Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 26 (2023), No. 6, pp. 1219–1229
DOI : 10.47974/JIM-1620

Explaining new parameters conjugate analysis based on the quadratic model

Basim Abbas Hassan *
Abdulameer A. Saad †
Department of Mathematics
College of Computer Science and Mathematics
University of Mosul
Mosul
Iraq

Abstract

The coefficient conjugate is often the subject of conjugate gradient techniques. In this work, we use the quadratic function to unconstrained optimization problems and construct an unique coefficient conjugate gradient method. Based on a novel coefficient conjugate, we provided a search direction and fascinating conjugate gradient approach. We demonstrate that, given the appropriate circumstances, the proposed technique is universally convergent. Our numerical findings show that this method is effective for unconstrained optimization problems.

Subject Classification: *90C30, 65K05, 49M37.*

Keywords: *Optimization, Conjugate gradient methods, Globally convergent, Numerical results.*

1. Introduction

In tackling iteration problems, the conjugate gradient method has several advantages. Typically, an unconstrained optimization problem is stated as follows:

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

where f is smooth-function, see [1]. Iteratively, the conjugate gradient approach is of the form:

* *E-mail:* **basimah@uomosul.edu.iq (Corresponding Author)**
† *E-mail:* **abdulameer2023@gmail.com**

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

while d_k is search direction and α_k is precise step size, in the quadratic case,

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

It is required to employ an iterative technique for generic nonlinear functions [2], [3]. In this study, we determine the step length using the standard Wolfe criteria:

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

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

where $0 < \delta < \sigma < 1$, [4] contains further information. The formula for computing search directions is as follows:

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

where β_k is chosen in such a way that d_k and d_{k+1} must satisfy the conjugacy property. Fletcher-Reeves (FR) [5] and Dai-Yuan (DY) [6] are two instances of well.known β_k:

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

Alternatively by using other formulas (e.g. see [7-10]), although some CG methods are theoretically effective, others are successful quantitatively. If the goal function is strictly convex quadratic, all of these techniques are equal. When applied to generic non-quadratic functions, however, they behave differently, as shown in [11]. The CG approach was then applied to generic unconstrained optimization problems in [5]. As CG techniques do not need the storage of matrices, they are now widely accepted to be particularly advantageous for addressing large-scale unconstrained optimization problems. The primary driver for the development of these methods was the need to hasten the Newton method's convergence. The quasi-Newton direction, according to its theory, approximates the direction offered by the quasi-Newton technique to replicate the quasi-Newton method (6). As a result, we aim to find a β_k parameter such that:

$$-Q_{k+1}^{-1} g_{k+1} = -g_{k+1} + \beta_k s_k \tag{8}$$

where Q_{k+1} is a Hessian matrix, see Nazareth [12]. The goal of these new works is to develop a search direction that meets the adequate descent requirement and has global convergence. So that the original conjugate gradient techniques, benefits can be maximized. Give a thorough overview of these strategies and explain their convergence characteristics. Several cutting-edge optimization techniques have been described in this area [13]–[16].

A brand-new conjugate gradient technique was created and investigated on the basis of the quadratic function. Numerical tests have proven that the recommended strategy is superior than the existing FR method.

2. Derivations of some new coefficient conjugate

Information about f acquired during iteration is used in more efficient methods. One method to make use of this data is to create a function estimate that we can easily reduce. A quadratic function with a minimum can be used for this purpose. Let's say we want to approximate f with:

$$f_{k+1} = f_k + s_k^T g_k + \frac{1}{2} s_k^T Q(x_k) s_k \tag{9}$$

where $Q(x_k)$ is the second-order derivatives Hessian matrix. To get the derivative of equation (9), as:

$$\nabla f_{k+1} = g_k + Q(x_k) s_k \tag{10}$$

We get, by butting (11) into (10).

