EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 15, No. 4, 2022, 1908-1916
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

A New Class of Optimization Methods Based on Coefficient Conjugate Gradient

Yoksal A. Laylani[1], Basim A. Hassan[2,*], Hisham M. Khudhur[2]

[1] *Department of Mathematics, College of Sciences, University of Kirkuk, Kirkuk, Iraq*
[2] *Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq*

Abstract. The coefficient conjugate serves as the foundation for many conjugate gradient methods. The quadratic model is used to derive a novel coefficient conjugate in this study. Its global convergence result might be produced under Wolfe line search circumstances. The conjugate gradient method's performance for unconstrained optimization problems is demonstrated through numerical tests.

2020 Mathematics Subject Classifications: 90C30, 65K05, 49M37
Key Words and Phrases: Optimization, Conjugate Gradient, Numerical results

1. Introduction

One of the most important iterative approaches for solving the unconstrained optimization issue is the Conjugate Gradient (CG) method, which is as follows:

$$f(x^*) = \min_{x \in R^n} f(x). \tag{1}$$

where $f : R^n \to R$ is a smooth function, see [7, 11, 12]. The goal function and its gradient are all that is required of the CG techniques in each iteration, see [13]. As a result, this strategy is particularly well suited to solving optimization problems. These approaches use the following iterative formula:

$$x_{k+1} = x_k + \alpha_k d_k. \tag{2}$$

If f is a quadratic, the one-dimensional minimizer along the ray may be calculated analytically as:

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \tag{3}$$

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v15i4.4575

Email addresses: yuksel@uokirkuk.edu.iq (Y. A. Laylani), basimah@uomosul.edu.iq (B. A. Hassan), hisham892020@uomosul.edu.iq (H. M. Khudhur)

Iterative procedures are required for generic non-linear functions, in [9] has further information about this. The Wolfe requirements are frequently employed in the convergence analysis and implementation of conjugate gradient techniques to obtain the step length α_k satisfying:

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

$$d_k^T g(x_k + \alpha_k d_k) \geq \sigma d_k^T g_k \tag{5}$$

where $0 < \delta < \sigma < 1$. More information is available in [8, 15]. In practice, the conjugate gradient search direction for the next iteration looks like this:

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{6}$$

where β_k is a scalar. There are two well-known methods for selecting β_k:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2},\ \beta_{k+1}^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k} \tag{7}$$

separately by the Fletcher-Reeves (FR) approach [2] and the Dai-Yan (DY) method [1]. It approaches have good convergence properties, but their numerical results aren't as good as the others [9]. Many variations of this method have been developed throughout the years, and some are extensively utilized in practice. Take, for example, Hideaki and Yasushi [14] and Basim [3]:

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

These algorithms have a very high numerical efficiency. The quadratic model has been established to improve efficiency for unconstrained problems, in order to maximize the benefits of the original conjugate gradient approaches. A few novel optimization approaches have been presented in this regard [4–6]. Using a conjugacy condition, we provide a modified conjugate gradient technique in this paper. The sufficient descent property is satisfied by the new search direction, which is independent of the line search and the convexity assumption on the objective function. Under the conventional assumptions, the new method has global convergence for generic functions. When dealing with unconstrained optimization issues, numerical experiments show that the new approach is superior.

2. Deriving new Coefficient Conjugate

As everyone is aware, the following outcomes are produced when the objective function to be decreased is quadratic and accurate line searches are used:

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

where Q is Hessian matrix. The conjugacy condition is what it's called. In [13] has further information. Because the $d_{k+1} = -g_{k+1} + \beta_k d_k$ for the conjugate gradient technique, the following is the result:

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

as a result, a coefficient conjugacy is obtained. We now utilize a quadratic model to derive the new conjugate gradient approach, which can be represented as:

$$f_{k+1} = f_k + s_k^T g_k + \frac{1}{2} s_k^T Q(x_k) s_k \tag{11}$$

We may calculate the first order derivative:

$$\nabla f_{k+1} = g_k + Q(x_k) s_k \tag{12}$$

Using (12) in (11) we get the following:

$$s_k^T y_k + f_{k+1} - f_k = \frac{1}{2} s_k^T Q(x_k) s_k \tag{13}$$

We may write the equation (13) by using (3), as follows:

$$d_k^T Q(x_k) s_k = \frac{1}{2} \frac{\alpha_k (g_k^T d_k)^2}{(s_k^T y_k + (f_{k+1} - f_k))} = \rho_k s_k^T y_k \tag{14}$$

