Journal of Education and Science
Vol. 29, No.4, 2019, pp.270-281
ISSN 1812-125X
http://edusj.mosuljournals.com

تطوير جديد لطريقة التدرج المترافق في الأمثلية الغير مقيدة

د. زياد محمد عبدالله

كلية علوم الحاسوب والرياضيات
جامعة تكريت
zeyaemoh1978@tu.edu.iq

DOI: 10.33899/edusj.2019.184741

Received	Accepted
15/ 05 / 2019	20 / 06 / 2019

الخلاصة

تختص هذه الدراسة بخوارزمية التدرج المترافق الغير خطية والتي تم استخدامها بشكل واسع في الامثلية خصوصا مسائل الامثلية ذات القيود العالية ولكونها لا تحتاج الى خزن في خوارزميات المصفوفات. ان الخوارزمية المقترحة هي تحسين لخوارزمية هيداكي خوارزمية التدرج المترافق. والتي تحقق شرط الانحدار الكافي وبمعلمة (HY) وياسوشي والتي تحسب باستخدام شرط الترافق. الخوارزمية المقترحة عادة تولد اتجاه البحث الانحدار والذي تحقق التقارب باستعمال بعض الفرضيات. ان الفكرة الأساسية من هذا العمل هو لإثبات التقارب الشمولي لطريقة التدرج المترافق اللاخطية. تكشف النتائج الاحصائية عن فعالية الخوارزمية المقترحة ولحل مشاكل الاختبار المعطى.

الكلمات المفتاحية: التدرج المترافق الغير خطي، الامثلية، هيداكي، ياسوشي (HY)، خوارزمية التدرج المترافق

New Modification Nonlinear Conjugate Gradient Method for Optimization

Zeyad Mohammed Abdullah
zeyaemoh1978@tu.edu.iq

Faculty of Computers Sciences and Mathematics, University of Tikrit, IRAQ

DOI: 10.33899/edusj.2019.184741

Received 15/ 05 / 2019 **Accepted** 20 / 06 / 2019

Abstract

This study proposes "a nonlinear Conjugate gradient algorithm that is widely used in optimization, especially for large scale optimization problems, because it does not require the storage of any matrices algorithm". This algorithm modifies Hideaki and Yasushi's (HY) "conjugate gradient algorithm". It satisfies "a parameterized sufficient descent condition with a parameter δ_k", which is calculated using the conjugacy condition. The new proposed algorithm always produces descent search directions and it is shown to be convergent under some assumptions. The main idea of this work is to prove the global convergence for the modification nonlinear conjugate gradient method. The statistical results reveal the effectiveness of the proposed algorithm for problems of the given test.

Keywords: Nonlinear Conjugate Gradient, optimization, Hideaki and Yasushi (HY) conjugate gradient algorithm.

1. Introduction:

Optimization is an important tool in various fields including engineering, production management, economy etc. This study considers the problem of unconstrained optimization represented in this equation:

$$\min\left\{f(x) \;\middle|\; x \in R^n\right\} \tag{1}$$

where "$f : R^n \to R$" is differentiable continuously and its gradient "$g(x) = \nabla f(x)$" is accessible [3]. Methods of conjugate gradient are related to the most famous iterative methods used for solving problems of large scale optimization. The iterative equation below is often used as a method of nonlinear conjugate gradient:

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

where “α_k” represents a positive scalar, which is known as “the step length” and defined through “a line search”; while “d_k” is created as follows:

$$“d_{k+1} = \begin{cases} -g_{k+1} & \text{if k} = 0 \\ -g_{k+1} + \beta_k d_k & \text{if k} > 0 \end{cases}” \tag{3}$$

In this equation, “β_k” is a scalar that determines the different methods of conjugate gradient. The best recognized methods of conjugate gradient are DY [4] and FR [6]. β_k formulas in these methods are given below:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad , \quad \beta_k^{DY} = \frac{\|g_{k+1}\|^2}{y_k^T d_k} \tag{4}$$