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

As a result:

$$d_k^T Q(x_k) d_k = \frac{1}{2} \frac{(g_k^T d_k)^2}{s_k^T y_k + (f_{k+1} - f_k)} \tag{12}$$

The resulting matrix $Q(x_k)$ should be as follows:

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

where I_n is identity matrix. Put $Q(x_k)$ in the equation (8), we get:

$$\beta_k = \left(1 - \frac{d_k^T d_k}{\rho_k s_k^T y_k}\right)\frac{g_{k+1}^T y_k}{s_k^T y_k} \quad \rho_k^1 = \frac{1}{2}\frac{(g_k^T d_k)^2}{(s_k^T y_k)((f_{k+1} - f_k) + s_k^T y_k)} \tag{14}$$

Using the main property, exact line-search in (11), then (14) reduces to:

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

and

$$\rho_k^3 = \frac{1}{2}\frac{(g_k^T d_k)^2}{(s_k^T y_k)((f_{k+1} - f_k) + \alpha_k g_k^T g_k)} \tag{16}$$

these formulas, also known as BD1, BD2 and BD3.The following are the steps of a new type of parameter CG-technique. The following is a summary of the new algorithms:

Stage 1. Set $x_1 \in R^n$, $d_1 = -g_1$ and $k = 0.$

Stage 2. If $\| g_k \| = 0$ then stop.

Stage 3. Evaluate. α_k satisfy-the (4) and (5).

Stage 4. Compute β_k as given in formula (17).

Stage 5. Generatethe search directionas $d_{k+1} = -g_{k+1} + \beta_k d_k.$

Stage 6. Set $k = k+1$ and go to step2.

3. Convergence analysis

To investigate New global convergence features, we start with the following assumptions :

1. The $\Omega = \{x \in R^n / f(x) \le f(x_1)\}$ level set is limited.
2. "The gradient of function g is Lipschitz continuous in some neighborhood Λ of $\Omega,$ i.e. there exists a constant L > 0 such that":

$$\| g(o) - g(\tau) \| \le L \| o - \tau \|, \ \forall \tau, o \in \Lambda \tag{17}$$

Theorem 1: *If the assumption holds, then the search-directions generated by (6) and (14) are descent directions.*

Proof: Using induction. Since $d_0 = -g_0$, we grt $g_0^T d_0 = -\| g_0 \|^2 < 0$. Assume-that $d_k^T g_k \le 0$ is true, and then multiply (6) by g_{k+1}, to get:

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \left(1 - \frac{d_k^T d_k}{\rho_k s_k^T y_k}\right) \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k^T g_{k+1} \tag{18}$$

Using $s_k = \alpha_k d_k$ in above equation, we get:

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \left(1 - \frac{s_k^T s_k}{\rho_k \alpha_k^2 s_k^T y_k}\right) \frac{y_k^T g_{k+1}}{s_k^T y_k} s_k^T g_{k+1} \tag{19}$$

Note that $s_k^T y_k \le L s_k^T s_k$ and $y_k^T g_{k+1} \le L\, s_k^T g_{k+1}$ from Lipschitz condition, by affixing:

$$d_{k+1}^T g_{k+1} \le -g_{k+1}^T g_{k+1} + \left(1 - \frac{1}{L \rho_k \alpha_k^2}\right) \frac{L\,(s_k^T g_{k+1})^2}{s_k^T y_k} \tag{20}$$

Obviously, it follows that:

$$d_{k+1}^T g_{k+1} \le -g_{k+1}^T g_{k+1} + \left(\frac{L \rho_k \alpha_k^2 - 1}{L \rho_k \alpha_k^2}\right) \frac{L\,(s_k^T g_{k+1})^2}{s_k^T y_k} \tag{21}$$

As a result of (21), we have:

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

the end of our proof.

A theorem crucial for determining global convergence was established by Dai et al. [17].

Lemma 1: *Let (2) generate sequences x_k, d_k fulfill the descent property, and α_k satisfy the convergence property (4-5). If:*

$$\sum_{k \ge 0} \frac{1}{\| d_{k+1} \|^2} = \infty \,, \tag{23}$$

Then:

$$\lim_{k \to \infty} \inf \| g_{k+1} \| = 0 \tag{24}$$

In [17] proves this lemma. We begin by introducing our main theorem for this work.

