

STRONG CONSISTENCY FOR AR MODEL WITH MISSING DATA

MYUNG SOOK LEE

ABSTRACT. This paper is concerned with the strong consistency of the estimators of the autocovariance function and the spectral density function for the autoregressive process in the case where only an amplitude modulated process with missing data is observed. These results will give a simple and practical sufficient condition for the strong consistency of those estimators. Finally, some examples are given to illustrate the application of main result.

1. Introduction

In this paper, the strong consistency of the estimators of the autocovariance function and the spectral density function for the autoregressive process X_t is investigated in the case where only an amplitude modulated process $Y_t = m_t X_t$ is observed.

In order to do this, a few new notations and a proposition are used.

Let X_t denote the value of the variable at time t , the p th order real valued autoregressive process (AR(p)) with autocovariance function $\sigma_X(l) = E[X_t X_{t+l}]$ and spectral density function $f_X(\lambda)$. The process can be written as

$$(1.1) \quad X_t = \sum_{j=1}^p \beta_j X_{t-j} + \epsilon_t, \quad t = 0, 1, 2, \dots$$

where ϵ_t is independently and identically distributed (i.i.d.) random variables with mean zero and variance σ^2 .

Received July 10, 2004.

2000 Mathematics Subject Classification: 62F12.

Key words and phrases: strong consistency, Stochastic process, stationary autoregressive process.

However in this autoregressive process, because the parameter $\beta = (\beta_1, \beta_2, \dots, \beta_p)^t$ and the variance σ^2 are unknown, the autocovariance function $\sigma_X(l)$ and the spectral density function $f_X(\lambda)$ are unknown. The problem of the time series analysis make inference about β , $\sigma_X(l)$ and $f_X(\lambda)$ in another appropriate way, on the basis of observations X_t , $t = 1, 2, \dots, N$.

When X_1, X_2, \dots, X_N are all observed, the following two estimators of the autocovariance function $\sigma_X(l)$ are known:

$$C_X(l) = \frac{1}{N-l} \sum_{t=1}^{N-l} X_t X_{t+l} \quad \text{and} \quad \tilde{C}_X(l) = \frac{1}{N} \sum_{t=1}^{N-l} X_t X_{t+l}.$$

Although only the first one is unbiased the second estimator is normally preferred since, in general, it has a smaller mean square error and is a positive semi-definite function. But $C_X(l)$ is not necessarily positive semi-definite. Note that the Fourier Transform of the positive semi-definite function $\tilde{C}_X(l)$ is a nonnegative function [13].

Assume that the process X_t is stationary, that is, all the zeros of polynomial $B(z)$ are outside the unit circle, where $B(z) = 1 - \sum_{j=1}^p \beta_j z^j$. Note that if X_t is stationary, there exists a sequence $\{\alpha_j\}_{j=0}^{\infty}$ of real numbers such that $X_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$ where $\{\alpha_0, \alpha_1, \dots, \alpha_n, \dots\}$ are related to $\{\beta_1, \beta_2, \dots, \beta_p\}$ by $A(z) = \sum_{j=0}^{\infty} \alpha_j z^j = \frac{1}{B(z)}$, $|z| \leq 1$, $\alpha_0 = 1$.

And since spectral density function of autoregressive process of order p is

$$f_X(\lambda) = \frac{\sigma^2}{2\pi |B(e^{i\lambda})|^2} = \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \sigma_X(l) e^{-il\lambda}$$

we can take $\tilde{f}_X(\lambda)$ as an estimator of the spectral density function $f_X(\lambda)$,

where $\tilde{f}_X(\lambda) = \frac{\tilde{\sigma}^2}{2\pi |\tilde{B}(e^{i\lambda})|^2}$ where $\tilde{B}(z) = 1 - \sum_{j=1}^p \tilde{\beta}_j z^j$, $\tilde{\beta}_j$ and $\tilde{\sigma}^2$ are

the estimator of β_j and σ^2 . Hence in this case the statistical purpose is to find the estimator of the coefficient vector β and a white noise variance σ^2 based on the observations X_1, X_2, \dots, X_N . Multiplying each side of