A modified method of DY that has been proposed by Andrei [2] is termed as A. The “direction d_{k+1}” in this method is defined by:

$$“d_{k+1} = -g_{k+1} + \beta_k^A s_k \quad ” \tag{5}$$

Where

$$\beta_k^A = \frac{\|g_{k+1}\|^2}{y_k^T s_k} - \delta_k \frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{[y_k^T s_k]^2} \tag{6}$$

The equation (6) is a descent direction with the use of a line search of standard Wolfe. Moreover, for ensuring the universal convergence of iterative scheme, d_{k+1} direction has to satisfy “the sufficient descent condition” below:

$$“g_{k+1}^T d_{k+1} \le -c\,\|g_{k+1}\|^2” \tag{7}$$

Here, “c” represents “a positive real-valued constant”. More performance profile is given in [9].

The modified method with a descent property and new algorithms are stated in the second section. The universal convergence of the new method is proved in the third section. Next, statistical results are reported in the fourth section for testing the proposed method and discussing the results. Lastly, the conclusions are presented at the end of this paper.

2. A modified HY method

This section recalls the method of HY conjugate gradient developed by Hideaki and Yasushi [7]. The HY method is defined as the following:

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

HY method is a useful method because of its good convergence property.

Being inspired by the idea of Andrei [2], a modified method of Hideaki and Yasushi's (HY) nonlinear conjugate gradient is proposed in this section. β_k^{ZMM} formula is presented as follows:

$$\beta_k^{ZMM} = \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} - \delta_k \frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2} \tag{9}$$

δ_k represents a positive parameter. It is assumed that $f_k - f_{k+1} \neq 0$, so that β_k^Z is well defined. The main notion lies in selecting a new "β_k^{ZMM}". Through searching a particular direction d_{k+1}; the new method will possess a sufficient descent property. This assumption is proved in the following theorem.

Theorem 1:

If "$f_k - f_{k+1} \neq 0$", there will be "$d_{k+1} = -g_{k+1} + \beta_k^Z d_k$," ($\beta_k^{ZMM}$ is specified in equation 9); so:

$$\text{“}\, g_{k+1}^T d_{k+1} \leq -\left(1 - \frac{1}{4\,\delta_k}\right)\|g_{k+1}\|^2. \,\text{”} \tag{10}$$

Under the condition $\frac{1}{4} < \delta_k < \infty$.

Proof:

Since "$d_0 = -g_0$", there is $g_0^T d_0 = -\|g_0\|^2 < 0$. Suppose that $g_k^T d_k < -c_1\|g_k\|^2$ for all $k \in n$. When multiplying (5) by g_{k+1}, with (9), there will be:

$$\text{“}\, g_{k+1}^T d_{k+1} \leq -\|g_{k+1}\|^2 + \frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} - \delta_k \frac{(g_{k+1}^T s_k)^2\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}. \,\text{”} \tag{11}$$

But

$$\frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} = \frac{[((2/\alpha_k)(f_k - f_{k+1})g_{k+1})/\sqrt{2\delta_k}]^T[\sqrt{2\delta_k}(g_{k+1}^T s_k)g_{k+1}]}{[(2/\alpha_k)(f_k - f_{k+1})]^2} \quad (12)$$

$$\text{“} \le \frac{\frac{1}{2}\left[\frac{1}{2\delta_k}((2/\alpha_k)(f_k - f_{k+1}))^2\|g_{k+1}\|^2 + 2\delta_k(g_{k+1}^T s_k)^2\|g_{k+1}\|^2\right]}{[(2/\alpha_k)(f_k - f_{k+1})]^2}.\text{”}$$

$$\le \frac{1}{4\delta_k}\|g_{k+1}\|^2 + \delta_k\frac{(g_{k+1}^T s_k)^2\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}$$

By substituting (12) into (11) using several algebra steps that can be reduced to (10).

The numerical algorithms (2), (5) and (9) reveal that the performance of δ_k is relatively different according to its different choices. Consequently, for obtaining an efficient algorithm, a procedure for calculating δ_k shall be presented. Principally, this depends on the condition of conjugacy “ $y_k^T d_{k+1} = 0$ ”.

