Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 26 (2023), No. 6, pp. 1207–1217
DOI : 10.47974/JIM-1619

Elastic conjugate gradient methods to solve iteration problems

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

Abstract

Using the conjugacy condition and second-order Taylor expansion, we present a new formula in this paper. The novel formula, unlike traditional conjugate gradient approaches, utilizes information on available function values and gradients. The worldwide convergence findings of the formula are discussed, and numerical data show that this technique works.

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

Keywords: *Conjugate gradient, Formula conjugate gradient, Global convergence.*

1. Introduction

Unconstrained optimization techniques address the iterative approach of getting the minimum value for the following issues:

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

where $f : R^n \to R$ is smooth function, see [1]. Conjugate gradient(CG) techniques are a common sort of iterative approach for solving (1) among the several iterative methods available. If g_k is the gradient of $f(x_k)$, then the approach for solving (1) iteratively is as:

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

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

the search direction denotes by d_k and the step size denotes by α_k, in the quadratic case,

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

Different line search criteria, such as Wolfe conditions, can also be used as an alternative:

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

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

where $0 < \delta < \sigma < 1$, see [2]. The search direction d_{k+1} is defined by:

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

where β_k is referred to as CG-coefficient. The CG-coefficient is an important component in conjugate gradient approaches in all subjects. [3]–[7] provide other equations for getting conjugate parameters. We will concentrate on the well-known Hestenes-Stiefel and Fletcher-Reeves approach, in which the formulas are:

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

where $y_k = g_{k+1} - g_k$. The Hestenes-Stiefel formula always meets the conjugacy condition:

$$d_{k+1}^T y_k = 0 \tag{8}$$

More details can be found in [8]. Wu and Chen (known as the WC technique) [9] obtained the following CG formula by modifying Hestenes-Stiefel to meet the descent property, as:

$$\beta_k^{WC} = \frac{y_{k+1}^T g_{k+1}}{d_k^T y_k} + \frac{2(f_k - f_{k+1}) + g_k^T s_k}{d_k^T y_k} \tag{9}$$

The conjugacy criterion of Perry [10] is satisfied by the following formula:

$$d_{k+1}^T y_k = -s_k^T g_{k+1} \tag{10}$$

As a result, efforts have been made to develop global and numerically efficient conjugate gradient algorithms. We also claim that the search instructions meet the necessary conditions for an acceptable descent if and only if:

$$g_{k+1}^T d_{k+1} \leq -c \left\| g_{k+1} \right\|^2 \tag{11}$$

where is a nonnegative constant. In [11] provides a more detailed performance profile that may be used to investigate the convergence analysis of our suggested technique. Finally, numerical studies show that Wu and Chen [9] appear to be more successful than the existing family of methods. The positive numerical findings prompted us to examine other changes that we felt would increase numerical performance. This method of approaching optimization problems has been published in several papers [12]–[16].

A unique parameter conjugate is obtained using the second-order Taylor expansion. The derived method looks into both theoretical and numerical results.

2. Our new formulas

Taylor expansion is the source of this derivation. Consider the quadratic model in the vicinity of the point:

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

The gradient of a second-order Taylor expansion form shows this:

$$\nabla f_{k+1} = g_k + Q(x_k) s_k \tag{13}$$

We got the following results using (13) in (12):

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

As a result of the preceding equation and equation (3), we may deduce:

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

where:

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

Equations (13) and (10), together, lead to the following conclusion:

$$d_{k+1}^T y_k = -\rho_k s_k^T y_k - s_k^T g_k \tag{17}$$

Furthermore, $d_{k+1} = -g_{k+1} + \beta_k s_k$ and (17) suggest that:

$$\beta_k d_k^T y_k = g_{k+1}^T y_k - \rho_k s_k^T y_k - s_k^T g_k \tag{18}$$

As a result,

$$\beta_k = \frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{\rho_k s_k^T y_k}{s_k^T y_k} - \frac{s_k^T g_k}{s_k^T y_k} \tag{19}$$

Exact line search in (14), followed by (16), results in:

$$\rho_k^2 = \frac{1}{2} \frac{(g_k^T s_k)^2}{s_k^T y_k(-s_k^T g_k + (f_{k+1} - f_k))} \tag{20}$$

and

$$\rho_k^3 = \frac{1}{2} \frac{(g_k^T s_k)^2}{s_k^T y_k(\alpha_k g_k^T g_k + (f_{k+1} - f_k))} \tag{21}$$

This formula is referred to as the BR. Below is the BR conjugate gradient algorithm.