(1.1) by X_{t-k} , $k = 0, 1, \dots, p$, and taking expectations, we obtain the Yule-Walker equations,

$$\begin{aligned} & \sigma_X(k) \\ = & \begin{cases} \beta_1 \sigma_X(k-1) + \beta_2 \sigma_X(k-2) + \dots + \beta_p \sigma_X(k-p), & k = 1, 2, \dots \\ \beta_1 \sigma_X(k-1) + \beta_2 \sigma_X(k-2) + \dots + \beta_p \sigma_X(k-p) + \sigma^2, & k = 0 \end{cases} \end{aligned}$$

or, in the matrix form $\Gamma_X \beta = \gamma_X$, where

$$\Gamma_X = \begin{pmatrix} \sigma_X(0) & \sigma_X(1) & \dots & \sigma_X(p-1) \\ \sigma_X(1) & \sigma_X(0) & \dots & \sigma_X(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_X(p-1) & \sigma_X(p-2) & \dots & \sigma_X(0) \end{pmatrix}$$

$\beta = (\beta_1, \beta_2, \dots, \beta_p)^t$, and $\gamma_X = (\sigma_X(1), \sigma_X(2), \dots, \sigma_X(p))^t$. Hence the variance of the whitenoise process satisfies the following

$$\sigma^2 = \sigma_X(0) - \beta_1 \sigma_X(-1) - \dots - \beta_p \sigma_X(-p) = \sigma_X(0) - \beta^t \gamma_X.$$

Note that $\sigma_X(-k) = \sigma_X(k)$. The equations can be used to determine $\sigma_X(0), \sigma_X(1), \dots, \sigma_X(p)$ from σ^2 and β .

On the other hand, if we replace the covariances $\sigma_X(l)$, $l = 0, 1, \dots, p$, by the corresponding sample covariances $\tilde{C}_X(l)$, we obtain a set of equations for the so-called Yule-Walker estimators $\tilde{\beta}$ and $\tilde{\sigma}^2$ of β and σ^2 , namely, $\tilde{\Gamma}_X \tilde{\beta} = \tilde{\gamma}_X$ and $\tilde{\sigma}^2 = \tilde{C}_X(0) - \tilde{\beta}^t \tilde{\gamma}_X$ where

$$\tilde{\Gamma}_X = \begin{pmatrix} \tilde{C}_X(0) & \tilde{C}_X(1) & \dots & \tilde{C}_X(p-1) \\ \tilde{C}_X(1) & \tilde{C}_X(0) & \dots & \tilde{C}_X(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{C}_X(p-1) & \tilde{C}_X(p-2) & \dots & \tilde{C}_X(0) \end{pmatrix}$$

$$\tilde{\gamma}_X = (\tilde{C}_X(1), \tilde{C}_X(2), \dots, \tilde{C}_X(p))^t, \quad \tilde{\beta} = (\tilde{\beta}_1, \tilde{\beta}_2, \dots, \tilde{\beta}_p)^t.$$

In this study the situation can be considered in which those data are missing. This situation sometimes arises, for example, at the time of recording instruments failure intermittently or lost observations due to clerical errors. In the frequency domain the spectral analysis has been tackled by Jones [9] and Parzen [12]. Also Scheinok [15] and Bloomfield [2] consider a spectral analysis when observations are missing at random.

For notational convenience the amplitude modulated sequence are introduced. Define the sequence $Y_t = m_t X_t$, $t = 0, \pm 1, \pm 2, \dots$, which is called an amplitude modulated version of $\{X_t\}$. Assume that m_t are independent of the X_t process. Then since $\sigma_Y(l) = \sigma_m(l)\sigma_X(l)$, the natural type of covariance estimate now has the form

$$R_X(l) = \frac{\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_t X_{t+l}}{\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l}} \quad \text{if} \quad \sum_{t=1}^{N-l} m_t m_{t+l} \neq 0.$$

Hence it can be written as

$$R_X(l) = \frac{R_Y(l)}{R_m(l)} \quad \text{if} \quad R_m(l) \neq 0$$

where $R_m(l) = \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l}$ and $R_Y(l) = \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_t X_{t+l}$.

The purpose of this paper is to develop the asymptotic properties of $R_Y(l)$, $R_X(l)$ and $f_X^*(\lambda)$, where

$$f_X^*(\lambda) = \frac{\sigma^{*2}}{2\pi |B^*(e^{i\lambda})|^2}$$

and $B^*(z) = 1 - \sum_{j=1}^p \beta_j^* z^j$, $\sigma^{*2} = R_X(0) - \beta^{*t} r_X$ and $R_X \beta^* = r_X$ where

$$R_X = \begin{pmatrix} R_X(0) & R_X(1) & \dots & R_X(p-1) \\ R_X(1) & R_X(0) & \dots & R_X(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_X(p-1) & R_X(p-2) & \dots & R_X(0) \end{pmatrix}$$

$r_X = (R_X(1), R_X(2), \dots, R_X(p))^t$, $\beta^* = (\beta_1^*, \beta_2^*, \dots, \beta_p^*)^t$.