When employing (9) in (5), the direction below is resulted:

$$\text{“} d_{k+1} = -g_{k+1} + \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})}s_k - \delta_k\frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}s_k. \text{”} \quad (13)$$

This could be written as:

$$\text{“} d_{k+1} = -\ Q_{k+1}g_{k+1}.\text{”}$$

where the matrix Q_{k+1} is :

$$\text{“} Q_{k+1} = I - \frac{s_k g_{k+1}^T}{(2/\alpha_k)(f_k - f_{k+1})} + \delta_k\frac{\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}s_k s_k^T. \text{”} \quad (14)$$

Now, by symmetrizing Q_{k+1} as :

$$\text{“} \bar{Q}_{k+1} = I - \frac{s_k g_{k+1}^T + g_{k+1}s_k^T}{(2/\alpha_k)(f_k - f_{k+1})} + \delta_k\frac{\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}s_k s_k^T. \text{”} \quad (15)$$

The following direction can be considered:

$$\text{“}\, d_{k+1} = -\bar{Q}_{k+1}\, g_{k+1} \,.\text{”} \tag{16}$$

In (16), Q_{k+1} is symmetrized as $\bar{Q}_{k+1}$ because the computed direction in (14) is similar to the methods of quasi-Newton. Nevertheless, this paper only uses the symmetry without modifying more “$\bar{Q}_{k+1}$” with the aim of satisfying the equation of “quasi-Newton”.

Based on the condition of conjugacy, “$y_k^T d_{k+1} = 0$”, namely:

$$\text{“}\, y_k^T \bar{Q}_{k+1}\, g_{k+1} = 0 \,,\text{”} \tag{17}$$

The next equation (18) results from the equation (17), as follows:

$$\text{“}\begin{array}{l} y_k^T g_{k+1} - \dfrac{y_k^T s_k \|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} - \dfrac{s_k^T g_{k+1}(y_k^T g_{k+1})}{(2/\alpha_k)(f_k - f_{k+1})} + \delta_k \dfrac{\|g_{k+1}\|^2 (y_k^T s_k)(s_k^T g_{k+1})}{[(2/\alpha_k)(f_k - f_{k+1})]^2} = 0 \\ \delta_k + \dfrac{y_k^T g_{k+1}[(2/\alpha_k)(f_k - f_{k+1})]^2}{\|g_{k+1}\|^2 (y_k^T s_k)(s_k^T g_{k+1})} - \dfrac{(2/\alpha_k)(f_k - f_{k+1})}{s_k^T g_{k+1}} - \dfrac{(y_k^T g_{k+1})(2/\alpha_k)(f_k - f_{k+1})}{\|g_{k+1}\|^2 (y_k^T s_k)} = 0 \end{array}\text{”} \tag{18}$$

it follows that:

$$\delta_k = -\frac{y_k^T g_{k+1}[(2/\alpha_k)(f_k - f_{k+1})]^2}{\|g_{k+1}\|^2 (y_k^T s_k)(s_k^T g_{k+1})} + \frac{(2/\alpha_k)(f_k - f_{k+1})}{s_k^T g_{k+1}} + \frac{(y_k^T g_{k+1})(2/\alpha_k)(f_k - f_{k+1})}{\|g_{k+1}\|^2 (y_k^T s_k)} \tag{19}$$

Therefore, using (19) in (9), there will be:

$$\beta_k^{ZMM} = \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} - \begin{bmatrix} -\dfrac{y_k^T g_{k+1}[(2/\alpha_k)(f_k - f_{k+1})]^2}{\|g_{k+1}\|^2 (y_k^T s_k)(s_k^T g_{k+1})} \\ +\dfrac{(2/\alpha_k)(f_k - f_{k+1})}{s_k^T g_{k+1}} \\ +\dfrac{(y_k^T g_{k+1})(2/\alpha_k)(f_k - f_{k+1})}{\|g_{k+1}\|^2 (y_k^T s_k)} \end{bmatrix} \frac{(g_{k+1}^T s_k)\|g_{k+1}\|^2}{[(2/\alpha_k)(f_k - f_{k+1})]^2}$$

