International Journal of Engineering & Technology, 7 (2.2) (2018) 84-86

International Journal of Engineering & Technology

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

Research Paper

An Optimization of the Autoregressive Model Using the Grid Search Method

Imam Tahyudin [1,2*], Hidetaka Nambo[3], Yoshitaka Goto[4]

[1,3,4] School of Natural Science and Technology, Division of Electrical Engineering and Computer Science, Kanazawa University, Japan
[2] Department of Information System, STMIK AMIKOM Purwokerto, Central Java, Indonesia
**Corresponding author E-mail: imam@blitz.ec.t.kanazawa-u.ac.jp*

Abstract

The purpose of this study is to find the parameters that can produce the best value on the model Autoregressive (AR). The parameter evaluation method used is the Maximum Likelihood Estimator (MLE) and using Grid Search optimization methods. The experimental data used in this study was a sunspot dataset. Based on our analysis, the best Autoregressive model was a 3rd order AR model.

Keywords: *Numerical optimization; Autoregressive model; MLE; Grid search method.*

1. Introduction

Time series is one of the studies in the field of statistics. It is usually expressed as a streak data interval. Time series data are found in a variety of fields such as in the economics field, the unemployment data, the cash flow data of hospital; in the financial field, examples such as the daily stock exchange average, the data distribution of dividends etc.; in the environmental field, for example on daily rainfall data, air quality readings, the phenomenon of El Nino, etc. In the fields of geology, a good example is river level prediction [1, 2]. The purpose of the analysis of time series of which is to describe and summarize a time series data, and to create an appropriate model for making predictions.

The use of time series which is improved by combining with another method has been done by Abdurrahman (2014) about time series algorithms which combined with Particle Swarm optimization method [3]. In addition, another research about PSO base neural network vs traditional classic models for seasonal time series forecasting is conducted by R. Adhikari et al. this research is using three PSO model to optimize two ANN models. The result that the proposed method is better than traditional classic time series model [4, 5]. Next, the paper "Time series prediction using PSO-optimized neural network and hybrid feature selection algorithm for IEEE load data" is presented by M. Sheikhan and N. Mohammadi. This research used two feature selection methods which are Genetic Algorithm and Ant Colony Optimization. And then, to analyze the data using ANN which is optimized by PSO method. The result shows that the proposed method has extremely improved of accuracy by using MAPE (mean absolute percentage error) parameter [6]. Furthermore, in 2015 conducted research on time series analysis through AR modeling. In this study shows various types of AR models such as univariate and multivariate AR models, a radial base function autoregressive model and so on [7].

The aim of this study is to optimize the time series, AR (Autoregressive) model [8-10] at Sunspot Series dataset with the evaluation parameter using maximum likelihood estimation (MLE) and the optimization, using Grid search method.

This research is arranged as follows: Section 2 describes proposed method; Section 3 presents results and discussion; and finally, the conclusion is given in Section 4.

2. Methodology

2.1. AR Model

Association Autoregressive is a form of regression but not linking the dependent variable, but a value connects earlier in time lag (the time interval) are manifold. Therefore, a model Autoregressive will declare a prediction as a function of previous values of a particular time series [11-13]. Autoregressive (AR) model with order p denoted by AR (p). The general form the model AR (p) is 79:

$$X_t = \sum_{r=1}^{p} \emptyset_r X_{t-r} + \epsilon_t \quad (1)$$

Where ∅_1,….,∅_r is a constant value and [8-10] is a sequence of random variables which are independent (uncorrelated) with an average of 0 and variance σ^2.

2.2. Numerical Optimization

2.2.1. Grid Search

The simplest approach to numerical optimization of data is using the Grid search method. For illustration, if we have a model AR (1) and the evaluation parameters using MLE, so the process of log likelihood AR (1) is shown in the following Equation 9.

$$L(\theta) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log\left(\frac{\sigma^2}{(1-\emptyset^2)}\right) - \frac{1}{2}\frac{-\left\{y_1-\left[\frac{c}{(1-\emptyset)}\right]\right\}^2}{\sigma^2} - \frac{T-1}{2}\log(2\pi)$$

Copyright © 2016 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

$$-\frac{T-1}{2}\log(\sigma^2) - \frac{1}{2}\sum_{t=2}^{T}\left[\frac{-\{y_t-[c+\phi y_{t-1}]\}^2}{\sigma^2}\right]$$

Then, the optimization solution using this method is to assume the process average is zero (c = 0) and the variance is 1 (σ^2= 1). Therefore, there is one unknown parameter which is the coefficient of autoregressive. Thus, the likelihood value can be obtained as follows 9.

$$\mathcal{L}(\phi) = \frac{T}{2}log(2\pi) + \frac{1}{2}log(1-\phi^2) - \frac{1}{2}(1-\phi^2)y_1^2 - \frac{1}{2}\sum_{t=2}^{T}(y_t - \phi y_{t-1})^2$$

2.3. Step of Research Analysis

Prepare the sunspot data series which is composed of two variables; variable time and the number of sunspots.

