Thailand Statistician
October 2022; 20(4): 892-904
http://statassoc.or.th
Contributed paper

A Comparison of Parameter Estimation Methods for the First-Order of Random Coefficient Autoregressive Model

Autcha Araveeporn*
Department of Statistics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.
*Corresponding author; e-mail: autcha.ar@kmitl.ac.th

Received: 8 May 2021
Revised: 29 October 2021
Accepted: 3 December 2021

Abstract

This paper compares the least-squares, maximum likelihood, and Bayesian methods for estimating an unknown parameter in the random coefficient autoregressive (RCA) model. The RCA model depends on the random coefficients and the time series data in terms of the autoregressive model. We mention estimating unknown parameters by using least-squares, maximum likelihood, and Bayesian methods. We concentrate on only the first-order models of the RCA model depending on the unknown parameter under time series data. The least-squares method is a widely used method by minimizing the sum of squared residuals and differential concerning unknown parameters. Next, the maximum likelihood method is another method that is well-known and often used for estimating parameters based on the likelihood function and observed data. Finally, the Bayesian method carries out Markov chain Monte Carlo (MCMC) method to generate samples from a posterior distribution, which, after being averaged, give the estimated value of the unknown parameter. We use a Gibbs sampling algorithm in our MCMC calculation. The efficiency of the three methods is to compare according to the average mean square error for simulation data. The least-squares method performs better than the maximum likelihood and Bayesian method except for the trend data for simulation data. The average mean square error of the least-squares method shows the minimum values that indicated their performance in most cases. Lastly, we try these methods with the series of days of the gold price per one-baht weight on one year as actual data. The result shows that the least-squares method still worked better than the maximum likelihood and Bayesian method, similar to the simulation of test data.

Keywords: Bayesian method, least-squares method, maximum likelihood method.

1. Introduction

In finance, the data are collected in the form of time series data applied for modeling from the past data and forecasting to the future data. The time series data are the ordered sequence of observations taken at regular intervals such as daily exchange rate, weekly stock prices, monthly oil consumptions, and annual growth rates. Usually, the time series data exhibits changing data as a trend, volatility, stationary, nonstationary, and random walk, especially when the time-series data are systematically collected over a long period. The time series models are beneficial models when users have serially correlated data. However, it is also one of the time series models, which helps forecast the values in future time.

A model widely used to fit the stationary data such as the autoregressive (AR) model, moving average (MA) model, and autoregressive moving average (ARMA) model. For nonstationary time series data, the autoregressive integrated moving average (ARIMA) model can use when the data has the trend. These models have some problems with over-specify the model and estimating the integration parameter. An alternative way to model by volatility is to use the conditional heteroscedastic autoregressive moving average (CHARMA) model (Tsay 1987). Nicholls and Quinn (1982) studied the random coefficient autoregressive (RCA) model. It is the class of autoregressive models whose coefficients are random. This model has unique among the non-linear model and time series data in that their analysis is quite tractable. It has been possible to find conditions stationarity and stability, to derive estimates of the unknown parameter. Aue et al. (2006) proposed the quasi-maximum likelihood method to estimate parameters of the RCA model of order one that derived the strong consistency and the asymptotic normality of the estimators. Wang and Ghosh (2008) used the Bayesian estimation and unit root test of the first-order estimate on the RCA model. Benmoumen et al. (2019) elaborated an algorithm to estimate p-order random coefficient autoregressive model parameters. This algorithm combines the quasi-maximum likelihood method, the Kalman filter, and the simulated annealing method.

The statistical principles associated with the least-squares and maximum likelihood methods underlie the parameter estimation. The least-squares method is the most popular method used to determine the best fit parameter by minimizing the sum of squares of the deviations of the values, such as the regression analysis. The least-squares method can apply to the time series model found in the paper of Khall and Moraes (2012), which used the linear least square method for time series analysis with application to the methane time series data. In addition, the maximum likelihood method involves defining a likelihood function of observing the data sample given the probability function. The modeling of time series can apply the maximum likelihood method for estimating parameters and forecasting future value. Andrews and Davis (2006) proposed the maximum likelihood estimation for autoregressive-moving average model in which all roots of the autoregressive polynomial are reciprocals of roots of the moving average polynomial and vice versa. To estimate an unknown parameter, the least-square and maximum likelihood methods can solve to estimate the parameter of several models depending on the past and present data for time series modeling.

Furthermore, the Bayesian method is treated the random variables distributed according to the probability distribution for estimating unknown parameter which relates on the prior distribution. From Bayes' theorem, the posterior distribution depends on the prior distribution and probability distribution of random variables or a hierarchical model. However, it is hard to see what the posterior distribution might look like, and it is impossible to solve the analytical problem because it contains several parameters.

This study is interested in estimating the RCA model parameter based on the least-squares, maximum likelihood, and Bayesian methods. The performance of these methods looks at the minimum of the average of the mean square error (MSE) method for several situations of simulated data and MSE for actual data.

