Modified Hestenes - Stiefel Conjugate Gradient (MHS-CG) Method for Solving Unconstrained Optimization

Osama M. T. Wais[1] , Khalil K. Abbo[2] and Ibrahim A. Saleh[3]

[1]Mathematics Department, College of Basic Education
University of Mosul, Mosul. Iraq
usama_al_kassaab@uomosul.edu.iq

[2] Mathematics Department, College of Basic Education
University of Telafer, Tall'Afar. Iraq
dr.khalil-khuder@uomosul.edu.iq

[3] Software Engineering Department, College of Computer Science and Mathematics
University of Mosul, Mosul. Iraq
i.hadedi@gmail.com

Received: 29.01.2022, Accepted: 12.09.2022, Published: 03.01.2023

Abstract — *The conjugate gradient technique is one of the most effective methods for solving and minimizing unconstrained optimization problems, and it is widely utilized. In this research, we introduce a novel nonlinear conjugate gradient approach with excellent convergence for unconstrained minimization problems that is based on the nonlinear conjugate gradient method. The new algorithm has the property of descent as well as global convergence. Results from the numerical evaluations demonstrate that the new technique is very efficient in practical computing and outperforms previous comparable approaches in a wide range of conditions.*
Keywords: Unconstrained optimization, Nonlinear conjugate gradient method, Global convergence.
Mathematics Subject Classification: 90C26, 65K10.

1 Introduction

Let us take the issue of unconstrained optimization of the n variables in the given description [1], [2]:

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

Cite as: *Osama M. T. Wais , Khalil K. Abbo and Ibrahim A. Saleh, Modified Hestenes - Stiefel Conjugate Gradient (MHS-CG) method for solving unconstrained optimization, Journal of Multidisciplinary Modeling and Optimization 4(2) (2021),32-42.*

where $f: R^n \to R$ is continuous differentiable function , It is among the most effective optimization techniques for achieving a problem solution that the conjugate gradient algorithms are used (1) . The conjugate gradient method has the form [3], [4]

$$x_{k+1} = x_k + \alpha_k d_k \quad , \quad k = 0,1,2,3, \dots \tag{2}$$

Where x_0 is an initial point, α_k is a step size , $g_k = \nabla f(x)$ and d_k can be taken as [5], [6]:

$$d_k = \begin{cases} -g_k & : \quad k = 0 \\ -g_k + \beta_k d_{k-1} & : \quad k \geq 1 \end{cases} \tag{3}$$

In the case of the nonlinear conjugate gradient approach, the step size α_k is often calculated by employing both the exact and inexact line search techniques. As a result, the Wolfe inexact line search using the following formula was used in this investigation. The conventional Wolfe line search needs a value of α_k in order to be satisfied [7].

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

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

Where $0 < \delta \leq \sigma < 1$. Line search accuracy would generally be improved with a lower value of σ; however, the processing time would be increased as a result. In order to make analysis easier, a stronger condition was developed (4)and [7]

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

in which the search for the strong Wolfe line (SWP) is determined by the requirements provided by (4) and (6). Generally, d_k is required to satisfy [8], [9], [10]

$$d_k^T g_k < 0 \tag{7}$$

Insuring d_k is a descent direction of $f(x)$ at x_k . In order to keep the convergence property, d_k is often require to satisfy the sufficient descent condition [11]

$$g_k^T d_k \leq -c\|g_k\|^2 \tag{8}$$

Where the constant $c > 0$.

Different β_k will determine different conjugate gradient methods. Some previous works in this field show that a proper choice of this parameter led us to a better numerical performance. Some famous classical choices of β_k can be found in Hestenes and Stiefel (HS) [12], Polak and Ribiere and Polyak (PRP)[13], Fletcher and Reeves (FR)[14], Fletcher (CD)[15], Dai and Yuan (DY)[8], and Liu and Storey (LS) [16] as shown below:

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k}, \quad \beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}, \beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}$$
$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{d_k^T g_k}, \quad \beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \qquad \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k} \tag{9}$$

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

It is possible to categorize the formulae shown above into two categories. PRP, HS, and LS are included in the first category. When it comes to finding the solution to large-scale functions, these formulae are considered to be among the most efficient CG methods available. They are able to do this because of an integrated automatic restart function that prevents them from being stuck during the computing process. This algorithmic convergence to a solution, on the other hand, may fail for certain problems, and the criteria for their convergence have not been shown under certain inexact line search conditions until recently. The FR, DY, and CD approaches are included in the other category. Despite the fact that these algorithms have excellent convergence qualities, their numerical performance is often suboptimal as a result of the jamming phenomenon.

