MATEMATIKA: MJIAM, 2021, Volume 37, Number 3, 147–155

The Numerical Calculation of Hybrid Conjugate Gradient Method Under Armijo Line Search and Its Application

[1]Norhaslinda Zullpakkal*, [2]Norrlaili Shapiee, [3]Syazwani Mohd Zokri and [4]Mohd Rivaie

[1,3,4]Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA (UiTM) Cawangan Terengganu,
Kampus Kuala Terengganu, 21080 Kuala Terengganu, Terengganu, Malaysia

[2]Tamhidi Centre, Universiti Sains Islam Malaysia (USIM)
71800 Nilai, Negeri Sembilan, Malaysia

*Corresponding author: lindazullpakkal@uitm.edu.my

Article history
Received: 21 August 2021
Received in revised form: 2 November 2021
Accepted: 9 November 2021
Published online: 15 December 2021

Abstract Conjugate gradient (CG) method is known due to its simplicity, global convergence and low memory requirement. To date, the research on CG method in Google Scholar has reached 1470000. Nowadays, the modification on hybrid CG method has become a focus among researchers. Thus, this paper introduces a new hybrid CG coefficient by combining two previous coefficients, Linda-Aini-Mustafa-Rivaie (LAMR) and Norrlaili-Rivaie-Mustafa-Ismail (NRMI). Since LAMR has a good performance under strong Wolfe while NRMI is quite good with exact line search, it is guaranteed that the new proposed hybrid CG method, NL will yield a good numerical analysis under Armijo line search. NL is compared to LAMR, NRMI and Abashar-Mustafa-Rivaie-Ismail (AMRI) to solve the unconstrained optimization problems. Based on the performance profile, NL coefficient is able to solve 58% problems with least iteration number and 52% problems with least CPU time. In order to test its capability, this NL coefficient is applied in regression analysis for data fitting. A real data set concerning Employees' Provident Fund (EPF) dividend rate has been chosen to construct the linear regression model. The linear model of NL coefficient is compared to the least square and Excel trendline methods. According to the relative error, it shows that NL coefficient is applicable to solve real-life problem which makes it a promising method.

Keywords hybrid; conjugate gradient method; Armijo line search; LAMR; regression analysis

Mathematics Subject Classification 49M37

1 Introduction

An optimization tool aims to find the maximum or minimum value for any optimization problems. Optimization problems can be either constrained or unconstrained. Due to the fact that constrained optimization problems can be adapted into unconstrained optimization problems with the help of famous methods such as Kuhn Tucker and Penalty functions, the unconstrained problem has been prioritized. The standard unconstrained optimization problem is the minimum function of x where x is an element of the real numbers, R^n and $f : R^n \to R$ is constantly differentiable. This problem is termed as,

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

This function is minimized by an iterative scheme,

$$x_{k+1} = x_k + \alpha_k d_k \text{ for } \quad k = 1, 2, ... \tag{2}$$

The next iteration point and current iteration point are represented by x_{k+1} and x_k, respectively. The notation of α_k indicates the stepsize while d_k is the search direction. The stepsize can be solved by exact or inexact line search. There are abundance of inexact line searches but Goldstein [1] , Wolfe [2] and Armijo [3] are the three main references. Armijo line search is introduced in 1966. It is stated to be the simplest line search where it is easy to be implemented in computation procedure. In this line search, with given positive constant $s$and $\rho, \mu \in (0, 1)$, the α_k is chosen by,

$$f(x_k + \alpha_k d_k) \le f(x_k) + \mu \alpha_k g_k^T d_k \tag{3}$$

in which $\alpha = \max\{s, s\rho, s\rho^2, ...\}$. The function value should be reduced in a proportional manner to both the α_k and the directional derivative $g_k^T d_k$. Condition (3) is also referred as the sufficient decrease condition. This line search tends to choose medium range of stepsize.

