International Journal of Engineering & Technology, 7 (4.30) (2018) 64-67

International Journal of Engineering & Technology

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

Research paper

Hierarchical Bayesian Estimation for Stationary Autoregressive Models Using Reversible Jump MCMC Algorithm

Suparman[1]*, Mohd Saifullah Rusiman[2]

[1]Department of Mathematics Education, University of Ahmad Dahlan, Indonesia
[2]Universiti Tun Hussein Onn Malaysia, Malaysia
**Corresponding author E-mail:suparman@pmat.uad.ac.id*

Abstract

The autoregressive model is a mathematical model that is often used to model data in different areas of life. If the autoregressive model is matched against the data then the order and coefficients of the autoregressive model are unknown. This paper aims to estimate the order and coefficients of an autoregressive model based on data. The hierarchical Bayesian approach is used to estimate the order and coefficients of the autoregressive model. In the hierarchical Bayesian approach, the order and coefficients of the autoregressive model are assumed to have a prior distribution. The prior distribution is combined with the likelihood function to obtain a posterior distribution. The posterior distribution has a complex shape so that the Bayesian estimator is not analytically determined. The reversible jump Markov Chain Monte Carlo (MCMC) algorithm is proposed to obtain the Bayesian estimator. The performance of the algorithm is tested by using simulated data. The test results show that the algorithm can estimate the order and coefficients of the autoregressive model very well. Research can be further developed by comparing with other existing methods.

Keywords: *Autoregressive model; Hierarchical Bayesian; Reversible jump MCMC.*

1. Introduction

An autoregressive model is a time series model that is often used to model data in different areas of life. The autoregressive model (AR) is a flexible model by setting the order and model parameters. Okada *et al.* [1] used the AR model to diagnose Parkinson's disease. Ramdane-Cherif *et al.* [2] applied the AR model to the eye tremor movement. The eye tremor movement is extracted from the recorded eye position signal. Kisi [3] used the AR model to predict stream flow. Zhao, Morgan, and Davis [4] used the AR model to classify the output from gas chromatography. Lee and Chon [5] used the AR model to model the extraction of respiratory rate. Figueiredo and Figueiras [6] used the AR model to detect damage. Kim, Faloutsos, and Yang [7] used the AR model to predict EEG data. Jayawardhana *et al.* [8] used the AR model to identify structural damage. Zhang, Qi, and Li [9] used the AR model to simulate dynamic light scattering (DLS) signals. Zhao *et al.* [10] used the AR model to predict channels in wireless networks. Dai, Liu, and Zhang [11] applied the AR model to the pre-earthquake ionospheric anomaly analysis. Yuewen *et al.* [12] used the AR model to predict the engine's exhaust gas main engine temperature. The AR model can predict the changing trend of smoke temperature. Song [13] used the AR model to identify the frequency of random signals. Kaewwit, Lursinsap, and Sophatsathit [14] used the AR model to determine the high accuracy of biometric electroencephalography (EEG). Padmavathi and Krishna [15] used the AR model to detect atrial fibration.

Let $x = (x_1, \dots, x_n)$ be n time series data where n denotes a number of observations. This time series is said to have a p-order AR model, written by AR (p), when this time series satisfies the stochastic equation as follows:

$$x_t = z_t + \sum_{i=1}^{p} \phi_i^{(p)} x_{t-i} \quad (1)$$

for i = 1,…, n. The random variable z_t is a random error at time t and z_t is assumed to have a normal distribution with mean 0 and variance σ^2. The vector $\phi^{(p)} = \left(\phi_1^{(p)}, \dots, \phi_p^{(p)}\right)$ denotes the coefficient vector of model AR (p). The AR(p) model is called stationary if and only if the equation $\phi(b) = 1 - \sum_{i=1}^{p} \phi_i^{(p)} b^i$ is zero for value b outside the circle with radius equal to one.

