The Maximum Likelihood Method of Random Coefficient Dynamic Regression Model

Autcha Araveeporn

Abstract—The Random Coefficient Dynamic Regression (RCDR) model is to developed from Random Coefficient Autoregressive (RCA) model and Autoregressive (AR) model. The RCDR model is considered by adding exogenous variables to RCA model. In this paper, the concept of the Maximum Likelihood (ML) method is used to estimate the parameter of RCDR(1,1) model. Simulation results have shown the AIC and BIC criterion to compare the performance of the the RCDR(1,1) model. The variables as the stationary and weakly stationary data are good estimates where the exogenous variables are weakly stationary. However, the model selection indicated that variables are nonstationarity data based on the stationary data of the exogenous variables.

Keywords—Autoregressive, Maximum Likelihood Method, Nonstationarity, Random Coefficient Dynamic Regression, Stationary.

I. Introduction

MOST data are collected in the form of time series that often exhibits nonstationarity and stationary models. The nonstationarity models might be caused by several aspects including changes in trend volatility and random walk. The heteroscedasticity or volatility has been modeled in the literature by various authors, for instance, [1], [2] evaluated risk in finance, [3] monitored the reliability of nonlinear prediction. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

There are several volatility models in time series, starting by [4] who introduced AutoRegressive Conditional Heteroscedastic model (ARCH) which was obtained the predictive variance for U.K. inflation rate. To obtain more flexibility, the ARCH model has been extended by [5] who produced the Generalized ARCH (GARCH) model. The GARCH model is allowed the past data time series and the past volatility in this model. To overcome some weakness of the GARCH model, [6] proposed the Exponential GARCH (EGARCH) model that is used the log condition variance to relax the positiveness constraint of coefficient model. [7] proposed the Conditional Heteroscadastic AutoRegressive Moving Average (CHARMA), in which is not similar to the GARCH model, but these two models possess similar second-order condition properties. The special case of CHARMA model that reduced to the Random Coefficient Autoregressive (RCA) model which was studied by [8].

A. Araveeporn is with the Department of Statistics, King Mongkut's Institute of Technology Ladkrabang, Bangkok, 10520, Thailand. e-mail: kaautcha@hotmail.com.

One of the common method is the estimation to practice, through Maximum Likelihood (ML) method that can be developed flexible statistics to a point estimation. The estimation of volatility model, [4] introduced a class of stochastic process called ARCH model. He derived the likelihood function of these processes and described the maximum likelihood estimators. [9] reviewed the penalized maximum likelihood estimation in nonparametric regression and density estimation. [5] extended the Engle's ARCH processes allowing for a much more flexible lag structure. [5] derived the conditionals for stationarity of this class of processes and also discussed maximum likelihood estimation of the linear regression model with GARCH errors. [10] presented a methodology for nonparametric Maximum Likelihood estimation of time-varying model.

Our research goal is to present the ML method to model nonstationarity and stationary data. We introduce the new class of Random Coefficient Dynamic Regression (RCDR) model that is extended from the RCA model.

In this paper, the RCDR model is developed from RCA model in Section II. Section III describes the parameter estimation procedure from the Maximum Likelihood (ML) method of RCDR(1,1) model and shows the properties of ML estimators. Section IV illustrates the results of simulation study and we discuss the results based on AIC and BIC criterion. A conclusion of the results is presented in Section V.

II. RCDR Model

In the case of univariate time series data, the RCA model is used the conditional variance to evolve with previous observations denoted RCA(p). The RCA model is written as

$$\begin{aligned} x_t &= \alpha + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sigma \varepsilon_t \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \end{aligned} \tag{1}$$

where

$$\begin{aligned} \underline{\beta}_t &= (\beta_{t1}, \ldots, \beta_{tp})' \\ \underline{\mu}_\beta &= (\mu_{\beta 1}, \ldots, \mu_{\beta p})' \end{aligned}$$

The ε_t and $\underline{\mu}_\beta$ are the sequences of independent of random vectors with mean zero and unit variance.

The RCA model consists of one variable but sometime it is not enough to estimate the coefficient of the time series model so the disadvantage can be improved the model by adding exogenous variables. The time series model can produce the time series dynamic modeling when the observations of time series data have correlated with exogenous variables, the dynamic modeling will help accurate the coefficient model.

