Asymtotic Distribution of the Bootstrap Parameter Estimator for the AR(p) Model

Bambang Suprihatin(✉), Endro Setyo Cahyono, and Novi Rustiana Dewi

Sriwijaya University, Indralaya, Indonesia
bambangs@unsri.ac.id

Abstract. This paper is the generalization of our two previous researches about asymptotic distribution of the bootstrap parameter estimator for the AR(1) and AR(2) models. We investigate the asymptotic distribution of the bootstrap parameter estimator of pth order autoregressive or AR(p) model by applying the delta method. The asymptotic distribution is the crucial property in inference of statistics. We conclude that the bootstrap parameter estimator of the AR(p) model is asymptotically converges in distribution to the p–variate normal distribution.

Keywords: Autocovariance function · Limiting distribution · Measurable function · Residuals bootstrap

1 Introduction

Consider the following stationary second order autoregressive AR(p) process:

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} \epsilon_t, \tag{1}$$

where ϵ_t is a zero mean white noise process with constant variance σ^2. Let the vector $\widehat{\boldsymbol{\theta}} = (\widehat{\theta_1}, \widehat{\theta_2}, \ldots, \widehat{\theta_p})^T$ is the estimator of the parameter vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_p)^T$ of (1) and $\widehat{\boldsymbol{\theta}}^{*}$ be the bootstrap version of $\widehat{\boldsymbol{\theta}}$. Studying of estimation of the unknown parameter involves: (i) what estimator should be used? (ii) having chosen a particular estimator, is this consistent? (iii) how accurate is the chosen estimator? (iv) what is the asymptotic behavior of such estimator? (v) what is the method used in proving the asymptotic properties?

Bootstrap is a general methodology for answering the second and third questions, while the delta method is one of tools used to answer the last two questions. Consistency theory is needed to ensure that the estimator is consistent to the actual parameter as desired, and thereof the asymptotic behavior of such estimator will be studied.

The consistency theories of parameter of autoregressive model have studied in [1,3,4], and for bootstrap version of the same topic, see *e.g.* [5–8,10]. They deal with the bootstrap approximation in various senses (*e.g.*, consistency

S. Misra et al. (Eds.): ICCSA 2019, LNCS 11620, pp. 12–22, 2019.
https://doi.org/10.1007/978-3-030-24296-1_2

of estimator, simulation results, limiting distribution, applying of Edgeworth expansions, etc.), and they reported that the bootstrap works usually very well. The accuracy of the bootstrapping method for autoregressive model studied in [2,9]. They showed that the parameter estimates of the autoregressive model can be bootstrapped with accuracy that outperforms the normal approximation. The asymptotic result for the AR(1) model has been exhibited in [11]. We concluded that the bootstrap parameter estimator for the AR(1) model converges in distribution to the normal distribution. A good perform of the bootstrap estimator is applied to study the asymptotic distribution of $\widehat{\theta}^*$ using the delta method. We describe the asymptotic distribution of the autocovariance function and investigate the bootstrap limiting distribution of $\widehat{\boldsymbol{\theta}}^*$. Section 2 reviews the asymptotic distribution of estimator of mean and autocovariance function for the autoregressive model. Section 3 describes the bootstrap and delta method. Section 4 deals with the main result, *i.e.* the asymptotic distribution of $\widehat{\boldsymbol{\theta}}^*$ by applying the delta method. Section 5 briefly describes the conclusions of the paper.

2 Estimator of Mean and Autocovariance for the Autoregressive Model

Suppose we have the observed values $X_1, X_2, \ldots, X_n$ from the stationary AR(p) process. Mean and autocovariance are two important statistics in investigating the consistency properties of the estimator $\widehat{\boldsymbol{\theta}} = \left(\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p\right)^T$ for the parameter θ of the AR(p) model. A natural estimators for parameters mean, covariance and correlation function are

$$\widehat{\mu}_n = \overline{X}_n = \frac{1}{n}\sum_{t=1}^{n} X_t,\ \widehat{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n-h}(X_{t+h} - \overline{X}_n)(X_t - \overline{X}_n),$$

and $\widehat{\rho}_n(h) = \widehat{\gamma}_n/\widehat{\gamma}_n(0)$ respectively. These all three estimators are consistent (see, *e.g.* [3,13]). The following theorem describes the property of the estimator $\overline{X}_n$, is stated in [3].