The search direction, d_k in the iterative scheme can be solved either by Steepest Descent, Newton's, Quasi Newton (QN), Conjugate Gradient (CG) or combination of these methods, hybrid CG-QN. The basic search direction of CG method is designated as,

$$d_k = \begin{cases} -g_k & \text{if } k = 0 \\ -g_k + \beta_k d_{k-1} & \text{if } k \ge 1 \end{cases} \tag{4}$$

where g_k is the gradient of function and β_k is the CG coefficient. As in (4), when $k = 0$, the search direction of CG method is same as SD method.

Various conjugate gradient methods have been proposed with difference choice of the parameter β_k. As stated in Shapiee *et al.* [4], the examples of well-known β_k are Hestenes-Stiefel (HS), Dai-Yuan (DY), Rivaie-Mustafa-Ismail-Leong (RMIL), Liu-Storey (LS), Conjugate Descent (CD), Polak-Ribiere (PR) and Fletcher-Reeves (FR). Guang-ming [5] proved that LS coefficient under Armijo line search satisfies the sufficient descent condition and it is globally convergent towards solution point. Ibrahim *et al.* [6] proposed a new BFGS-CG method under Armijo line search and it is proven to be globally convergent. Yin *et al.* [7] also proposed a modified Polak-Ribiere-Polyak (PRP) CG method based on the modified secant equation under Armijo line search. Kamandi *et al.* [8] proposed new CG-like method under the Armijo condition satisfies sufficient descent condition. Zhang *et al.* [9] modified PRP coefficient and test the method by using Armijo line search. Since modified CG methods and Armijo line search are still being an interest among researchers, it has been a good idea to extend the research. The modified CG coefficients Norrlaili-Rivaie-Mustafa-Ismail (NRMI), Abashar-Mustafa-Rivaie-Ismail (AMRI) and Linda-Aini-Mustafa-Rivaie (LAMR) proposed by Shapiee *et al.* [10], Abashar *et al.* [11] and Zull *et al.* [12] respectively are used for comparison are listed below,

$$\beta_k^{NRMI} = \frac{g_k^T(g_k - g_{k-1})}{g_{k-1}^T(g_k - d_{k-1})},$$
$$\beta_k^{AMRI} = \frac{g_k^T(g_k) - \frac{\|g_k\|}{\|g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{d_{k-1}^T(d_{k-1})},$$
$$\beta_k^{LAMR} = \frac{g_k^T(\frac{\|d_{k-1}\|}{\|d_{k-1}-g_k\|}g_k - g_{k-1})}{\frac{\|d_{k-1}\|}{\|d_{k-1}-g_k\|}\|d_{k-1}\|^2}.$$

This paper is organized by the modification of the new hybrid CG method in the next section, followed by the numerical results and implementation and conclusion.

2 New Hybrid Conjugate Gradient Method and Algorithm

This new coefficient is inspired by previous researchers, Dai et al., Yang et al., Liu and Jiu, Malik et al. and Sulaiman et al. [13-17] which proposed the hybrid CG methods. Thus, a new hybrid CG method is presented and denoted as Norrlaili-Linda (NL). This new coefficient is the combination of two modified CG coefficients which are NRMI and LAMR. This new method is stated as below,

$$\beta_k^{NL} = \begin{cases} \beta_k^{NRMI} & \text{if } 0 \le \beta_k^{LAMR} \le \beta_k^{NRMI} \\ \beta_k^{LAMR} & \text{otherwise} \end{cases},$$

The algorithm for computing the proposed method is referred from Rivaie *et al.* [18],

Step 1 Process of initialization. Given x_0, set $k = 0$.

Step 2 Computation of $\beta_k^{NL}, \beta_k^{LAMR}, \beta_k^{NRMI}$ and β_k^{AMRI}.

Step 3 Computation of d_k based on CG method. If $\|g_k\| = 0$, then stop.

Step 4 Computation of α_k under Armijo line search in (3).

Step 5 A new point is updated based on the iterative scheme in (2).

Step 6 Check convergent test, $f(x_{k+1}) < f(x_k)$ and stopping criteria, $\|g_k\| \le 10^{-6}$.