If the AR model is matched against the data, generally the order and the AR model coefficients are unknown. Methods to estimate the AR model order have been proposed by several authors, for example: [1] and [16]. Okada *et al.* [1] used the Akaike information criterion (AIC) to estimate the AR modeling order. Khorshidi and Karimi [16] compared various criteria, namely the final prediction error (FPE) criterion and the AIC, to estimate the AR model order. Likewise, methods for estimating AR model parameters have been proposed by several authors, for example: [16] and [17]. Khorshidi and Karimi [16] used the Least-Squares-Forward (LSF) method to estimate the AR model parameters. Chen *et at.* [17] used Hubor's M-estimation theory to estimate the AR model parameters. But in the various parameter estimation methods that is proposed by the authors, the order model is often assumed to be known.

This paper proposes the estimation of AR order and AR model parameters simultaneously that satisfy the condition for stationarity. The stationary AR model is very useful for forecasting. This paper aims to estimate parameter values $\left(p, \phi^{(p)}, \sigma^2\right)$ of the AR model simultaneously based on observational data $x = (x_1, \dots, x_n)$.

Copyright © 2018 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.

2. Research Method

This research uses hierarchical Bayesian approach. The AR order model, the AR model coefficients, and the error variance are considered as random variables having a certain distribution. This distribution is known as the prior distribution. Determination of the prior distribution for the parameters $\left(p, \phi^{(p)}, \sigma^2\right)$ is done in the following way: The prior distribution of the order p is chosen by the binomial distribution with the parameter λ. The conditional distribution of the coefficient $\phi^{(p)}$ if known p is a uniform distribution at the interval of $(-1,1)^p$. The prior distribution of the error variance σ^2 is the inverted gamma distribution with parameters 1 and $\beta/2$. Hierarchically, the prior distribution of λ is the uniform distribution at the interval (0,1). The prior distribution of β is Jeffrey's distribution. Then the prior distribution of parameters $\left(p, \phi^{(p)}, \sigma^2\right)$ is combined with the probability function of x to obtain the posterior distribution of parameters $\left(p, \phi^{(p)}, \sigma^2\right)$. Let $\pi\left(p, \phi^{(p)}, \sigma^2\right)$ express the prior distribution for parameters $\left(p, \phi^{(p)}, \sigma^2\right)$ and let $f\left(x\middle|p, \phi^{(p)}, \sigma^2\right)$represents the likelihood function for data x, then the posterior distribution for the parameters $\left(p, \phi^{(p)}, \sigma^2\right)$ can be expressed as follows:

$$\pi\left(p, \phi^{(p)}, \sigma^2\middle|x\right) \propto f\left(x\middle|p, \phi^{(p)}, \sigma^2\right)\ \pi\left(p, \phi^{(p)}, \sigma^2\right)$$

The posterior distribution is proportional to the multiplication of likelihood function and prior distribution. Since the order p is not known, the form of posterior distribution is very complicated. The determination of the Bayes estimator cannot be done analytically. Therefore the Bayes estimator is determined using the reversible jump MCMC algorithm [18]. Reversible jump MCMC algorithm allows the transformation from one AR model to another AR model. Transformation is not just from one AR model to another AR model that has the same order, but the transformation from one AR model to another AR model that has a different order. In other words, the transformation is done in a space that has different dimensions. The performance of reversible jump MCMC algorithm is tested using simulated data.

3. Results and Discussion

Let $s = (x_{p+1}, \dots, x_n)$ be the realization of the AR(p) model. If the value $s_0 = (x_1, \dots, x_p)$ is known, the likelihhod function of s can be written more or less as follows:

$$L\left(s\middle|p, \phi^{(p)}, \sigma^2\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{(n-p)/2} \exp - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n} g^2\left(t, p, \phi^{(p)}\right) \quad (2)$$

Where

$$g^2\left(t, p, \phi^{(p)}\right) = x_t - \sum_{i=1}^{p} \phi_i^{(p)} x_{t-i} \quad (3)$$

For $t = p+1, \dots, n$ with initial value $x_1 = \cdots = x_p = 0$. Let S_p be the stationarity region. Let $\phi^{(p)} = \left(\phi_1^{(p)}, \dots, \phi_p^{(p)}\right)$and $r^{(p)} = (r_1, \dots, r_p)$. By using transformation

$$F: \phi^{(p)} \in S_p \rightarrow r^{(p)} \in (-1,1)^p \quad (4)$$

Then the model AR $(x_t)_{t\in Z}$ is stationary if and only if $(r_1, \dots, r_p) \in (-1,1)^p$ [19]. Further likelihood function for x can be rewritten as follows :

$$L\left(s\middle|p, \phi^{(p)}, \sigma^2\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{(n-p)/2} \exp - \frac{1}{2\sigma^2}\sum_{t=p+1}^{n} g^2\left(t, p, F^{-1}\left(\phi^{(p)}\right)\right) \quad (5)$$