Theorem 1. *If $\{X_t\}$ is stationary process with mean μ and autocovariance function $\gamma(\cdot)$, then as $n \to \infty$,*

$$Var\left(\overline{X}_n\right) = E\left(\overline{X}_n - \mu\right)^2 \to 0 \quad \textit{if}\ \gamma(n) \to 0,$$

and

$$nE\left(\overline{X}_n - \mu\right)^2 \to \sum_{j=-\infty}^{\infty} \gamma(h) \quad \textit{if} \sum_{j=-\infty}^{\infty} |\gamma(h)| < \infty.$$

It is not a loss of generality to assume that $\mu_X = 0$. Under some conditions (see, *e.g.*, [13]), the sample autocovariance function can be written as

$$\widehat{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n-h} X_{t+h}X_t + O_p(1/n). \tag{2}$$

The asymptotic behavior of the sequence $\sqrt{n}\,(\widehat{\gamma}_n(h) - \gamma_X(h))$ depends only on $n^{-1}\sum_{t=1}^{n-h} X_{t+h}X_t$. Note that a change of $n-h$ by n or vice versa, is asymptotically negligible, so that, for simplicity of notation, to study the behavior of (2) we can equivalently study the average

$$\widetilde{\gamma}_n(h) = \frac{1}{n}\sum_{t=1}^{n} X_{t+h}X_t. \tag{3}$$

Both (2) and (3) are unbiased estimators of $E(X_{t+h}X_t) = \gamma_X(h)$, under the condition that $\mu_X = 0$. Their asymptotic distribution then can be derived by applying a central limit theorem to the averages $\overline{Y}_n$ of the variables $Y_t = X_{t+h}X_t$. As in [13], the autocovariance function of the series Y_t can be written as

$$V_{h,h} = \kappa_4(\varepsilon)\gamma_X(h)^2 + \sum_g \gamma_X(g)^2 + \sum_g \gamma_X(g+h)\gamma_X(g-h), \tag{4}$$

where $\kappa_4(\varepsilon) = E(\varepsilon_1^4) - 3\big(E(\varepsilon_1^2)\big)^2$, the fourth cumulant of ε_t. The following theorem is due to [13] that gives the asymptotic distribution of the sequence $\sqrt{n}\,(\widehat{\gamma}_n(h) - \gamma_X(h))$.

Theorem 2. *If $X_t = \mu + \sum_{j=-\infty}^{\infty}\psi_j\varepsilon_{t-j}$ holds for an i.i.d. sequence ε_t with mean zero and $E\big(\varepsilon_t^4\big) < \infty$ and numbers ψ_j with $\sum_j |\psi_j| < \infty$, then*

$$\sqrt{n}\,(\widehat{\gamma}_n(h) - \gamma_X(h)) \to_d N(0, V_{h,h}).$$

3 Bootstrap and Delta Method

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a population with common distribution F, and let $T(X_1, X_2, \ldots, X_n; F)$ be the specified random variable or statistic of interest, possibly depending upon the unknown distribution F. Let F_n denote the empirical distribution function of the random sample $X_1, X_2, \ldots, X_n$, *i.e.*, the distribution putting probability $1/n$ at each of the points $X_1, X_2, \ldots, X_n$. A bootstrap sample is defined to be a random sample of size n drawn from F_n, say $X^* = X_1^*, X_2^*, \ldots, X_n^*$. The bootstrap sample at first bootstrapping is usually denoted by X^{*1}. In general, the bootstrap sample at Bth bootstrapping is denoted by X^{*B}. The bootstrap data set $X^{*b} = X_1^{*b}, X_2^{*b}, \ldots, X_n^{*b}$, $b = 1, 2, \ldots, B$ consists of members of the original data set $X_1, X_2, \ldots, X_n$, some appearing zero times, some appearing once, some appearing twice, etc. The bootstrap method is to approximate the distribution of $T(X_1, X_2, \ldots, X_n; F)$ under F by that of $T(X_1^*, X_2^*, \ldots, X_n^*; F_n)$ under F_n.