If it fulfill both of the conditions, then stop. Alternatively, return to Step 2 by $k = k + 1$.

3 Numerical Results and Discussion

This numerical experiment is important to ensure the efficiency and robustness of the methods. The four CG coefficients, $\beta_k^{NL}, \beta_k^{LAMR}, \beta_k^{NRMI}$ and β_k^{AMRI} are tested with eleven test functions which are referred from Andrei et al. [19]. Three different initial points and small-scale variables are chosen which are listed in Table 1.

These coefficients are computed with Armijo line search utilizing MatlabR2019a subroutine programming. Typically, the number of iterations (NOI) and CPU time are considered for numerical analysis. Interpretation of the results is integral to creating the performance profile

Table 1: Sample Test Functions

No.	Test Functions	Initial Points	Variables
1	Booth	(4,4), (12,12), (20,20)	2
2	Dixon And Price	(8,8), (16,16), (20,20)	2
3	Generalized Quartic	(4,4), (8,8), (16,16)	2
4	Power	(2,2), (8,8), (12,12)	2
5	Six Hump	(4,4), (8,8), (20,20)	2
6	Zettl	(2,2), (12,12), (16,16)	2
7	Fletcher	(2,2), (6,6), (12,12)	2, 4
8	Diagonal 4	(4,4), (8,8), (12,12)	2, 4, 10
9	Extended Denschnb	(12,12), (16,16), (20,20)	2, 4, 10
10	Extended Himmelblau	(4,4), (12,12), (20,20)	2, 4, 10
11	Shallow	(4,4), (12,12), (16,16)	2, 4, 10

by Sigmaplot. As an overview, the performance profile by Dolan and More [20] is used to show the ratio of a specific solver's best time to the best time for all solvers. It is the easiest way as this application is able to evaluate the effectiveness and robustness of the tested methods in graphs which are shown as in Figure 1 and Figure 2. According to Mohamed *et al.* [21], the convergence rate determines a method's effectiveness while the total of test problems solved determine its robustness. As illustrated in the graphs, the top right curve represents the most robust CG method as well as its ability to solve a greater number of test problems, whereas the top left curve represents the fastest method for the least amount of NOI or CPU time.

Based on Figure 1 and Figure 2, NL coefficient is obviously the most effective one as it has the least NOI and CPU time. Then, it is followed by AMRI, LAMR and NRMI. By a closer inspection of Figure 1, NL coefficient is able to solve 58% problems with least NOI and 52% problems with least CPU time. AMRI, LAMR and NRMI coefficients are able to solve 30%, 13%, 7% problems with least NOI. However, in term of CPU time in Figure 2, NRMI is unable to solve any problem with least CPU time, but AMRI and LAMR solve 43% and 6% respectively. Besides, both figures show that LAMR is able to overtake AMRI at certain points. Last but not least, these four coefficients are able to solve all the test problems where these coefficients approach 1.0 at the right side.

4 Implementation

To prove its capability, the new hybrid CG method is applied to regression analysis. The proposed CG method is compared with Excel Trendline (ET) and Least Square (LS) methods based on the relative error. The proposed method is applied for a real data set for practical application. The real data set is taken based on the real-life problems which is obtained from the website of Employees' Provident Fund (EPF), www.kwsp.gov.my.