2. The RCA Model

The general class of random coefficient autoregressive (RCA) model of order p, that is given by

$$x_t = \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-1} + \varepsilon_t, t = 2,3,...,n. \quad (1)$$

Wang and Ghosh (2008) suggested $\underline{\beta}_{ti} = \underline{\mu}_\beta + \Omega_\beta \underline{u}_t$, where α is the scalar of constant, $\underline{\beta}_{ti} = \underline{\mu}_\beta + \Omega_\beta \underline{u}_t$, is a sequence of independent random vectors with mean $\underline{\mu}_\beta = (\mu_{t1}, \mu_{t2},...,\mu_{tp})'$ and covariance matrix Ω_β. It is assumed that ε_t's are the sequence of i.i.d. (independent and identically distributed random variables) from distribution mean zero and unit variance.

In this paper, we focus on the simplicity case study of the first order on RCA(1) following

$$x_t = \alpha + \beta_t x_{t-1} + \varepsilon_t, \; t = 2,3,...,n, \text{ and } \beta_t = \mu_\beta + \sigma_\beta u_t, \quad (2)$$

where β_t's are i.i.d. random variables with mean μ_β, and variance σ_β^2, ε_t's are i.i.d. random variables with mean zero and variance σ_ε^2. The RCA(1) model can be rewritten as

$$x_t = \alpha + \beta_t x_{t-1} + \varepsilon_t = \alpha + \mu_\beta x_{t-1} + u_t, \quad (3)$$

where $u_t = \sigma_\beta v_t x_{t-1} + \varepsilon_t$, then v_t is a random variable with mean zero and variance one and independent of ε_t.

3. Method of Parameter Estimation

To estimate the parameter of the RCA(1) model, we propose the concept of least-squares criterion, maximum likelihood method, and Bayesian method based on the MCMC method.

3.1. Least-squares method

The first estimated method, we propose the least-squares criterion to estimate parameter $\theta = (\alpha, \mu_\beta)$ by minimizing sum of squared residuals. Let γ_t be the information set up to time t, and $u_t = x_t - \mu_\beta x_{t-1}$, then it can see that $E(u_t | \gamma_{t-1}) = 0$, $E(u_t^2 | \gamma_{t-1}) = \sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2$, and $Var(u_t | \gamma_{t-1}) = \sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2$.

Given a sample $x_1, x_2, \ldots, x_n$, the parameter $\theta = (\alpha, \mu_\beta)$ is to estimate by minimizing $\sum_{t=1}^{n} u_t^2$ with respect to $\theta = (\alpha, \mu_\beta)$, thus the least-squares estimator $\hat{\theta}_{LS} = (\hat{\alpha}_{LS}, \hat{\mu}_{\beta,LS})$ is given by

$$\sum_{t=2}^{n} (u_t)^2 = \sum_{t=2}^{n} \left(x_t - \alpha - \mu_\beta x_{t-1} \right)^2. \quad (4)$$

Differential concerning parameter $\hat{\alpha}_{LS}$, $\hat{\mu}_{\beta,LS}$:

$$\frac{\partial}{\partial \alpha} \sum_{t=1}^{n} (u_t)^2 = \frac{\partial}{\partial \alpha} \sum_{t=1}^{n} \left(x_t - \alpha - \mu_\beta x_{t-1} \right)^2 = 0,$$

and
$$\frac{\partial}{\partial \mu_\beta} \sum_{t=1}^{n} (u_t)^2 = \frac{\partial}{\partial \mu_\beta} \sum_{t=1}^{n} \left(x_t - \alpha - \mu_\beta x_{t-1} \right)^2 = 0.$$

Then, we get

$$\hat{\alpha}_{LS} = \frac{\sum_{t=1}^{n} x_t}{n} - \mu_\beta \frac{\sum_{t=2}^{n} x_{t-1}}{n}, \tag{5}$$

and
$$\hat{\mu}_{\beta,LS} = \frac{\sum_{t=2}^{n} x_t x_{t-1} - \alpha \sum_{t=2}^{n} x_{t-1}}{\sum_{t=2}^{n} x_{t-1}^2}. \tag{6}$$

From (6), let us replace in (5) and the solution of $\hat{\alpha}_{LS}$ is

$$\hat{\alpha}_{LS} = \frac{\sum_{t=2}^{n} x_{t-1}^2 \sum_{t=1}^{n} x_t - \sum_{t=2}^{n} x_t x_{t-1} \sum_{t=2}^{n} x_{t-1}}{\left(n \sum_{t=2}^{n} x_{t-1}^2 - \left(\sum_{t=2}^{n} x_{t-1} \right)^2 \right)},$$

or $\hat{\mu}_{\beta,LS}$ can rewrite as

$$\hat{\mu}_{\beta,LS} = \frac{n \sum_{t=2}^{n} x_t x_{t-1} - \sum_{t=1}^{n} x_t \sum_{t=2}^{n} x_{t-1}}{\left(n \sum_{t=2}^{n} x_{t-1}^2 - \left(\sum_{t=2}^{n} x_{t-1} \right)^2 \right)}.$$

For RCA(1) model, it can be fitted model as $\hat{x}_t = \hat{\alpha}_{LS} + \hat{\mu}_{\beta,LS} x_{t-1}$ $, t = 2, 3, \ldots n.$