Let a functional T is defined as $T(X_1, X_2, \ldots, X_n; F) = \sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$, where $\widehat{\boldsymbol{\theta}}$ is the estimator for the coefficient $\boldsymbol{\theta}$ of a stationary AR(p) model. The bootstrap version of T is $T(X_1^*, X_2^*, \ldots, X_n^*; F_n) = \sqrt{n}(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}})$, where $\widehat{\boldsymbol{\theta}}^*$ is a bootstrap version of $\widehat{\boldsymbol{\theta}}$ computed from sample bootstrap $X_1^*, X_2^*, \ldots, X_n^*$. The residuals bootstrapping procedure for the time series data to obtain $X_1^*, X_2^*, \ldots, X_n^*$ was proposed in [6]. In bootstrap view, the key of bootstrap terminology says that the population is to the sample as the sample is to the bootstrap samples. Therefore, when we want to investigate the asymptotic distribution of bootstrap estimator $\widehat{\boldsymbol{\theta}}^*$, we investigate the distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}})$ contrast to the distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$. Thus, the bootstrap is a device for estimating $P_F\big(\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq x\big)$ by $P_{F_n}\big(\sqrt{n}(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}}) \leq x\big)$. We propose the delta method in estimating for such distribution.

The delta method consists of using a Taylor expansion to approximate a random vector of the form $\boldsymbol{\phi}(T_n)$ by the polynomial $\boldsymbol{\phi}(\theta) + \boldsymbol{\phi}'(\theta)(T_n - \theta) + \cdots$ in $T_n - \theta$. This method is useful to deduce the limit law of $\boldsymbol{\phi}(T_n) - \boldsymbol{\phi}(\theta)$ from that of $T_n - \theta$, which is guaranteed by the following theorem, as stated in [12].

Theorem 3. *Let $\boldsymbol{\phi} : \mathbf{D}_\phi \subset \mathbf{R}^k \to \mathbf{R}^m$ be a map defined on a subset of $\mathbf{R}^k$ and differentiable at $\boldsymbol{\theta}$. Let T_n be random vector taking their values in the domain of ϕ. If $r_n(T_n - \theta) \to_d T$ for numbers $r_n \to \infty$, then $r_n\big(\boldsymbol{\phi}(T_n) - \boldsymbol{\phi}(\theta)\big) \to_d \boldsymbol{\phi}_\theta^{'}(T)$. Moreover, $\Big| r_n\big(\boldsymbol{\phi}(T_n) - \boldsymbol{\phi}(\theta)\big) - \boldsymbol{\phi}_\theta^{'}\big(r_n(T_n - \theta)\big)\Big| \to_p 0$.*

Assume that $\widehat{\boldsymbol{\theta}}_n$ is a statistic, and that $\boldsymbol{\phi}$ is a given differentiable map. The bootstrap estimator for the distribution of $\boldsymbol{\phi}(\widehat{\theta}_n - \boldsymbol{\phi}(\theta)$ is $\boldsymbol{\phi}(\widehat{\theta}_n^* - \boldsymbol{\phi}(\widehat{\theta}_n)$. If the bootstrap is consistent for estimating the distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})$, then it is also consistent for estimating the distribution of $\sqrt{n}\big(\boldsymbol{\phi}(\widehat{\theta}_n) - \boldsymbol{\phi}(\theta)\big)$, as given in the following theorem. The theorem is due to [12].

Theorem 4 (Delta Method For Bootstrap). *Let $\phi : \mathbf{R}^k \to \mathbf{R}^m$ be a measurable map defined and continuously differentiable in a neighborhood of θ. Let $\widehat{\boldsymbol{\theta}}_n$ be random vector taking their values in the domain of $\boldsymbol{\phi}$ that converge almost surely to θ. If $\sqrt{n}\big(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\big) \to_d T$, and $\sqrt{n}\big(\widehat{\boldsymbol{\theta}}_n^* - \widehat{\boldsymbol{\theta}}_n\big) \to_d T$ conditionally almost surely, then both $\sqrt{n}\big(\boldsymbol{\phi}(\widehat{\theta}_n) - \boldsymbol{\phi}(\theta)\big) \to_d \boldsymbol{\phi}_\theta^{'}(T)$ and $\sqrt{n}\big(\boldsymbol{\phi}(\widehat{\theta}_n^*) - \boldsymbol{\phi}(\widehat{\theta}_n)\big) \to_d \boldsymbol{\phi}_\theta^{'}(T)$ conditionally almost surely.*

4 Main Result

We now address our main result, which is summarized in the following theorem.

