**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 27, No. 1, July 2022, pp. 86~91
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v27.i1.pp86-91 

# Using a new type of formula conjugate on the gradient methods

**Basim A. Hassan, Ranen M. Sulaiman**
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq

**Article Info**

***Article history:***

Received Apr 7, 2022
Revised May 10, 2022
Accepted Jun 2, 2022

***Keywords:***

Conjugate gradient
Convergence property
Descent property
Formula conjugate gradient
Numerical results

**ABSTRACT**

Unconstrained optimization problems, such as energy minimization, can be solved using the conjugate gradient method. For its major characteristic, the optimal formula conjugate encompasses all conjugate gradient algorithms. In conjugate gradient approaches, the formula conjugate is typically the focus point and it's playing a very important role for conjugate gradient approaches. To offer the essential descent criteria in this work, we devised a novel formula based on the second order Taylor which have the descent property too. Our research focused on our suggested method's-convergence property with Wolfe condition is established and numerical performance. Comparison to FR-method, the new algorithm shows significant improvement in numerical results.



***Corresponding Author:***

Basim A. Hassan
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul
Mosul, Iraq
Email: basimah@uomosul.edu.iq

## 1. INTRODUCTION

The main goal of this study is to present an efficient optimization approach for minimizing a function's objective function $f: R^n \rightarrow R$, see [1]. The following equation gives the usual form of iterations for discovering the extreme points of the function:

$$x_{k+1} = x_k + \alpha_k d_k \quad (1)$$

certainly, $\alpha_k$ and $d_k$ are the most significant aspects of an optimization model (1), and they determine the effectiveness of a technique. Hassan [2] has further information. The length of a step can be determined by:

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \quad (2)$$

typically, the Wolfe line search criteria [3], [4] are used to determine the step length $\alpha_k$ in (1):

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \delta \alpha_k g_k^T d_k \quad (3)$$

$$d_k^T g(x_k + \alpha_k d_k) \geq \sigma d_k^T g_k \quad (4)$$

where $0 < \delta < \sigma < 1$. The vector direction is computed as follows in the general conjugate gradient method:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \quad (5)$$

notice that (4) is a generic update formula for the conjugate gradient parameter-based search direction computation. The key point in various literatures (refer to [5]-[8]) is to pick the scalar $\beta_k$, which leads to a variety of conjugate gradient techniques depending on the $\beta_k$. Dai *et al.* [1] and Fletcher and Reeves [9] devised the conjugate gradient approach, which goes like this:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{g_k^T g_k} \quad , \quad \beta_k^{DR} = \frac{\|g_{k+1}\|^2}{d_k^T y_k} \tag{6}$$

denote $y_k$ to be the gradient change. Many conjugate gradient methods have been created, and numerous publications have been written on the subject, addressing both computational and theoretical aspects. Iiduka and Narushima [10], $\beta_k$ is calculated as:

$$\beta_k^{HY} = \frac{\|g_{k+1}\|^2}{2/\alpha_k(f_k - f_{k+1})} \tag{7}$$

as a result, the parameter's value has changed $\beta_k$ in [2] is as (8).

$$\beta_k^{B} = \frac{\|g_{k+1}\|^2}{(f_k - f_{k+1})/\alpha_k - g_k^T d_k/2} \tag{8}$$

accelerated gradient descent techniques were classified as a subclass of gradient descent methods during the creation of gradient models [11]. Scalar $\beta_k$ may be chosen in a variety of ways, and a good choice leads to improved convergence. These are descent-direction approaches as well.

In this study, we describe a conjugate gradient technique based on the conjugacy criterion. Our recommended strategy ensures enough descent using any line search. We can establish global convergence of our suggested approach under correct conditions if the line search fits the Wolfe criteria. The numbers suggest that the technique we propose is feasible.