We generate a new formula β_k by plugging (14) into (10):

$$\beta_k = \frac{g_{k+1}^T y_k}{\rho_k s_k^T y_k},\ \rho_k^1 = \frac{1}{2} \frac{\alpha_k (g_k^T d_k)^2}{s_k^T y_k (s_k^T y_k + (f_{k+1} - f_k))} \tag{15}$$

Because f is a quadratic function and exact line search is used, the following is the result:

$$\beta_k = \frac{\|g_{k+1}\|^2}{\rho_k s_k^T y_k},\ \rho_k^1 = \frac{1}{2} \frac{\alpha_k (g_k^T d_k)^2}{s_k^T y_k (s_k^T y_k + (f_{k+1} - f_k))} \tag{16}$$

Using and exact line search in (13), then (16) reduces to:

$$\rho_k^2 = \frac{1}{2} \frac{\alpha_k (g_k^T d_k)^2}{s_k^T y_k (-s_k^T g_k + (f_{k+1} - f_k))} \tag{17}$$

and

$$\rho_k^3 = \frac{1}{2} \frac{\alpha_k (g_k^T d_k)^2}{s_k^T y_k (\alpha_k g_k^T g_k + (f_{k+1} - f_k))} \tag{18}$$

As a result, our proposal, the so-called BN, becomes clear. The proposed technique is provided with a new algorithm based on the aforementioned.

2.1. Algorithm

Step 1. Given$x_1 \in R^n$. Set $k = 1$ and $d_1 = -g_1$.

Step 2. Let the step size α_k satisfying the (4) and (5).

Step 3. Let $x_{k+1} = x_k + \alpha_k d_k$. If $\|g_{k+1}\| \leq 10^{-6}$, then stop.

Step 4. Update β_k by the (16), then d_{k+1} by (6).

Step 5. Set $k = k + 1$ and go to Step2.

The descent condition exposes a useful characteristic for the new formula conjugate gradient technique. Additional properties are highly crucial in the convergence proved in this investigation. The preceding material is summarized in the following theorem.

Theorem 1. *If we use the new way to create $\{x_k\}$ and $\{d_k\}$, we get:*

$$d_{k+1}^T g_{k+1} < 0, \, and \, d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k \tag{19}$$

Proof. It goes without saying that if $d_k = -g_k$ then $d_1^T g_1 < 0$. Assume that $d_k^T g_k < 0$ for any k. From (6) and (17), it is simple to deduce:

$$\begin{aligned} & d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \beta_k d_k^T g_{k+1} \\ & = -\beta_k \left[\rho_k s_k^T y_k\right] + \beta_k d_k^T g_{k+1} \\ & = \beta_k [d_k^T g_{k+1} - \rho_k s_k^T y_k] \end{aligned} \tag{20}$$

We may deduce the following from equations (15) and (20):

$$d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k \tag{21}$$

It is obvious that $d_k^T g_k < 0$, thus we get:

$$d_{k+1}^T g_{k+1} < 0 \tag{22}$$

The proof is finished.Other methods, can be proven in the same way.

3. Convergence Analysis

We look at how new algorithms are converged. On the objective function, the following assumptions are made.

- $D = \{x | f(x) \leq f(x_0)\}$ is a bounded level set.
- In some neighborhood D that contains L_0, the gradient is Lipschitz continuous; that is, L exists such that:

$$\|g(\upsilon) - g(\omega)\| \leq L \|\upsilon - \omega\| \, \forall \upsilon, \, \omega \in D \tag{23}$$

For further information, read [1, 10] and Zoutendijk [16] achieved the following significant finding.