Theorem 5. *Let $\widehat{\boldsymbol{\theta}} = \big(\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p\big)^T$ be the estimator of $\boldsymbol{\theta} = \big(\theta_1, \theta_2, \ldots, \theta_p\big)^T$ of the stationary AR(p) process, and $\widehat{\boldsymbol{\theta}}^* = \big(\widehat{\theta}_1^*, \widehat{\theta}_2^*, \ldots, \widehat{\theta}_p^*\big)^T$ would be the bootstrap version of $\widehat{\boldsymbol{\theta}}$. The sequence of random variables $\sqrt{n}\big(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}}\big)$ converges in distribution to the normal distribution with mean $\mathbf{0}$ and covariance matrix as defined in (11).*

Proof. The model of an AR(p) process is:

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \cdots + \theta_p X_{t-p} + \varepsilon_t. \tag{5}$$

The Yule-Walker equation for (5) as follows:

$$\widehat{M} \begin{pmatrix} \widehat{\theta}_1 \\ \widehat{\theta}_2 \\ \vdots \\ \widehat{\theta}_p \end{pmatrix} = \begin{pmatrix} \sum_{t=2}^n X_t X_{t-1} \\ \sum_{t=3}^n X_t X_{t-2} \\ \vdots \\ \sum_{t=p+1}^n X_t X_{t-p} \end{pmatrix},$$

with

$$\widehat{M} = \begin{pmatrix} \sum_{t=1}^n X_t^2 & \sum_{t=2}^n X_t X_{t-1} & \cdots & \sum_{t=p}^n X_t X_{t-p+1} \\ \sum_{t=2}^n X_t X_{t-1} & \sum_{t=1}^n X_t^2 & \cdots & \sum_{t=p-1}^n X_t X_{t-p+2} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=p}^n X_t X_{t-p+1} & \sum_{t=p-1}^n X_{t-p+2} & \cdots & \sum_{t=1}^n X_{t^2} \end{pmatrix}.$$

From the Yule-Walker equation, by applying the moment method we obtain

$$\begin{pmatrix} \widehat{\theta}_1 \\ \widehat{\theta}_2 \\ \vdots \\ \widehat{\theta}_p \end{pmatrix} = \widehat{M}^{-1} \begin{pmatrix} \sum_{t=2}^n X_t X_{t-1} \\ \sum_{t=3}^n X_t X_{t-2} \\ \vdots \\ \sum_{t=p+1}^n X_t X_{t-p} \end{pmatrix}. \tag{6}$$

From (6), the vector of estimator $\widehat{\boldsymbol{\theta}} = \left(\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p\right)^T$ can be expressed as a measurable function $\phi = \left(\phi_1, \phi_2, \ldots, \phi_p\right)^T : \mathbf{R}^{p+1} \to \mathbf{R}^p$, with

$$\widehat{\theta}_i \equiv \phi_i\left(\sum_{t=1}^n X_t^2, \sum_{t=2}^n X_t X_{t-1}, \ldots, \sum_{t=p+1}^n X_t X_{t-p}\right).$$

For each function $\phi_i : \mathbf{R}^{p+1} \to \mathbf{R}$, $i = 1, 2, \ldots, p$ can be described from the system (6) and be written as $\phi_i \equiv \phi_i(u_1, u_2, \ldots, u_{p+1})$. The function ϕ_i is differentiable and its derivative matrix is

$$\phi_i' = \left(\frac{\partial \phi_i}{\partial u_1} \; \frac{\partial \phi_i}{\partial u_2} \; \cdots \; \frac{\partial \phi_i}{\partial u_{p+1}} \right).$$

By using Theorem 3, we obtain
$\sqrt{n}\Big(\phi\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2, \ldots, \frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}X_t\Big) - \phi\big(\gamma_X(0), \ldots, \gamma_X(p)\big)\Big)$

$$= \phi'_{(\gamma_X(0),\gamma_X(1),\ldots,\gamma_X(p))} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}X_t - \gamma_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1)$$

$$= \begin{pmatrix} \phi'_{1(\gamma_X(0),\gamma_X(1),\ldots,\gamma_X(p))} \\ \phi'_{2(\gamma_X(0),\gamma_X(1),\ldots,\gamma_X(p))} \\ \vdots \\ \phi'_{p(\gamma_X(0),\gamma_X(1),\ldots,\gamma_X(p))} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}X_t - \gamma_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1)$$