3.1. Hierarchical Bayesian estimator

The determination of the prior distribution of the parameters mentioned above is as follows:

a) The prior distribution for order p is binomial distributed with parameter λ

$$\pi(p|\lambda) = C_{p_{max}}^{p} \lambda^p (1-\lambda)^{p_{max}-p} \quad (6)$$

b) For order p is determined, the distribution for coefficient vector $r^{(p)}$ is uniform distribution on the interval $(-1,1)^p$

c) The distribution for variance σ^2 is inverted gamma distribution with parameters α/2 and β/2

$$\pi(\sigma^2|\alpha, \beta) = \frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)} (\sigma^2)^{-(1+\alpha/2)} \exp - \frac{\beta}{2\sigma^2} \quad (7)$$

Here the distribution for hyperparameter λ is uniform distribution on the interval (0,1), the value of α is taken equals 2, and the distribution for parameter β is Jeffrey's distribution. Let $H_1 = \left(p, r^{(p)}, \sigma^2\right)$ and $H_2 = (\lambda, \beta)$. Thus the prior distribution for parameters H_1 and H_2 can be presented as follows:

$$\pi(H_1, H_2) = \pi(p|\lambda)\pi\left(r^{(p)}\middle|p\right)\pi(\sigma^2|\alpha, \beta)\pi(\lambda)\pi(\beta) \quad (8)$$

$$= C_{p_{max}}^{p} \lambda^p (1 - \lambda)^{p_{max}-p} \left(\frac{1}{2}\right)^p \frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)} (\sigma^2)^{-(1+\alpha/2)} \exp - \frac{\beta}{2\sigma^2}\frac{1}{\beta} \quad (9)$$

According to Bayes's Theorem, posterior distributions for parameters H_1 and H_2 can be expressed as

$$\pi(H_1, H_2|s) \propto L(s|H_1)\pi(H_1, H_2)$$

Posterior distribution is a combination of likelihood function and prior distribution that is assumed before the sample is taken. In this case the posterior distribution $\pi(H_1, H_2|s)$ has a very complicated form so that the Bayes estimator cannot be determined by analysis. Therefore the reversible jump MCMC algorithm is proposed to determine the Bayes estimator.

3.2. Reversible jump MCMC algorithm

Let $M = (H_1, H_2)$. In general, the MCMC method is a sampling method by creating a homogeneous Markov chain $M_1, \dots, M_n$ that satisfies aperiodic and irreducible properties such that $M_1, \dots, M_n$ can be considered as a random variable following the distribution $\pi(H_1, H_2|s)$. Thus, the $M_1, \dots, M_n$ can be used to estimate the parameter M. To realize it Gibbs Hybrid algorithm is adopted. It consists of two stages: (1) the distribution simulation of $\pi(H_1|H_2, s)$ and (2) the simulation distribution of $\pi(H_2|H_1, s)$. The Gibbs algorithm [20] is used to simulate the distribution $\pi(H_2|H_1, s)$. The reversible jump MCMC algorithm is used to simulate the distribution $\pi(H_1|H_2, s)$.

The distribution simulation $\pi(H_2|H_1, s)$ is done in the following way: The conditional distribution H_2 is known to H_1 and s can be expressed as

$$\pi(H_2|H_1, s) \propto \lambda^p (1-\lambda)^{p_{max}-p} (\beta/2)^{\alpha/2} \exp - \frac{\beta}{2\sigma^2}\frac{1}{\beta}$$