If $m_t \equiv 1$ (when X_t is always observed), then $\tilde{C}_X(l) \equiv R_X(l)$ and $f_X(\lambda) \equiv f_X^*(\lambda)$.

A number of examples of $\{m_t\}$ have been considered in the literature and these reviewed by Dunsmuir and Robinson [4] and [5]. Stochastic $\{m_t\}$ have been considered by Scheinok [15] and Bloomfield [2]. Scheinok

considers the case where $\{m_t\}$ is a sequence of independent Bernoulli trials while Bloomfield generalize this to include the dependence in the m_t .

Almost all of the papers cited in the previous paragraph are concerned with nonparametric spectral estimation for amplitude modulated processes. And Dunsmuir and Robinson [4] consider the strong consistency of $R_X(l)$ under the some conditions.

In this paper, the strong consistency of the estimator $f_X^*(\lambda)$ is investigated about the spectral density function $f_X(\lambda)$ under the model (1.1). In order to this, the strong consistency of $R_X(l)$ is also investigated under the different conditions.

Now the necessity of basic assumptions and a well-known theorem is stated necessary earlier in this paper. Thus this is used as a proposition without proof.

DEFINITION 1.1. A sequence X_t will be said to be asymptotically stationary if and only if the following limits exists a.s.:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N X_t, \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^{N-l} X_t X_{t+l} \text{ exists for all integers } l.$$

DEFINITION 1.2. If X_t is stationary or asymptotically stationary, the time series is said to be ergodic if the sample covariance function $\tilde{C}_X(l)$ is consistent at the quadratic mean estimator of $\sigma_X(l) = E[X_t X_{t+l}]$. That is, the stochastic process X_t is called ergodic if its ensemble average equal appropriate time average (i.e., with probability 1, any statistic of X_t can be determined from a single sample $X_t(\zeta)$).

In order for this to be the case it is necessary and sufficient that for each l , $\lim_{N \rightarrow \infty} \text{Var} \tilde{C}_X(l) = 0$.

The following notation is used: $\sigma_m(l) = E[m_t m_{t+l}]$. Furthermore, throughout this paper the following assumptions are readed for the model (1.1).

Assumptions:

A1. $\{m_t\}, \{X_t\}$ are independent

A2. $m_t, t = 1, 2, \dots$, is asymptotically stationary with

$$\mu = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t \quad \text{a.s.}, \quad \sigma_m(l) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \quad \text{a.s.}$$

A3. m_t 's are independent and $\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t^4 < \infty$.

A4. $X_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$, $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ where ϵ_t is strictly stationary and ergodic and $E[\epsilon_t^4] < \infty$.

PROPOSITION 1.3. Let ξ_1, ξ_2, \dots be a sequence of square integrable random variable such that $E[\xi_i] = E[\xi_i \xi_j] = 0 (i < j, i, j = 1, 2, \dots)$, $\sum_{i=1}^{\infty} \frac{E[\xi_i^2] \log^2 i}{i^2} < \infty$. Then $\zeta_n = \frac{\xi_1 + \xi_2 + \dots + \xi_n}{n} \xrightarrow{\text{a.s.}} 0$

2. Strong consistency

In this section the sufficient condition for the strong consistency of $R_X(l)$ and $f_X^*(\lambda)$ will be proved. It should be first proved that $R_Y(l)$ is a strongly consistent estimator of $\sigma_Y(l)$.

THEOREM 2.1. Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. random variables with mean zero and variance σ^2 . And let A1 - A4 hold. Then $R_X(l)$ is strongly consistent estimator of $\sigma_X(l)$, i.e., $R_X(l) \xrightarrow{\text{a.s.}} \sigma_X(l)$.

Proof. Note that $X_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$. Let

$$X_{1t} = \sum_{j=0}^M \alpha_j \epsilon_{t-j}, \quad X_{2t} = \sum_{j=M+1}^{\infty} \alpha_j \epsilon_{t-j}$$

and $Y_{it} = m_t X_{it}$, $i = 1, 2$. Then $R_Y(l) = \sum_{i=1}^2 \sum_{j=1}^2 R_{ij}(l)$ where $R_{ij}(l) =$