3.2. Maximum likelihood method

The maximum likelihood method extends from the least-squares method by considering the observations $\{x_1, \ldots, x_n\}$ and the probability distribution function. From (2), we know that $x_t = \alpha + \beta_t x_{t-1} + \varepsilon_t$. Then we assume that the observed data is the normal distribution, then $E(x_t \mid x_{t-1}) = \alpha + \mu_\beta x_{t-1}$ and $Var(x_t \mid x_{t-1}) = \sigma_\beta^2 x_{t-1} + \sigma_\varepsilon^2$. Now, the likelihood function is defined by

$$L(\theta) = L(\theta \mid x_t x_{t-1}) = \prod_{t=2}^{n} f(x_t \mid x_{t-1}) = \left(\frac{1}{2\pi} \right)^{n/2} \prod_{t=2}^{n} \left(\sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2 \right)^{-1/2} \exp\left\{ -\frac{1}{2} \sum_{t=2}^{n} \frac{(x_t - \alpha - \mu_\beta x_{t-1})^2}{\sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2} \right\}.$$

It is more convenient to work by setting parameter $\nu = \frac{\sigma_\beta^2}{\sigma_\varepsilon^2}$, so it can be written by

$$L(\theta) = \left(\frac{1}{2\pi} \right)^{n/2} \sigma^2 \prod_{t=2}^{n} \left(1 + \nu x_{t-1}^2 \right)^{-1/2} \exp\left\{ -\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^{n} \frac{(x_t - \alpha - \mu_\beta x_{t-1})^2}{1 + \nu x_{t-1}^2} \right\}. \tag{7}$$

Take ln on likelihood function following;

$$\ln L(\theta) = -\frac{n}{2} \ln(2\pi) - \frac{\sigma_\varepsilon^2}{2} \sum_{t=2}^{n} \ln(1 + \nu x_{t-1}^2) - \frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^{n} \frac{(x_t - \alpha - \mu_\beta x_{t-1})^2}{1 + \nu x_{t-1}^2}. \tag{8}$$

Differential with respect to parameter α and μ_β,

$$\frac{\partial}{\partial \alpha} \ln L(\theta) = -\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^{n} \frac{(x_t - \alpha - \mu_\beta x_{t-1})}{1 + \nu x_{t-1}^2}, \quad \frac{\partial}{\partial \mu_\beta} \ln L(\theta) = -\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^{n} \frac{(x_t - \alpha - \mu_\beta x_{t-1}) x_{t-1}}{1 + \nu x_{t-1}^2}.$$

Let $\dfrac{\partial}{\partial(\alpha,\mu_\beta)}\ln L(\theta)=0.$ The $\hat{\alpha}$ is approximated by

$$\frac{\partial}{\partial\alpha}\ln L(\theta)=\sum_{t=2}^{n}\frac{x_t}{1+\nu x_{t-1}^2}-\alpha\sum_{t=2}^{n}\frac{1}{1+\nu x_{t-1}^2}-\mu_\beta\sum_{t=2}^{n}\frac{x_{t-1}}{1+\nu x_{t-1}^2}=0,$$

then

$$\hat{\alpha}=\frac{\sum_{t=2}^{n}\frac{x_t}{1+\nu_{t-1}^2}-\mu_\beta\sum_{t=2}^{n}\frac{x_{t-1}}{1+\nu x_{t-1}^2}}{\sum_{t=2}^{n}\frac{1}{1+\nu x_{t-1}^2}}. \tag{9}$$

Similarly, the $\hat{\mu}_\beta$ is approximated by

$$\frac{\partial}{\partial\mu_\beta}\ln L(\theta)=\sum_{t=2}^{n}\frac{x_t\,x_{t-1}}{1+\nu x_{t-1}^2}-\alpha\sum_{t=2}^{n}\frac{x_{t-1}}{1+\nu x_{t-1}^2}-\mu_\beta\sum_{t=2}^{n}\frac{x_{t-1}^2}{1+\nu x_{t-1}^2}=0,$$

then

$$\hat{\mu}_\beta=\frac{\sum_{t=2}^{n}\frac{x_t x_{t-1}}{1+\nu x_{t-1}^2}-\alpha\sum_{t=2}^{n}\frac{x_{t-1}}{1+\nu x_{t-1}^2}}{\sum_{t=2}^{n}\frac{x_{t-1}^2}{1+\nu x_{t-1}^2}}. \tag{10}$$

From (9) and (10), it can be rewritten as

$$\hat{\alpha}=\frac{c_1-\hat{\mu}_\beta c_2}{c_3} \text{ and } \hat{\mu}_\beta=\frac{c_4-\hat{\alpha}\ c_2}{c_5},$$

where

$$c_1=\sum_{t=2}^{n}\frac{x_t}{1+\nu x_{t-1}^2},\ c_2=\sum_{t=2}^{n}\frac{x_{t-1}}{1+\nu x_{t-1}^2},\ c_3=\sum_{t=2}^{n}\frac{1}{1+\nu x_{t-1}^2},\ c_4=\sum_{t=2}^{n}\frac{x_{t-1}^2}{1+\nu x_{t-1}^2},\ c_5=\sum_{t=2}^{n}\frac{x_t x_{t-1}}{1+\nu x_{t-1}^2}.$$

Finally, we obtain the estimator, $\hat{\alpha}_{ML}=\dfrac{c_1c_5-c_2c_4}{c_3c_5-c_2^2}$ and $\hat{\mu}_{\beta,ML}=\dfrac{c_3c_5-c_1c_2}{c_3c_4-c_2^2}$. For observed fitting values of RCA(1) model, it can be written as $\hat{x}_t=\hat{\alpha}_{ML}+\hat{\mu}_{\beta,ML}x_{t-1}$, $t=2,3,...n.$

3.3. Bayesian method

For the Bayesian method, the Markov chain Monte Carlo (MCMC) method allows estimating the shape of the posterior distribution. Morton and Finkenstadt (2005) used the Markov chain Monte Carlo method to model the susceptible-infected-recovered model for infectious disease. For convenience, the Bayesian estimator has computed the parameter on the model from the mean of posterior distribution by the MCMC method (Gilks et al. 1996). We also carry out the Gibbs sampling algorithm (Geman and Geman 1984) from the MCMC method by rjags package of R program to estimate an unknown parameter.

