t_1 = -g_1 \Longrightarrow t_1^T g_1 = -\|g_1\|^2 \le 0$, and this satisfies the descent property. Now we have to prove the descent for all $k \ge 1$, by multiplying (3) by g_{i+1}:

$$t_{i+1}^T g_{i+1} = -\|g_{i+1}\|^2 + \beta_i^{\wedge}\, t_i^T g_{i+1} \tag{27}$$

if an exact line search is used then $t_i^T g_{i+1} = 0 \Longrightarrow t_{i+1}^T g_{i+1} = -\|g_{i+1}\|^2 \le 0$, but if we used inexact line search then (27) yield:

$$t_{i+1}^T . g_{i+1} = -\|g_{i+1}\|^2 + \left[\frac{y_i^T g_{i+1}}{t_i^T y_i} - \frac{t_i^T g_{i+1}}{t_i^T y_i} . \left(6\frac{(f_i - f_{i+1} + s_i^T g_{i+1})}{t_i^T y_i} - 2\omega\right)\right](t_i^T g_{i+1}) \tag{28}$$

from (SWC) and (10) and the equality:

$$y_i^T g_{i+1} < \|g_{i+1}\|^2 \tag{29}$$

and (8), (9) and since:

$$t_i^T y_i \ge t_i^T (g_{i+1} - g_i) \ge (\text{ş} - 1) g_i^T t_i \tag{30}$$

then:

$$t_{i+1}^T g_{i+1} \le -\|g_{i+1}\|^2 + \left[\frac{\|g_{i+1}\|^2}{(\text{ş}-1)g_i^T t_i} - \frac{(-p_2 g_i^T t_i)}{t_i^T y_i}\left(\frac{-6(f_{i+1} - f_i - s_i g_{i+1})}{(\text{ş}-1)g_i^T t_i} - 2\omega\right)\right](-p_2\, g_i^T t_i)$$

$$\le -\|g_{i+1}\|^2 + \left[\frac{\|g_{i+1}\|^2}{(\text{ş}-1)g_i^T t_i} + \frac{(p_2 g_i^T t_i)}{t_i^T y_i}\left(\frac{-6(p_1\omega g_i^T t_i - \omega p_2 g_i^T t_i)}{(\text{ş}-1)g_i^T t_i} - 2\omega\right)\right](-p_2 g_i^T t_i)$$

$$\le -\|g_{i+1}\|^2 - p_2 g_i^T t_i \frac{\|g_{i+1}\|^2}{(\text{ş}-1)g_i^T t_i} - \frac{\omega(p_2 g_i^T t_i)^2}{t_i^T y_i}\left(\frac{-6(p_1 - p_2)}{(\text{ş}-1)} - 2\right)$$

$$\le -\|g_{i+1}\|^2 - p_2 \frac{\|g_{i+1}\|^2}{(\text{ş}-1)}$$

$$\le -(1 + \frac{p_2}{(\text{ş}-1)})\|g_{i+1}\|^2 \tag{31}$$

let $\kappa = \frac{p_2}{(\S-1)}$ and $0 < \kappa < 1 <$ is negative, then $(1 + \kappa) = m$ is a positive number then (31) satisfies $t_{i+1}^T g_{i+1} \leq -m \|g_{i+1}\|^2$, which completes the proof. To state the global convergence for the new algorithms, you should see [13], [25] which is containing the Zoutindijk condition.

4. THE GLOBAL CONVERGENCE PROPERTY

Theorem (2): suppose that acceptance (1) holds, consider the algorithm (1) satisfies Wolfe condition, then:

$$\lim_{k\to\infty} \inf \| g_{i+1}\| = 0 \tag{32}$$

Proof: If the theorem is not true, then $\exists \; 3 > 0$, s.t $\| g_{i+1}\| > 3, \forall k$, then (3) can be written as:

$$g_{i+1}^T + t_{i+1} = \beta_i^{\wedge} t_i \tag{33}$$

by squaring both sides of (33) and rearranging it yields:

$$\|t_{i+1}\|^2 = -\|g_{i+1}\|^2 - 2\, g_{i+1}^T t_{i+1} + (\beta_i^{\wedge})^2 \|t_i\|^2$$
$$= (\beta_i^{\wedge})^2 \|t_i\|^2 - 2\, g_{i+1}^T t_{i+1} - \|g_{i+1}\|^2 \tag{34}$$

