International Journal of Engineering & Technology, 7 (2.15) (2018) 94-97

International Journal of Engineering & Technology

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

Research Paper

Estimating the unemployment rate using least square and conjugate gradient methods

Nur Syarafina Mohamed[1]*, Mustafa Mamat[2], Mohd Rivaie[3], Nur Hamizah Abdul Ghani[2], Norhaslinda Zull[2], Syazni Syoid[2]

[1]Technical Foundation, Malaysian Institute of Industrial Technology, Universiti Kuala Lumpur, Pasir Gudang, Johor, Malaysia
[2]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Terengganu, Malaysia
[3]Department of Computer Science and Mathematics, Universiti Teknologi MARA, Terengganu, Malaysia
**Corresponding author E-mail: nursyarafina@unikl.edu.my*

Abstract

Unemployment rate is one of the major issues among Malaysian citizens. The unemployment rate indicates the percentage of the total workforce who are actively seeking employment and currently unemployed. In this paper, a data of unemployment rate of a state in Malaysia from year 2000 until 2015 is collected. The statistics data is extracted by Labour Force Survey Malaysia (LFSM) which was conducted monthly by using household approach targeted to working ages between 15 to 64 years old. An estimation data for year 2016 can be forecasted by using discrete least square method of numerical analysis and conjugate gradient method in unconstrained optimization. These methods have been chosen based on its simplicity and accuracy. The calculations are based on linear and quadratic models for each the method together with their errors. Results showed that the conjugate gradient method is comparable with the least square method.

Keywords: *Conjugate gradient method; Least square method; Unconstrained optimization; Workforce.*

1. Introduction

In economics, finance, trade, meteorology and medicine, regression analysis is often used as a prediction tool [1]. With the purpose of forecasting and understanding the relationship between dependent and independent variables, the Least Square (LS) method is famously chosen in this area due to its relative absolute error approach [2]. On the other hand, the LS method is used because of the data fitting usage in finding a function. Data fitting is the process of fitting models to data and analyzing the accuracy of the fit. For linear fitting, the Least Squares method is relatively simple to use because it merely need to employ 2×2 matrix formation.

The Least Squares method is also useful for comparisons between multiple regression models [3-6]. The classical regression model is defined by

$$y = h\left(x_1, x_2, ..., x_p + \varepsilon\right) \tag{1}$$

where y is the response variable, x_i is the predictor variable with $i = 1, 2, ..., p, p > 0$ is an integer constant and ε is the error term. The function $h\left(x_1, x_2, ..., x_p\right)$ describes the type of relationship between y and $x = \left(x_1, x_2, ..., x_p\right)$. Thus, the following linear regression model

$$y = a_0 + a_1x_1 + a_2x_2 + ... + a_px_p + \varepsilon \tag{2}$$

is the regression model with the simplest version, where $a_0, a_1, ..., a_p$ is defined as the regression parameters.

In regression analysis, the most important task is to estimate the parameters $a = \left(a_0, a_1, ..., a_p\right)$. By using the Least Squares method, the problem is defined as

$$\min\ E(a) = \sum_{i=1}^{n}\left(y_i - a_0 + y_1x_{i1} + y_2x_{i2} + ... + y_px_{ip}\right)^2 \tag{3}$$

where y_i is the data estimation of the i^{th} response variable and $x_{i1}, x_{i2}, ..., x_{ip}$ are p data evaluation of the response variable. Given that m is the number of data. Then, if the dimension of p and m is small, the parameters $a = \left(a_0, a_1, ..., a_p\right)$ can be acquired from a multivariate process of calculus thus it is not difficult to see that the problem defined is the same as the unconstrained optimization problem [7-8]. Consider an optimization problem [9],

$$\min_{x\in R^n} f(x)\,. \tag{4}$$

$f: R^n \rightarrow R$ is known as a continuously differentiable function which is bounded from below. An iterative formula which is known by

$$x_{k+1} = x_k + \alpha_k d_k, \qquad k = 0, 1, 2, ... \tag{5}$$

is generated starting from its initial guess at point x_0. From (5),

x_k is the current iterate point, $\alpha_k > 0$ is a step size which is obtained by one dimensional search. In this paper, the exact line search is used,

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k). \tag{6}$$

The d_k in (6) is known as the search direction used for Conjugate Gradient (CG) methods which is defined by,