In Bayesian estimation for RCA(1) model, we proposed a three-level hierarchical model. At the first level is the conditional distribution of the data x_t's given the observed random variables x_{t-1}, coefficient α,β_t, and σ_ε^2. The second level consists of the conditional distribution β_t given the parameter μ_β and σ_β^2. Finally the last level shows the prior distribution of $\theta=\left(\alpha,\mu_\beta\right)^\top$.

Consequently, given the sample variables $x_1, x_2, \ldots, x_n,$ we are able to express the RCA(1) model in the following hierarchical structure,

$$x_t \mid x_{t-1}, \alpha, \beta_t, \sigma_\varepsilon^2 \sim N\left(\alpha + \beta_t x_{t-1}, \sigma_\varepsilon^2\right), \ \beta_t \mid \mu_\beta, \sigma_\beta^2 \sim N\left(\mu_\beta, \sigma_\beta^2\right), \ \left(\alpha, \mu_\beta\right) \sim p\left(\alpha, \mu_\beta\right), \tag{11}$$

where $p(\cdot)$ is the prior density of θ which reflects our prior about the unknown parameters. Following (11), we can express the likelihood function of θ as,

$$L\left(\alpha, \mu_\beta \mid x_1, x_2, \ldots, x_n, \varepsilon_1, \ldots, \varepsilon_n\right) = \phi\left(x_1; \alpha, \sigma_\varepsilon\right) \prod_{i=2}^{n} \phi\left(x_t; \alpha + \beta_t x_{t-1}, \sqrt{\sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2}\right), \tag{12}$$

where $\phi(x; \mu, \sigma)$ denotes the density function of a normal distribution with mean μ and standard deviation $\sigma.$ Therefore, the joint posterior density of the parameters is given by,

$$f\left(\theta \mid x_1, x_2, \ldots, x_n\right) \propto L\left(\theta \mid x_1, x_2, \ldots, x_n\right) p(\theta),$$

where $p(\theta)$ is a prior density of $\theta.$ For the parameter estimation of RCA(1) model, the prior distribution of $\theta = \left(\alpha, \mu_\beta\right)^\top$ are considered to be a continuous random variable in the set of real numbers following normal distribution. From the hierarchical structure in (12), the joint posterior density can be written as

$$f\left(\alpha \mid x_1, x_2, \ldots, x_n\right) \propto \int f\left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2 \mid x_1, x_2, \ldots, x_n\right) d\mu_\beta \, d\sigma_\beta^2 \, d\sigma_\varepsilon^2,$$

and

$$f\left(\mu_\beta \mid x_1, x_2, \ldots, x_n\right) \propto \int f\left(\alpha, \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2 \mid x_1, x_2, \ldots, x_n\right) d\alpha \, d\sigma_\beta^2 \, d\sigma_\varepsilon^2.$$

The joint posterior distribution of α and $\mu_\beta,$ via the normal distribution so that features of distribution are often possible to specify lower and upper bounds for these parameters which can be accurately determined.

To manage Bayesian analysis for RCA(1), we interested in the properties of the density of $\theta = \left(\alpha, \mu_\beta\right)^\top$. Deriving the joint posterior density for θ amounts to integrating out the unobserved coefficients α and $\mu_\beta.$ We can perform the likelihood function to obtain posterior estimator. The complicated likelihood function, we used the so-called Markov chain Monte Carlo (MCMC) method (Gilk el al. 1996) to generate samples from the posterior distribution of $\theta = \left(\alpha, \mu_\beta\right)^\top$. We will carry out the Gibbs sampler (Gelfand et al. 1990), a widely used MCMC method, to obtain the parameter from the posterior distribution using the software R on package rjags.

The condition densities of parameter in RCA(1) model is $f\left(\alpha \mid \mu_\beta, \sigma_\beta^2, \sigma_\varepsilon^2, \underline{x}\right),$ and $f\left(\mu_\beta \mid \alpha, \sigma_\beta^2, \sigma_\varepsilon^2, \underline{x}\right)$ as the full conditional densities of α and $\mu_\beta,$ respectively. The Gibbs sampling algorithm is

1. Initialize $\alpha^{(0)}$ and $\mu_\beta^{(0)},$ for $k = 1, 2, \ldots, m + M.$
2. Draw $\alpha^{(k)}$ from $f\left(\alpha \mid \mu_\beta^{(k-1)}, \sigma_\beta^{2(k-1)}, \sigma_\varepsilon^{2(k-1)}, \underline{x}\right).$
3. Draw $\mu_\beta^{(k)}$ from $f\left(\mu_\beta \mid \alpha^{(k)}, \sigma_\beta^{2(k-1)}, \sigma_\varepsilon^{2(k-1)}, \underline{x}\right).$

where $m = 2{,}000$ are burn-in and $M = 5{,}000$ are the number of samples generated after burn-in. Repeating the above sampling steps, we obtain a discrete-time Markov chain $\left\{\left(\alpha^{(k)}, \mu_\beta^{(k)}\right); k = 1, 2, \ldots\right\}$ whose stationary distribution is the joint posterior density of the parameters.