Essentially, we will extend the RCA model by adding the exogenous variables of y_t denoted as

$$\begin{aligned} x_t &= \alpha_t + \sum_{i=1}^{p} \beta_{ti} x_{t-i} + \sigma\varepsilon_t \\ \underline{\beta}_t &= \underline{\mu}_\beta + \Sigma_\beta^{1/2} \underline{u}_t \\ \alpha_t &= \sum_{j=1}^{q} \eta_j y_{t-j} + \varepsilon_t \end{aligned} \tag{2}$$

The ε_t and $\underline{\mu}_\beta$ are the sequences of independent of random vectors with mean zero and unit variance, so the model (2) is called Random Coefficient Dynamic Regression (RCDR) Model or RCDR(p,q) model.

We will consider the simplified case of RCDR model with $p = q = 1$ and $\sigma = 1$; denoted by the RCDR$(1, 1)$, and we can rewrite as

$$\begin{aligned} x_t &= \alpha_t + \beta_t x_{t-1} + \varepsilon_t \\ \beta_t &= \mu_\beta + \sigma_\beta u_t \\ \alpha_t &= \eta y_{t-1} + \varepsilon_t \end{aligned} \tag{3}$$

where β_t's are iid random variables with mean μ_β and variance σ_β^2, ε_t's are iid random variables with mean 0 and variance σ^2, and β_t's and ε_t's are independent.

The parameters of RCDR$(1, 1)$ consist of the intercept term η, the mean μ_β and variance σ_β^2 of the coefficient β_t and the variance σ^2 of the ε_t, or defined as $\theta = (\eta, \mu_\beta, \sigma_\beta^2, \sigma^2)'$. In the literature, there is the RCA(1) with the slight modifications to model setup that used the nature of problem at hand might motivate to assume the two random variables β_t and ϵ_t to be correlated, see [11].

III. Parameter Estimation for RCDR$(1, 1)$

The method of maximum likelihood has been widely used in estimation. For any set of observations, $x_1, \ldots, x_n$, time series or not, the likelihood function $L(\theta)$ is define to be the joint probability density of obtaining the data actually observed. However, it is considered as a function of the unknown parameters in the model with the observed data held fixed.

To estimate parameter of RCDR(1,1) model, we propose the maximum likelihood method to estimate parameter $\theta = (\eta, \mu_\beta, \sigma_\beta^2, \sigma^2)'$. The time series data $\{x_t\}$ and $\{y_t\}$from (3) obtain following:

$$E(x_t|\ x_{t-1}) = \alpha_t + \mu_\beta\ x_{t-1}$$

$$Var(x_t|\ x_{t-1}) = (1 + \eta^2)\sigma^2 + \sigma_\beta^2 x_{t-1}^2$$

The maximum likelihood method consider the likelihood function from (3) to estimate μ_β as

$$\begin{aligned} Ł(\theta) &= Ł(\theta|x_t\ x_{t-1}) = \prod_{t=2}^{n} f(x_t|\ x_{t-1}) \\ &= \left(\frac{1}{2\pi}\right)^{n/2} \prod_{t=2}^{n} \left[(1+\eta^2)\sigma^2 + \sigma_\beta^2 x_{t-1}^2\right]^{-1/2} \\ &\quad \exp\left\{-\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \alpha_t - \mu_\beta x_{t-1})^2}{(1+\eta^2)\sigma^2 + \sigma_\beta^2 x_{t-1}^2}\right\} \end{aligned} \tag{4}$$

Constructing the new likelihood function by setting parameter, let $\eta^2 = \omega, \tau = \frac{\sigma_\beta^2}{\sigma^2}$ and substitute $\alpha_t = \eta\ y_{t-1}$, so it show that

$$\begin{aligned} Ł(\theta) &= \left(\frac{1}{2\pi}\right)^{n/2} \prod_{t=2}^{n} [\sigma^2]^{-1/2} \left[(1+\omega) + \tau x_{t-1}^2\right]^{-1/2} \\ &\quad \exp\left\{-\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]}\right\} \end{aligned} \tag{5}$$

The ln likelihood function following;

$$\begin{aligned} \ln L(\theta) &= -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 \\ &\quad -\frac{1}{2}\sum_{t=2}^{n} \ln\ \left[(1+\omega) + \tau x_{t-1}^2\right] \\ &\quad -\left\{\frac{1}{2}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]}\right\} \end{aligned} \tag{6}$$