$$\beta_k^{ZMM} = \frac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} - \begin{bmatrix} -\dfrac{y_k^T g_{k+1}}{y_k^T s_k} \\ +\dfrac{\|g_{k+1}\|^2}{(2/\alpha_k)(f_k - f_{k+1})} \\ +\dfrac{(y_k^T g_{k+1})(g_{k+1}^T s_k)}{(y_k^T s_k)(2/\alpha_k)(f_k - f_{k+1})} \end{bmatrix}$$

$$\beta_k^{ZMM} = \frac{y_k^T g_{k+1}}{(y_k^T s_k)} - \frac{(y_k^T g_{k+1})(g_{k+1}^T s_k)}{(y_k^T s_k)(2/\alpha_k)(f_k - f_{k+1})} = \frac{y_k^T g_{k+1}}{(y_k^T s_k)}\left[1 - \frac{(g_{k+1}^T s_k)}{(2/\alpha_k)(f_k - f_{k+1})}\right] \quad (20)$$

Now, the new algorithm can be outlined in the next section.

3.1 Outline of the new algorithms:

Step 1. Initialization: Choose $x_1 \in R^n$ and the parameters $0 < \delta < \sigma$.

Calculate " $f(x_1)$ and g_1". Consider "$d_1 = -g_1$" and develop "the initial guess $\alpha_1 = 1/\|g_1\|$".

Step 2. Test the continuance of iterations. If $\|g_{k+1}\| \le 10^{-6}$, then stop.

Step 3. Line search. Compute $\alpha_{k+1} > 0$ to determine whether it is satisfying the Wolfe line search Condition:

$$\text{“ } f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \text{ ”} \quad (21)$$

$$\text{“ } d_k^T g(x_k + \alpha_k d_k) \ge \sigma\, d_k^T g_k \text{ ”} \quad (22)$$

Next, update these variables $x_{k+1} = x_k + \alpha_k d_k$. Calculate "$f_{k+1}$, g_{k+1}" and "$s_k = x_{k+1} - x_k$" and "$y_k = g_{k+1} - g_k$".

Step 4. I Calculation of direction. Calculate "$d_{k+1} = -g_{k+1} + \beta_k s_k$", where

β_k^{ZMM} is calculated as in (9). When satisfying "the restart of Powell criterion $|g_{k+1}^T g_k| \ge 0.2\,\|g_{k+1}\|^2$", then develop "$d_{k+1} = -\ g_{k+1}$"; if not, then set "$k = k+1$" and continue with the second step.

3. Convergence analysis

Consider these assumptions:

i- "The level set $L = \{x \in R^n | f(x) \le f(x_0)\}$ is bounded".

ii- In a part, "U and L, $f(x)$" are continuously differentiable and "their gradient id Lipschitz is continuous", i.e. there is a constant $L > 0$, so:

$$\text{“}\|g(x_{k+1}) - g(x_k)\| \le L\|x_{k+1} - x_k\|,\ \forall x_{k+1}, x_k \in U \text{ ”} \quad (23)$$

Further details are found in [8,10]. Based on the assumptions on " f, ", there is "a constant", then "a constant $\Gamma > 0$" exists; so:

$$\|g_{k+1}\| > \Gamma \quad (24)$$

for all $x \in L$.

Dai et al. [5] verified that for any method of "conjugate gradient" with "strong Wolfe line search", the result is as follows:

Lemma (1):

Suppose that the assumptions (i) and (ii) are held, then consider the methods of "conjugate gradient" (2) and (5), where "d_{k+1}" is "a descent direction" and α_k is obtained by "the strong Wolfe line search" in (3) and (4). If

$$\text{“}\sum_{k\geq 0} \frac{1}{\|d_{k+1}\|^2} = \infty \text{ ,”} \tag{25}$$

then

$$\text{“}\liminf_{k\to\infty} \|g_{k+1}\| = 0 \text{”.} \tag{26}$$

For "uniformly convex functions", it can be proved that the direction norm created by (5) and (20) is bounded above. So, based on Lemma 1, the following result can be proved.