Since this distribution is a gamma distribution with parameters α/2 and $\frac{1}{2\sigma^2}$, the Gibbs algorithm can be used to simulate the distribution of $\pi(H_2|H_1,s)$.
The distribution simulation $\pi(H_1|H_2,s)$ is done in the following way: The conditional distribution of H_1 if it is known $H_2$and s is

$$\pi\left(p,r^{(p)}\middle|H_1,s\right)=\int_{R^+}\pi(H_1|H_2,s)\,d\sigma^2 \quad (10)$$

Let $v=\frac{\alpha}{2}+\frac{n-p_{max}}{2}$ and $w=\frac{\beta}{2}+\frac{1}{2}\sum_{t=p_{max}+1}^{n}g^2(t,p,F^{-1}(r^{(p)}))$. Since $w=\frac{\beta}{2}+\frac{1}{2}\sum_{t=p_{max}+1}^{n}g^2(t,p,F^{-1}(r^{(p)}))$ then

$$\pi\left(p,r^{(p)}\middle|H_1,s\right)\propto C_{p_{max}}^{p}\lambda^p(1-\lambda)^{p_{max}-p}\left(\frac{1}{2}\right)^p\frac{(\beta/2)^{\alpha/2}}{\Gamma(\alpha/2)}\frac{1}{\beta}\frac{\Gamma(v)}{w^v}$$

On the other hand, the distribution $\pi(H_1|H_2,s)$ can be expressed as the product of the distribution of $\pi\left(p,r^{(p)}\middle|H_1,s\right)$ and the distribution of $\pi\left(\sigma^2\middle|p,r^{(p)},H_2,s\right)$, i.e:

$$(H_1|H_2,s)=\pi\left(p,r^{(p)}\middle|H_1,s\right)\pi\left(\sigma^2\middle|p,r^{(p)},H_2,s\right) \quad (11)$$

Furthermore, to simulate the distribution of $\pi(H_1|H_2,s)$, a hybrid algorithm is used. It consists of two stages: (a) The distribution simulation $\pi\left(\sigma^2\middle|p,r^{(p)},H_2,s\right)$, (b) The distribution simulation $\pi\left(p,r^{(p)}\middle|H_1,s\right)$. Gibbs algorithm is used to simulate the distribution $\pi\left(\sigma^2\middle|p,r^{(p)},H_2,s\right)$.

The distribution simulation $\pi\left(p,r^{(p)}\middle|H_1,s\right)$ is done by using the reversible jump MCMC algorithm. The reversible jump MCMC algorithm uses three types of transformations, namely: birth of the order, death of the order, and change in the coefficient

3.2.1. Birth / death of the order

The birth of the order from the AR(p) model to the AR(p+1) model is done by adding coefficients. Let p be the actual value for the order and $r^{(p)}=(r_1,\ldots.,r_p)$ is the actual value for the AR(p) model coefficient. As in Suparman and Doisy [21], the random variable u is chosen according to the triangular distribution with mean 0

$$g(u)=\begin{cases}u+1, & -1<u<0\\ 1-u, & 0<u<1\end{cases} \quad (12)$$

The coefficient vector $r^{(p)}$ is completed with random variable u, so the proposed new coefficient vector is $r^{(p+1)}=(r_1,\ldots.,r_p,u)$. The acceptance/rejection probability corresponding to the birth order is $\alpha_N=\min\{1,r_N\}$ where

$$r_N=\frac{\pi\left(p+1,r^{(p+1)}\middle|H_2,s\right)}{\pi\left(p,r^{(p)}\middle|H_2,s\right)}\frac{q\left(p+1,r^{(p+1)};p,r^{(p)}\right)}{q\left(p,r^{(p)};p+1,r^{(p+1)}\right)} \quad (13)$$

In contrast, the death of the order from the AR(p+1) model to the AR(p) model is done by removing the last coefficient. Let p+1 be the actual value of the order and $r^{(p+1)}=(r_1,\ldots.,r_p,r_{p+1})$ is the actual value for the AR(p+1) model coefficient. The coefficient r_{p+1} is removed. So the proposed new coefficient vector is $r^{(p)}=(r_1,\ldots.,r_p)$. The probability of acceptance/rejection corresponding to order death is $\alpha_D=\min\{1,r_N^{-1}\}$.