The next step is differentiable from (6) with respect to μ_β,η, and σ^2

$$\frac{\partial \ln L(\theta)}{\partial \mu_\beta} = \sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})x_{t-1}}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]} \tag{7}$$

$$\frac{\partial \ln L(\theta)}{\partial \eta} = \sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})y_{t-1}}{\sigma^2[(1+\omega) + \tau x_{t-1}^2]} \tag{8}$$

$$\begin{aligned} \frac{\partial \ln L(\theta)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} \\ &\quad + \frac{1}{2\sigma^4}\sum_{t=2}^{n} \frac{(x_t - \eta\ y_{t-1} - \mu_\beta x_{t-1})^2}{(1+\omega) + \tau x_{t-1}^2} \end{aligned} \tag{9}$$

Now we get

$$\frac{\partial \ln L(\theta)}{\partial(\mu_\beta, \eta, \sigma^2)} = 0$$

We obtain the estimators;

$$\hat{\mu}_\beta = \frac{a_1 - \hat{\eta}\ a_2}{a_3} \tag{10}$$

$$\hat{\eta} = \frac{a_4 - \hat{\mu}_\beta\ a_2}{a_5} \tag{11}$$

$$\hat{\sigma}^2 = (n)^{-1} \sum_{t=2}^{n} \frac{(x_t - \hat{\eta}\ y_{t-1} - \hat{\mu}_\beta\ x_{t-1})^2}{(1+\omega) + \tau x_{t-1}^2} \tag{12}$$

where

$$\begin{aligned} a_1 &= \sum_{t=2}^{n} \frac{x_t x_{t-1}}{(1+\omega)+\tau x_{t-1}^2}, a_2 = \sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{(1+\omega)+\tau x_{t-1}^2} \\ a_3 &= \sum_{t=2}^{n} \frac{x_{t-1}^2}{(1+\omega)+\tau x_{t-1}^2}, a_4 = \sum_{t=2}^{n} \frac{x_t y_{t-1}}{(1+\omega)+\tau x_{t-1}^2} \\ a_5 &= \sum_{t=2}^{n} \frac{y_{t-1}^2}{(1+\omega)+\tau x_{t-1}^2} \end{aligned}$$

The ML estimates $\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2$, and $\hat{\sigma}_\beta^2$ ca be obtained by calculating $\hat{\tau}$, where $\hat{\tau}$ is the minimizer of the following function of τ,

$$g(\tau) = \ln(\sigma^2) + \sum_{t=2}^{n} \ln((1+\omega)+\tau x_{t-1}^2)$$

That is, we have profiled the log-likelihood as a function of τ only.

The ML estimates $\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2$, and $\hat{\sigma}_\beta^2$ are obtained by

$$\begin{aligned} \ln L(\hat{\theta}) &= -\inf_{(\theta)} \ln L(\theta) \\ \ell(\hat{\eta}, \hat{\mu}_\beta, \hat{\sigma}^2, \hat{\sigma}_\beta^2) &= -\inf_{(\eta,\mu_\beta,\sigma^2,\sigma_\beta^2)} \ln L(\eta, \mu_\beta, \sigma^2, \sigma_\beta^2) \end{aligned}$$

$$\hat{\sigma}^2 = (n)^{-1} \sum_{t=2}^{n} \frac{(x_t - \hat{\eta}\ y_{t-1} - \hat{\mu}_\beta x_{t-1})^2}{(1+\hat{\omega}) + \hat{\tau} x_{t-1}^2}$$

$$\hat{\sigma}_\beta^2 = \hat{\sigma}^2 \hat{\tau}$$

and

$$\hat{\omega} = \hat{\eta}^2$$

A. The Properties of ML Estimators

For the point estimation, we might consider properties as if the sample sizes becomes infinite. In this section, we will look at the properties of ML estimators : consistency and asymptotic efficiency.