$$= \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1(p+1)} \\ A_{21} & A_{22} & \ldots & A_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \ldots & A_{p(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}X_t - \gamma_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1),$$

where A_{ij}, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, p + 1$ are the constants depend on $\gamma_X(0), \gamma_X(1), \ldots, \gamma_X(p)$. Precisely, for every $i = 1, 2, \ldots, p$, it holds

$$A_{ij} = \frac{\partial \phi_i}{\partial u_j}\bigg|_{(\gamma_X(0),\gamma_X(1),\ldots,\gamma_X(p))}, \quad j = 1, 2, \ldots, p + 1. \tag{7}$$

According Theorem 2, the limiting distribution for $\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2,$ $\frac{1}{n}\sum_{t=2}^{n} X_t X_{t-1}, \ldots, \frac{1}{n}\sum_{t=p+1}^{n} X_t X_{t-p}\Big)^T$ is

$$\sqrt{n}\left(\begin{pmatrix} \frac{1}{n}\sum_{t=1}^{n} X_t^2 \\ \frac{1}{n}\sum_{t=2}^{n} X_t X_{t-1} \\ \vdots \\ \frac{1}{n}\sum_{t=p+1}^{n} X_t X_{t-p} \end{pmatrix} - \begin{pmatrix} \gamma_X(0) \\ \gamma_X(1) \\ \vdots \\ \gamma_X(p) \end{pmatrix}\right)$$

$$\to_d N_{p+1}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} V_{0,0} & V_{0,1} & \cdots & V_{0,p} \\ V_{1,0} & V_{1,1} & \cdots & V_{1,p} \\ \vdots & \vdots & \ddots & \vdots \\ V_{p,0} & V_{p,1} & \cdots & V_{p,p} \end{pmatrix}\right). \tag{8}$$

By Theorem 3, if $(Z_1, Z_2, \ldots, Z_{p+1})^T$ having multivariate normal distribution as in (8), then

$$\begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1(p+1)} \\ A_{21} & A_{22} & \ldots & A_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \ldots & A_{p(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^2 - \gamma_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}X_t - \gamma_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}X_t - \gamma_X(p)\Big) \end{pmatrix}$$

$$\to_d \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1(p+1)} \\ A_{21} & A_{22} & \ldots & A_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \ldots & A_{p(p+1)} \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_{p+1} \end{pmatrix}$$

$$\sim N_p \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \tau_{12} & \cdots & \tau_{1p} \\ \tau_{21} & \tau_2^2 & \cdots & \tau_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{p1} & \tau_{p2} & \cdots & \tau_p^2 \end{pmatrix} \right),$$

with

$$\begin{aligned} \tau_i^2 &= Var(A_{i1}Z_1 + A_{i2}Z_2 + \cdots + A_{i(p+1)}Z_{p+1}) \\ &= \sum_{j=1}^{p+1} A_{ij}^2 Var(Z_j) + 2 \sum_{1 \le k < j \le p+1} A_{ik}A_{ij} Cov(Z_k, Z_j) \\ &= \sum_{j=1}^{p+1} A_{ij}^2 V_{j-1,j-1} + 2 \sum_{1 \le k < j \le p+1} A_{ik}A_{ij} V_{k-1,j-1} \\ \tau_{ik} &= Cov(A_{i1}Z_1 + \cdots + A_{i(p+1)}Z_{p+1}, A_{k1}Z_1 + \cdots + A_{k(p+1)}Z_{p+1}), \end{aligned}$$

for every $i \neq k$, $i, k = 1, 2, \ldots, p$. Thus, according Theorem 3 can be concluded that

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) = \sqrt{n}\Big(\phi\Big(\sum_{t=1}^{n} X_t^2, \ldots, \sum_{t=p+1}^{n} X_{t-p}X_t\Big) - \phi\big(\gamma_X(0), \ldots, \gamma_X(p)\big)\Big)$$

$$\to_d N_p \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^2 & \tau_{12} & \cdots & \tau_{1p} \\ \tau_{21} & \tau_2^2 & \cdots & \tau_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{p1} & \tau_{p2} & \cdots & \tau_p^2 \end{pmatrix} \right).$$