## 2. DERIVING NEW COEFFICIENT CONJUGATE

The conjugacy condition is well recognized to play a crucial role in convergence analysis and numerical computation. The conjugacy requirement is used in this new strategy. The conjugate gradient approach is such that the conjugacy requirement holds in the situation when f is quadratic model and V is determined using the precise line search, namely:

$$d_{k+1}^T Q d_k = 0 \tag{9}$$

as a result of adding (5) to the previous equation:

$$\beta_k = \frac{g_{k+1}^T Q s_k}{d_k^T Q s_k} \tag{10}$$

the new conjugate gradient formula is derived using the second order Taylor formula. Allow me to explain:

$$f(x) = f(x_{k+1}) + g_{k+1}^T(x - x_{k+1}) + \frac{1}{2}\,(x - x_{k+1})^T Q(x_{k+1})(x - x_{k+1}) \tag{11}$$

the $f(x)$ has the following gradient as:

$$g_{k+1} = g_k + Q(x_{k+1}) s_k \tag{12}$$

using (3) with exact line search in (11), we get:

$$s_k^T Q(x_{k+1}) s_k = (f_{k+1} - f_k) + 3/2\, \alpha_k g_k^T g_k \tag{13}$$

To write (13), we have (14).

$$d_k^T Q(x_{k+1}) s_k = (f_{k+1} - f_k)/\alpha_k + 3/2\, g_k^T g_k \tag{14}$$

putting (15) in (10), we obtain:

$$\beta_k = \frac{g_{k+1}^T y_k}{(f_{k+1} - f_k)/\alpha_k + 3/2 g_k^T g_k} \quad (15)$$

we obtain the following results when we use exact line search.

$$\beta_k = \frac{\|g_{k+1}\|^2}{(f_{k+1} - f_k)/\alpha_k + 3/2 g_k^T g_k} \quad (16)$$

As a result of our novel formula, the so-called new method. Using this way, we will obtain the new algorithm which generates different formulae by using the second order Taylor formula and conjugacy condition. However, we may now present a new algorithm as follows, employing our suggested new parameter (16) as:
**Step 1:** Let $x_1 \in R^n$ and $d_1 = -g_1$. Put $k = 1$.
**Step 2:** If $\|g_{k+1}\| \leq 10^{-6}$, then stop.
**Step 3:** Estimate $\alpha_k$ and which of the three fulfill (3-4) and variables should be updated $x_{k+1} = x_k + \alpha_k d_k$.
**Step 4:** Calculate the value of $\beta_k$, as defined in (16).
**Step 5:** Put $k = k + 1$ and go to 5.
**Theorem (2.2):** for the new approach, we have the following in the form of (1) and (5), with the new parameter given by (16).

$$d_{k+1}^T g_{k+1} < 0 \text{ and } d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k \quad (17)$$

**Proof:** from evident that $d_1 = -g_1$ satisfies $d_1^T g_1 < 0$. Let that $d_k^T g_k < 0$ hold for any $k$. Straight from (5) and (16) we obtained that:

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \beta_k d_k^T g_{k+1} = -\beta_k((f_{k+1} - f_k)/\alpha_k + 3/2\, g_k^T g_k)) + \beta_k d_k^T g_{k+1} \quad (18)$$

using (12) in above equation, we get:

$$d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k \quad (19)$$

since $d_k^T g_k < 0$, then we get:

$$d_{k+1}^T g_{k+1} < 0 \quad (20)$$

so the proof is completed.

## 3. CONVERGENCE ANALYSIS

Each method in the numerical optimization methods has different convergence properties and needed some assumptions. Our recommended strategy ensured enough descent using any line search. In the study of innovative conjugate gradient procedure' global convergence, the assumptions listed below are commonly utilized. Assumption A:
- The bounded level set $\Omega = \{u/f(u) \leq f(u_0))\}$.
- The function f is smooth and its gradient is Lipschitz continuous in some neighborhood $\Omega$ of $L_0$, i.e.,where $L > 0$ constant, such that:

$$\|g(v) - g(z)\| \leq L\|v - z\|, \forall v, z \in \Omega \quad (21)$$

for more details see [12]. In order to prove the convergence property of our new algorithm and some algorithem in optimization, we use the proof idea of a researcher Zoutendijk. The Zoutendijk condition is a lemma that is widely used to check the conjugate gradient technique's global convergence. Zoutendijk was the first to say it in [13].

**Lemma (3.1):** assume that Assumption A is true. The sequence $\{x_k\}$ produced by any iteration algorithm thus satisfies:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{22}$$