2 The Conjugacy Condition (General Review)

The search direction d_k for many unconstrained optimization methods, including the Newton-like (QN), memoryless BFGS method, and the limited memory BFGS method, can be written as

$$d_{k+1} = -H_{k+1} g_{k+1} \tag{10}$$

So, H is a positively defined symmetric matrix of the type that satisfies the equation (QN):

$$H_{k+1} y_k = s_k \tag{11}$$

where s_k represents the step. Using (2.1) and (2.2), we get :

$$d_{k+1}^T y_k = -(H_{k+1} g_{k+1})^T y_k = -g_{k+1}^T (H_{k+1} y_k) = -g_{k+1}^T s_k \tag{12}$$

Which is called Perry Conjugacy Condition .
Since in this case, $g_{k+1}^T s_k = 0$, if the search is an exact line search (ELS), then the previous relationship leads to the conjugate condition. However, applied numerical algorithms usually adopt inexact line search (ILS) rather than exact line search (ELS). For this reason, it seems more appropriate to replace the conjugate condition

$$d_{k+1}^T y_k = 0. \tag{13}$$

Under the following conditions Dai and Liao (DL) generalized the Conjugacy condition defined in (2.3) to the following:

$$d_{k+1}^T y_k = -t g_{k+1}^T s_k \tag{14}$$

Since $t \geq 0$ is a scalar quantity.
And to make sure that the search direction d_k in (3) satisfies the conjugate condition (14) by multiplying (3) by y_k and using (14), we get :

$$\beta_{k+1}^{DL} = \frac{g_{k+1}^T (y_k - t s_k)}{d_k^T y_k} \tag{15}$$

It is obvious that if $t = 1$ in the above eq. we get β^{Pr} Perry method :

$$\beta_{k+1}^{DL} = \beta_{k+1}^{HS} - t\frac{g_{k+1}^T s_k}{d_k^T y_k} \tag{16}$$

From the point of view of the global convergence of general functions, Dai and Liao proposed a modification of formula (16) that restricts the first term to non-negative values.

$$\beta_{k+1}^{DL+} = \max\left\{\frac{g_{k+1}^T y_k}{d_k^T y_k}, 0\right\} - t\frac{g_{k+1}^T s_k}{d_k^T y_k} \tag{17}$$

The step length α_k can be obtained by using any of the search line formulas. The strong Wolfe conditions (4) and (6) are often used in conjugate gradient methods .Dai and Liao demonstrated global convergence, and for more information, see [11].

3 Modified Hestenes - Stiefel Conjugate Gradient (OKI1-CG)

By using the Conjugacy condition to Dai and Liao from (12) and (14) we get:

$$-y_k^T g_{k+1} + \beta y_k^T s_k = -t s_k^T g_{k+1} \tag{18}$$

and assuming the value of $t = \frac{|s_k^T g_{k+1}|}{s_k^T y_k}$ we get:

$$-y_k^T g_{k+1} + \beta y_k^T s_k = -\frac{|s_k^T g_{k+1}|}{s_k^T y_k} s_k^T g_{k+1} \tag{19}$$

$$\beta y_k^T s_k = y_k^T g_{k+1} - \frac{|s_k^T g_{k+1}| s_k^T g_{k+1}}{s_k^T y_k}. \tag{20}$$

By dividing both sides of equation (20) by $(y_k^T s_k)$ we get our new β:

$$\beta^{OKI\,1} = \frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}. \tag{21}$$