Furthermore, analog with the asymptotic distribution for the random vector $\sqrt{n}\left(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)$, we do the same for the random vector $\sqrt{n}\left(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}}\right)$, with $\widehat{\boldsymbol{\theta}}^*$ is the bootstrap version of $\widehat{\boldsymbol{\theta}}$. Let $T(X_1, X_2, \ldots, X_n; F)$ be a statistic. Let $\widehat{F}_n$ be the emphirical distribution function of $X_1, X_2, \ldots, X_n$, *i.e.* the distribution taking probability of $1/n$ for each $X_1, X_2, \ldots, X_n$. Bootstrap sample $X_1^*, X_2^*, \ldots, X_n^*$ can be obtained by using the residuals bootstrap. The function T is defined as the random variable

$$T(X_1, X_2, \ldots, X_n; F) = \sqrt{n}\big(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}\big),$$

with $\widehat{\boldsymbol{\theta}} = \big(\widehat{\theta}_1, \widehat{\theta}_2, \ldots, \widehat{\theta}_p\big)^T$ is the estimator for the coefficient $\boldsymbol{\theta}$ of AR(p) model. The bootstrap version of T is

$$T(X_1^*, X_2^*, \ldots, X_n^*; \widehat{F}_n) = \sqrt{n}\big(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}}\big),$$

with $\widehat{\boldsymbol{\theta}}^*$ is the bootstrap version of $\widehat{\boldsymbol{\theta}}$ which is computed using bootstrap sample $X_1^*, X_2^*, \ldots, X_n^*$. Bootstrap method is a tool for estimating the distribution $P_F\big(\sqrt{n}(\widehat{\theta}_i - \theta_i) \le x\big)$ using distribution $P_{\widehat{F}_n}\big(\sqrt{n}\big(\widehat{\theta}_i^* - \widehat{\theta}_i\big) \le x\big)$.

The stationary of an autoregressive process $\{X_t\}$ infer that X_t can be expressed as linear process and the residuals bootstrap yielding the sequence of i.i.d. $\{\varepsilon_t^*\}$, hence the Theorem 2 can be applied. According Theorem 2, the multivariate central limit of $\Big(\frac{1}{n}\sum_{t=1}^n X_t^{*2}, \frac{1}{n}\sum_{t=2}^n X_t^* X_{t-1}^*, \ldots, \frac{1}{n}\sum_{t=p+1}^n X_t^* X_{t-p}^*\Big)^T$ is

$$\sqrt{n}\left(\begin{pmatrix} \frac{1}{n}\sum_{t=1}^n X_t^{*2} \\ \frac{1}{n}\sum_{t=2}^n X_t^* X_{t-1}^* \\ \vdots \\ \frac{1}{n}\sum_{t=p+1}^n X_t^* X_{t-p}^* \end{pmatrix} - \begin{pmatrix} \widehat{\gamma}_X(0) \\ \widehat{\gamma}_X(1) \\ \vdots \\ \widehat{\gamma}_X(p) \end{pmatrix}\right)$$

$$\to_d N_{p+1}\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} V_{0,0}^* & V_{0,1}^* & \cdots & V_{0,p}^* \\ V_{1,0}^* & V_{1,1}^* & \cdots & V_{1,p}^* \\ \vdots & \vdots & \ddots & \vdots \\ V_{p,0}^* & V_{p,1}^* & \cdots & V_{p,p}^* \end{pmatrix}\right). \tag{9}$$

By applying the *plug-in* principle on the estimator $\widehat{\boldsymbol{\theta}}$, we obtain the bootstrap estimator $\widehat{\boldsymbol{\theta}}^*$. As in the previous process, the estimator $\widehat{\boldsymbol{\theta}}^* = \Big(\widehat{\theta}_1^*, \widehat{\theta}_2^*, \ldots, \widehat{\theta}_p^*\Big)^T$ can be expressed as a measurable functional $\boldsymbol{\phi} = (\phi_1, \phi_2, \ldots, \phi_p)^T : \mathbf{R}^{p+1} \to \mathbf{R}^p$. Each component of estimator $\widehat{\theta}_i^*, i = 1, 2, \ldots, p$ can be expressed as a measurable function $\phi_i : \mathbf{R}^{p+1} \to \mathbf{R}$,

$$\widehat{\theta}_i^* \equiv \phi_i\Bigg(\sum_{t=1}^n X_t^{*2}, \sum_{t=2}^n X_t^* X_{t-1}^*, \ldots, \sum_{t=p+1}^n X_t^* X_{t-p}^*\Bigg).$$