**Theorem (3.2):** the sequence $\{x_k\}$ created by the new approach meets the following conditions:

$$\lim_{k\to\infty} inf \left\|g_k\right\| = 0 \tag{23}$$

**Proof:** assume that (23) is false by induction. The (8) is changed to $d_{k+1} + g_{k+1} = \beta_k d_k$, and when both sides are squared, we get:

$$\|d_{k+1}\|^2 + \left\|g_{k+1}\right\|^2 + 2d_{k+1}^T g_{k+1} = (\beta_k)^2 \|d_k\|^2 \tag{24}$$

putting (24), results in:

$$\|d_{k+1}\|^2 = \frac{\left(d_{k+1}^T g_{k+1}\right)^2}{\left(d_k^T g_k\right)^2} \|d_k\|^2 - 2d_{k+1}^T g_{k+1} - \left\|g_{k+1}\right\|^2 \tag{25}$$

taking (25) and dividing it by $\left(d_{k+1}^T g_{k+1}\right)^2$, we have:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} = \frac{\|d_k\|^2}{(d_k^T g_k)^2} = \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} - \frac{2}{d_{k+1}^T g_{k+1}} \tag{26}$$

we may deduce the following from the equation:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \le \frac{\|d_k\|^2}{\left(d_k^T g_k\right)^2} - \left(\frac{\|g_{k+1}\|}{\left(d_{k+1}^T g_{k+1}\right)^2} + \frac{1}{\|g_{k+1}\|^2}\right) + \frac{1}{\|g_{k+1}\|} \le \frac{\|d_k\|^2}{\left(d_k^T g_k\right)^2} + \frac{1}{\|g_{k+1}\|^2} \tag{27}$$

Hence,

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \le \sum_{i=1}^{k+1} \frac{1}{\|g_i\|^2} \tag{28}$$

let there exists $c_1 > 0$ such that $\left\|g_k\right\| \ge c_1$ for all $k \in n$ , then:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \le \frac{k+1}{c_1^2} \tag{29}$$

using the assumption and (29) as a guide, we can see that:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty \tag{30}$$

we know that $\lim_{k\to\infty} inf \|g_k\| = 0$ holds because of Lemma 1.

## 4. NUMERICAL RESULTS

The results of numerical experiments are presented in this section. The FR-Algorithm was compared to the new conjugate gradient algorithm. In Fortran, both algorithms are written. Andrei [14] is where you'll find the test problems. The following papers and references are available to interested readers (Ref. [15]-[19]). We chose (15) unconstrained optimization problems in extended, and we investigated numerical experiments with the amount of variables n=100;1,000 for each test function. As a termination condition, we apply the inequality $\|g_{k+1}\| \le 10^{-6}$. Many papers have proposed this method for optimization problems also [20]-[26]. The Wolfe condition test results were collected and reported in Table 1, where each column has the following meanings; $\delta=0.0001$ and $\sigma=0.5$. The problem's name is problem, the problem's size is dim, the number of iterations (NOI), and the number of function evaluations are all terms used to describe the problem (NOF). In terms of iterations and function evaluations, Table 1 indicates how many difficulties these methods have resolved.

Table 1 compares the new methods to the Fletcher and Reeves [1] convex optimization technique; as demonstrated in Table 2. The new algorithm saves NI, NR, and NF overall when compared to the conventional Fletcher and Reeves (FR) methodology. Particularly for our collection of test issues, which showed the performance improvement of the new algorithm.

