International Journal of Engineering & Technology, 7 (2.14) (2018) 25-27

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

The application of new conjugate gradient methods in estimating data

Syazni Shoid [1]*, Norrlaili Shapiee[1], Norhaslinda Zull[1], Nur Hamizah Abdul Ghani[1], Nur Syarafina Mohamed[1], Mohd Rivaie[2], Mustafa Mamat[1]

[1]Faculty of Informatics and Computing,
Universiti Sultan Zainal Abidin (Unisza), Terengganu, Malaysia
[2]Department of Computer Sciences and Mathematics,
Universiti Teknologi MARA (UiTM) Terengganu, Malaysia
**Corresponding author E-mail: syaznishoid@yahoo.com*

Abstract

Many researchers are intended to improve the conjugate gradient (CG) methods as well as their applications in real life. Besides, CG become more interesting and useful in many disciplines and has important role for solving large-scale optimization problems. In this paper, three types of new CG coefficients are presented with application in estimating data. Numerical experiments show that the proposed methods have succeeded in solving problems under strong Wolfe Powell line search conditions.

Keywords: *Conjugate Gradient Coefficient, Inexact Line Search, Least Squares, Regression, Strong Wolfe Powell*

1. Introduction

The conjugate gradient (CG) method become attentive because this method does not require matrix storage and their iteration cost is low compared to other methods. Consider the CG method as a solution to the numerical problem for the following constrained problem,

$$\min \quad f(x) \text{ subject to } x \in X \ , \tag{1}$$

where $f(x)$ is an objective function, the $x \in R^n$ is a decision variable and $X \in R^n$ is a constraint set. If $x = R^n$, then optimization problem (1) can be express as an unconstrained optimization problem,

$$\min_{x \in R^n} \quad f(x) \ . \tag{2}$$

The CG problems is solved iteratively, $x_0 \in R^n$ is the starting point, using a recurring formula,

$$x_{k+1} = x_k + \alpha_k d_k \qquad k=0,1,2,..., \tag{3}$$

where x_k is the current iteration point and the $\alpha_k > 0$ is a step-size, obtained through some line search methods.
Among the line searches used in practice is an inexact line searches. One of the most common and popular of inexact line searches is Wolfe line search[13]. This line search introduced two conditions as follow,

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

$$g_{k+1}^T d_k \ge \sigma g_k^T d_k \tag{5}$$

where $0 < \mu < \sigma < 1$. By replacing (5) with condition (6), then this is called as strong Wolfe line search,

$$\left| g_{k+1}^T d_k \right| \le -\sigma g_g^T d_k \ . \tag{6}$$

The CG methods also determines the search direction d_k with

$$d_k = \begin{cases} -g_k & \text{if} \quad k = 0 \\ -g_k + \beta_k d_{k-1} & \text{if} \quad k \ge 1 \end{cases} \tag{7}$$

where $g_k = \nabla f(x_k)$, and $\beta_k \in R$ is a coefficients which determines the difference between the CG method. The examples of β_k suggested by early researchers are Fletcher and Reeves (FR) method[1], Polak Ribiere and Rolyak (PR) method[2], Hestenes and Stiefel (HS) method[3], Rivaie, Mustafa, Ismail and (RMIL) method[4], Syazni, Rivaie, Mustafa and Ismail (SRMI) method[5], Norrlaili, Rivaie, Mustafa and Ismail (NRMI) method[6], Rivaie, Abashar, Mustafa and Ismail (RAMI) method[7]. Here are the formulas,

$$\beta_k^{FR} = \frac{g_k^T g_k}{\|g_{k-1}\|^2} \tag{8}$$

$$\beta_k^{PRP} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \tag{9}$$

$$\beta_k^{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \tag{10}$$

$$\beta_k^{RMIL} = \frac{g_k^T (g_k - g_{k-1})}{\|d_{k-1}\|^2} \tag{11}$$

$$\beta_k^{SRMI} = \frac{\dfrac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} + \dfrac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k - d_{k-1})}}{2} \tag{12}$$

$$\beta_k^{NRMI} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}{}^T (g_k - d_{k-1})} \tag{13}$$

$$\beta_k^{RAMI} = \frac{g_{k+1}^T\left(g_{k+1} - \frac{\|g_{k+1}\|}{\|g_k\|} g_k\right)}{d_k^T (d_k - g_{k+1})} \quad (14)$$

where $g_{k-1} = \nabla f(x_{k-1})$ and $\|.\|$ is the Euclidian norm of vectors. In addition, there are many other researchers who conduct studies on CG[8-12].