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

where g_k is the gradient of $f(x)$ at the point x_k. The β_k used in this paper is obtained from [9-11]. This β_k is known as

$$\beta_k^{SMR} = \max\left\{ 0, \frac{\|\mathrm{g}_k\|^2 - |g_k^T g_{k-1}|}{\|d_{k-1}\|^2} \right\} \tag{8}$$

where SMR stands for Syarafina-Mustafa-Rivaie. This new CG coefficient is known to possess the global convergence properties and has a superior performance [9-11]. The g_k and g_{k-1} denote the gradients of $f(x)$ at the point x_k and x_{k-1} respectively. CG method is chosen due to its simplicity and easy to implement. There are several of CG methods introduced by previous researchers. Details explanation of them can be found from [12-19]. In this paper, both of the methods mentioned earlier are going to be used in estimating the unemployment rate which will be discuss later in the next sections.

2. Derivation process

By comparing the total least square error, the LS method involves determining the best approximating line [5]. Consider a set of data, (x_i, y_i) where x is said to be exact values if and only if y values have errors. The error is defined as

$$E_i = (a_0 + a_1 x_i) - y_i \tag{9}$$

The strategy to fit the “best” line through the data would be to minimize the sum of the residual error squares for all the available data.

$$\min \sum_{i=1}^{n} E_i^2 = \sum_{i=1}^{n} ((a_0 + a_1 x) - y_i)^2 \tag{10}$$

$$\min \sum_{i=1}^{n} E_i^2 = \sum_{i=1}^{n} ((a_0 + a_1 x + a_2 x^2) - y_i)^2 \tag{11}$$

Differentiate (10) with respect to a_0 and a_1 and (11) with respect to a_0, a_1 and a_2. Solve them simultaneously, then the general formula to find the linear and quadratic discrete least square models can be described as;

$$\begin{bmatrix} n & \sum_{i=1}^{n} x \\ \sum_{i=1}^{n} x & \sum_{i=1}^{n} x^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} n & \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 \\ \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i^3 & \sum_{i=1}^{n} x_i^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} x_i^2 y_i \end{bmatrix} \tag{13}$$

The calculation process by using the LS method and CG method are explained briefly by using the algorithm in section 2.1 and 2.2.

2.1. Least square method

Step 1: Identifying formula for both linear and quadratic model from (12) and (13).
Step 2: Identifying variables and data summation based on (12) and (13).
Step 3: Calculation of a_0, a_1 and a_2.
Step 4: Generating equations.
Step 5: Calculate error by using| (Exact value-Approximate value) / Exact value | and (12).
Step 6: Model estimation.

One can observe that there exists a linear relationship between the year and the value of unemployment rate, the regression equation is given by $y = a_0 + a_1 x$ with a_0 and a_1 denoting the regression parameters. Thus,

$$\min_{x \in R^3} f(a) = \sum_{i=1}^{n} [y_i - a(1, x_i, x_i^2)^T]^2 \tag{14}$$

2.2. Conjugate gradient method

Step 1: Initialization. Set $k = 0$ and select x_0. Identifying formula from (12) and (13) for both linear and quadratic.
Step 2: Compute β_k^{SMR} from (8).
Step 3: Compute search directions d_k based on (7). If $\|g_k\| = 0$, then stop.
Step 4: Solve α_k using the exact line search from (6).
Step 5: Updating new initial point using (5).
Step 6: Convergence test and stopping criteria. If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| \leq \varepsilon$ then stop. Otherwise, go to Step 2 with $k = k + 1$.

3. Statistical problem

The data set from Table 1 shows the unemployment rate in Terengganu from year 2000 until 2015. The statistic is derived from the Labour Force Survey (LFS) [21], where the data provided was up until 2015 only. The data is collected monthly by using a household approach, within the working age which is 15-64 years old. From Table 1, the $x-$variable denotes the year whilst the $y-$variable denotes the rate of unemployment. For data fitting process, only the data from 2000 to 2014 is considered. The data for year 2015 is reserved for error calculation.

Table 1: Unemployment rate in Terengganu for year 2000 to 2015

Number of Data (x)	Years	Unemployment Rate (y)
1	2000	3.0
2	2001	2.7
3	2002	3.2
4	2003	2.9
5	2004	3.2

6	2005	3.1
7	2006	3.6
8	2007	2.6
9	2008	3.4
10	2009	3.8
11	2010	3.7
12	2011	3.2
13	2012	3.0
14	2013	3.4
15	2014	4.2
16	2015	4.0

Next, the data in Table 1 is used to formulate the linear and quadratic models in (12) and (13) for LS and CG methods.3.1.