Table 1. The FR and new methods numerical results

| P. No | n | FR algorithm | | | New .algorithm | | |
|---|---|---|---|---|---|---|---|
| | | NI | NR | NF | NI | NR | NF |
| Trigonometric | 100 | 19 | 12 | 35 | 18 | 11 | 33 |
| | 1000 | 38 | 23 | 65 | 37 | 23 | 65 |
| Extended Rosenbrock | 100 | 47 | 18 | 93 | 40 | 23 | 88 |
| | 1000 | 78 | 45 | 131 | 38 | 21 | 83 |
| Extended Beale | 100 | 32 | 15 | 52 | 16 | 9 | 30 |
| | 1000 | 22 | 10 | 42 | 13 | 7 | 25 |
| Generelized Tridiegonal 1 | 100 | 25 | 11 | 43 | 23 | 8 | 44 |
| | 1000 | 46 | 28 | 741 | 40 | 23 | 528 |
| Extended Tridiagonal 1 | 100 | 32 | 13 | 64 | 16 | 7 | 33 |
| | 1000 | 77 | 46 | 169 | 15 | 6 | 30 |
| Generalized Tridiagonal 2 | 100 | 37 | 8 | 67 | 40 | 15 | 59 |
| | 1000 | 73 | 27 | 115 | 49 | 19 | 77 |
| Extended Himmelblau | 100 | 12 | 5 | 25 | 10 | 6 | 19 |
| | 1000 | 14 | 6 | 29 | 10 | 6 | 19 |
| Extended PSC1 | 100 | 15 | 9 | 31 | 8 | 6 | 17 |
| | 1000 | 8 | 6 | 17 | 7 | 5 | 15 |
| Q. Diagonal Perturbed | 100 | 124 | 41 | 231 | 48 | 9 | 87 |
| | 1000 | 445 | 196 | 711 | 197 | 46 | 352 |
| Extended Wood | 100 | 71 | 35 | 110 | 26 | 11 | 50 |
| | 1000 | 47 | 15 | 84 | 34 | 12 | 65 |
| Extended Hiebart | 100 | 101 | 40 | 217 | 83 | 47 | 190 |
| | 1000 | 101 | 40 | 214 | 86 | 49 | 194 |
| ARWHEAD (CUTE) | 100 | 9 | 4 | 18 | 8 | 5 | 16 |
| | 1000 | 12 | 7 | 82 | 8 | 5 | 16 |
| NONDIA (CUTE) | 100 | 13 | 7 | 25 | 11 | 6 | 22 |
| | 1000 | 15 | 7 | 29 | 9 | 6 | 19 |
| DIXMAANE (CUTE) | 100 | 121 | 65 | 218 | 89 | 34 | 143 |
| | 1000 | 345 | 169 | 634 | 252 | 86 | 400 |
| Partial Perturbed Q | 100 | 74 | 21 | 123 | 79 | 20 | 120 |
| | 1000 | 370 | 88 | 616 | 251 | 66 | 436 |
| **Total** | | **2423** | **1017** | **5031** | **1561** | **597** | **3275** |

Table 2. The performance percentage for the new CG algorithm compared with FR method

| | NF | NR | NI |
|---|---|---|---|
| FR method | 100% | 100% | 100% |
| New method | 64.42% | 58.70% | 65.09% |

## 5. CONCLUSION

In this study, we offer a novel nonlinear CG-algorithm based on the strictly convex quadratic function described by (10) that is globally convergent, functional, and fulfills the descent property under certain assumptions. The novel types presented in this research are successful, according to computer studies. The new algorithm was sufficient and effective.

## REFERENCES


[1] Y. H. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y. Yuan, "Convergence properties of nonlinear conjugate gradient methods," S*IAM Journal on Optimization*, vol. 10, no. 2, pp. 345–358, Feb. 2000, doi: 10.1137/S1052623494268443.

[2] B. A. Hassan, "A new formula for conjugate parameter computation based on the quadratic model," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 13, no. 3, pp. 954-961, 2019, doi: 10.11591/ijeecs.v13.i3.pp954-961.

[3] P. Wolfe, "Convergence Conditions for Ascent Methods," *SIAM review*, vol. 11, no. 2, 1969, doi: 10.1137/1011036.

[4] P. Wolfe, "Convergence Conditions for Ascent Methods. II: Some Corrections," *SIAM review*, vol. 13, no. 2, 1971, doi: 10.1137/1013035.

[5] R. Fletcher, "Practical methods of optimization," in *John Wiley & Sons*, 1987. [Online]. Available: https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118723203.fmatter.

[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, pp. 409-436, 1952, doi: 10.6028/JRES.049.044.

[7] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: Theory," *Journal of optimization theory and applications*, vol. 69, no. 1, pp. 129–137, Apr. 1991, doi: 10.1007/BF00940464.

[8] E. Polak and G. Ribiere, "Note on the convergence of conjugate direction methods (in French: Note sur la convergence de méthodes de directions conjuguées)," *ESAIM: Mathematical Modeling and Numerical Analysis-Mathematical Modeling and Numerical Analysis*, vol. 3, no. R1, pp. 35–43, 1969.

[9] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The computer journal*, vol.7, pp. 149-154, 1964, doi: 10.1093/COMJNL/7.2.149.

[10] H. Iiduka and Y. Narushima, "Conjugate gradient methods using value of objective function for unconstrained optimization," *Optimization Letters*, vol. 6, no. 5, pp. 941–955, Jun. 2012, doi: 10.1007/s11590-011-0324-0.

[11] P. Stanimirovic and M. Miladinovic, "Accelerated gradient descent methods with line search," *Numerical Algorithms*, vol. 54, pp. 503–520, Aug. 2010, doi: 10.1007/s11075-009-9350-8.

[12] Y.-H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on optimization*, vol. 10, Jun. 1999, doi: 10.1137/S1052623497318992.