Theorem 2:

Suppose that the assumption is held, then consider methods (2) and (5), where β_k^Z is "a descent direction" specified by (20) and "α_k" is found by "the Wolfe line search". Suppose that "f" is "a uniformly convex function" on L, namely, there is "a constant $\mu, M > 0$"; so:

$$\text{“}(\nabla f(x) - \nabla f(y))^T (x - y) \geq M\|x - y\|^2 \text{”} \tag{27a}$$

or equally, there is a constant $\mu > 0$; so:

$$z^T G\, z \geq \mu\|z\|^2 \tag{27b}$$

for any "$x, y \in L,$" then:

$$\text{“}\lim_{k\to\infty} \|g_{k+1}\| = 0 \text{”.} \tag{28}$$

Proof:

Based on (27), it follows that $y_k^T s_k \geq \mu\|s_k\|^2$. Since there is "a descent direction", it follows that "$g_{k+1}^T s_k < y_k^T s_k + g_k^T s_k < y_k^T s_k$". By employing "Wolfe condition" (22), the result is:

$$y_k^T s_k = g_{k+1}^T s_k - g_k^T s_k \geq (\sigma - 1) g_k^T s_k = -(1-\sigma) g_k^T s_k \tag{29}$$

By employing Wolfe condition (21), there will be:

$$\text{“}\frac{1}{(1/\alpha_k)(f_k - f_{k+1})} \leq \frac{1}{-\delta g_k^T d_k}\text{”}. \tag{30}$$

From (29) and (30), there is:

$$\begin{aligned} 1 - \frac{(g_{k+1}^T s_k)}{(2/\alpha_k)(f_k - f_{k+1})} &\leq 1 - \frac{(1-\sigma) g_k^T s_k}{-2\delta g_k^T d_k} = 1 - \frac{(1-\sigma)\alpha_k g_k^T d_k}{-2\delta g_k^T d_k} \\ &\leq 1 - \frac{(g_{k+1}^T s_k)}{(2/\alpha_k)(f_k - f_{k+1})} \\ &\leq 1 + \frac{(1-\sigma)\alpha_k}{2\delta} \end{aligned} \tag{31}$$

Therefore,

$$\text{“}\begin{aligned} \left|\beta_k^{ZMM}\right| &\leq \frac{L\|g_{k+1}\|\|s_k\|}{\mu\|s_k\|^2}\left[1 + \frac{(1-\sigma)\alpha_k}{2\delta}\right] \\ &\leq \frac{L\Gamma}{\mu\|s_k\|}\left[1 + \frac{(1-\sigma)\alpha_k}{2\delta}\right] \end{aligned}\text{”} \tag{32}$$

Hence:

$$\text{“}\begin{aligned} \|d_{k+1}\| &\leq \|g_{k+1}\| + \left|\beta_k^{ZMM}\right|\|s_k\| \\ &\leq \Gamma + \frac{L\Gamma}{\mu}\left[1 + \frac{(1-\sigma)\alpha_k}{2\delta}\right] \\ &\leq \left(1 + \frac{L}{\mu}\left[1 + \frac{(1-\sigma)\alpha_k}{2\delta}\right]\right)\Gamma, \end{aligned}\text{”} \tag{33}$$

Specifically, (25) is true. Consequently, (28) is obtained from Lemma 1 which corresponds to (30) for “uniformly convex functions”.

4. Numerical Results

Fifteen problems of classical unconstrained optimization were studied respectively [1] in order to test the new proposed algorithm. **Fortran** 90 was used to implement all tests. Firstly, $\delta = 0.001$ and $\sigma = 0.9$ were set. If $\|g_{k+1}\| \leq 10^{-6}$**,** then the process would stop. Table 1 shows the statistical results of modified HY and FR method.