$$\frac{1}{N} \sum_{t=1}^{N-l} Y_{it} Y_{j(t+l)}$$

So by the Cauchy-Schwarz inequality,

$$\begin{aligned} |R_{ij}(l)| &= \left| \frac{1}{N} \sum_{t=1}^{N-l} Y_{it} Y_{j(t+l)} \right| \\ &\leq \left(\frac{1}{N} \sum_{t=1}^{N-l} m_t^2 X_{it}^2 \right)^{\frac{1}{2}} \left(\frac{1}{N} \sum_{t=1}^{N-l} m_{t+l}^2 X_{j(t+l)}^2 \right)^{\frac{1}{2}} \\ &\leq \left(\frac{1}{N} \sum_{t=1}^N m_t^4 \right)^{\frac{1}{2}} \left(\frac{1}{N} \sum_{t=1}^N X_{it}^4 \right)^{\frac{1}{4}} \left(\frac{1}{N} \sum_{t=1}^N X_{j(t+l)}^4 \right)^{\frac{1}{4}}. \end{aligned}$$

Note that $\overline{\lim}_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t^4 = K < \infty$ by A3. Since X_{1t}, X_{2t} are ergodic,

$$\overline{\lim}_{N \rightarrow \infty} |R_{ij}(l)| \leq K(E[X_{it}^4])^{\frac{1}{4}}(E[X_{jt}^4])^{\frac{1}{4}} \text{ a.s..}$$

Note that for sufficiently large M, $E[X_{2t}^4]$ converges to 0 since $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$. Therefore, by choosing M sufficiently large, it can make $R_{ij}(l)$ converges to 0 with probability 1 for sufficiently large sample size, if i or j = 2. And since $\{X_t\}$ and $\{m_t\}$ are independent, $\sigma_Y(l) = \sigma_m(l)\sigma_X(l)$. Furthermore, $R_Y(l) - \sigma_Y(l)$ becomes

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_t X_{t+l} - \sigma_m(l)\sigma_X(l) \\ &= \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_t X_{t+l} - \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sigma_X(l) \\ & \quad - \left[\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} - \sigma_m(l) \right] \sigma_X(l). \end{aligned}$$

Since $\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l}$ converges to $\sigma_m(l)$ with the probability 1, by A2, we only need to show that

$$\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_t X_{t+l} - \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sigma_X(l) \xrightarrow{\text{a.s.}} 0.$$

Note that $R_Y(l) = \sum_{i=1}^2 \sum_{j=1}^2 R_{ij}(l)$ and $R_{ij}(l)$ converges to 0 with probability 1 if i or j = 2. Hence we only need to show that

$$(2.1) \quad \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_{1t} X_{1(t+l)} - \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sigma_X(l) \xrightarrow{\text{a.s.}} 0.$$

Note that

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_{1t} X_{1(t+l)} - \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sigma_X(l) \\ &= \frac{1}{N} \sum_{j=0}^M \sum_{k=0}^M \alpha_j \alpha_k \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \\ & \quad - \frac{\sigma^2}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sum_{j=M-(l-1)}^M \alpha_j \alpha_{j+l} - \frac{\sigma^2}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sum_{j=M+1}^{\infty} \alpha_j \alpha_{j+l} \end{aligned}$$

for sufficiently large M. Here we used the following equations:

$$\begin{aligned} \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} X_{1t} X_{1(t+l)} &= \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \sum_{j=0}^M \alpha_j \epsilon_{t-j} \sum_{k=0}^M \alpha_k \epsilon_{t+l-k} \\ &= \frac{1}{N} \sum_{j=0}^M \sum_{k=0}^M \alpha_j \alpha_k \sum_{t=1}^{N-l} m_t m_{t+l} \epsilon_{t-j} \epsilon_{t+l-k} \\ E[X_{1t} X_{1(t+l)}] &= \sum_{j=0}^M \sum_{k=0}^M \alpha_j \alpha_k \sigma^2 \delta_{l,(k-j)}, \\ E[X_{1t} X_{2(t+l)}] &= \sum_{j=M-(l-1)}^M \alpha_j \alpha_{j+l} \sigma^2 \\ E[X_{2t} X_{1(t+l)}] &= 0, \\ E[X_{2t} X_{2(t+l)}] &= \sum_{j=M+1}^{\infty} \alpha_j \alpha_{j+l} \sigma^2. \end{aligned}$$

But since $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$, $\sum_{j=0}^{\infty} |\alpha_j \alpha_{j+l}| < \infty$, by Cauchy-Schwarz's inequality, $\sum_{j=M+1}^{\infty} \alpha_j \alpha_{j+l} \rightarrow 0$ which shows that the last two terms tends to 0 as M goes to infinity. Hence we only need to show that

$$(2.2) \quad \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \xrightarrow{\text{a.s.}} 0.$$