1) Consistency
The consistency of

$$\theta = (\eta_1, \ldots, \eta_q, \mu_{\beta 1}, \ldots, \mu_{\beta p}, \sigma_{\beta 1}^2, \ldots, \sigma_{\beta p}^2, \sigma^2)'$$

will be shown by examining

$$\lim_{n\to\infty} P_\theta(|W_n - \theta| \geq \epsilon) = 0$$

where $W_n = W_n(x_1, \ldots, x_n)$ is a consistent sequence of estimators of parameter θ. Recall that, for an estimator W_n, Chebychev's Inequality states

$$P_\theta(|W_n - \theta| \geq \epsilon) \leq \frac{E_\theta[(W_n-\theta)^2]}{\epsilon^2}$$

For the second term, we have

$$\begin{aligned} & E\left(\frac{(x_t - \sum_{j=1}^q \eta_j\ y_{t-j} - \sum_{i=1}^p \mu_{\beta i} x_{t-i})^2}{(1+\sum_{j=1}^q \omega_j) + \sum_{i=1}^p \tau_i x_{t-i}^2}\right) \\ &= \left|E\frac{(x_t - \sum_{j=1}^q \eta_j\ y_{t-j} - \sum_{i=1}^p \mu_{\beta i} x_{t-i})^2}{(1+\sum_{j=1}^q \omega_j) + \sum_{i=1}^p \tau_i x_{t-i}^2}\right| \\ &\leq \left|E\frac{(x_t - \sum_{j=1}^q \eta_j\ y_{t-j} - \sum_{i=1}^p \mu_{\beta i} x_{t-i})^2}{(1+\sum_{j=1}^q \omega_j) + \sum_{i=1}^p \tau_i x_{t-i}^2}\right| \\ &< \infty \end{aligned}$$

Therefore, we can conclude that

$$\begin{aligned} \left(E\frac{(x_t - \sum_{j=1}^q \eta_j\ y_{t-j} - \sum_{i=1}^p \mu_{\beta i} x_{t-i})^2}{(1+\sum_{j=1}^q \omega_j) + \sum_{i=1}^p \tau_i x_{t-i}^2}\right) &= \frac{\sigma^2}{n} \\ &< \infty \end{aligned}$$

Moreover, (13) says that the sample size becomes infinite, the estimators will be arbitrarily close to the parameter with zero in probability.

2) Asymptotic Efficiency [12]
Let $x_1, \ldots, x_n$ be iid $f(x|\theta)$, let $\hat{\theta}$ denote the ML estimator of θ, and let W_n be a continuous function of θ.

$$\sqrt{n}\ [W_n - \theta] \to n[0, \upsilon(\theta)],$$

where $\upsilon(\theta)$ is the Cramér-Rao Lower Bound. That is, W_n is a consistent and asymptotically efficient estimator of θ.
Under the property of consistency, the variance of estimator is

$$\begin{aligned} Var(\hat{\theta}) &= E_\theta(W_n - \theta)^2 = \frac{\sigma^2}{n} \\ &\approx \frac{1}{E_\theta\left(\frac{\partial \ln L(\theta)}{\partial \theta}\right)^2} \end{aligned} \tag{13}$$

Suppose that

$$\sqrt{n}\left(\frac{W_n - \theta}{\sigma}\right) \to \text{Z in distribution}$$

where $Z \sim \text{Normal}(0,1)$. By applying Slutsky's Theorem we conclude

$$W_n - \theta = \left(\frac{\sigma}{\sqrt{n}}\right)\left(\sqrt{n}\frac{W_n-\theta}{\sigma}\right) \to \lim_{n\to\infty}\left(\frac{\sigma}{\sqrt{n}}\right) Z = 0$$

so $W_n - \theta \to 0$ in distribution. We know that convergence in distribution to a point is equivalent to convergence in probability, so W_n is consistent estimator of θ.

IV. A Simulation Study

The simulation study to estimate parameter $\theta = (\eta, \mu_\beta, \sigma_\beta^2, \sigma^2)'$ for the performance of ML method. At the beginning, we generate data $y_t, t = 1, 2, \ldots, n$ from the AR(1) model by taking $\eta = 0.1,\ 0.5$ and 0.9 following;

$$\alpha_t = \eta y_{t-1} + \varepsilon_t : AR(1) \tag{14}$$

To illustrate the implication of AR model, Figure 1 shows the 100 sample sizes for each 3 coefficients ($\eta = 0.1,\ 0.5$ and 0.9). It should be noted that $\eta = 0.1$ is stationary, $\eta = 0.5$ is weakly stationary, and $\eta = 0.9$ is the nonstationarity case.