We rate that if exact line search used the $\beta^{OKI\,1}$ reduces to the β^{HS} method format if line search is exact and objective function is quadratic then (2.12) reduces to the β^{FR} method. the our new β_k which is known as β_k^{OKI1} where (OKI denotes Osama , Khalil and Ibrahim) and the algorithm is given by :

Algorithm 1

Stage 1: Initialization. Given $x_0 \in R^n$, set $k = 0$, compute $f_0 = f(x_0)$, $g_0 = \nabla f(x_0)$ and $d_0 = -g_0$.
Stage 2 : If $\|g_k\| = 0$,then stop ; otherwise continue .
Stage 3 : Compute α_k on the basis of (4) and (6) .
Stage 4 : Update a new point on the basis of (2).
Stage 5 : Compute β_k on the basis of (21)
Stage 6 : Compute d_k on the basis of (3) .
Stage 7:Convergent test and stopping criteria.$\|g_k\| \le \epsilon$ then stop. Otherwise go to Stage 2 with $k = k + 1$.

4 Descent property

In this section we prove that our algorithm defined in equations (3) and (21) generates descent direction for all iteration according to the following theorem .

Assumption 1([3],[6],[17]) *Assume that :*

I. *The level set* $S = \{x \in R^n : f(x) \le f(x_1)\}$ *is bounded*
II. *In a neighborhood* N *of* S. *The function* f *is continuously differentiable and its gradient is Lipschitz continuous. i.e. there exists a constant* $L > 0$ *such that*
$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\| \qquad (22)$$
Under these assumptions on f *there exists a constant* $\gamma \ge 0$ *such that* $\|g(x)\| \le \gamma$ *for all* $x \in S$.

Theorem 1 *Suppose that assumption (1) hold and consider the algorithm* $x_{k+1} = x_k + \alpha_k d_k$ *where* d_k *generated by (3) and (21), and* α_k *satisfies the Strong Wolfe Conditions (4) and (6) with* $\theta < \frac{1}{2}$. *Assume that* $\|s_k\|$ *tends to zero, also let there exists constants* $\gamma_1, \gamma_2 \ge 0$ *such that*

$$\|g_k\|^2 \ge \gamma_1 \|s_k\|^2, \qquad (23)$$
$$\|g_{k+1}\|^2 \ge \gamma_2 \|s_k\|. \qquad (24)$$

Then where d_k *generated by (3) and (21) satisfies the sufficient descent condition.*

Proof. The proof is by induction
For $k = 0$ then $d_1 = -g_1$ and $g_1^T d_1 \le -c\|g_1\|^2$, $0 < c < 1$. Let

$$g_k^T d_k \le -c\|g_k\|^2 \qquad (25)$$

Then for $k + 1$ we have

$$d_{k+1} = -g_{k+1} + \left(\frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}\right) s_k. \qquad (26)$$

Multiply both sides by g_{k+1}^T we have

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \left(\frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}\right) s_{k+1}^T g_k \qquad (27)$$
$$\le -\|g_{k+1}\|^2 + \left(\frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}\right) \|s_k\|\|g_k\| \qquad (28)$$
$$\le -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}\|s_k\|\|g_k\|}{y_k^T s_k} \qquad (29)$$

$\because y_k^T s_k \ge -(1-\sigma) g_k^T d_k \ge c\|g_k\|$, with Lipschitz and using (23) and (24) we have

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{L\,\|s_k\|^2\|g_{k+1}\|^2}{c\,\|g_k\|^2} \le -\|g_{k+1}\|^2 + \frac{L\,\|s_k\|^3\gamma_2}{c\,\gamma_1\|s_k\|^2} \tag{30}$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{L\,\gamma_2\,\|s_k\|}{c\,\gamma_1} \tag{31}$$

$\because \|s_k\| \to 0$, $\exists\, 0 < \delta < 1$ then $\frac{L\,\gamma_2\,\|s_k\|}{c\,\gamma_1} \le \delta\|g_{k+1}\|^2$ and by substituting in (31) we have

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \delta\|g_{k+1}\|^2 \tag{32}$$

$$g_{k+1}^T d_{k+1} \le -(1+\delta)\|g_{k+1}\|^2 = -c\,\|g_{k+1}\|^2 \tag{33}$$

5 Global convergence

Next we will show that CG method with $\beta_{k+1}^{OKI\,1}$ converges globally. We need the following Lemma for the convergence of the proposed new algorithm .