1. Model the sunspot series datasets using the AR model.
2. Evaluation of the AR model parameters using the MLE.
3. Optimize the AR model using the Grid Search method.

3. Results and Discussion

The data used is the benchmark data, i.e. Sunspot dataset series which is the total annual count of sunspots from 1700 to 1988. The number of data points used was 317. Data analysis was performed using a MacBook Pro with the specification (2.7 GHz Intel Core i5, 8 GB 1867 MHz, with DDR3 memory). Running the software package MATLAB R2016b.

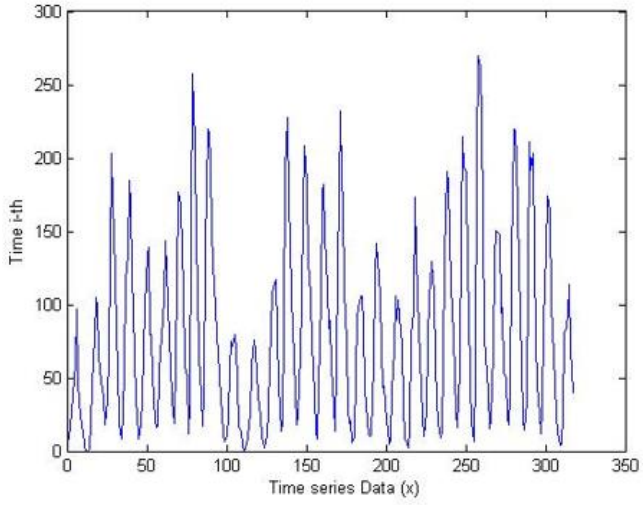


Fig. 1: Sunspot Series Dataset

Figure 1 shows the sunspot frequency from 0 (1700 AD) to 320 (1988 AD) in this time series. The number of sunspots reached its maximum value of 260 at time step 275 while the lowest value reached was 0 at several time steps. The average number of sunspots was 75.

Application of the optimization process used three AR models. The result from the first model is shown in Figure 2. In this figure, the maximum likelihood reached a value of about -1590 at iteration 60 and decrease thereafter. The likelihood estimate started at -2550, gradually rises until it reached the peak at around iteration 60. However, after that, it fell back significantly up to the 100th iteration when the likelihood value was approximately -2450.

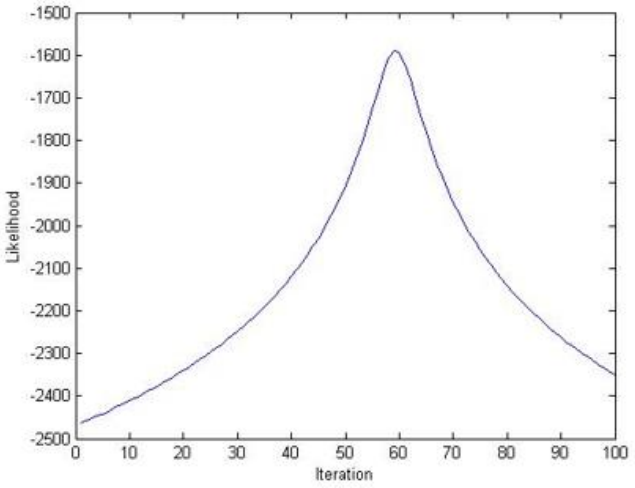


Fig. 2: AR model order 1

Figure 3 shows the likelihood estimate of a second AR model where the likelihood values showed results similar to the first AR model, again the maximum value was achieved after about 60 iterations and the likelihood value of was -1515. However, the results for the third AR model cannot be shown because it has enormous dimensions. Nevertheless, according to the results of MATLAB, the output of third AR model likelihood value is better than the previous AR model, which is about -1508.