4. Results and discussion

In this section, unemployment rate is estimated by using both LS and CG methods. By solving (12) and (13) simultaneously, the values of a_0, a_1 and a_2 are obtained. Thus, the approximate functions for linear and quadratic LS methods are given as $f(x) = 2.823809524 + 0.055357142\,x$ and $f(x) = 3.002330759 + 0.033004345\,x + 3.642894008\,e^{-6}x^2$ respectively. Therefore, by using the approximate functions for LS method, the estimated rate of unemployment both for linear and quadratic models respectively for year 2015 are 3.71 and 3.53.

Implementing CG method to the problem stated in Table 1, the optimization problems obtained from (10) and (11) are formed around the first to the fifteenth data by using the MATLAB subroutine program. Thus, the obtained functions below are used as a test problem for linear and quadratic optimization model respectively.

$$f(a_0,a_1) = 15a_0^2 + 240a_0a_1 - 98a_0 + 1240a_1^2 - 815a_1 + 162.64 \tag{15}$$

$$\begin{aligned} f(a_0,a_1,a_2) = {} & 15a_0^2 + 240a_0a_1 + 2480a_0a_2 \\ & - 98a_0 + 1240a_1^2 + 28800a_1a_2 \\ & - 815a_1 + 178312a_2^2 - 8611.4a_2 \\ & + 162.64. \end{aligned} \tag{16}$$

Two different Trend Lines (TL) of linear and quadratic type are constructed based on the data from Table 1. These linear and quadratic (TL) are often referred to as lines of best fit which indicates as the data behaviours in order to determine if there are certain patterns. Figures 1 and 2 show the graphs of linear and quadratic models respectively which are automatically plotted by using Microsoft Excel software.

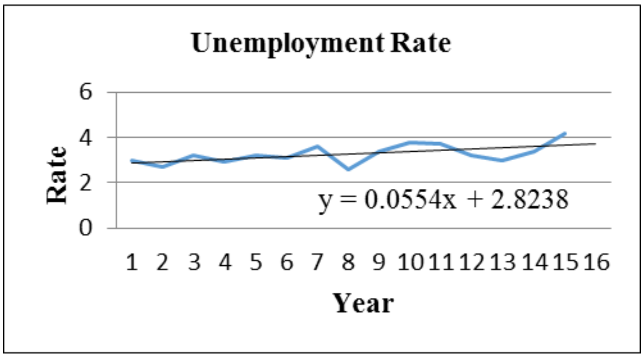

Fig. 1: Linear TL for unemployment rate in Terengganu

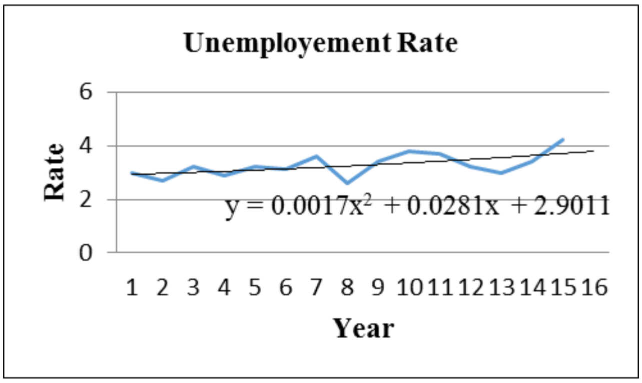

Fig. 2: Quadratic TL for unemployment rate in Terengganu

The equations of the TL are automatically calculated from Microsoft Excel software. Even though both lines in Figures 1 and 2 are plotted for linear and quadratic respectively, they are quite identical due to small differences between each data. The approximate functions obtained by both linear and quadratic models for both methods are compared with the TL plotted. The data for year 2015 is estimated by using all functions obtained. Then the relative errors of the data generated from each models are computed by comparing the actual data with the estimated data. From this relative errors obtained, the method efficiency can be determined. The actual data for year 2015 is 4.0 while the estimation point obtained for each models is presented in Table 2.

Table 2: Estimation points and relative errors

Models	Estimation Point	Relative Error
Linear LS	3.7095238	0.072619050
Quadratic LS	3.5313329	0.117166775
Linear TL	3.7095238	0.072619050
Quadratic TL	3.7868125	0.053296875
Linear CG	3.7095238	0.072619050
Quadratic CG	3.7868130	0.053296750

From Table 2, the relative errors of the actual data versus the predicted data for all methods are calculated. The relative error for the data generated from Quadratic CG is slightly smaller than the other models'. Thus, it could be said that Quadratic CG approach is most suitable to be applied to the data mentioned in Table 1.