The function ϕ_i for every $i = 1, 2, \ldots, p$ can be determined from the system (6) and be written as $\phi_i \equiv \phi_i(u_1, u_2, \ldots, u_{p+1})$. The function ϕ_i is differentiable and its derivative matrix is

$$\phi_i' = \left(\frac{\partial \phi_i}{\partial u_1} \; \frac{\partial \phi_i}{\partial u_2} \; \cdots \; \frac{\partial \phi_i}{\partial u_{p+1}} \right).$$

By applying Theorems 3 and 4, we obtain

$$\sqrt{n}\Big(\boldsymbol{\phi}\Big(\tfrac{1}{n}\textstyle\sum_{t=1}^{n} X_t^{2*}, \ldots, \tfrac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^{*}X_t^{*}\Big) - \boldsymbol{\phi}\big(\widehat{\gamma}_X(0), \ldots, \widehat{\gamma}_X(p)\big)\Big)$$

$$= \boldsymbol{\phi}'_{(\widehat{\gamma}_X(0),\widehat{\gamma}_X(1),\ldots,\widehat{\gamma}_X(p))} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^{2*} - \widehat{\gamma}_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*}X_t^{*} - \widehat{\gamma}_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^{*}X_t^{*} - \widehat{\gamma}_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1)$$

$$= \begin{pmatrix} \phi'_{1(\widehat{\gamma}_X(0),\widehat{\gamma}_X(1),\ldots,\widehat{\gamma}_X(p))} \\ \phi'_{2(\widehat{\gamma}_X(0),\widehat{\gamma}_X(1),\ldots,\widehat{\gamma}_X(p))} \\ \vdots \\ \phi'_{p(\widehat{\gamma}_X(0),\widehat{\gamma}_X(1),\ldots,\widehat{\gamma}_X(p))} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^{2*} - \widehat{\gamma}_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*}X_t^{*} - \widehat{\gamma}_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^{*}X_t^{*} - \widehat{\gamma}_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1)$$

$$= \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1(p+1)} \\ B_{21} & B_{22} & \ldots & B_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p1} & B_{p2} & \ldots & B_{p(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^{2*} - \widehat{\gamma}_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*}X_t^{*} - \widehat{\gamma}_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^{*}X_t^{*} - \widehat{\gamma}_X(p)\Big) \end{pmatrix} + \mathbf{o}_p(1),$$

where B_{ij}, $i = 1, 2, \ldots, p$, $j = 1, 2, \ldots, p+1$ are constants depend on $\widehat{\gamma}_X(0), \widehat{\gamma}_X(1), \ldots, \widehat{\gamma}_X(p)$. More precisely, for every $i = 1, 2, \ldots, p$,

$$B_{ij} = \left.\frac{\partial \phi_i}{\partial u_j}\right|_{(\widehat{\gamma}_X(0),\widehat{\gamma}_X(1),\ldots,\widehat{\gamma}_X(p))}, \quad j = 1, 2, \ldots, p+1. \tag{10}$$

If $(W_1, W_2, \ldots, W_{p+1})^T$ having multivariate normal distribution as in (9), then by Theorem 4, we obtain

$$\begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1(p+1)} \\ B_{21} & B_{22} & \ldots & B_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p1} & B_{p2} & \ldots & B_{p(p+1)} \end{pmatrix} \begin{pmatrix} \sqrt{n}\Big(\frac{1}{n}\sum_{t=1}^{n} X_t^{2*} - \widehat{\gamma}_X(0)\Big) \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^{*} X_t^{*} - \widehat{\gamma}_X(1)\Big) \\ \vdots \\ \sqrt{n}\Big(\frac{1}{n}\sum_{t=p+1}^{n} X_{t-p}^{*} X_t^{*} - \widehat{\gamma}_X(p)\Big) \end{pmatrix}$$

$$\rightarrow_d \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1(p+1)} \\ B_{21} & B_{22} & \ldots & B_{2(p+1)} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p1} & B_{p2} & \ldots & B_{p(p+1)} \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ \vdots \\ W_{p+1} \end{pmatrix}$$