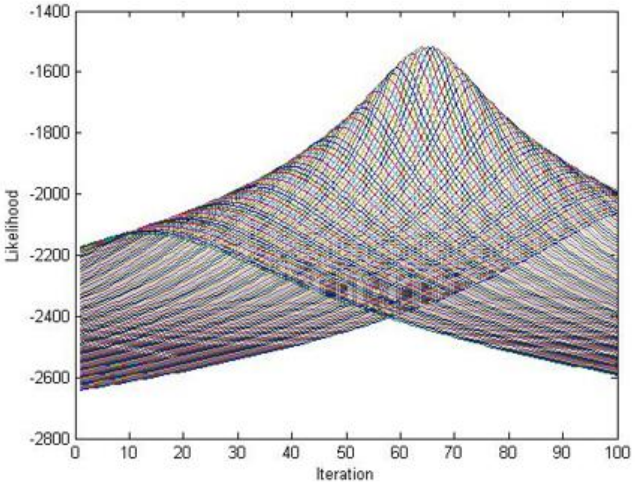


Fig. 3: AR model order 2

Therefore, based on analysis of the results from the first AR model a coefficient values a(1) = 0.9 and the likelihood value of -1589,404 was estimated. The second AR model had coefficients of a(1) = 1.5 and a(2) = -0.6 and the estimated value was -1515.4435. For the third AR model, the coefficient values obtained, respectively were a (1) = 1.5, a(2) = -0.7 and a (3) = 0.1 with a likelihood value of -1508.6877.

Based on the results of the analysis, it can be chosen the best models with the greatest likelihood value can be chosen. Because the third AR model, which had the largest likelihood value (-1508.6877) is most likely the best model.

4. Conclusion

This study analyzed the time series based on AR models using historical sunspot dataset by using MLE method for parameter evaluation and optimization methods using Grid Search. The best result had the parameter values of a (1) = 1.5, a (2) = -0.7, a (3) = 0.1, and the likelihood value was -1508.6877. Therefore, the third order AR model was selected as the best model in this study. This research

can be developed further by comparing this method with the other evaluation and optimization methods such as Steepest Ascent, Newton Raphson, or Particle Swarm Optimization.

References

[1] C. Chatfield. (2002). *Time-Series Forecasting*.

[2] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. (2008). *Time Series Analysis Forecasting and Control 4th Edition*.

[3] H. M. ALBEHADILI, ABDURRAHMAN, and N. E. ISLAM, "An Algorithm for Time Series Prediction Using Particle Swarm Optimization (PSO) " *International Journal of Scientific Knowledge,* vol. 4 No. 6, pp. 26-33, 2014.

[4] R. Adhikari and R. K. Agrawal. (2008). *An Introductory Study on Time Series Modeling and Forecasting*.

[5] R. Adhikari, R. K. Agrawal, and L. Kant, "PSO based Neural Networks vs. Traditional Statistical Models for Seasonal Time Series Forecasting," in *2013 3rd IEEE Int. Adv. Comput. Conf.*, 2013, pp. 719–725.

[6] M. Sheikhalishahi, V. Ebrahimipour, H. Shiri, H. Zaman, and M. Jeihoonian, "A hybrid GA–PSO approach for reliability optimization in redundancy allocation problem," *Int. J. Adv. Manuf. Technol.*, vol. 68, pp. 317–338, 2013.

[7] Haviluddin, R. Alfred, J. H. Obit, M. H. A. Hijazi, and A. A. A. Ibrahim, "A Performance Comparison of Statistical and Machine Learning Techniques in Learning Time Series Data," *Advanced Science Letters*, pp. 3037-3041, 2015.

[8] A. S. Ahmar, "A Comparison of α-Sutte Indicator and ARIMA Methods in Renewable Energy Forecasting in Indonesia," *International Journal of Engineering & Technology*, vol. 7, pp. 9-11, 2018.

[9] A. S. Ahmar, S. Guritno, A. Rahman, I. Minggi, M. Arif Tiro, M. Kasim Aidid, *et al.*, "Modeling Data Containing Outliers using ARIMA Additive Outlier (ARIMA-AO)," *Journal of Physics: Conf. Series*, vol. 954, 2018.

[10] A. Rahman and A. S. Ahmar, "Forecasting of primary energy consumption data in the United States: A comparison between ARIMA and Holter-Winters models," in *AIP Conference Proceedings* vol. 1885, ed, 2017.

[11] S. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting Methods and Applications, 3rd Edition*, 3rd Edition ed. New York: John Wiley, New York, 1998.

[12] R. Fildes and S. Makridakis, "The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting," *Int. Stat. Rev. / Rev. Int. Stat.*, vol. 63, pp. 289–308, 1995.

[13] Haviluddin and A. Jawahir, "Comparing of ARIMA and RBFNN for short-term forecasting," *International Journal of Advances on Intelligent Informatics (IJAIN)*, vol. 1, (1), March 2015, pp. 15-22, 2015.