Lemma 1 [15] *Suppose that assumption(1) hold. Consider any conjugate gradient method in from (2) and (3), where d_k is a descent direction and α_k is obtained by the strong Wolfe line search. If*

$$\sum_{k>1} \frac{1}{\|d_{k+1}\|^2} = \infty, \tag{34}$$

then we have

$$lim_{k\to\infty}\, g_k = 0. \tag{35}$$

Theorem 2 *Suppose that Assumption (1) hold. Consider the algorithm $x_{k+1} = x_k + \alpha_k d_k$ where d_k is descent direction and α_k is obtained by the Strong Wolfe Conditions (4) and (6), and there exists the nonnegative constant w_1 such that*

$$\|g_{k+1}\|^2 \le w_1\|s_k\| \tag{36}$$

and the function f is uniformly convex function i.e. $\exists$ a constant $\mu \ge 0$ such that for all $x, y \in S$

$$y_k^T s_k \ge \mu\,\|s_k\|^2, \tag{37}$$

then

$$lim_{k\to\infty}\, g_k = 0. \tag{38}$$

Proof. From (21) and (36) we have

$$\left|\beta^{OKI\,1}\right| = \left|\frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{\left(s_k^T g_{k+1}\right)^2}{\left(s_k^T y_k\right)^2}\right| \le \left|\frac{y_k^T g_{k+1}}{y_k^T s_k}\right| + \left|\frac{\left(s_k^T g_{k+1}\right)^2}{\left(s_k^T y_k\right)^2}\right| \tag{39}$$

$$\le \frac{\|y_k\|\|g_{k+1}\|}{\mu\,\|s_k\|^2} + \frac{\|s_k\|^2\,\|g_{k+1}\|^2}{\mu^2\,\|s_k\|^4} \tag{40}$$

$$\le \frac{L\|s_k\|\gamma}{\mu\,\|s_k\|^2} + \frac{\|s_k\|^2\,\|g_{k+1}\|^2}{\mu^2\,\|s_k\|^4} \le \frac{L\gamma}{\mu\,\|s_k\|} + \frac{w_1\|s_k\|^3}{\mu^2\,\|s_k\|^4} \tag{41}$$

$$\le \frac{L\gamma}{\mu\,\|s_k\|} + \frac{w_1}{\mu^2\,\|s_k\|} \tag{42}$$

then $\left|\beta^{OKI\,1}\right| \le \frac{L\gamma}{\mu\,\|s_k\|} + \frac{w_1}{\mu^2\,\|s_k\|}$

$$\|d_{k+1}\| \le \|g_{k+1}\| + |\beta^{OKI\,1}|\|s_k\| \le \gamma + \frac{L\gamma}{\mu} + \frac{w_1}{\mu^2} = \frac{\gamma\mu^2 + L\gamma\mu + w_1}{\mu^2} \tag{43}$$

and

$$\sum \frac{1}{\|d_{k+1}\|} = \sum \frac{\mu^2}{\gamma\mu^2 + L\gamma\mu + w_1} = \infty$$

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

For general non-linear functions the global convergence proof of the our algorithm is based on the Zoutendijk condition (see [18]) combined with sufficient descent condition hold and $\|d_k\|$ is bounded. Suppose that the level set S is bounded and function f is bounded from below. Additionally, assume that there exists $\bar{\gamma} \ge 0$ such that $\|g_k\| \ge \bar{\gamma}$.