Algorithm BR.

Initialization: Given $x_0 \in R^n$ and the parameters $\varepsilon > 0$, $\delta \in (0,1)$, $\sigma \in (\delta,1)$,

Set $k = 0$, $d_0 = -g_0$.

Stage 1. If $\| g_k \| \le \varepsilon$ then stop.

Stage 2. Compute α_k by a suitable line search.

Stage 3. New point $x_{k+1} = x_k + \alpha_k d_k$, and compute β_k by (18).

Stage 4. Calculate the search direction $d_{k+1} = -g_{k+1} + \beta_k s_k$.

Stage 5. Let $k = k + 1$ and go to stage 2.

3. Convergence analysis:

This section focuses on the BR method's convergence feature. We make the following assumptions:

i. $f(x)$ is bounded on the set $\Psi = \{x \in R^n : f(x) \le f(x_0)\}$.

ii. g is Lipschitz continuous, i.e. a positive constant L exists such that:

$$\|g(z) - g(u)\| \le L \|z - u\|, \ \forall\ z, u \in R^n \tag{22}$$

There is a constant that lies behind these function assumptions $¥ \ge 0$ such that $\| \nabla f(x) \| \le ¥$. See [2]. The adequate descent requirement plays a vital effect.

Theorem 1: *If* d_{k+1} *is created using a novel algorithm,* $d_{k+1}^T g_{k+1} \le -c \| g_{k+1} \|^2$ *holds.*

Proof: Since $d_0 = -g_0$, we obtained $g_0^T d_0 = -\| g_0 \|^2 \le 0$. Let $g_k^T d_k < 0$ for all $k \in n$. When we multiply (6) by g_{k+1}^T, we get:

$$d_{k+1}^T g_{k+1} = -\| g_{k+1} \|^2 + \left[-\rho_k - \frac{s_k^T g_k}{s_k^T y_k} + \frac{g_{k+1}^T y_k}{s_k^T y_k} \right] s_k^T g_{k+1} \tag{23}$$

We have the following from (23) and (13):

$$d_{k+1}^T g_{k+1} = -\| g_{k+1} \|^2 + \left[\frac{g_{k+1}^T y_k}{s_k^T y_k} - \frac{s_k^T g_{k+1}}{s_k^T y_k} \right] s_k^T g_{k+1} \tag{24}$$

By Lipschitz condition, we note that:

$$y_k^T g_{k+1} \le L\ s_k^T g_{k+1} \tag{25}$$

From equation (24) and equation (25), we obtained:

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + \left[L\ \frac{s_k^T g_{k+1}}{s_k^T y_k} - \frac{s_k^T g_{k+1}}{s_k^T y_k} \right] s_k^T g_{k+1} \tag{26}$$

Taking advantage of the inequality:

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 + [L - 1] \frac{(s_k^T g_{k+1})^2}{s_k^T y_k} \tag{27}$$

As a result of (27), we get:

$$d_{k+1}^T g_{k+1} \le -\|g_{k+1}\|^2 < 0 \tag{28}$$

The next lemma proved in [17] for any conjugate gradient algorithm using the Wolfe conditions.

Lemma 1: Any conjugate gradient technique (2) uses α_k from the Wolfe line search and d_k as a descent. a taken to be a descent direction. Assuming assumptions (i) and (ii).

If :
$$\sum_{k>1} \frac{1}{\| d_{k+1} \|^2} = \infty \tag{29}$$

then
$$\lim_{k\to\infty}(\inf \| g_{k+1} \|) = 0 \tag{30}$$

In [17] is where you'll find the proof.

The convergence property result of the new Algorithm is shown in the next theorem.