Regression is used to determine the relationship between two or more variables and use this relationship to make estimation. The linear regression model is defined by

$$y = a_0 + a_1 x \quad (15)$$

where a_0 and a_1 are the regression parameters. The value of parameters is calculated by method of least squares that minimized the problem to solve the above regression model

$$\min Q = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i)]^2 \quad (16)$$

Then, the least square problems is transform to an unconstrained optimization problem as

$$\min_{x \in R^2} f(a) = \sum_{i=1}^{n} [y_i - a(1, x_i)^T]^2 \quad (17)$$

The quadratic regression model is defined by

$$y = a_0 + a_1 x + a_2 x^2 \quad (18)$$

where a_0, a_1 and a_2 are the regression parameters. Similar to the problem above, the least squares method can be used to solve the problem as

$$\min Q = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + a_2 x^2)]^2 \quad (19)$$

Similar transformation of least squares problem into an unconstrained optimization problem yields the following as

$$\min_{x \in R^3} f(a) = \sum_{i=1}^{n} [y_i - a(1, x_i, x_i{}^2)^T]^2 \quad (20)$$

The structure of this paper is organized as follows. In the section 2, new types of CG coefficients are presented. In section 3, we presented the data analysis, while in Section 4, some numerical outcome and discussion corresponding to these new β_k is given. Lastly, our conclusion is presented in Section 5.

2. New types of conjugate gradient coefficients

We used three types of β_k formulas named as β^{SYRM}, β^{SRM} and β^{NRMI+}. The formulas are shown below,

$$\beta_k^{SYRM} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_k - g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{d_{k-1}^T (d_{k-1} - g_k)} \quad (21)$$

The SYRM is classified as a modified CG method. The original idea is from RAMI in (14) where a new formula for numerator has been proposed and the original formula for denominator as the RAMI formula has been retained.

$$\beta_k^{SRM} = \begin{cases} \dfrac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, & g_k^T g_k > \left|g_k^T g_{k-1}\right| \\ \dfrac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_k - g_{k-1}\|}\left|g_k^T g_{k-1}\right|}{d_{k-1}^T (d_{k-1} - g_k)}, & else \end{cases} \quad (22)$$

The SRM is classified as hybrid CG method. For this method, we combine two methods: HS and SYRM method with certain conditions.

$$\beta_k^{NRMI+} = \max\left\{0, \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k - d_{k-1})}\right\} \quad (23)$$

The last method is the NRMI+ method which is derived from NRMI in (13). We extend the original method by assuming it as a nonnegative value.

The following algorithm is a common algorithm of CG method used in this study.

ALGORITHM 1: The basic of Conjugate Gradient algorithm

Step 1: Given an initial point x_0 and set $k = 0$.

Step 2: Computing conjugate gradient coefficients, β_k based on (21) until (23).

Step 3: Computing search direction

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

If $g_k = 0$, terminate the execution of the algorithm.

Step 4: Computing step size α_k by inexact line search rule Solve, (4) and (6).

Step 5: Updating new point

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

Step 6: Convergent test and stopping criteria.

If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| < \varepsilon$, then terminate.

Otherwise go to Step 1 with $k = k + 1$.

3. Data analysis

Table 1 below shows the actual data obtained from Traffic Statistic Branch, Bukit Aman PDRM[14]. The data represent the number of road death from 2004 to 2014.

Table.1: The Number of Road Death

Number Of Data (x)	Years	Road Death Index(y)
1	2004	4.51
2	2005	4.18
3	2006	3.98
4	2007	3.73
5	2008	3.63
6	2009	3.55
7	2010	3.4
8	2011	3.21
9	2012	3.05
10	2013	2.9
11	2014	2.66

However, for the purpose of computation, only data from 2004 to 2013 are considered while data for 2014 are reserved for error calculation. In this experiment, the function used to solve the problem above is considered linear and quadratic cases and solved by SYRM, SRM and NRMI methods. Furthermore, for the purpose of comparison, sum squares method is also applied to solve the problem. Then, the results are analyzed by finding the relative error of the CG and sum squares methods using the formula shown below,

$$\text{Relative error} = \left|\frac{\text{exact value} - \text{approximate value}}{\text{exact value}}\right|$$

Calculation of relative error is important to determine the best method to estimate the number of road death for 2014.

4. Results and discussion

The data is implemented in SYRM, SRM, NRMI+ and sum square smethods. All calculations involve the use of Microsoft Excel and Matlab 12 subroutine program. The results of linear and quadratic regression model for each method are shown below in Table 2 and Table 3 respectively.