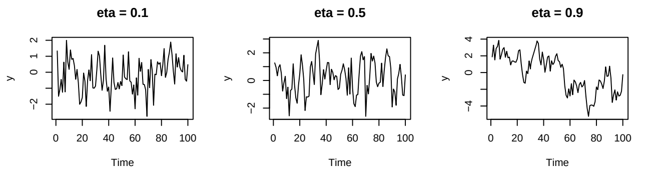

Fig. 1. The time series plot for data generated from AR(1) of $y_t(\eta = 0.1, 0.5$ and $0.9)$

Next, we consider the RCDR$(1, 1)$ model where y_t are generated from the AR(1). Therefore we obtain the x_t in term of RCDR$(1, 1)$ written as

$$x_t = \eta y_{t-1} + \beta_t x_{t-1} + \varepsilon_t \tag{15}$$

In Figures 3-4, we present the data generating in 6 cases :
1) $\sigma^2 = 1$, $\mu_\beta = 0.5$ and $\sigma^2_\beta = 0.25$
2) $\sigma^2 = 1$, $\mu_\beta = 0.995$ and $\sigma^2_\beta = 0.01$
3) $\sigma^2 = 1$, $\mu_\beta = 0.1$ and $\sigma^2_\beta = 0.99$
4) $\sigma^2 = 1$, $\mu_\beta = -0.995$ and $\sigma^2_\beta = 0.01$
5) $\sigma^2 = 1$, $\mu_\beta = -0.1$ and $\sigma^2_\beta = 0.99$
6) $\sigma^2 = 1$, $\mu_\beta = 0$ and $\sigma^2_\beta = 1$

It should be noted that Case 2 is the nonstationarity case and the Case 4 tends to be around its mean value of 0 as the stationary process. For Case 1, 3, 5, and 6, we can't define the character of the time series plot when there are slightly different figures that depended on the η under y_t. In order

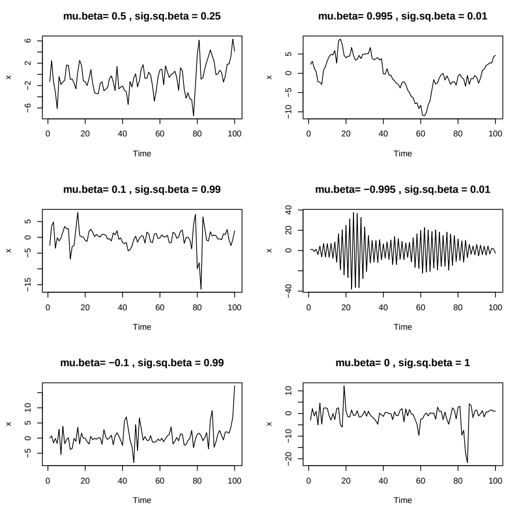

Fig. 2. The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.1)$

to assess the model performance, it is customary to use some type of model selection criteria such as Akaike Information Criterion (AIC) introduced by [13] and Bayesian Information Criterion (BIC) studied by [14]. In our simulation studies we explore the performance of the RCDR(1,1) model in picking up the true models. AIC and BIC are defined as,

$$\begin{aligned} AIC(\theta) &= -2\ln L(\theta) + 2m \\ BIC(\theta) &= -2\ln L(\theta) + m\ln(n) \end{aligned}$$