Theorem 2: *Assume Assumptions is correct. If f is a uniformly. convex. function, that is, if a constant* $\mu > 0$ *exists such that:*

$$(\nabla f(z) - \nabla f(u))^T \ge \mu \| z-u \|^2,\ \forall z,u \in R^n \tag{31}$$

then:

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

Proof: By definition β_k and (15) that:

$$\left|\beta_k\right| = \left| \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{v_k^T g_{k+1}}{d_k^T y_k} \right| \tag{33}$$

$\| y_k \| \le L\| s_k \|$ is obtained by using (19), while $s_k^T y_k \ge \mu \| s_k \|^2$ is obtained by using (29). When we apply the Cauchy inequality to the previous inequalities, then (34), we obtain:

$$|\beta_k| \le \frac{L\| s_k \| \| g_{k+1} \|}{\mu \| s_k \|^2} + \frac{\| s_k \| \| g_{k+1} \|}{\mu \| s_k \|^2} \le \left(\frac{L+1}{\mu} \right) \frac{¥}{\| s_k \|} \tag{34}$$

As a result of utilizing (32) in (6), we get:

$$\| d_{k+1} \| \leq \| g_{k+1} \| + |\beta_k| \, \| s_k \| \leq ¥ + \left(\frac{L+1}{\mu} \right) \frac{¥}{\| s_k \|} \| s_k \| \leq ¥ + \frac{¥L + ¥}{\mu} \leq \frac{¥(\mu + L + 1)}{\mu} \quad (35)$$

demonstrating that (30) is correct.

4. Numerical results

Compares the numerical results of the BR and FR algorithms. The Fortran language was used to create the algorithms. We utilized the following parameters in our implementation: $\delta_1 = 0.001$ and $\delta_2 = 0.9$, with $\| g_{k+1} \| \leq 10^{-6}$ as the termination criterion. The performance of the algorithms is displayed using Andrei's [18] performance profile. NI and NF stand for the number of iterations and objective function evaluations, respectively, and are used to compare numerical performance. The breadth of optimization study is seen in [19]-[25]. This technique is shown to be efficient for the test issues using numerical behavior comparisons with FR-CG (see Table 1 and Table 2).

Table 1

Nnumrical results for FR and New methods

		FR algorithm		**BR1 (16)**		**BR2 (20)**		**BR3 (21)**	
P. No	**n**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
1	100	47	93	41	84	39	80	39	80
	1000	78	131	53	98	47	96	67	125
2	100	43	88	37	80	34	70	31	66
	1000	46	92	39	82	35	72	35	72
3	100	61	1024	34	58	70	1279	35	57
	1000	Fail	Fail	Fail	Fail	Fail	Fail	Fail	Fail
4	100	37	67	39	61	42	64	37	60
	1000	73	115	61	95	59	95	58	92
5	100	124	231	83	163	83	163	83	163
	1000	445	711	240	455	232	439	232	439
6	100	71	110	37	71	37	71	37	71
	1000	47	84	53	93	40	73	40	72
7	100	32	65	22	52	22	52	22	52
	1000	53	116	42	95	44	99	42	95
8	100	13	25	14	28	14	27	13	25

Contd...

	1000	15	29	11	23	11	23	11	23
9	100	74	123	91	135	81	125	74	116
	1000	370	616	270	450	290	485	275	463
10	100	69	1202	60	1003	30	54	Fail	Fail
	1000	98	1967	50	566	74	1341	52	685
11	100	49	80	27	45	22	42	22	41
	1000	129	166	22	43	17	31	16	31
12	100	12	25	14	27	14	27	14	27
	1000	11	23	12	23	12	23	12	23
13	100	122	156	20	37	21	39	21	39
	1000	130	166	22	40	21	38	22	42
14	100	22	46	22	46	22	46	22	46
	1000	26	52	25	52	25	51	25	51
15	100	112	147	38	65	38	64	37	65
	1000	110	145	40	66	38	63	40	67
Total		2450	5771	1459	3133	1484	5078	1414	3188

Problems numbers indicant for: "1. is the Extended Rosenbrock, 2. is the Extended White & Holst, 3. is the Hager, 4. is the Generalized Tri diagonal 2, 5. is the Quadratic Diagonal Perturbed, 6. is the Extended Wood, 7. is the Extended Quadatic 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 quarticd GQ2".