A Markov chain of parameters of α and μ_β, are constructed by computing the mean sampling from the joint posterior density as standard distribution following:

$$\hat{\alpha}_{Bayes} = \frac{\sum_{k=1}^{M} \alpha^{(k)}}{M} \quad \text{and} \quad \hat{\mu}_{\beta,Bayes} = \frac{\sum_{k=1}^{M} \mu_\beta^{(k)}}{M}.$$

For observed fitting values of RCA(1) model, it can be written as

$$\hat{x}_t = \hat{\alpha}_{Bayes} + \hat{\mu}_{\beta,Bayes} x_{t-1} \quad, t = 2, 3, \ldots n.$$

4. Simulation Study

This study's objective is to estimate parameters $\theta = (\alpha, \mu_\beta)$ from RCA(1) by using the least-squares, maximum likelihood, and Bayesian methods. The results have shown to compare the average estimators in the sample sizes 100, 300, and 500. The mean square error (MSE) evaluates the difference between the estimated values and simulated values. We also computed the MSE as the criterion defined following:

$$MSE = \frac{\sum_{t=2}^{n} (x_t - \hat{x}_t)^2}{n-1},$$

where x_t denotes the simulated values, and $\hat{x}_t$ denotes the estimated values. The simulation study is divided in three parts. At the first process, we generated data $x_t, t = 1, 2, \ldots, n$ from RCA(1) as

$$x_t = \alpha + \mu_\beta x_{t-1} + u_t, \quad u_t \sim N(0, \sigma_u^2), \quad \sigma_u^2 = \sigma_\varepsilon^2 + \sigma_\beta^2 x_{t-1}^2.$$

The parameter of RCA(1) model is mentioned on four cases as;

1. $\alpha = 0.5,\ \mu_\beta = 1, \sigma_\varepsilon^2 = 1, \sigma_\beta^2 = 0,$
2. $\alpha = 0,\ \mu_\beta = 0.99, \sigma_\varepsilon^2 = 1, \sigma_\beta^2 = 0.01,$
3. $\alpha = 0,\ \mu_\beta = 0.01, \sigma_\varepsilon^2 = 1, \sigma_\beta^2 = 0.99,$
4. $\alpha = 0,\ \mu_\beta = 0.6, \sigma_\varepsilon^2 = 1, \sigma_\beta^2 = 0.4.$

The goal of the four cases is to see the different data depending on the different parameters, and the type of data is essential to fit the parameter of these methods. Figures 1-3 depict the 100, 300, and 500 sample sizes.

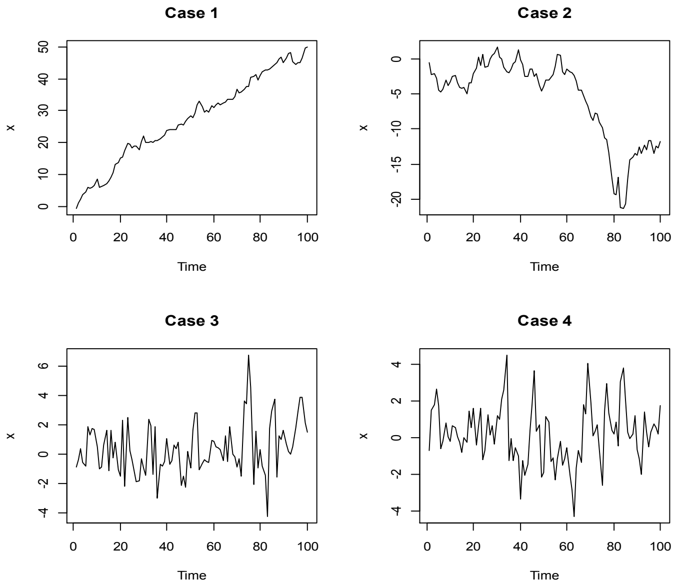

Figure 1 The time series plot for generated data (100 sample sizes)

Figure 1 shows the generating data of 100 sample sizes. It should note that Case 1 is the trend data, Case 2 displays as the random walk, and Cases 3 and 4 also tend to oscillate around their mean zero.