5. Conclusion

The data of unemployment rate in Terengganu is analyzed and applied for LS and CG methods and used to predict the estimated data for the year 2015. From this implementation, it is found that both LS and CG methods can solve the problem for both linear and quadratic models under exact line search. In conclusion, the estimated data obtained by quadratic CG method has the smallest relative error. Thus, it can be said that the SMR coefficient is usable in statistical analysis for forecasting study under CG method.

Acknowledgement

The authors are grateful to the editor and the anonymous reviewers for their comments and suggestions which improved this paper substantially.

References

[1] Moyi AU, Leong WJ & Saidu I (2014), On the application of three-term conjugate gradient method in regression analysis. *International Journal of Computer Applications* 102, 1–4.

[2] Chapra SC & Canale RP (2006), *Numerical methods for engineers*, McGraw-Hill.

[3] Armstrong JS & Fildes R (1995), Correspondence on the selection of error measures for comparisons among forecasting methods. *Journal of Forecasting* 14, 67–71.

[4] Makridakis S, Anderson S, Carbone R, Hibson M, Lewandoski R, Newton J, Parzen E & Wikler R (1984), *The forecasting accuracy of major times series methods*, John Wiley and Sons.

[5] Narula S & Wellington J (1977), An algorithm for the minimum sum of weighted absolute errors regression. *Communication in Statistics* 6, 341–352.

[6] Khoshgoftaar TM, Bhattacharyya BB & Richardson GD (1992), Predicting software errors during development using nonlinear regression models: A comparative study. *IEEE Transactions on Reliability* 41, 390–395.

[7] Burden RL & Faires JD (2011), *Numerical analysis: Interpolation and polynomial approximation*, Cengage Learning.

[8] Yuan G & Wei Z (2009), New line search methods for unconstrained optimization. *Journal of the Korean Statistical Society* 38, 29–39.

[9] Mohamed NS, Mamat M, Susilawati F & Rivaie M (2016), A new coefficient of conjugate gradient method for nonlinear unconstrained optimization. *Jurnal Teknologi* 78, 131–136.

[10] Mohamed NS, Mamat M & Rivaie M (2016) Solving a large scale nonlinear unconstrained optimization with exact line search direction by using new coefficient of conjugate gradient methods. *AIP Conference Proceedings* 1787, 1–7.

[11] Mohamed NS, Mamat M & Rivaie M (2017) A new nonlinear conjugate gradient coefficient under strong Wolfe-Powell line search. *AIP Conference Proceedings* 1870, 1–7.

[12] Hamoda M, Mamat M, Rivaie M, Salleh Z & Amani Z (2015), A new nonlinear conjugate gradient coefficient for unconstrained optimization. *Applied Mathematical Sciences* 9, 1813–1822.

[13] Abashar A, Mamat M, Rivaie M, Mohd I & Omer O (2014), The proof of sufficient descent condition for a new type of conjugate gradient methods. *AIP Conference Proceedings* 1602, 293–303.

[14] Ghani NHA, Rivaie M & Mamat M (2016), A modified form of conjugate gradient method for unconstrained optimization problems. *AIP Conference Proceedings* 1739, 1–8.

[15] Hajar N, Mamat M, Rivaie M & Jusoh I (2016) A new type of descent conjugate gradient method with exact line search. *AIP Conference Proceedings* 1739, 1–8.

[16] Zoutendijk G (1970), Nonlinear programming computational methods. *Integer and Nonlinear Programming* 143, 37–86.

[17] Shoid S, Rivaie M & Mamat M (2016), A modification of classical conjugate gradient method using strong Wolfe line search. *AIP Conference Proceedings* 1739, 1–8.

[18] Shapiee N, Rivaie M & Mamat M (2016), A new classical conjugate gradient coefficient with exact line search. *AIP Conference Proceedings* 1739, 1–8.

[19] Abidin ZZ, Mamat M, Rivaie M & Mohd I (2014), A new steepest descent method. *AIP Conference Proceedings* 1602, 273–278.

[20] Buonaccorsi JP & Elkinton JS (2002), Regression analysis in a spatial-temporal context: Least square, generalized least square and the use of the bootstrap. *Journal of Agricultural, Biological and Environment Statistics* 7, 4–20.

[21] Department of Statistics Malaysia (DOSM) (2017), *LFS, Unemployment rate in Malaysia from 1982-2014*. DOSM.