Table 2

Algorithm ratio New HS cost.

	FR algorithm	**BR1 (16)**	**BR2 (20)**	**BR3 (21)**
NI	100 %	59.55 %	60.57 %	57.71%
NF	100 %	54.28 %	87.99 %	55.24 %

Conclusion

The goal function value was incorporated in the new equations, although the value decreased when compared to traditional CG approaches. The formula's global convergence findings are explained, and numerical results demonstrate that it is a very efficient formula. Clearly, we have found new proposed algorithm best FR algorithm in about (40-43%) NI and (13-46%) NF. The performance percentage of our new proposed algorithm is best than the FR algorithm.

References

[1] W. J. Leong and M. A. Hassan, Scaled memoryless symmetric rank one method for large-scale optimization, *Appl Math Comput*, 218(2), 413–418 (2011), http:// doi: 10.1016/j.amc.2011.05.080.

[2] P. Wolfe, Convergence conditions for ascent methods, *SIAM review*, 11(2), 226–235 (1969), http:// doi: 10.1137/1011036.

[3] Y.-H. Dai and 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.

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

[5] R. Fletcher, Practical methods of optimization. John Wiley & Sons, 2013.

[6] Y. Liu and 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.

[7] E. Polak and 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.

[8] M. R. Hestenes and 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] C. Wu and G. Chen, New type of conjugate gradient algorithms for unconstrained optimization problems, *Journal of Systems Engineering and Electronics,* 21(6), 1000–1007 (2010), http://doi: 10.3969/j.issn.1004-4132.2010.06.012.

[10] A. Perry, A modified conjugate gradient algorithm, *Oper Res*, 26(6) 1073–1078 (1978), http://doi: 10.1287/opre.26.6.1073.

[11] G. Zhou, A descent algorithm without line search for unconstrained optimization, *Appl Math Comput*, 215(7), 2528–2533 (2009), http:// doi: 10.1016/j.amc.2009.08.058.

[12] B. A. Hassan and R. M. Sulaiman, Using a New Type Quasi-Newton Equation for Unconstrained Optimization, in 2021 7th International Engineering Conference "Research & Innovation amid Global Pandemic"(IEC), 118–122 (2021). http:// doi: 10.1109/I EC52205. 2021. 9476089.

[13] B. A. Hassan and R. M. 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(3), 1830–1836 (2021).

[14] B. A. Hassan, K. Muangchoo, F. Alfarag, A. H. Ibrahim, and 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.pp389-397.

[15] B. A. Hassan and M. A. Kahya, A new class of quasi-Newton updating formulas for unconstrained optimization, *Journal of Interdisciplinary Mathematics,* 24(8), 2355–2366 (2021), http:// doi: 10.1080/09720502.2021.1961980.

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

[17] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y. Yuan, Convergence properties of nonlinear conjugate gradient methods, *SIAM Journal on Optimization*, 10(2), 345–358 (2000), http:// doi: 10.1137/ S1052623494 268443.

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

[19] B. A. Hassan, Z, M. Abdullah, S. A. Husssin, Some a New Conjugate Gradient Methods For Solving Unconstrained Optimization Problems, *Journal of Information and Optimization Sciences*, 43(4), 893–903 (2022) http:// doi: 10.1080/ 02522667.2022.2103300.

[20] B. A. Hassan, Z, M. Abdullah, S. A. Husssin, A New Family of Conjugate Gradient Methods to Solving Unconstrained Optimization

Problems, *Journal of Information and Optimization Sciences*, 41(4), 959–972 (2022), http:// doi: 10.1080/ 02522667.2022.2103299.

[21] 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.

[22] W. Gao, L. Shi, M. R. Farahani, Szeged Related Indices of TUAC6[p,q], *Journal of Discrete Mathematical Sciences and Cryptography*. 20(2), 553-563 (2017), http://doi:10.1080/09720529.2016.1228312.

[23] W. Gao, 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.

[24] 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 (May 2020), http://doi:10.1080/02522667.2020.1745380.

[25] S. Ediz,İdris 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.