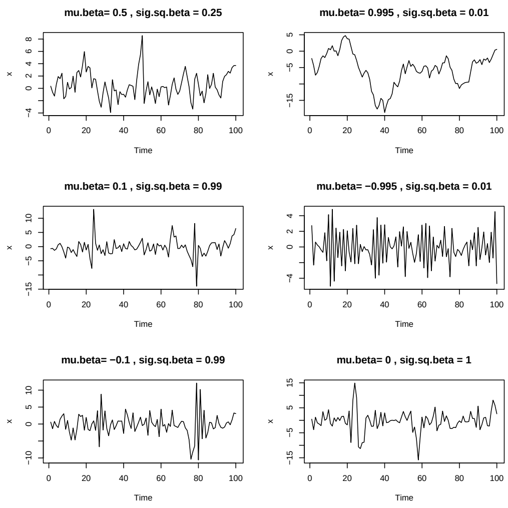

Fig. 3. The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.5)$

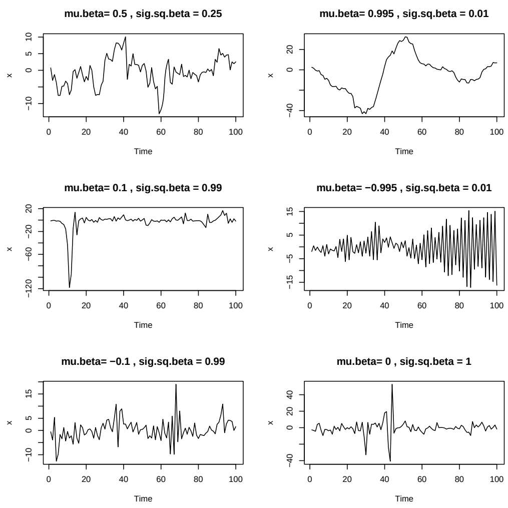

Fig. 4. The time series plot for data generated from RCDR(1,1) of $x_t(\eta = 0.9)$

where $\ln L(\cdot)$ is the log-likelihood function, m is the number of parameters in the model, and n is the number of the sample sizes.

The selection of the chosen model is then made by considering the smallest AIC and BIC in each case.

The results of the simulations were carried out using R program that was used to generate data and performed the parameter values from ML method. We simulated data with the sample sizes $n = 100$ and 500, and repeated the data generation for model fitting 500 times. Tables I- III show various Monte Carlo (MC) of the estimates obtained by taking $\eta = 0.1, 0.5$ and 0.9 in 6 cases.

The third and the fourth columns of these tables represent the AIC and BIC criterion to perform in picking up the chosen model when the estimators are fitted. From Tables I, it appears that both AIC and BIC are performing reasonably well when the data are generated from true parameter $\sigma^2 = 1, \mu_\beta = 0, \sigma^2_\beta = 1$ at sample sizes $n = 100$, and $\sigma^2 = 1, \mu_\beta = 0.1, \sigma^2_\beta = 0.99$ at sample sizes $n = 500$ when $\eta = 0.1$. However, the $\eta = 0.5$, model is good fit at parameter $\sigma^2 = 1, \mu_\beta = 0.1, \sigma^2_\beta = 0.99$. On the other hand, the $\eta = 0.9$, parameters $\sigma^2 = 1, \mu_\beta = -0.1, \sigma^2_\beta = 0.99$ prefer RCDR model at sample sizes $n = 100$, but the sample sizes $n = 500$ is fitted in parameter $\sigma^2 = 1, \mu_\beta = -0.995, \sigma^2_\beta = 0.01$. For each η, the small sample sizes are fitted better than the large sample sizes.

TABLE I
THE AVERAGE OF AIC AND BIC FOR DIFFERENT RCDR MODELS($\eta = 0.1$, SAMPLE SIZES N = 100, 500 AT 500 REPLICATIONS)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	157.963	168.383
	$\mu_\beta = 0.5, \sigma^2_\beta = 0.25$	n=500	893.249	910.107
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.326	169.745
	$\mu_\beta = 0.995, \sigma^2_\beta = 0.01$	n=500	894.414	911.273
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	158.258	168.679
	$\mu_\beta = 0.1, \sigma^2_\beta = 0.99$	n=500	**892.983**	**909.841**
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.266	168.687
	$\mu_\beta = -0.995, \sigma^2_\beta = 0.01$	n=500	893.110	909.968
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	158.181	168.602
	$\mu_\beta = -0.1, \sigma^2_\beta = 0.99$	n=500	893.029	909.888
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	**157.920**	**168.341**
	$\mu_\beta = 0, \sigma^2_\beta = 1$	n=500	893.483	990.341

TABLE II
THE AVERAGE OF AIC AND BIC FOR DIFFERENT RCDR MODELS($\eta = 0.5$, SAMPLE SIZES N = 100, 500 AT 500 REPLICATIONS)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	158.078	168.498
	$\mu_\beta = 0.5, \sigma^2_\beta = 0.25$	n=500	893.443	910.302
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.341	169.762
	$\mu_\beta = 0.995, \sigma^2_\beta = 0.01$	n=500	893.975	910.833
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	**157.879**	**168.300**
	$\mu_\beta = 0.1, \sigma^2_\beta = 0.99$	n=500	**893.264**	**910.123**
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.088	168.509
	$\mu_\beta = -0.995, \sigma^2_\beta = 0.01$	n=500	893.275	910.133
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	157.940	168.361
	$\mu_\beta = -0.1, \sigma^2_\beta = 0.99$	n=500	893.302	910.161
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	158.399	168.820
	$\mu_\beta = 0, \sigma^2_\beta = 1$	n=500	893.353	910.211