Table.2: The Linear Regression Model

Method	Result
SYRM	4.52466666666665-0.165575757575809x
SRM	4.52466666666661-0.165575757575749x
NRMI+	4.52466666666661-0.165575757575749x
Least Squares	4.52466666666666-0.165575757575757x

Table.3: The Quadratic Regression Model

Method	RESULT
SYRM	FAIL
SRM	4.659666663129-0.233075766305x+ 0.006136364734x^2
NRMI+	4.659666663121-0.233075766325x+ 0.006136364736x^2
Least Squares	4.659666666667-0.233075757576x+ 0.006136363636x^2

As we can see, all methods can solve the problems for linear and quadratic cases except SYRM, where it failed to solve quadratic case. The relative error for each data is defined by comparing the actual data and the approximation data. The model that gives the smallest sum of relative error is considered the best of approximation function. Table 4 and table 5, show the results of the relative error of each method for linear and quadratic cases respectively.

Table.4: The Relative Error for linear

METHOD	RELATIVE ERROR
SYRM	0.01072100313498
SRM	0.01072100313479
NRMI+	0.01072100313479
Least Squares	0.01072100313480

Table.5: The Relative Error for quadratic

METHOD	RELATIVE ERROR
SYRM	FAIL
SRM	0.01467085291184
NRMI+	0.01467085292636
Least Squares	0.01467084639498

Based on the above results, we can say that the linear case gives the smallest relative error value compared to the quadratic case. In the case of linear, we can see the SRM and NRMI+ method is comparable which gives the smallest relative error value compared to SYRM and least squares methods. Thus, SRM and NRMI+ method are the best method to estimate the number of road deaths in 2014. As a result when x=11, then y=2.70.

5. Conclusion

The data is applied to the CG methods which are SYRM, SRM, NRMI+ method and sum squares methods. Then, it is found that all methods can solve the problem for a linear case. Instead, for quadratic case only the SYRM method fails to solve the problem. As a conclusion, the number of road death for 2014 is estimated as 2.70.

Acknowledgement

We would like to thank The Ministry of Higher Education of Malaysia for their sponsorship MyPhD KPM.

References

[1] Fletcher R & Reeves C, Function minimization by conjugate gradients (1964), *Comput J* 7, 149-154.

[2] Polak E & Ribiere G, Note sur la convergence de directions conjugees (1969), *Rev.Francaise Informat Recherche Operationelle,3E Annee* 16, 35-43.

[3] Hestenes MR & Steifel E, Method of conjugate gradient for solving linear equations (1952), *J,Res.Nat.Bur.Stand* 49, 409-436.

[4] Rivaie M, Mamat M, Leong WJ & Ismail M, A new conjugate gradient coefficient for large scale nonlinear unconstrained optimization (2012), *Int. Journal of Math. Analysis* 6, 23 1131-1146.

[5] Shoid S, Rivaie M & Mamat M, A modification of classical conjugate gradient method using strong Wolfe line search (2016), *AIP Conf. Proc.*, 1739, 020071, doi: 10.1063/1.4952562.

[6] Shapiee N, Rivaie M & Mamat M, A new classical conjugate gradient coefficient with exact line search (2016), *AIP Conf. Proc.* 1739,020082, doi: 10.1063/1.4952562.

[7] Rivaie M, Abashar A, Mamat M & Mohd I, The conjugence properties of a new type of conjugate gradient methods (2014), *App. Math. Science* 8, no. 1, 33-44.

[8] Ghani NHA, Rivaie M & Mamat M, A modified form of conjugate gradient method for unconstrained optimization problems (2016), *AIP Conf. Proc.* 1739, 020076, doi: 10. 1063/1.4952556.

[9] Hajar N, Mamat M, Rivaie M & Jusoh I, A new type of desent conjugate gradient method with exact line search (2016), *AIP Conf. Proc.* 1739, 020089, doi: 10. 1063/1.4952569.

[10] Khadijah W, Rivaie M, Mamat M & Jusoh I, A spectral KRMI conjugate gradient method under the strong-Wolfe line search (2016), *AIP Conf. Proc.* 1739, 020072, doi: 10. 1063/1.4952552.

[11] Zoutendjik G, Nonlinear programming computational methods (1970), *in: J. Abadie (Ed.), Interger and Nonlinear Programming, North-Holland, Amsterdan*, pp.37-86.

[12] Al-Baali M, Descent property and global convergence of Fletcher-Reeves method with inexact line search (1985), *IMA. J. Numer. Anal* 5, 121-124.

[13] Wolfe P, Convergence conditions for ascent method. II: some corrections (1971), *SIAM Rev* 13(2), 185-188.

[14] Traffic Statistics Branch, Bukit Aman PDRM, RTD Corporate Division, MIROS Headquarters, Road death index 2004-2014,available online: http://www.jkjr.gov.my/ms/maklumat_keselamatan/ststik-kmlgn-jln-rya/1858-indeks-kematian-jalan-raya-2004-2014.html), 24 jun 2015.