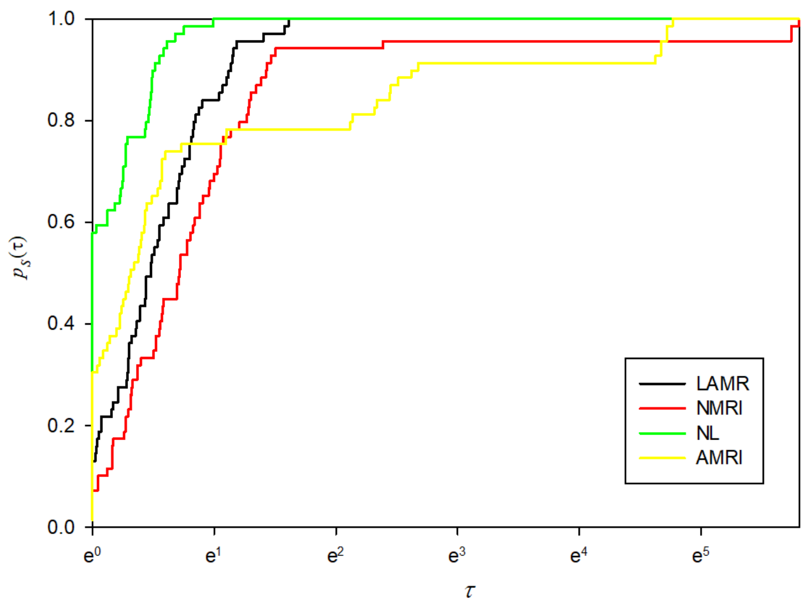

Figure 1: Performance Profile Based on Iteration Number

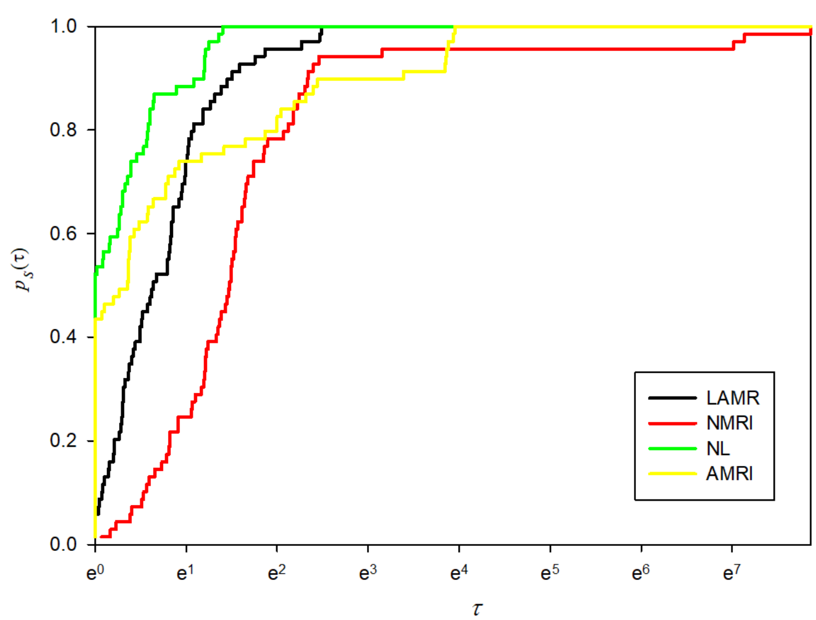

Figure 2: Performance Profile Based on CPU Time

Table 2: EPF Dividend Rate from 2001 to 2017

Number of data (x)	Year	EPF Dividend Rate (y)
1	2001	5.00
2	2002	4.25
3	2003	4.50
4	2004	4.75
5	2005	5.00
6	2006	5.15
7	2007	5.80
8	2008	4.50
9	2009	5.65
10	2010	5.80
11	2011	6.00
12	2012	6.15
13	2013	6.35
14	2014	6.75
15	2015	6.40
16	2016	5.70
17	2017	6.90

Based on Table 2, the data from 2001–2016 are analysed by forming linear model while the last data which is in year 2017 is excluded as it will be used to estimate the accuracy of the regression model. A trendline and its linear model are constructed by Microsoft Excel based on the data from 2001–2016 as shown in Figure 3.

Using Microsoft Excel, the linear model for Trendline Method is shown below:

$$y = 0.1338970588x + 4.3462500000.$$

Next, the linear model of least square method is generated automatically using the Matlab software:

$$y = 0.1338970588x + 4.3462500000.$$

The implementation of the new hybrid CG method for data fitting has been done using the Matlab software. The NL coefficient is compared with NRMI, AMRI and LAMR same initial point which is (10,10). The result for NOI, CPU Time and linear model for each CG coefficient is shown in Table 3.