Theorem 2: *Suppose that each and every assumption is true. Let's use the New approach to generate the sequences* $\{x_k\}, \{d_k\}$ *and* $\| g(x) \| \le \Gamma.$ *If step size* α_k *meets Wolfe's requirements, we have:*

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

Proof: From the direction of search (6) and the definition of β_k by (17) we get:

$$\begin{aligned}
\|d_{k+1}\| &= \| g_{k+1} \| + \left| \left(1 - \frac{d_k^T d_k}{\rho_k s_k^T y_k} \right) \frac{g_{k+1}^T y_k}{s_k^T y_k} \right| \|s_k\| \\
&= \| g_{k+1} \| + \left| (1-\omega) \frac{g_{k+1}^T y_k}{s_k^T y_k} \right| \|s_k\| \\
&\le \| g_{k+1} \| + |(1-\omega)| \frac{\| g_{k+1} \| \|y_k\|}{\|s_k\| \|y_k\|} \|s_k\| \\
&\le (2-\omega) \| g_{k+1} \|
\end{aligned} \tag{26}$$

where $\omega = d_k^T d_k / \rho_k s_k^T y_k$. Therefore,

$$\sum_{k\ge1} \frac{1}{\| d_k \|^2} \ge \left(\frac{1}{2-\omega} \right) \frac{1}{\Gamma} \sum_{k\ge1} 1 = \infty \tag{27}$$

Appling Lemma 1, implies that $\lim_{k\to\infty} \inf \| g_k \| = 0.$

4. Numerical results

Below, we'll examine a few numerical instances. It is important that the methods suggested in Theorems 1 and 2 result in a solution. 15 samples from the CUTEr [18] library were used to assess the numerical-performance of the-suggested techniques. Table 1 contains a list of all the issues and aspects that were looked at throughout the study. The terms total iterations, total restarts, and total function evaluations, respectively, are denoted by NOI, IRS, and NOF. Testing was done on the conjugate gradient algorithms FR and New Formula (5). The inequality $\| g_{k+1} \| \le 10^{-6}$ is used as a termination condition. Both tactics were put to the test using the Wolfe line search. The area of optimization study is much broader than in [19-22].

Their relative performance characteristics are shown in Table 1 and Table 2. As seen in Table 1 and Table 2, New performs well and is comparable to the FR technique.

Table 1

The FR and New approach numerical results.

		FR algorithm		BD1 (14)		BD2 (15)		BD3 (16)	
P. No	**n**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
1	100	47	93	41	91	40	88	36	77
	1000	78	131	37	83	39	91	38	87
2	100	43	88	43	94	43	93	44	96
	1000	46	92	39	89	46	96	51	103
3	100	61	1024	32	52	30	49	32	50
	1000	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
4	100	37	67	38	59	37	60	44	67
	1000	73	115	64	100	67	109	68	109
5	100	124	231	127	254	127	254	83	166
	1000	445	711	Fail	Fail	Fail	Fail	Fail	Fail
6	100	71	110	58	112	55	103	43	85
	1000	47	84	64	121	59	110	63	125
7	100	32	65	21	50	21	50	21	50
	1000	53	116	37	91	37	92	34	86
8	100	13	25	12	23	12	23	12	23
	1000	15	29	12	24	19	39	12	25
9	100	74	123	87	132	90	140	91	137
	1000	370	616	304	505	305	502	307	508
10	100	69	1202	48	637	28	48	49	602
	1000	98	1967	45	389	51	599	35	142
11	100	49	80	19	35	15	28	21	39
	1000	129	166	12	24	12	24	12	24
12	100	12	25	11	23	11	23	11	23
	1000	11	23	11	23	11	23	11	23
13	100	122	156	14	28	14	28	15	29
	1000	130	166	15	30	15	29	16	30
14	100	22	46	22	46	22	46	22	46
	1000	26	52	26	52	26	52	23	54
15	100	112	147	40	62	47	72	39	62
	1000	110	145	39	63	41	66	46	71
Total		2074	7184	1318	3292	1320	2937	1279	2939