V. CONCLUSION

We have proposed a new estimators for RCDR(1,1) model by using ML method. Through a Monte Carlo simulation study, we evaluated the performance of estimator to fit RCDR(1,1) model. The proposed estimator of stationary and weakly stationary data is a good performance when the exogenous variables are weakly stationary data. However, the estimator of the nonstationarity data is good fit when the exogenous variables is stationary process. Therefore, the proposed estimator $\eta, \sigma^2, \mu_\beta$, and σ^2_β are based on the correlation between 2 variables, so the correlation is changed while the model fitting is changed too.

TABLE III
THE AVERAGE OF AIC AND BIC FOR DIFFERENT RCDR MODELS($\eta = 0.9$, SAMPLE SIZES N = 100, 500 AT 500 REPLICATIONS)

Case	Parameters	n	AIC	BIC
Case 1	$\eta = 0.1, \sigma^2 = 1$	n=100	158.674	169.095
	$\mu_\beta = 0.5, \sigma^2_\beta = 0.25$	n=500	893.654	910.512
Case 2	$\eta = 0.1, \sigma^2 = 1$	n=100	159.376	169.796
	$\mu_\beta = 0.995, \sigma^2_\beta = 0.01$	n=500	894.817	911.676
Case 3	$\eta = 0.1, \sigma^2 = 1$	n=100	158.571	168.991
	$\mu_\beta = 0.1, \sigma^2_\beta = 0.99$	n=500	893.640	910.499
Case 4	$\eta = 0.1, \sigma^2 = 1$	n=100	158.513	168.934
	$\mu_\beta = -0.995, \sigma^2_\beta = 0.01$	n=500	**893.496**	**910.355**
Case 5	$\eta = 0.1, \sigma^2 = 1$	n=100	**158.387**	**168.808**
	$\mu_\beta = -0.1, \sigma^2_\beta = 0.99$	n=500	893.777	910.635
Case 6	$\eta = 0.1, \sigma^2 = 1$	n=100	158.419	168.840
	$\mu_\beta = 0, \sigma^2_\beta = 1$	n=500	893.676	910.535

APPENDIX A
THE CONDITION OF ML ESTIMATORS

To use the RCDR(1,1) model to verify that a function of parameters has a local maximum at estimators, it must be shown that the following three conditions hold [12].

1) The first-order partial derivatives are 0.

$$\frac{\partial \ln L(\theta)}{\partial(\mu_\beta, \eta, \sigma^2)} = 0$$

2) At least one second-order partial is negative.
This condition can see after the first-order partial derivatives from (7) and (8).

$$\frac{\partial^2 \ln L(\theta)}{\partial \eta^2} = -\sum_{t=2}^{n} \frac{y^2_{t-1}}{\sigma^2[(1+\omega) + \tau x^2_{t-1}]}$$

$$\frac{\partial^2 \ln L(\theta)}{\partial \mu^2_\beta} = -\sum_{t=2}^{n} \frac{x^2_{t-1}}{\sigma^2[(1+\omega) + \tau x^2_{t-1}]}$$

3) The Jacobian of the second-order partial derivatives is positive.
For the *ln* likelihood function, the second-order partial derivatives can be written in the symmetric matrix and denoted

$$V = \begin{pmatrix} V_{11} & V_{12} & V_{13} \\ V_{21} & V_{22} & V_{23} \\ V_{31} & V_{32} & V_{33} \end{pmatrix}$$

where each element V_{ij} of V given by

$$V_{11} = \frac{\partial^2 \ln L(\theta)}{\partial \eta^2} = -\sum_{t=2}^{n} \frac{y^2_{t-1}}{\lambda_t}$$

$$V_{12} = V_{21} = \frac{\partial^2 \ln L(\theta)}{\partial \eta \partial \mu_\beta} = -\sum_{t=2}^{n} \frac{x_{t-1} y_{t-1}}{\lambda_t}$$

$$V_{13} = V_{31} = \frac{\partial^2 \ln L(\theta)}{\partial \eta \partial \sigma^2} = -\sum_{t=2}^{n} \frac{y_{t-1}}{\lambda^2_t}$$