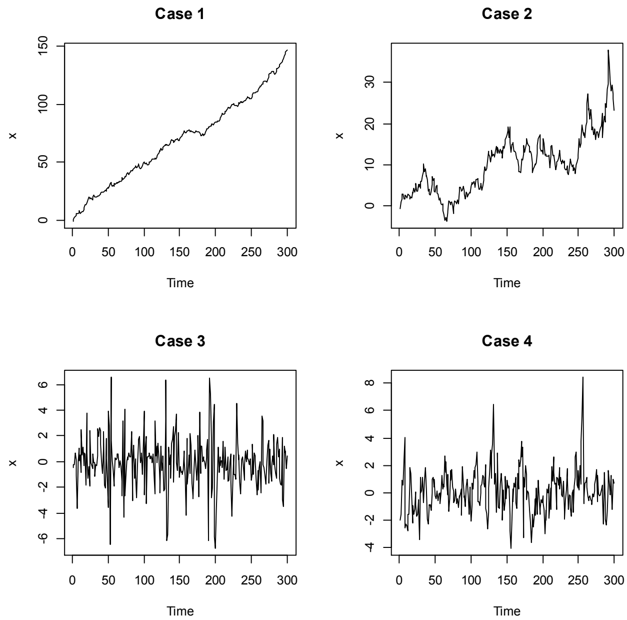

Figure 2 The time series plot for generated data (300 sample sizes)

Also, the generated data of 300 sample sizes provide the trend and random walk process in Cases 1 and 2, but the stationary process is displayed in Cases 3 and 4 in Figure 2.

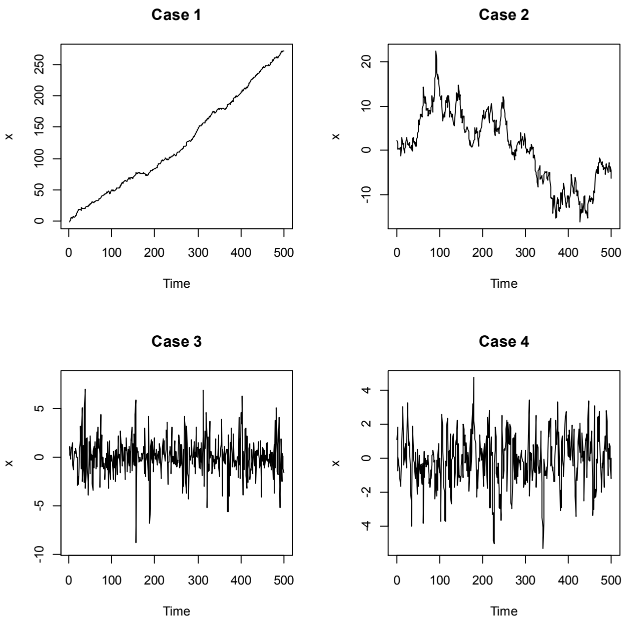

Figure 3 The time series plot for generated data (500 sample sizes)

The impact of the large sample sizes (500) also identified the oscillation in Cases 3 and 4 provided in Figure 3.

In the second part, we obtain the estimator $\hat{\theta}_{LS} = (\hat{\alpha}_{LS}, \hat{\mu}_{\beta,LS})$ from least-squares method, $\hat{\theta}_{ML} = (\hat{\alpha}_{ML}, \hat{\mu}_{\beta,ML})$ from maximum likelihood method and $\hat{\theta}_{Bayes} = (\hat{\alpha}_{Bayes}, \hat{\mu}_{\beta,Bayes})$ from Bayesian method based on MCMC method. From least-squares method, we get x_t from $\hat{x}_t = \hat{\alpha}_{LS} + \hat{\mu}_{\beta,LS} x_{t-1}$, $t = 2,3,...,n,$ The estimated values from maximum likelihood method are computed by and x_t from Bayesian analysis is approximated by $\hat{x}_t = \hat{\alpha}_{ML} + \hat{\mu}_{\beta,ML} x_{t-1}$, $t = 2,3,...n,$ and $\hat{x}_t = \hat{\alpha}_{Bayes} + \hat{\mu}_{\beta,Bayes} x_{t-1}$, $t = 2,3,...,n.$

Finally part, we simulated data at 500 replications from RCA(1). We also compute the Monte Carlo mean and standard deviation for each parameter and sample size. To see this, we compute the average of MSE (AMSE) to compare the valuable estimation between the least-squares, maximum likelihood, and Bayesian methods, as shown in Tables 1-3.

From Table 1, the AMSE of the least-squares method shows the minimum values for Cases 2, 3, and 4. When the sample sizes are increasing, the standard deviation of the parameter is decreasing too. Meanwhile, the absolute standard deviation of the least-squares method presents the lowest values; thus, this method is consistent for estimating parameters.

Table 1 The mean (sd) and AMSE of parameter estimation on least-squares method

Case	Parameter	Sample sizes		
		$n = 100$	$n = 300$	$n = 500$
1	$\alpha = 0.5$	0.5672	0.5177	0.5098
		(0.1932)	(0.1078)	(0.0884)
	$\mu_\beta = 1$	0.9979	0.9997	0.9999
		(0.0071)	(0.0013)	(0.0006)
	AMSE	0.9889	0.9943	0.9950
2	$\alpha = 0$	0.0115	–0.0035	0.0007
		(0.3444)	(0.1525)	(0.1079)
	$\mu_\beta = 0.99$	0.9322	0.9690	0.9763
		(0.0474)	(0.0189)	(0.0126)
	AMSE	**1.3587**	**1.5577**	**1.7165**
3	$\alpha = 0$	0.0063	–0.0014	0.0003
		(0.1092)	(0.0583)	(0.0472)
	$\mu_\beta = 0.01$	–0.0007	0.0053	0.0078
		(0.1081)	(0.0595)	(0.0463)
	AMSE	**1.0875**	**1.0917**	**1.0931**
4	$\alpha = 0$	0.0071	–0.0038	-0.0004
		(0.1640)	(0.0885)	(0.0723)
	$\mu_\beta = 0.6$	0.5282	0.5632	0.5752
		(0.1257)	(0.0935)	(0.0764)
	AMSE	**2.5280**	**2.5440**	**2.5409**