dividing both sides of (34) by $(g_{i+1}^T t_{i+1})^2$:

$$\frac{\|t_{i+1}\|^2}{(g_{i+1}^T t_{i+1})^2} = \frac{(\beta_i^{\wedge})^2 \|t_i\|^2}{(g_{i+1}^T t_{i+1})^2} - \frac{2}{g_{i+1}^T t_{i+1}} - \frac{\|g_{i+1}\|^2}{(g_{i+1}^T t_{i+1})^2}$$
$$= \frac{(\beta_i^{\wedge})^2 \|t_i\|^2}{(g_{i+1}^T t_{i+1})^2} - \left(\frac{\|g_{i+1}\|}{g_{i+1}^T t_{i+1}} - \frac{1}{\|g_{i+1}\|}\right)^2 + \frac{1}{\|g_{i+1}\|^2}$$
$$\leq \frac{(\beta_i^{\wedge})^2 \|t_i\|^2}{(g_{i+1}^T t_{i+1})^2} + \frac{1}{\|g_{i+1}\|^2} \tag{35}$$

now from (26), (29), (30) and the (8)-(10) and that $t_i^T g_i = -\|t_i\|^2$ we have:

$$\beta_i^{\wedge} \leq \frac{\|g_{i+1}\|^2}{(\S-1) g_i^T t_i} + \frac{p_2 g_i^T t_i}{-(\S-1)\|t_i\|^2}\left(\frac{6(p_1\omega\|t_i\|^2 - p_2\omega\|t_i\|^2)}{(1-\S)\|t_i\|^2} + 2\omega\right) \tag{36}$$

$$\leq \frac{\|g_{i+1}\|^2}{-(\S-1)\|t_i\|^2} + \frac{p_2 g_i^T t_i}{-(\S-1)\|t_i\|^2}\left(\frac{6(p_1\omega - p_2\omega)}{(1-\S)} + 2\omega\right) \tag{37}$$

$$\leq \frac{\|g_{i+1}\|^2}{-(\S-1)\|t_i\|^2} \tag{38}$$

by squaring (38), we have $(\beta_i^{\wedge})^2 = \left(\frac{\|g_{i+1}\|^2}{-(\S-1)\|t_i\|^2}\right)^2$ and sub it in (35):

$$\frac{\|t_{i+1}\|^2}{(g_{i+1}^T t_{i+1})^2} \leq \frac{\|g_{i+1}\|^4}{(-(\S-1))^2 \|t_i\|^4} \cdot \frac{\|t_i\|^2}{(g_{i+1}^T t_{i+1})^2} + \frac{1}{\|g_{i+1}\|^2}$$
$$\leq \frac{1}{\|t_i\|^2} + \frac{1}{\|g_{i+1}\|^2} = \frac{1}{D_1} + \frac{1}{\gamma} = D_2 \tag{39}$$

since $\|t_1\|^2 = -g_1^T t_1 = \|g_1\|^2$ by noting that $\frac{\|t_i\|^2}{(g_0^T t_0)^2} = \frac{1}{\|t_i\|^2}$, then (39) yields that:

$$\frac{\|t_{i+1}\|^2}{(g_{i+1}^T t_{i+1})^2} \leq \sum_{i=1}^{k} \frac{1}{\|g_i\|^2} \quad \forall k \to \frac{1}{D_2} \sum_{k\geq 1} 1 = \infty$$

this contradiction to Zoutendijk condition and with this contradiction, we complete the proof of the theorem.

5. NUMERICAL FACTS

The primary goal of this work is to compire and compute the proposed method's execution for a set of test functions against the well-known HS routine. These test experiments were collected by Andrei [26]. We select (20) large-scale test problems and consider two dimensions (n=100, n=1000) for each test.

The stop criterion is $\|g_{i+1}\| \leq 10^{-6}$, all codes were written in Fortran 90. We denote the number of iterations as (NuI) and (NuF), (NuR) as the number of evaluation functions and restarts. All of these results are reported in the Table 1 while the percentage performance with respect to (NuI), (NuF) and (NuR) is denoted as 92.75%, 72.53%, 61.71% respectively.