3.2.2. Change of the coefficient

The change of coefficient from AR(p) to AR(p) is done by changing each coefficient. Let $r^{(p)}=(r_1,\ldots.,r_p)$ is the actual value for the coefficients. For i=1,...,p, take the random variable $u_i=sin(r_i+s)$ with s taken according to the uniform distribution on the interval [-π/10,π/10]. So the new coefficient vector is $r^{*(p)}=(r_1^*,\ldots,r_i^*=u_i,\ldots,r_p^*)$. The acceptance/rejection probability corresponding to the coefficient change is $\alpha_C=min\{1,r_C\}$ where

$$r_C=\frac{\pi\left(p,r^{*(p)}\middle|H_2,s\right)}{\pi\left(p,r^{(p)}\middle|H_2,s\right)}\frac{q\left(p,r^{*(p)};p,r^{(p)}\right)}{q\left(p,r^{(p)};p,r^{*(p)}\right)}$$

3.3. Simulation study

The reversible jump MCMC algorithm is used to identify the AR model order and the AR model parameter for the simulated data. A simulation study is conducted to find out whether the performance of the reversible jump MCMC algorithm worked well or not.
To know the performance of reversible jump MCMC algorithm simulation study is conducted. Figure 1 is an AR simulation data made according to the equation (1) with n = 250, order p = 3, $\phi^{(3)}=\left(\phi_1^{(3)}=-0.36,\ \phi_2^{(3)}=-0.24,\ \phi_3^{(3)}=0.81\right)$and $\sigma^2=4$.

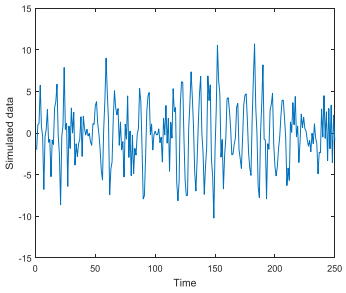

Fig 1. : Simulated Data

The reversible jump MCMC algorithm is implemented in this simulation data to estimate the AR model order, AR model coefficients, and error variance. Figure 2 shows the histogram of the AR model order.

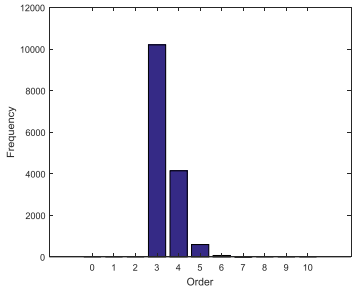

Fig 2.: Histogram of the AR Order

Figure 2 shows that the mode of AR order is reached in order 3. This means that the estimator for AR order is p = 3. After it is determined that the most suitable AR model is AR (3) then the estimator for the AR coefficient and corresponding error variance is determined, i.e:
$\hat{\phi}^{(3)}=\left(\hat{\phi}_1^{(3)}=-0.36,\quad \hat{\phi}_2^{(3)}=-0.26,\quad \hat{\phi}_3^{(3)}=0.82\right)$ and $\hat{\sigma}^2=3.79$. Table 1 summarizes the comparison between AR order estimators, AR coefficient estimators, and error variance estimators with AR-order values, AR coefficients, and error variance.

Table 1: Comparison between the value of parameters and the value of estimators

Value of Parameters	Value of Estimators
p = 3	$\hat{p}=3$
$\phi^{(3)}=(-0.36,-0.24,0.81)$	$\hat{\phi}^{(3)}=(-0.36,-0.26,0.82)$
$\sigma^2=4$	$\hat{\sigma}^2=3.79$

Table 1 shows that the reversible jump MCMC algorithm can estimate the AR model order, AR model coefficients, variance error very well.

4. Conclusion

The above description is a review of the theory of the reversible jump MCMC algorithm to estimate the order of AR model, AR model coefficient, and error variance. The simulation study shows that the algorithm can estimate AR model parameters very well. The proposed algorithm has the advantage that the resulting estimation is an AR model that verifies the condition of the stationarity. Another advantage is that the algorithm can estimate parameters $\left(p, \phi^{(p)}, \sigma^2\right)$ simultaneously.

Research can be further developed in comparison with existing estimation methods to determine effectiveness. Research may also be developed on the replacement of assumptions for errors, such as AR models with not normally distributed errors.