Lemma 1. *Allow all assumptions to be true. Consider any iteration algorithm that uses the Wolfe line search to get α_k. Then:*

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \tag{24}$$

Theorem 2. *Assume that all of your assumptions are correct. If (19) is satisfied by formula β_k, we have:*

$$\lim_{k\to\infty} \inf \|g_k\| = 0 \tag{25}$$

Proof. Assume that (25) isn't true by induction. (6) is rewritten as $d_{k+1}+g_{k+1} = \beta_k d_k$, and upon squaring both sides, we obtain:

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

Applying (21), yields:

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

Equation (27) is divided into $(d_{k+1}^T g_{k+1})^2$, we get:

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} = \frac{\|d_k\|^2}{(d_k^T g_k)^2} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} - \frac{2}{d_{k+1}^T g_{k+1}} \tag{28}$$

The equation becomes, by completing the square:

$$\begin{aligned} \frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} &\le \frac{\|d_k\|^2}{(d_k^T g_k)^2} - \left(\frac{\|g_{k+1}\|}{(d_{k+1}^T g_{k+1})} + \frac{1}{\|g_{k+1}\|^2} \right) + \frac{1}{\|g_{k+1}\|^2} \\ &\le \frac{\|d_k\|^2}{(d_k^T g_k)^2} + \frac{1}{\|g_{k+1}\|^2} \end{aligned} \tag{29}$$

Hence,

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} \le \sum_{i=1}^{k+1} \frac{1}{\|g_i\|^2} \tag{30}$$

Suppose that there exists $c_1 > 0$ such that $\|g_k\| \ge c_1$ for all $k \in n$.Then :

$$\frac{\|d_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2} < \frac{k+1}{c_1^2} \tag{31}$$

Assume and note in Equation (31) that:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty \tag{32}$$

Based on Lemma 1, we get $\lim_{k\to\infty} \inf \|g_k\| = 0$ holds. Other methods, can be proven in the same way.

4. Numerical Results

The outcomes of mathematical experiments are presented in this section. The FR-Algorithm was contrasted with the innovative conjugate gradient algorithm. Fortran was used to create both algorithms. Section 1. contains the exam questions. We investigated numerical experiments for 15 extended unconstrained optimization problems with the number of variables for each test function. We use the inequality as a termination condition. Table 1 contains the findings of the Wolfe condition test, with the following definitions for each column: The terminology used to describe the issue comprises problem, problem size, dim, number of iterations (NI), and number of function evaluations (NF). Table 1 shows how many challenges these strategies have addressed in terms of iterations and function evaluations.

Problem numbers indicant for : 1. is the Trigonometric, 2. is the Extended Rosenbrock, 3. is the Hager, 4. is the Extended Tridiagonal 1, 5. is the Generalized Tridiagonal 2, 6. is the Extended PSC1, 7. is the Extended Tridiagonal 2, 8. is EDENSCH (CUTE), 9. is the STAIRCASE S1, 10. is the DENSCHNA (CUTE), 11. is the DENSCHNC (CUTE), 12. is the Extended White & Holst, 13. is the Extended Block-Diagonal BD2, 14. is the Generalized quartic GQ2, 15. is the Extended Beale.

Table 1: The numerical results of the FR and New methods

		FR algorithm		BN1 (16)		BN2 (17)		BN3 (18)	
P. No.	n	NI	NF	NI	NF	NI	NF	NI	NF
1	100	19	35	18	33	18	33	18	34
	1000	38	65	32	57	34	62	38	65
2	100	47	93	40	80	40	82	42	90
	1000	78	131	37	78	34	74	39	85
3	100	61	1024	47	665	25	43	28	46
	1000	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
4	100	32	64	10	21	10	21	13	27
	1000	77	129	16	31	16	31	15	31
5	100	37	67	40	63	40	63	44	70
	1000	73	115	65	102	67	107	61	98
6	100	15	31	8	17	8	17	8	17
	1000	8	17	7	15	7	15	7	15
7	100	40	65	35	56	37	58	36	56
	1000	43	68	49	382	39	59	42	65
8	100	69	1202	45	637	38	364	26	48
	1000	98	1967	65	1279	43	440	33	257
9	100	671	1066	441	689	480	763	612	982
	1000	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
10	100	20	33	10	19	10	19	10	19
	1000	19	35	9	18	9	18	9	18
11	100	49	80	15	28	14	26	15	26
	1000	129	166	13	26	13	26	13	26
12	100	43	88	38	86	33	75	37	79
	1000	46	92	34	77	35	76	40	89
13	100	122	156	12	23	12	23	12	23
	1000	130	166	12	23	12	23	12	23
14	100	112	147	34	57	37	61	34	58
	1000	110	145	38	58	40	60	35	55
15	100	32	52	12	24	12	24	16	32
	1000	22	42	12	24	12	24	19	38
Total		2240	7341	1194	4668	1175	2687	1314	2472

The novel procedure saves NI and NF over time when compared to the conventional Fletcher and Reeves (FR) methodology, notably for our set of test problems, as demonstrated in Table 2. Table 1 compares the new methods to the Fletcher and Reeves [2] convex optimization strategy.