Theorem 3 *Consider the algorithm $x_{k+1} = x_k + \alpha_k d_k$ where d_k generated by (3), (21) and α_k computed by Strong Wolfe Conditions (4) and (6). also assume that assumption (1) holds. Where level set S is bonded and objective function bounded from below. If $\|g_k\|^2 \ge w\|s_k\|^2$ then either $g_k = 0$ for some k or $lim_{k\to\infty}\|g_k\| = 0$.*

Proof. We have

$$|\beta^{OKI\,1}| = \left|\frac{y_k^T g_{k+1}}{y_k^T s_k} - \frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}\right| \le \left|\frac{y_k^T g_{k+1}}{y_k^T s_k}\right| + \left|\frac{(s_k^T g_{k+1})^2}{(s_k^T y_k)^2}\right| \tag{44}$$

$$\le \frac{\|y_k\|\|g_{k+1}\|}{y_k^T s_k} + \frac{\|s_k\|\|g_{k+1}\|}{y_k^T s_k} \tag{45}$$

$\because y_k^T s_k \ge -(1-\sigma) g_k^T s_k$ and $g_k^T s_k \le -c\,\|g_k\|^2$ then
$y_k^T s_k \ge (1-\sigma)c\|g_k\|^2$ that is

$$|\beta^{OKI\,1}| \le \frac{\|y_k\|\|g_{k+1}\|}{(1-\sigma)c\|g_k\|^2} + \frac{\|s_k\|\|g_{k+1}\|}{(1-\sigma)c\|g_k\|^2} \le \frac{L\|s_k\|\|g_{k+1}\|}{(1-\sigma)c\|g_k\|^2} + \frac{\|s_k\|\|g_{k+1}\|}{(1-\sigma)c\|g_k\|^2} \tag{46}$$

$\because \|g_k\|^2 \ge w\|s_k\|^2$ and $\|g_{k+1}\| \le \bar{\gamma}\|s_k\|$

$$|\beta^{OKI\,1}| \le \frac{L\|s_k\|\bar{\gamma}}{(1-\sigma)cw\|s_k\|^2} + \frac{\|s_k\|\,\bar{\gamma}}{(1-\sigma)cw\|s_k\|^2} = \left(\frac{L\,\bar{\gamma} + \bar{\gamma}}{(1-\sigma)cw}\right)\left(\frac{1}{\|s_k\|}\right) \tag{47}$$

$$\therefore \|d_{k+1}\| \le \|g_{k+1}\| + |\beta^{OKI\,1}|\|s_k\| \le \bar{\gamma} + \frac{L\,\bar{\gamma} + \bar{\gamma}}{(1-\sigma)cw} \tag{48}$$

Since S is bounded and objective function is bounded from below with Wolfe Condition $(f_{k+1} - f_k \le \rho\alpha_k g_k^T d_k)$ it follows that the Zoutendijk condition holds i.e.

$$\sum \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{49}$$

By descent property Theorem 1

$$\sum_{i=1}^{\infty} \frac{c^4}{\|s_k\|^2} \leq \sum_{i=1}^{\infty} \frac{\|g_k\|^4}{\|s_k\|^2} \leq \sum_{i=1}^{\infty} \frac{1}{w^2} \frac{g_k^T d_k}{\|s_k\|^2} < \infty \quad (50)$$

which contradicts with (43). Hence $\lim_{k\to\infty} \|g_k\| = 0$.

6 Numerical results

Here, we will present the numerical results of the suggested new approach, which were achieved by the use of the new formula for the β_k^{OKI1} conjugation coefficients and the Wolfe (4) and (6) conditional set of test functions in the unconstrained optimization derived from Andrei (2008)[19]. The practical side of unconstrained optimization algorithms is always required since it is complimentary to the theoretical side in the computation of these algorithms. If we want to fully grasp the algorithm's potential, we must put it to practical use by testing it on a variety of non-linear unconstrained situations. Many test functions have been chosen to assess the performance of the proposed method, and they have been included in this article and are explained in detail in the Appendix to this study. The functions are chosen for dimensions n=100,...,1000, and by comparing the performance of the new suggested algorithms with the HS and FR algorithms, it is determined that the algorithms are superior[20]. The terminating condition is $\|g_k\| = 10^{-6}$ for the g_k variable. All of the code is written in the FORTRAN language with double precision and the F77 default compiler parameters. In most cases, the test functions begin with a standard starting point and subsequently provide summary numerical findings, which are shown in Matlab's figures (1), (2), and (3).
Figure (1) shows that the OKI1 algorithm requires the fewest function computations, followed by the HS and FR algorithms.
Regarding the number of iterations in Figure (2), we see that the curve is located at the top, indicating that it requires the fewest repetitions.
Regarding time, Figure (3) shows that the OKI1 algorithm is clearly better to the HR and FR algorithms.