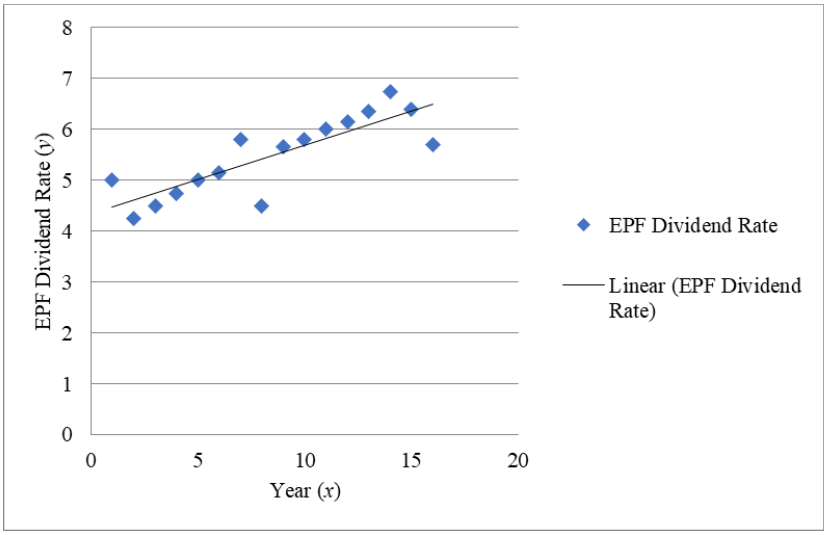

Figure 3: Linear Graph of EPF Dividend Rate versus Years

Table 3: Linear Models of CG Coefficients

Methods	CPU Time	NOI	Linear Models of CG Coefficients
NL	0.0139	61	$y = 0.1338970591\ x + 4.3462499975$
NRMI	1.2106	2742	$y = 0.1338970590\ x + 4.3462500011$
LAMR	0.0654	145	$y = 0.1338970638\ x + 4.3462499442$
AMRI	0.1389	356	$y = 0.1338970538\ x + 4.3462500540$

The linear models for each coefficient are used to calculate the relative error of Year 2017 using the following formula which is cited from Shoid *et al.* [22],

$$\text{relative error} = \frac{|\text{exact value} - \text{approximate value}|}{|\text{exact value}|}.$$

The estimated value and relative error for each method are shown in Table 4. Based on Table 4, Excel Trendline and Least Square methods yield the same estimated value. However, by taking to 7^{th} decimal places, all of these methods produce the same relative error which is 0.0402174. It shows that the new proposed CG method, NL is relevant for real life problem.

5 Conclusion

It is concluded that the new hybrid CG method is the most efficient and robust method. It is proven that NL coefficient is numerically convenient under Armijo line search with the least NOI and CPU time. Therefore, this coefficient is promising due to its applicability of solving the regression model.

Table 4: Estimated Value and Relative Error of Each Method

Methods	Estimated Value	Relative Error
NL	6.6225000022	0.04021739
NRMI	6.6225000041	0.04021739
LAMR	6.6225000288	0.04021739
AMRI	6.6224999686	0.04021740
Excel Trendline	6.6224999996	0.04021739
Least Square	6.6224999996	0.04021739

Acknowledgments

The authors extend their sincere gratitude to the editors and to the referees for the recommendations and constructive comments.

References

[1] Goldstein, A. A. On steepest descent. Journal of the Society for Industrial and Applied Mathematics, Series A: Control. 1965. 3(1): 147-151.

[2] Wolfe, P. Convergence conditions for ascent methods. SIAM review. 1969. 11(2): 226-235.

[3] Armijo, L. Minimization of functions having Lipschitz continuous first partial derivatives. Pacific Journal of mathematics. 1966. 16(1): 1-3.