Problems-numbers indicant for : "1. is the Extended Rosenbrock, 2. is the Extended White & Holst, 3. is the Hager, 4. is the Generalized Tridiagonal 2, 5. is the Quadratic Diagonal Perturbed, 6. is the Extended Wood, 7. is the Extended Quadratic Penalty QP2, 8. is NONDIA (CUTE), 9. is the Partial Perturbed Quadratic, 10. is the EDENSCH (CUTE), 11. is the DENSCHNC(CUTE), 12. is the DENSCHNB (CUTE), 13. is the Extended Block-Diagonal BD2, 14. is the FLETCHCR (CUTE), 15. is the Generalized quartic GQ2".

Clearly, we have found new proposed algorithm best FR algorithm in about (37-39%) NI and (40-55%) NF. The performance percentage of our new proposed algorithm is best than of the FR algorithm.

Table 2

Algorithm ratio New and FR expense.

	FR algorithm	BD1 (14)	BD2 (15)	BD3 (16)
NI	100 %	63.54 %	63.64 %	61.66%
NF	100 %	45.82%	40.88%	40.91%

Conclusion

In this study, a new conjugate gradient method called the quadratic model is used to find the parameter. Traditional methods determine the search direction based on the gradient of the goal function. Nonlinear conjugate gradient techniques that take into consideration objective function information are presented in this work. Numerical examples show that one of the ways is just as effective as or even more effective than conventional techniques. Future research will focus on finding strategies to speed up these processes as well as new convergence analysis.

References

[1] Z. Dai, F. Wen, Some improved sparse and stable portfolio optimization problems, *Finance Research Letters*, 27, 46-52 (2018), http://doi:10.1016/j.frl.2018.02.026.

[2] H.N. Jabbar, B.A. Hassan, Two-versions of descent conjugate gradient methods for large-scale unconstrained optimization, *Indonesian Journal of Electrical Engineering and Computer Science*, 22(3), 1643 (2021), http://doi:10.11591/ijeecs.v22.i3.1643-1649.

[3] S. Wright, J. Nocedal, others, Numerical optimization, *Springer Science*, 35(1), 67-68, 7 (1999), http:// doi:10.1007/978-0-387-40065-5.

[4] Y.H. Dai, L. Z Liao, New conjugacy conditions and related nonlinear conjugate gradient methods, *Al Math Optim,* 43(1), 87-101 (2001), http://doi:10.1007/s002450010019.

[5] R. Fletcher, C. M. Reeves, Function minimization by conjugate gradients, *Comput J,* 7(2), 149-154 (1964), http://doi:10.1093/comjnl/7.2.149.

[6] Y.-H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM Journal on Optimization,* 10(1), 177-182 (1999), http://doi:10.1137/S1052623497318992.

[7] R. Fletcher, Practical methods of optimization. John Wiley & Sons (2013).

[8] M. R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Res. Natl. Bur. Stand*. (1934)., 49(6), 409-501 (1952), http://doi:10.6028/jres.049.044.

[9] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms, part 1: theory*, J Optim Theory Appl*, 69(1), 129-137 (1991), http://doi:org/10.1007/BF00940464.

[10] E. Polak, G. Ribiere, Note sur la convergence de méthodes de directions conjuguées, *Rev. française d'informatique Rech. opérationnelle*. Série rouge, 3(16), 35-43 (1969), http://doi:10.1051/m2an/196903R100351.

[11] Y.-H. Dai, Y. Yuan, Nonlinear conjugate gradient methods, Shanghai Science and Technology Publisher, Shanghai, 2000.

[12] L. Nazareth, A relationship between the BFGS and conjugate gradient algorithms and its implications for new algorithms, *SIAM J Numer Anal*, 16(5), 794-800 (1979), http://doi:10.1137/0716059.

[13] B.A. Hassan, M. S. Younis, M. W. Taha, A. H. Ibrahim, A New Type of Step Sizes for Unconstrained Optimization, in *Journal of Physics: Conference Series*, 2021, 1999(1), 12099. http://doi:10.1088/1742-6596/1999/1/012099.