Table 2 The mean (sd) and AMSE of parameter estimation on maximum likelihood method

Case	Parameter	Sample sizes		
		$n = 100$	$n = 300$	$n = 500$
1	$\alpha = 0.5$	0.5871	0.5244	0.5132
		(0.1975)	(0.1096)	(0.0889)
	$\mu_\beta = 1$	0.9973	0.9997	0.9999
		(0.0073)	(0.0013)	(0.0006)
	AMSE	**0.9884**	**0.9942**	**0.9949**
2	$\alpha = 0$	0.0110	–0.0030	0.0031
		(0.3181)	(0.1348)	(0.0969)
	$\mu_\beta = 0.99$	0.9365	0.9728	0.9793
		(0.0492)	(0.0196)	(0.0133)
	AMSE	1.3609	1.5594	1.7185
3	$\alpha = 0$	0.0067	–0.0014	0.0026
		(0.1088)	(0.0581)	(0.1399)
	$\mu_\beta = 0.01$	0.0007	0.0046	–0.0008
		(0.1079)	(0.0591)	(0.1657)
	AMSE	1.0879	1.0919	1.1911
4	$\alpha = 0$	0.00845	–0.0011	0.0009
		(0.1305)	(0.0683)	(0.0550)
	$\mu_\beta = 0.6$	0.5795	0.5924	0.5947
		(0.1132)	(0.0630)	(0.0487)
	AMSE	2.6053	2.5816	2.5686

The maximum likelihood is a good performance at Case 1 as the trend data in Table 2. The AMSE has represented the minimum values for all sample sizes.

Table 3 The mean (sd) and AMSE of parameter estimation on Bayesian method

Case	Parameter	Sample sizes		
		$n = 100$	$n = 300$	$n = 500$
1	$\alpha = 0.5$	0.5876	0.5244	0.5130
		(0.1975)	(0.1094)	(0.0889)
	$\mu_\beta = 1$	0.9969	0.9995	1.0021
		(0.0243)	(0.0216)	(0.0226)
	AMSE	1.4461	4.4543	11.9705
2	$\alpha = 0$	0.0115	−0.0035	0.0006
		(0.3506)	(0.1531)	(0.1082)
	$\mu_\beta = 0.99$	0.9327	0.9674	0.9771
		(0.0532)	(0.0310)	(0.0258)
	AMSE	1.3825	1.5894	1.7649
3	$\alpha = 0$	0.0063	−0.0014	0.0002
		(0.1092)	(0.0583)	(0.0472)
	$\mu_\beta = 0.01$	−0.0008	0.0023	0.0057
		(0.1112)	(0.0629)	(0.0519)
	AMSE	1.0889	1.0922	1.0936
4	$\alpha = 0$	0.0069	−0.0038	−0.0006
		(0.1647)	(0.0887)	(0.0725)
	$\mu_\beta = 0.6$	0.5279	0.5622	0.5746
		(0.1276)	(0.0949)	(0.0798)
	AMSE	2.5298	2.5460	2.5428

From Table 3, the Bayesian AMSE's are larger than the other method, but this method outperforms the maximum likelihood method at Case 4 as it wildly oscillates around its mean zero. It can seem that the prior distribution of the Bayesian method is not affected to estimate parameters on the RCA model.

5. Application in Actual Data

In this section, we will consider applying the RCA(1) model using the least-squares, maximum likelihood, and Bayesian methods that we developed from the previous section. As a set of actual data, we play on the time series data, namely the daily gold price per one-baht weight (one-baht weight: 15.244 grams) from October 2019 to September 2020, which consists of 314 records and shown in Figure 4. The gold price data is interesting in this period because it gathered momentum during the coronavirus pandemic and looked like the simulation data on Case 2. These data are obtained from https://lgp.go.th/index.php?p=22&month=09&year=2019.

Figure 4 shows the random walk process of 314 records and looks like the increasing trend. The modeling steps as follows. At the first, we focus on the daily gold price per one-baht weight denoted $x_t, t = 1,2,3,...,314$. Then, we estimate the parameters from three methods and obtain $\hat{\alpha}_\theta = \hat{\alpha}_{LS}$, $\hat{\alpha}_{ML}$, $\hat{\alpha}_{Bayes}$ and $\hat{\mu}_{\beta,\theta} = \hat{\mu}_{\beta,LS}, \hat{\mu}_{\beta,ML}, \hat{\mu}_{\beta,Bayes}$. Finally, we use these estimators for predicting future values on October 2020 (27 days) as $\hat{x}_t = \hat{\alpha}_\theta + \hat{\mu}_{\beta,\theta} x_{t-1}$, $t = 315,316,...,341$.