$$V_{22} = \frac{\partial^2 \ln L(\theta)}{\partial \mu^2_\beta} = -\sum_{t=2}^{n} \frac{x^2_{t-1}}{\lambda_t}$$

$$V_{23} = V_{32} = \frac{\partial^2 \ln L(\theta)}{\partial \mu_\beta \partial \sigma^2} = -\sum_{t=2}^{n} \frac{x_{t-1}}{\lambda^2_t}$$

$$V_{33} = \frac{\partial^2 \ln L(\theta)}{\partial(\sigma^2)^2} = \sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n}\sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3}$$

where $\lambda_t = \sigma^2[(1+\omega)+\tau x_{t-1}^2]$ and $u_t = x_t - \eta\, y_{t-1} - \mu_\beta x_{t-1}$.

$$V_{11}V_{22}V_{33} = \sum_{t=2}^{n} \frac{y_{t-1}^2}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}^2}{\lambda_t} \left[\sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n}\sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3}\right]$$

$$V_{12}V_{23}V_{31} = -\sum_{t=2}^{n} \frac{x_{t-1}y_{t-1}}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2}$$

$$V_{13}V_{21}V_{32} = -\sum_{t=2}^{n} \frac{x_{t-1}y_{t-1}}{\lambda_t} \sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2} \sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2}$$

$$V_{31}V_{22}V_{13} = -\left(\sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t^2}\right)^2 \sum_{t=2}^{n} \frac{x_{t-1}^2}{\lambda_t}$$

$$V_{32}V_{23}V_{11} = -\left(\sum_{t=2}^{n} \frac{x_{t-1}}{\lambda_t^2}\right)^2 \sum_{t=2}^{n} \frac{y_{t-1}^2}{\lambda_t}$$

$$V_{12}V_{21}V_{33} = \left(\sum_{t=2}^{n} \frac{x_{t-1}y_{t-1}}{\lambda_t}\right)^2 \left[\sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n}\sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^3}\right]$$

Hence, we can compute the Jacobian by

$$V_{11}V_{22}V_{33} + V_{12}V_{23}V_{31} + V_{13}V_{21}V_{32}$$
$$-V_{31}V_{22}V_{13} - V_{32}V_{23}V_{11} - V_{12}V_{21}V_{33} > 0$$

REFERENCES

[1] T.G. Anderson and J. Lund, *Estimating continuous time stochastic volatility models of the short term interest rate*, Journal of Econometrics, 77, p. 343–377, 1977.

[2] A.R. Gallant and G. Tauchen, *Estimation of continuous time models for stock returns and interests rates*, Macroeconomic Dynamics, 1, p. 135–168, 1997.

[3] Q. Yao and H. Tong, *Quantifying the influence of initial values on nonlinear prediction*, Journal of the Royal Statistical Society, Series B, 56, p. 701–725, 1994.

[4] R.F. Engle, *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation*, Econometrica, 50, p. 987–1007, 1982.

[5] T. Bollerslev, *Generalized autoregressive conditional heteroscedasticity*, Journal of Econometrics, 31, p. 307–327, 1986.

[6] D.B. Nelson, *Conditional heteroskedasticity in asset returns : a new approach*, Econometrica, 59, p. 347–370, 1991.

[7] R. Tsay, *Conditional heteroscedastic time series models*, Journal of the American Statistical Association, 82, p. 590–604, 1987.

[8] D.F. Nicholls and B.G. Quinn, *Random coefficient autoregressive models: An introduction.*, Springer- Verlag Inc, New York, 1982.

[9] B.W. Silverman, *Penalized maximum likelihood estimation*, Encyclopaedia of Statistical Sciences, 6, p. 664–667, 1985.

[10] R. Dahlhous and P. Polonik, *Nonparametric quasi-maximum likelihood estimation for Gaussian locally stationary process*, The Annals of Statistics, 34, p. 2790–2824, 2006.

[11] C. Hwang and B.P. Carlin, *Parameter estimation for generalized random coefficient autoregressive process*, Journal of Statistical Planning and Inference, 68, p. 323–337, 1998.

[12] G. Casella and R.L. Berger, *Statistical Inference*, CA : Duxbury, 2002.

[13] H. Akaike, *A new look at the statistical model transactions*, IEEE Transactions on Automatic Control, 19, p. 716–723, 1984.

[14] G. Schwarz, *Estimating the dimension of a model*, Annals of Statistics, 6, p. 461–464, 1978.