[14] B.A. Hassan, K. Muangchoo, F. Alfarag, A.H. Ibrahim, A.B. Abubakar, An improved quasi-Newton equation on the quasi-Newton methods

for unconstrained optimizations, *Indonesian Journal of Electrical Engineering and Computer Science*, 22(2), 389-397 (2020), http://doi:10.11591/ijeecs. v22.i2.389-397.

[15] B.A. Hassan, Z. M. Abdullah, H. N. Jabbar, A descent extension of the Dai-Yuan conjugate gradient technique, *Indonesian Journal of Electrical Engineering and Computer Science*, 16(2), 661 (2019), http://doi:10.11591/ijeecs. v16.i2.661-668.

[16] B.A. Hassan, H. O. Dahawi, A. S. Younus, A new kind of parameter conjugate gradient for unconstrained optimization, *Indonesian Journal of Electrical Engineering and Computer Science*, 17(1), 404 (2020), http://doi:10.11591/ijeecs.v17.i1.404-411.

[17] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, *SIAM Journal on Optimization*, 10(2), 345-358 (2000), http://doi:10.1137/S1052623494268443.

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

[19] W. Gao, M. R. Farahani, The hyper-Zagreb index for an infinite family of nanostar dendrimer, *Journal of Discrete Mathematical Sciences and Cryptography*. 20(2), 515-523 (2017), http://doi:10.1080/09720529.2016.1220088.

[20] W. Gao, L. Shi, M.R. Farahani, Szeged Related Indices of $TUAC_6[p,q]$, *Journal of Discrete Mathematical Sciences and Cryptography*. 20(1), 553-563, 2017, http://doi:10.1080/09720529.2016.1228312.

[21] W. Gao and M. R. Farahani, The Zagreb topological indices for a type of Benzenoid systems jagged-rectangle, *Journal of Interdisciplinary Mathematics*, 20(5), 1341-1348 (2017), http://doi:10.1080/09720502.2016.1232037.

[22] W. Sajjad, M.S. Sardar, M. Cancan, S. Ediz, A. Q. Baig, On modified eccentric connectivity index of nanotube, *Journal of Information and Optimization Sciences*, 41(4), 959-972, (2020), http://doi:10.1080/02522667.2020.1745380.

[23] S. Ediz, İ. Çiftçi, M.Cancan, M.R.Farahani, On k-total distance degrees and k-total Wiener polarity index. *Journal of Information and*

Optimization Sciences. 42(7), 1469-1477 (2021). https://doi.org/10.1080/02522667.2021.1896652.

[24] F. Afzal, Shah Rukh, D. Afzal, M.R.Farahani, M.Cancan, S. Ediz. A study of non-commutative Von-Neumann regular rings. *Journal of Information and Optimization Sciences*. 42(7), 1425-1436 (2021). https://doi.org/10.1080/02522667.2021.1896650.

[25] M.S. Sardar, M. Alaeiyan, M.R.Farahani, M.Cancan, S. Ediz. Resistance distance in some classes of rooted product graphs obtained by Laplacian generalized inverse method. *Journal of Information and Optimization Sciences* 42(7), 1447-1467 (2021). https://doi.org/10.1080/02522667.2021.1899210.

[26] M. Nadeem, S. Ahmad, M.K. Siddiqui, M.A. Ali, M.R.Farahani, A.J.M. Khalaf. On some applications related with algebraic structures through different well known graphs. *Journal of Discrete Mathematical Sciences and Cryptography.* 24(2), 451-471 (2021). https://doi.org/10.1080/09720529.2021.1885806.

[27] A. Alsinai, H.M.U. Rehman, Y. Manzoor, M. Cancan, Z. Taş, M.R. Farahani, Sharp upper bounds on forgotten and SK indices of cactus graph, *Journal of Discrete Mathematical Sciences and Cryptography*, 1-22 (2022). https://doi.org/10.1080/09720529.2022.2027605.