[4] Shapiee, N., Rivaie, M., Mamat, M. and Ghazali, P. L. A new family of conjugate gradient coefficient with application. International Journal of Engineering & Technology. 2018. 7(3.28): 36-43.

[5] Guang-ming, Z. Convergence properties of a conjugate gradient method with Armijo-type line searches. Chinese Journal of Engineering Mathematics. 2008. 25: 405-410.

[6] Ibrahim, M. A. H., Mamat, M. and Leong, W. J. The hybrid BFGS-CG method in solving unconstrained optimization problems. Abstract and Applied Analysis. 2014.

[7] Yin, J., Wang, L. and Jiang, X. A modified PRP conjugate gradient method with Armijo line search for large-scale unconstrained optimization. In *2017 36th Chinese Control Conference (CCC).* IEEE. 2017: 2568-2571.

[8] Kamandi, A. and Amini, K. A new conjugate gradient-like method with sufficient descent condition and its global convergence based on the Armijo line search. *arXiv preprint arXiv:1904.06321.* 2019.

[9] Zhang, M., Zhou, Y. and Wang, S. A modified nonlinear conjugate gradient method with the Armijo line search and its application. *Mathematical Problems in Engineering.* 2020. ISSN:1024-123X.

[10] Shapiee, N., Rivaie, M., Mamat, M. and Mohd, I. A new modification of Hestenes-Stiefel method with descent properties. In *AIP Conference Proceedings.* American Institute of Physics. 2014. 1602(1): 520-526.

[11] Abashar, A., Mamat, M., Rivaie, M., Mohd, I. and Omer, O. The proof of sufficient descent condition for a new type of conjugate gradient methods. In *AIP Conference Proceedings.* American Institute of Physics. 2014. 1602(1): 296-303.

[12] Zull, N., Shoid, S., Ghani, N., Mohamed, N., Rivaie, M. and Mamat, M. A conjugate gradient method with descent properties under strong Wolfe line search. *Journal of Physics: Conference Series.* 2017. 890(1): 012105.

[13] Dai, Y.-h. and Yuan, Y. An efficient hybrid conjugate gradient method for unconstrained optimization. *Annals of Operations Research.* 2001. 103(1): 33-47.

[14] Yang, X., Luo, Z. and Dai, X. A Global Convergence of LS-CD Hybrid Conjugate Gradient Method. *Adv. Numerical Analysis.* 2013. 517452:1-517452:5.

[15] Liu, J. and Li, S. New hybrid conjugate gradient method for unconstrained optimization. *Applied Mathematics and Computation.* 2014. 245: 36-43.

[16] Malik, M., Abas, S. S., Mamat, M. and Mohammed, I. S. A new hybrid conjugate gradient method with global convergence properties. *International Journal of Advanced Science and Technology.* 2020. 29(5): 199-210.

[17] Sulaiman, I., Mamat, M., Waziri, M., Yakubu, U. and Malik, M. The convergence properties of a new hybrid conjugate gradient parameter for unconstrained optimization models. In *Journal of Physics: Conference Series.* IOP Publishing. 2021. 1734(1): 012012.

[18] Rivaie, M., Mamat, M., June, L. W. and Mohd, I. A new class of nonlinear conjugate gradient coefficients with global convergence properties. *Applied Mathematics and Computation.* 2012. 218(22): 11323-11332.

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

[20] Dolan, E. D. and Moré, J. J. Benchmarking optimization software with performance profiles. *Mathematical Programming.* 2002. 91(2): 201-213.

[21] Mohamed, N. S., Mamat, M., Rivaie, M. and Shaharudin, S. M. A comparison on classical-hybrid conjugate gradient method under exact line search. *International Journal of Advances in Intelligent Informatics.* 2019. 5(2): 150-168.

[22] Shoid, S., Zull, N., Rivaie, M., Mamat, M., Liza Ghazali, P. and Afendee Mohamed, M. A New Hybrid of Conjugate Gradient Method with Descent Properties. *International Journal of Engineering & Technology.* 2018. 7(3.28): 4.