Table 2: Ratio of algorithm New cost to HS cost

	FR algorithm	BN1 (16)	BN2 (17)	BN3 (18)
NI	100%	53.30 %	52.45 %	58.66 %
NF	100 %	63.58 %	36.60 %	33.67 %

5. Final Thoughts

In this study, we provide a novel, globally convergent, functional nonlinear CG method, and fulfills the descent property under certain assumptions, and is based on the strictly convex quadratic function represented by (10) According to computer studies, the unique kinds given in this study are successful.

References

[1] Yu-Hong Dai and Yaxiang Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, 10:177–182, 1999.

[2] Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. *The computer journal*, 7:149–154, 1964.

[3] Basim A. Hassan. A new formula for conjugate parameter computation based on the quadratic model. *Indonesian Journal of Electrical Engineering and Computer Science*, 3:954–961, 2019.

[4] Basim A. Hassan, Hussein O Dahawi, and Azzam S Younus. A new kind of parameter conjugate gradient for unconstrained optimization. *Indonesian Journal of Electrical Engineering and Computer Science*, 17:404, 2020.

[5] Basim A. Hassan, Kanikar Muangchoo, Fadhil Alfarag, Abdulkarim Hassan Ibrahim, and Auwal Bala Abubakar. An improved quasi-newton equation on the quasi-newton methods for unconstrained optimizations. *Indonesian Journal of Electrical Engineering and Computer Science*, 22:389–397, 2020.

[6] Basim A. Hassan and RM Sulaiman. A new class of self-scaling for quasi-newton method based on the quadratic model. *Indonesian Journal of Electrical Engineering and Computer Science*, 21:1830–1836, 2021.

[7] Hawraz N Jabbar and Basim A Hassan. Two-versions of descent conjugate gradient methods for large-scale unconstrained optimization. *Indonesian Journal of Electrical Engineering and Computer Science*, 22:1643, 2021.

[8] Xian-zhen Jiang and Jin-bao Jian. A sufficient descent dai–yuan type nonlinear conjugate gradient method for unconstrained optimization problems. *Nonlinear Dynamics*, 72:101–112, 2013.

[9] Jorge Nocedal and Stephen J Wright. Numerical optimization, springer series in operations research. *Siam J Optimization*, 2006.

[10] G Ribière and E Polak. Note sur la convergence de directions conjugées. *Rev. Francaise Informat Recherche Opertionelle*, 16:35–43, 1969.

[11] Ahmet SAHINER, İdris A Masoud ABDULHAMID, and Shehab A IBRAHEM. A new filled function method with two parameters in a directional search. *Journal of Multidisciplinary Modeling and Optimization*, 2:34–42, 2019.

[12] Ahmet Sahiner and Shehab A Ibrahem. A new global optimization technique by auxiliary function method in a directional search. *Optimization Letters*, 13:309–323, 2019.

[13] Wei Xue, Junhong Ren, Xiao Zheng, Zhi Liu, and Yueyong Liang. A new dy conjugate gradient method and applications to image denoising. *IEICE TRANSACTIONS on Information and Systems*, 101:2984–2990, 2018.

[14] N Yasushi and I Hideaki. Conjugate gradient methods using value of objective function for unconstrained optimization optimization letters. *V*, 6:941–955, 2011.

[15] Gaohang Yu, Jinhong Huang, and Yi Zhou. A descent spectral conjugate gradient method for impulse noise removal. *Applied mathematics letters*, 23:555–560, 2010.

[16] G Zoutendijk. Nonlinear programming, computational methods. *Integer and nonlinear programming*, pages 37–86, 1970.