In Table 1, the results of algorithms are written as NI/NR/NF, referring to the iteration number, the restart calls number and function evaluations, respectively. The dimension of the test problems is denoted by Dim.

Table 1: "A Comparison of different CG-algorithms with different test functions and different dimensions"

		Algorithm of FR			Algorithm of ZMM		
P. No.	**Dim**	**NI**	**NR**	**NF**	**NI**	**NR**	**NF**
1	100	47	18	93	37	21	78
	1000	78	45	131	35	19	78
2	100	43	18	88	36	20	79
	1000	46	19	92	32	16	69
3	100	32	15	52	13	7	26
	1000	22	10	42	15	8	28
4	100	10	6	27	9	6	25
	1000	24	16	191	23	15	49
5	100	32	13	64	7	4	15
	1000	77	46	129	13	7	26
6	100	37	8	67	42	17	63
	1000	73	27	115	61	27	98
7	100	15	9	31	8	6	17
	1000	8	6	17	7	5	15
8	100	180	60	313	72	25	137
	1000	F	F	F	85	25	164
9	100	89	32	174	67	32	154
	1000	107	40	211	72	37	180
10	100	124	41	231	47	6	85
	1000	445	196	711	176	32	307
11	100	71	35	110	28	11	53
	1000	47	15	84	29	10	56
12	100	101	40	217	80	51	180
	1000	101	40	214	82	50	180
13	100	32	12	65	23	12	54
	1000	53	22	116	37	20	88
14	100	130	49	196	113	35	175
	1000	364	119	593	376	108	589
15	100	121	65	218	78	24	126
	1000	345	169	634	210	65	344
Total		2848	1191	5226	1828	696	3374

Fail: The algorithm fails to converge.

Problems numbers demonstrate that: 1. is the Extended Rosenbrock, 2. is the Extended White & Holst, 3. is the Extended Beale, 4. is the Penalty, 5. is the Extended Tridiagonal 1, 6. is the Generalized Tridiagonal 2, 7. is the Extended PSC1, 8. is the Extended Powell, 9. is the Extended Maratos, 10. is the Quadratic Diagonal Perturbed, 11. is the Extended Wood, 12. is the Extended Hiebert, 13. is the Extended Quadratic Penalty QP2, 14. is the Quadratic QF2, 15. is the DIXMAANE (CUTE).

The performance of the two algorithms is shown in Table 1. Under the same computing environment, some conditions indicate that this modified method of conjugate gradient significantly outperforms the previous methods of conjugate gradient.

5. Conclusions

This study has confirmed that the new modified algorithm is effective in the computational solving of unconstrained optimization problems, and the statistical results reveal the effectiveness of the proposed algorithm for problems of the given test.

References

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

[2] Andrei, N. (2008). Another hybrid conjugate gradient algorithm for unconstrained optimization. *Numerical Algorithms*, *47*(2), 143-156.

[3] Li, C., Fang, L., & Cao, X. (2013). Global convergence of a kind of conjugate gradient method. *Indonesian Journal of Electrical Engineering and Computer Science*, *11*(1), 544-549.

[4] Dai, Y. H., & Yuan, Y. (1999). A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on optimization*, *10*(1), 177-182.

[5] Dai, Y., Han, J., Liu, G., Sun, D., Yin, H., & Yuan, Y. X. (2000). Convergence properties of nonlinear conjugate gradient methods. *SIAM Journal on Optimization*, *10*(2), 345-358.

[6] Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal*, *7*(2), 149-154.

[7] Iiduka, H., & Narushima, Y. (2012). Conjugate gradient methods using value of objective function for unconstrained optimization. *Optimization Letters*, *6*(5), 941-955.

[8] Hager, W. W., & Zhang, H. (2006). A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, *2*(1), 35-58.

[9] Amini, K., Ahookhosh, M., & Nosratipour, H. (2014). An inexact line search approach using modified nonmonotone strategy for unconstrained optimization. *Numerical Algorithms*, *66*(1), 49-78.

[10] Zhang, L., Zhou, W., & Li, D. (2006). Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search. *Numerische Mathematik*, *104*(4), 561-572.