Table 1. Numerical results of a new algorithm

Functions	dim	$\hat{\beta_i}$			HS		
		NuI	NuR	NuF	NuI	NuR	NuF
Extended Beale	100	14	8	26	13	7	26
	1000	17	10	32	17	10	32
Penalty	100	9	6	25	9	6	25
	1000	22	13	47	61	53	1290
Diagonal 2	100	58	21	101	61	19	103
	1000	204	69	351	207	59	339
Generalized Tridiagonal 1	100	22	6	44	22	6	44
	1000	27	12	50	27	134	50
Extended Tridiagonal 1	100	7	4	15	7	4	15
	1000	13	7	26	13	7	26
Extended Three Expo Terms	100	17	9	25	18	10	26
	1000	14	9	25	13	7	24
Generalized Tridiagonal 2	100	41	15	61	41	15	61
	1000	52	20	83	62	23	97
Diagonal 4	100	4	2	8	4	2	8
	1000	4	2	8	4	2	8
Extended Powell	100	55	18	103	56	19	104
	1000	79	22	149	76	26	140
Quadratic Diagonal Perturbed	100	49	10	89	73	18	129
	1000	182	29	324	193	31	343
Extended Wood WOODS	100	23	8	46	23	8	46
	1000	24	10	47	25	10	49
Himmelbh	100	6	3	13	6	3	13
	1000	6	3	13	6	3	13
Nondia	100	15	8	30	16	8	31
	1000	11	6	22	12	7	25
Dqdrtic	100	6	1	13	7	1	15
	1000	7	1	15	7	1	15
Dixmaanb	100	10	10	18	10	10	18
	1000	11	11	19	11	11	19
Liarwhd	100	17	10	31	17	10	31
	1000	22	11	47	23	12	52
Extended Block-Diagonal BD2	100	11	7	21	11	7	21
	1000	12	8	24	12	8	24
Diagonal 7	100	3	3	9	3	3	9
	1000	4	4	11	4	3	11
Generalized quartic GQ2	100	36	14	57	37	13	57
	1000	33	11	53	31	9	55
Denschna	100	9	6	17	9	6	17
	1000	9	6	18	9	6	18
Total		1165	433	2116	1256	597	3429

6. CONCLUSION

We have suggest a recent memoryless algorithm depending on scaling the (BFGS) formula. Where this newly method produces a sufficient descent direction while this property depends on the type of line search that is used in the algorithm which is important. We proposed the global convergence and the numerical results which are produced in the previous section showing the percentage of the method efficiency.

REFERENCES

[1] S. S. Rao, "*Engineering optimization: theory and practice*," John Wiley & Sons, 2019.

[2] M. Al-Baali, "Descent property and global convergence of the fletcher—reeves method with inexact line search," *IMA Journal of Numerical Analysis*, vol. 5, no. 1, pp. 121–124, 1985, doi: 10.1093/imanum/5.1.121.

[3] R. Fletcher, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, Feb. 1964, doi: 10.1093/comjnl/7.2.149.

[4] E. Polak and G. Ribiere, "Note on the convergence of conjugate direction methods," *French journal of computer science and operational research. Red Series*, vol. 3, no. 16, pp. 35–43, 1969, doi: 10.1051/m2an/196903r100351.

[5] B. T. Polyak, "The conjugate gradient method in extremal problems," *USSR Computational Mathematics and Mathematical Physics*, vol. 9, no. 4, pp. 94–112, Jan. 1969, doi: 10.1016/0041-5553(69)90035-4.