[13] G. Zoutendijk, "Nonlinear programming, computational methods," in *Integer and Nonlinear Programming*, J. Abadie, ed., North-Holland, Amsterdam, pp. 37-86, 1970.

[14] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, 2008.

[15] B. A. Hassan, "A new type of quasi-newton updating formulas based on the new quasi-newton equation," *Numerical Algebra, Control & Optimization*, vol. 10, no. 2, pp. 227-235, Jun. 2020, doi: 10.3934/naco.2019049.

[16] B. A. Hassan, "A modified quasi-Newton methods for unconstrained optimization," *Italian Journal of Pure and Applied Mathematics*, vol. 42, no. 2019, pp. 504-511, 2019.

[17] B. A. Hassan and R. Sulaiman, "A new class of self-scaling for quasi-newton method based on the quadratic model," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 21, pp. 1830-1836, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1830-1836.

[18] B. A. Hassan, F. Alfarag, A. Ibrahim, and A. Abubakar, "An improved quasi-Newton equation on the quasi-Newton methods for unconstrained optimizations," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 22, pp. 389–397, doi: 10.11591/ijeecs.v22.i2.pp997-1005.

[19] H. N. Jabbar and B. A. Hassan, "Two-versions of descent conjugate gradient methods for large-scale unconstrained optimization," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 22, no. 3, pp. 1643-1649, 2021, doi: 10.11591/ijeecs.v22.i3.pp1643-1649.

[20] I. M. Sulaiman, N. A. Bakar, M. Mamat, B. A. Hassan, M. Malik, and A. M. Ahmed, "A new hybrid conjugate gradient algorithm for optimization models and its application to regression analysis," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 23, no. 2, pp. 1100-1109, 2021, doi: 10.11591/ijeecs.v23.i2.pp1100-1109.

[21] M. Malik, A. Abubakar, S. Ibrahim, M. Mamat, S. Abas, and S. Firman, "A new three-term conjugate gradient method for unconstrained optimization with applications in portfolio selection and robotic motion control," *IAENG International Journal of Applied Mathematics*, vol. 51, pp. 471–486, Aug. 2021.

[22] M. Malik, M. Mamat, S. Abas, S. Ibrahim, and S. Firman, "Performance analysis of new spectral and hybrid conjugate gradient methods for solving unconstrained optimization problems," *IAENG International Journal of Applied Mathematics*, vol. 48, no. 1, pp. 66–79, Feb. 2021.

[23] L. Ladislav and V. Jan, "Nonlinear conjugate gradient methods, programs and algorithms of numerical mathematics," in *Proceedings of Seminar. Institute of Mathematics AS CR*, Dec. 2015, pp.130--135.

[24] B. A. Hassan, H. Alrehan, and A. Aladool, "A new kind of parameter conjugate gradient for unconstrained optimization," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 17, no. 1, pp 404-411, 2020, doi: 10.11591/ijeecs.v17.i1.pp404-411.

[25] B. A. Hassan and E. S. Al-Rawi, "A modified Newton's method for solving functions of one variable," *Italian Journal of Pure and Applied Mathematics*, Dec. 2021.

[26] W. W. Hager and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific Journal of Optimization*, vol. 2, pp. 35–58, Aug. 2006.


## BIOGRAPHIES OF AUTHORS

**Basim A. Hassan** is currently a Professor in Department of Mathematics, College of Computer Science and Mathematics, University of Mosul. He obtained his M.Sc and Ph.D degrees in Mathematics from the University of Mosul, in 2000 and 2010, respectively with specialization in optimization. To date, he has published more than 80 research paper in various international journals and conferences. He currently works on iterative methods. His research interest in applied mathematics, with a field of concentration of optimization include conjugate gradient, steepest descent methods, Broyden's family and quasi-Newton methods with application in signal recovery and image restoration. He can be contacted at email: basimah@uomosul.edu.iq.

**Ranen M. Sulaiman** is currently a Professor in Department of Mathematics, College of Computer Science and Mathematics, University of Mosul. He obtained his M.Sc degrees in Mathematics from the University of Mosul, in 2020 and 2021, respectively with specialization in optimization. To date, he has published more than 4 research paper in various international journals and conferences. He currently works on iterative methods. His research interest in applied mathematics, with a field of concentration of optimization include conjugate gradient, steepest descent methods, Broyden's family and quasi-Newton methods. He can be contacted at email: ranen.sulyman@uomosul.edu.iq.