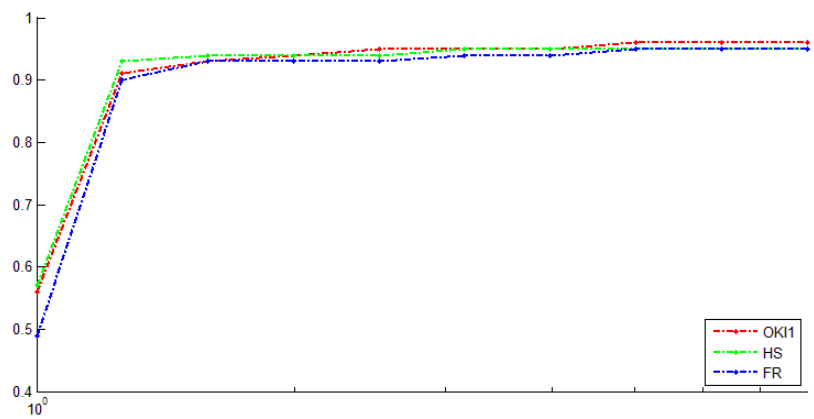

Figure 1: Performance based on function

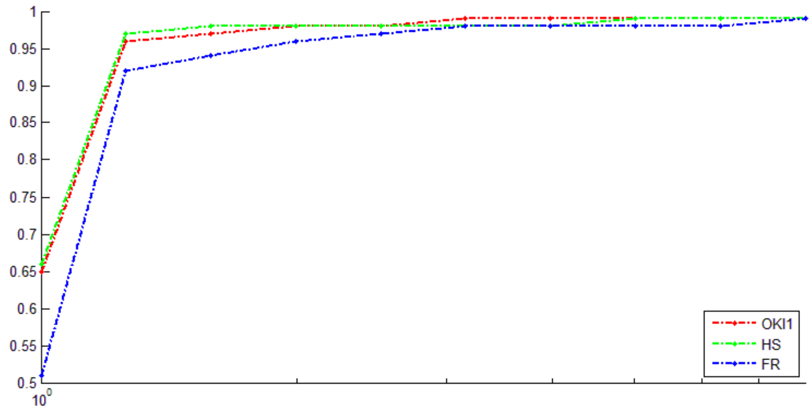

Figure 2: Performance based on iteration

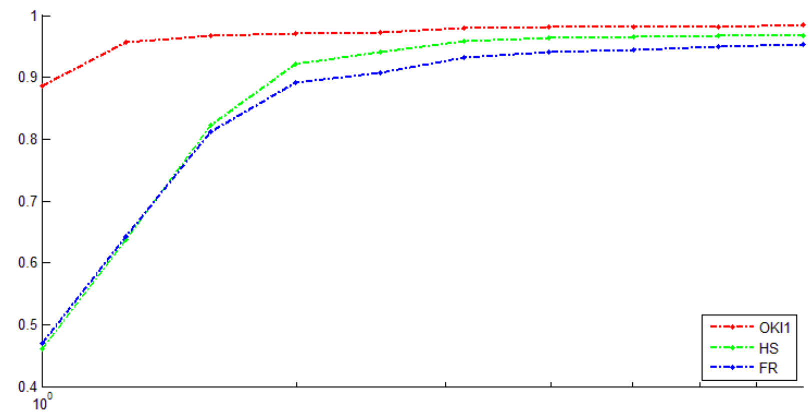

Figure 3: Performance based on time

7 Conclusions

Using the (Hestenes-Stiefel HS) technique, we present a novel Conjugate Gradient Method in this study. We investigate the quality of these formulae from a scientific standpoint, and we demonstrate the attribute of descent and convergence by using a number of assumptions to support our findings. We also looked into the features of the matrices and compared their performance to that of the (HS and FR) approaches, which yielded positive findings for us.