[6] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, p. 409, Dec. 1952, doi: 10.6028/jres.049.044.
[7] Y. Dai, "Convergence properties of the fletcher-reeves method," *IMA Journal of Numerical Analysis*, vol. 16, no. 2, pp. 155–164, 1996, doi: 10.1093/imanum/16.2.155.
[8] M. Al-Baali, Y. Narushima, and H. Yabe, "A family of three-term conjugate gradient methods with sufficient descent property for unconstrained optimization," *Computational Optimization and Applications*, vol. 60, no. 1, pp. 89–110, May 2014, doi: 10.1007/s10589-014-9662-z.
[9] N. Andrei, "A new three-term conjugate gradient algorithm for unconstrained optimization," *Numerical Algorithms*, vol. 68, no. 2, pp. 305–321, Apr. 2014, doi: 10.1007/s11075-014-9845-9.
[10] S. Babaie-Kafaki, "Two modified scaled nonlinear conjugate gradient methods," *Journal of Computational and Applied Mathematics*, vol. 261, pp. 172–182, May 2014, doi: 10.1016/j.cam.2013.11.001.
[11] X. Dong, H. Liu, and Y. He, "A self-adjusting conjugate gradient method with sufficient descent condition and conjugacy condition," *Journal of Optimization Theory and Applications*, vol. 165, no. 1, pp. 225–241, 2014, doi: 10.1007/s10957-014-0601-z.
[12] Y. Dong, "A practical PR+ conjugate gradient method only using gradient," *Applied Mathematics and Computation*, vol. 219, no. 4, pp. 2041–2052, Nov. 2012, doi: 10.1016/j.amc.2012.08.047.
[13] G. Zoutendijk, "Nonlinear programming, computational methods," *Integer and nonlinear programming*, pp. 37-86, 1970.
[14] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21–42, Feb. 1992, doi: 10.1137/0802003.
[15] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, no. 2, pp. 226–235, Apr. 1969, doi: 10.1137/1011036.
[16] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific Journal of Mathematics*, vol. 16, no. 1, pp. 1–3, Jan. 1966, doi: 10.2140/pjm.1966.16.1.
[17] Z. Dai and F. Wen, "Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property," *Applied Mathematics and Computation*, vol. 218, no. 14, pp. 7421–7430, Mar. 2012, doi: 10.1016/j.amc.2011.12.091.
[18] H. Liu, "A mixture conjugate gradient method for unconstrained optimization," *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, Apr. 2010, doi: 10.1109/iitsi.2010.57.
[19] L. Zhang, W. Zhou, and D. Li, "Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search," *Numerische Mathematik*, vol. 104, no. 4, pp. 561–572, Sep. 2006, doi: 10.1007/s00211-006-0028-z.
[20] G. Yuan and X. Lu, "A modified PRP conjugate gradient method," *Annals of Operations Research*, vol. 166, no. 1, pp. 73–90, Aug. 2008, doi: 10.1007/s10479-008-0420-4.
[21] S. S. Oren, "Self-scaling variable metric (SSVM) algorithms: part ii: implementation and experiments," *Management Science*, vol. 20, no. 5, pp. 863–874, Jan. 1974, doi: 10.1287/mnsc.20.5.863.
[22] S. S. Oren and D. G. Luenberger, "Self-scaling variable metric (SSVM) Algorithms: Part i: Criteria and sufficient conditions for scaling a class of algorithms," *Management Science*, vol. 20, no. 5, pp. 845–862, Jan. 1974, doi: 10.1287/mnsc.20.5.845.
[23] A. Y. Al-Bayati, "A new family of self-scaling variable metric algorithms for unconstrained optimization," *Journal of Education and Science, Mosul University*, vol. 12, pp. 25-54, 1991.
[24] M. C. Biggs, "A note on minimization algorithms which make use of non-quardratic properties of the objective function," *IMA Journal of Applied Mathematics*, vol. 12, no. 3, pp. 337–338, 1973, doi: 10.1093/imamat/12.3.337.
[25] A. M. Qasim, Z. F. Salih, and B. A. Hassan, "A new conjugate gradient algorithms using conjugacy condition for solving unconstrained optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 3, p. 1647, Dec. 2021, doi: 10.11591/ijeecs.v24.i3.pp1647-1653.
[26] N. Andrei, "An unconstrained optimization test functions collection," *Adv. Model. Optim*, vol. 10, no. 1, pp. 147-161, 2008.

BIOGRAPHIES OF AUTHORS

Aseel M. Qasim holds a Bachelor's degree in Mathematics from the University of Mosul in 2004 and a Master's degree in Mathematics from the University of Mosul as well in 2007. She has a number of researches published in local and international journals with impact factors. Her research interests include numerical optimization in particular and mathematics in general. She can be contacted at the email: aseel.albazaz@uomosul.edu.iq.

Zinah F. Salih received her B.Sc. degree in Mathematics from the University of Mosul in 2005, and the M.Sc. degree in Mathematics from the University of Mosul as well in 2007. She has written several articles in local and international journals with impact factors. Her research interests include in mathematical modeling and optimization. She can be contacted at the email: zn_f2020@uomosul.edu.iq.