$$\sim N_p \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^{2*} & \tau_{12}^{*} & \cdots & \tau_{1p}^{*} \\ \tau_{21}^{*} & \tau_2^{2*} & \cdots & \tau_{2p}^{*} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{p1}^{*} & \tau_{p2}^{*} & \cdots & \tau_p^{2*} \end{pmatrix} \right), \tag{11}$$

where

$$\begin{aligned} \tau_i^{2*} &= Var(B_{i1}W_1 + B_{i2}W_2 + \cdots + B_{i(p+1)}W_{p+1}) \\ &= \sum_{j=1}^{p+1} B_{ij}^2 Var(W_j) + 2 \sum_{1 \le k < j \le p+1} B_{ik}B_{ij}Cov(W_k, W_j) \\ &= \sum_{j=1}^{p+1} B_{ij}^2 V_{j-1,j-1}^{*} + 2 \sum_{1 \le k < j \le p+1} B_{ik}B_{ij}V_{k-1,j-1}^{*} \\ \tau_{ik}^{*} &= Cov(B_{i1}W_1 + \cdots + B_{i(p+1)}B_{p+1}, B_{k1}W_1 + \cdots + B_{k(p+1)}W_{p+1}), \end{aligned}$$

for every $i \neq k$, $i, k = 1, 2, \ldots, p$. Hence, by Theorem 4 we conclude that

$$\sqrt{n}\left(\widehat{\theta}^{*} - \widehat{\theta}\right) = \sqrt{n}\Big(\phi\Big(\sum_{t=1}^{n} X_t^{2*}, \ldots, \sum_{t=p+1}^{n} X_{t-p}^{*}X_t^{*}\Big) - \phi\big(\widehat{\gamma}_X(0), \ldots, \widehat{\gamma}_X(p)\big)\Big)$$

$$\rightarrow_d N_p \left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_1^{2*} & \tau_{12}^{*} & \cdots & \tau_{1p}^{*} \\ \tau_{21}^{*} & \tau_2^{2*} & \cdots & \tau_{2p}* \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{p1}^{*} & \tau_{p2}^{*} & \cdots & \tau_p^{2*} \end{pmatrix} \right),$$

completing the proof.

5 Conclusions

We conclude that the bootstrap parameter estimators of the AR(p) model are asymptotic and converge in distribution to the p–variate normal distribution.

References

1. Bibi, A., Aknounche, A.: Yule-Walker type estimators in periodic bilinear model: strong consistency and asymptotic normality. J. Stat. Methods Appl. **19**, 1–30 (2010)
2. Bose, A.: Edgeworth correction by bootstrap in autoregressions. Ann. Statist. **16**, 1709–1722 (1988)
3. Brockwell, P.J., Davis, R.A.: Time Series: Theory and Methods. Springer, New York (1991)
4. Brouste, A., Cai, C., Kleptsyna, M.: Asymptotic properties of the MLE for the autoregressive process coefficients under stationary Gaussian noise. Math. Methods Statist. **23**, 103–115 (2014)
5. Efron, B., Tibshirani, R.: An Introduction to the Bootstrap. Chapman & Hall, New York (1993)
6. Freedman, D.A.: On bootstrapping two-stage least-squares estimates in stationary linear models. Ann. Statist. **12**, 827–842 (1984)
7. Hardle, W., Horowitz, J., Kreiss, J.P.: Bootstrap methods for time series. Int. Stat. Rev. **71**, 435–459 (2003)
8. Politis, D.N.: The impact of bootstrap methods on time series analysis. Stat. Sci. **18**, 219–230 (2003)
9. Sahinler, S., Topuz, D.: Bootstrap and jackknife resampling algorithms for estimation of regression parameters. J. Appl. Quant. Methods **2**, 188–199 (2007)
10. Singh, K.: On the asymptotic accuracy of Efron's bootstrap. Ann. Statist. **9**, 1187–1195 (1981)
11. Suprihatin, B., Guritno, S., Haryatmi, S.: Asymptotic distribution of the bootstrap parameter estimator for AR(1) process. Model Assist. Statist. Appl. **10**, 53–61 (2015)
12. Van der Vaart, A.W.: Asymptotic Statistics. Cambridge University Press, Cambridge (2000)
13. Van der Vaart, A.W.: Lecture Notes on Time Series. Vrije Universiteit, Amsterdam (2012)