Acknowledgement

We would like to thank to University of Ahmad Dahlan, Indonesia who has provided a grant to present this paper at the 4th International Conference on the Application of Science and Mathematics in Malaysia.

References

[1] Okada K, Hando S, Teranishi M, Matsumoto Y & Fukumoto I (2001), Analysis of pathological tremors using the autoregression model. Frontiers Med. Biol. Engng 11(3), pp. 221-235.

[2] Ramdane-Cherif Z, Nait-Ali A, Motsch JF & Krebs MO (2004), An autoregressive (AR) model applied to eye tremor movement, clinical application in Schizophrenia. *Journal of Medical Systems*, 28(5), pp. 489-495.

[3] Kisi O (2005), Daily river forecasting using artificial neural networks and auto-regressive models. *Turkish J. Eng. Env. Sci* 29, pp. 9-20.

[4] Zhao W, Morgan JT & Davis CE (2008), Gas chromatography data classification based on complex coefficients of an autoregressive model. *Journal of sensors*, pp. 1-9.

[5] Lee J & Chon K H (2010), Respiratory rate extraction via an autoregressive model using the optimal parameter search criterion. *Annals of Biomedical Engineering* 38(10), pp. 3218-3225.

[6] Figueiredo E & Figueiras J (2011), Influence of the autoregressive model order on damage detection. *Computer-Aided Civil and Infrastructure Engineering* 26, pp. 225-238.

[7] Kim SH, Faloutsos C & Yang H-J (2013), Coercively adjusted autoregression model for forecasting in epilepsy. *EEG Computational and Mathematical Methods in Medicine*, pp 1-12.

[8] Jayawardhana M, Zhu X, Liyanapathirana R & Gunawardana U (2015), Statistical damage sensitive feature for structural damage detection using AR model coefficients. *Advances in Structural Engineering* 18(10), pp. 1551-1562.

[9] Zhang Y, Qi X & Li Q (2014), Simulation of dynamic light scattering based on AR model. *Applied Mechanics and Materials* 571-572, pp 840-844.

[10] Zhao N, Yu F R, Sun H, Yin H, Nallanathan A & Wang G (2015), Interference alignment with delayed channel state information and dynamic AR-model channel prediction in wireless networks. *Wireless Netw* 21, pp. 1227-1242.

[11] Dai X, Liu J & Zhang H (2015). Application of AR model in the analysis of pre-earthquake Ionospheric anomalies. *Mathematical Problem in Engineering*, pp. 157-184.

[12] Yuewen Z, Yongjiu Z, Zhufeng L & Peng Z (2015), Prediction of ship main engine exhaust gas temperature using AR model. *Applied Mechanics and Materials* 697, pp. 244-248.

[13] Song C (2016), Random signal frequency identification based on AR model spectral estimation. *International Journal on Smart and Intelligent Systems* 9(2).

[14] Kaewwit C, Lursinsap C & Sophatsathit P (2017), High accuracy EEG biometrics identification using ICA and AR model. *Journal of ICT* 16(2), pp. 354-373.

[15] Padmavathi K & Krishna K S R (2015), Detection of atrial fibrillation using autoregressive modeling International. *Journal of Electrical and Computer Engineering* 5(1), pp. 64-70.

[16] Khorshidi S & Karimi M (2009), Finite sample FPE and AIC criteria for autoregressive model order selection using same-realization predictions. *Journal on Advances in Signal Processing*, pp. 1-7.

[17] Chen H, Yang J, Liu C & Wang J (2014), Parameters estimation for autoregressive process. *Applied Mechanics and Materials* 543-547, pp. 1711-1716.

[18] Green P J (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* 82(4), pp. 711-732.

[19] Barndorff-Nielsen O & Schou G (1973), On the parametrization of autoregressive models by partial autocorrelation. *Journal of Multivariate Analysis* 3, pp. 408-419.

[20] Geman S & Geman D (1984), Stochastic relation, Gibbs distribution and the Bayesian restoration of image. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 6, pp. 721-741.

[21] Suparman & Doisy M (2014), Hierarchical Bayesian of ARMA models using simulated annealing algorithm. *Telkomnika* 12(1) pp. 87-96.