The mean square error (MSE) evaluates the difference between the real values and forecasting values. We also compute the MSE as the criterion defined following:

$$MSE = \frac{\sum_{t=1}^{27}(x_t - \hat{x}_t)^2}{27},$$

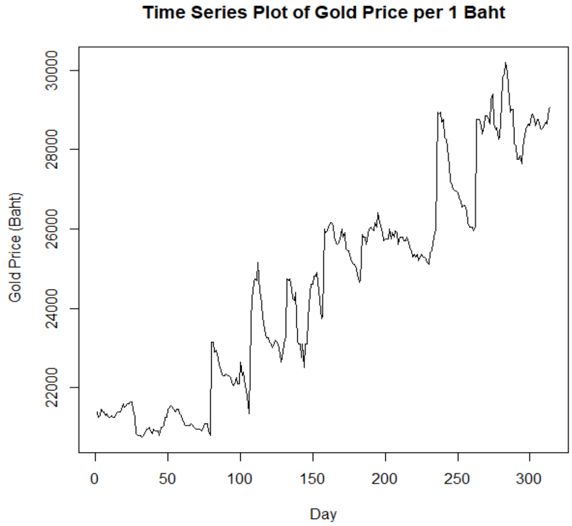

Figure 4 The time series plot for daily gold price per one-baht weight since October, 2019 to September 2020

where x_t denotes the real values, and $\hat{x}_t$ denotes the forecasting values. Figure 5 gives the plot of the gold price where the dashed line is presented of the least-squares method, the dotted line is presented the maximum likelihood method, and the line is presented the Bayesian method.

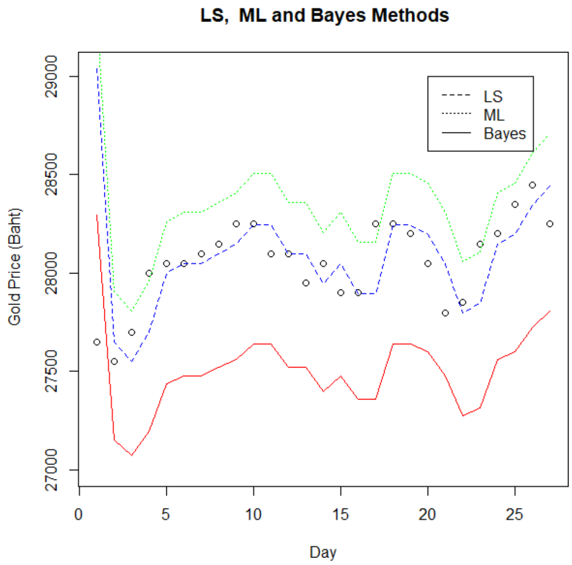

Figure 5 The plot of daily gold price per one-baht weight and forecasting of least-squares (LS), maximum likelihood (ML), and Bayesian (Bayes) methods

Compared with Figure 5, and it can see that the forecasting values of the least-squares method are relatively close to the observed series. Therefore, we should be more convinced by the MSE of the least-squares method given 93,460.5, the MSE of maximum likelihood given 372,465.2, and the Bayesian method is maximum as 176,833.8.

6. Conclusions

This paper studied the least-squares, maximum likelihood, and Bayesian methods for estimating the first order in RCA or called RCA(1) model. Through a Monte Carlo simulation, we evaluated the performance of these methods, showed the mean and standard deviation of parameters, and played the AMSEs of various data and sample sizes. It appears that most cases of the least-squares method perform well in picking up the correct model to see the AMSE is minimum. It is indicated that the RCA(1) model is affected on past observed data more than the informative prior to Bayesian Analysis. The maximum likelihood method outperforms the other methods for the trend data, but it is slightly different from the least-squares method. We are also interested in the power of estimating by the mean square error for application in actual data. We can see that the least-squares method outperforms the maximum likelihood and Bayesian method similar to the simulation study results. We would recommend fitting RCA(1) model on time series data by the least-squares method where stationary and non-stationary data are expected.

References

Andrews B, Davis, RA. Maximum likelihood estimation for all-pass time series models. J Multivar Anal. 2006; 97: 1638-1659.

Aue A, Horvath L, Steinebach J. Estimation in random coefficient autoregressive models. J Time Ser Anal. 2006; 27: 61-76.

Benmoumen M, Allal J, Salhi I. Parameter estimation for p-order random coefficient autoregressive (RCA) models based on Kalman filter. J Appl Math. 2019; 1-5.

Gelfand A, Hills S, Racine PA, Smith A. Illustration of Bayesian inference in normal data models using Gibbs sampling. J Am Stat Assoc. 1990; 85: 972-985.

Geman S, Geman D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. IEEE Tran Pattern Anal Mach Intell. 1984; 6: 721-741.

Gilks W, Richardson S, Speigelhalter D. Markov chain Monte Carlo in practice, interdisciplinary statistics. Suffolk: Chapman & Hall; 1996, 102-103.

Khall MAK, Moraes FP. Linear least squares method for time series analysis with an application to a methane time series. J Air Waste Manag Assoc. 2012; 45: 62-63.

Morton A, Finkenstadt BF. Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods. J R Stat Ser C Appl Stat. 2005; 54(3): 575-594.

Nicholls DF, Quinn BG. Random coefficient autoregressive models: an introduction. New York: Springer-Verlag; 1982, 45-56.

Tsay C. Conditional heteroscedastic time series models. J Am Stat Assoc.1987; 82: 590-604.

Wang D, Ghosh SK. Bayesian estimation and unit root tests for random coefficient autoregressive models. Model Assist Stat Appl. 2008; 3: 281-295.