Conflict of Interest Declaration

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

Ethics Committee Approval and Informed Consent

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

References

[1] O. M. T. Wais, Two modified spectral conjugate gradient methods for optimization, Coll. Basic Educ. Res. J., 14 (4) 2018, 533–553.

[2] H. M. Khudhur and K. K. Abbo, A new conjugate gradient method for learning fuzzy neural networks, J. Multidiscip. Model. Optim., 3 (2) 2021, 57–69.

[3] K. K. Abbo and O. M. T. Waiss, Investigation on extended conjugate gradient non-linear methods for solving unconstrained optimization, J. Educ. Sci., 26(1) 2013, 74–78.

[4] Z. M. Abdullah, M. Hameed, M. K. Hisham, and M. A. Khaleel, Modified new conjugate gradient method for unconstrained optimization, Tikrit J. Pure Sci., 24 (5) 2019, 86–90.

[5] K. K. Abbo, Y. A. Laylani, and H. M. Khudhur, Proposed new scaled conjugate gradient algorithm for unconstrained optimization, Int. J. Enhanc. Res. Sci. Technol. Eng., 5(7) 2016.

[6] H. N. Jabbar, K. K. Abbo, and H. M. Khudhur, Four--term conjugate gradient (cg) method based on pure conjugacy condition for unconstrained optimization, Kirkuk Univ. J. Sci. Stud., 13(2) 2018, 101–113.

[7] M. A. Wolfe, Numerical methods for unconstrained optimization: An introduction. Van Nostrand Reinhold New York, 1978.

[8] Y. H. Dai and Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, SIAM J. Optim., 10(1) 1999, 177–182.

[9] H. M. Khudhur, Numerical and analytical study of some descent algorithms to solve unconstrained optimization problems, University of Mosul, 2015.

[10] A. Al-Bayati, B. Hassan, and H. Jabbar, "A new class of nonlinear conjugate gradient method for solving unconstrained minimization problems," in *ICCISTA 2019 - IEEE International Conference on Computing and Information Science and Technology and their Applications 2019*, 2019, pp. 1–4, doi: 10.1109/ICCISTA.2019.8830657.

[11] Y. Dai and L. Liao, New conjugacy conditions and related non-linear CG methods, Appl. Math Optim., 43(1) 2001, 87-101.

[12] M. R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Res. Natl. Bur. Stand., 49(6) 1952, 409–436.

[13] E. Polak and G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, ESAIM: Math. Model. Numer. Anal., 3(R1) 1969, 35–43.

[14] R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients, Comput. J., 7(2) 1964, 149–154.

[15] R. Fletcher, Practical Methods of Optimization, John Wiley & Sons, New York, 1987.

[16] Y. Liu and C. Storey, Efficient generalized conjugate gradient algorithms, part 1: theory, J. Optim. Theory Appl., 69(1) 1991, 129–137.

[17] B. A. Hassan, O. M. T. Wais, and A. A. Mahmood, A class of descent conjugate gradient methods for solving optimization problems, Appl. Math. Sci., 13(12) 2019, 559–567.

[18] J. Nocedal and S. Wright, Numerical Optimization. Springer Science & Business Media, New York, 2006.

[19] N. Andrei, An unconstrained optimization test functions collection, Adv. Model. Optim, 10(1) 2008, 147–161.

[20] E. D. Dolan and J. J. Moré, Benchmarking optimization software with performance profiles, Math. Program., 91 (2) 2002, 201–213.

Osama M. T. Wais, ORCID: https://orcid.org/ 0000-0002-7725-2260
Khalil K. Abbo, ORCID: https://orcid.org/ 0000-0001-5858-625x
Ibrahim A. Saleh, ORCID: https://orcid.org/0000-00020358-633x