When $k = j + l$, we have

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \\ &= \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j}^2 - \sigma^2) \\ &\leq \frac{A^2}{N} \sum_{t=1}^N (\epsilon_{t-j}^2 - \sigma^2) = A^2 \left(\frac{1}{N} \sum_{t=1}^N \epsilon_{t-j}^2 - \sigma^2 \right) \\ &\xrightarrow{\text{a.s.}} 0 \quad \text{by the strong law of large numbers} \end{aligned}$$

since ϵ_t^2 are i.i.d. random variables where $A = \max_t |m_t|$. And if $k = j + l$, $\epsilon_{t-j} \epsilon_{t+l-k} = \epsilon_{s-j} \epsilon_{s+l-k}$ if $t < s$. And $\epsilon_{t-j} \epsilon_{t+l-k}$ are independent since ϵ_t are i.i.d.. Let $\xi_t = m_t m_{t+l} \epsilon_{t-j} \epsilon_{t+l-k}$. Then $E[\xi_t] = 0$. And

$$\begin{aligned} E[\xi_t \xi_s] &= E[(m_t m_{t+l} \epsilon_{t-j} \epsilon_{t+l-k})(m_s m_{s+l} \epsilon_{s-j} \epsilon_{s+l-k})] \\ &= E[m_t m_{t+l} m_s m_{s+l}] E[\epsilon_{t-j} \epsilon_{t+l-k}] E[\epsilon_{s-j} \epsilon_{s+l-k}] = 0 \quad \text{if } t < s. \\ E[\xi_t^2] &= E[m_t^2 m_{t+l}^2 \epsilon_{t-j}^2 \epsilon_{t+l-k}^2] = E[m_t^2] E[m_{t+l}^2] E[\epsilon_{t-j}^2] E[\epsilon_{t+l-k}^2] < \infty \end{aligned}$$

by A3. Furthermore, $\sum_{t=1}^{\infty} \frac{E[\xi_t^2] \log^2 t}{t^2} < \infty$. And hence by proposition 1.3,

$$\begin{aligned} & \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \\ &= \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \epsilon_{t-j} \epsilon_{t+l-k} \xrightarrow{\text{a.s.}} 0. \end{aligned}$$

Hence

$$\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \xrightarrow{\text{a.s.}} 0$$

as N tends to infinity. Thus $R_Y(l) \xrightarrow{\text{a.s.}} \sigma_Y(l)$. And note that by A2, $\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \xrightarrow{\text{a.s.}} \sigma_m(l)$. Therefore, by the Slutsky theorem, $R_X(l)$ is a strongly consistent estimator of $\sigma_X(l)$. This completes the proof. \square

Next we will prove that $f_X^*(\lambda)$ is a strongly consistent estimator of $f_X(\lambda)$.

THEOREM 2.2. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. random variables with mean 0 and variance σ^2 . And let A1 – A4 hold. Assume that $B(e^{i\lambda})$ is nonzero, $-\pi < \lambda \leq \pi$. Then $f_X^*(\lambda)$ is a strongly consistent estimator of $f_X(\lambda)$, i.e., $f_X^*(\lambda) \xrightarrow{\text{a.s.}} f_X(\lambda)$.*

To prove this theorem, we begin with the following lemma.

LEMMA 2.3. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. random variables with mean zero and variance σ^2 . And let A1 – A4 be hold. Then*

$$(a) \quad R_X \xrightarrow{\text{a.s.}} \Gamma_X \quad (b) \quad R_X^{-1} \xrightarrow{\text{a.s.}} \Gamma_X^{-1}.$$

Proof. See Appendix. □

COROLLARY 2.4. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. the random variables with mean zero and variance σ^2 . And let A1 – A4 be hold. Then*

$$(a) \quad \beta_j^* \xrightarrow{\text{a.s.}} \beta_j, \quad j = 1, 2, \dots, p \quad (b) \quad \sigma^{*2} \xrightarrow{\text{a.s.}} \sigma^2.$$

Proof. See Appendix. □

PROOF OF THEOREM 2.2. Note that

$$f_X^*(\lambda) = \frac{\sigma^{*2}}{2\pi|B^*(e^{i\lambda})|^2} \quad \text{and} \quad f_X(\lambda) = \frac{\sigma^2}{2\pi|B(e^{i\lambda})|^2}.$$

Hence, to prove that $f_X^*(\lambda)$ is a strongly consistent estimator of the $f_X(\lambda)$, we only need to show that $B^*(z) \xrightarrow{\text{a.s.}} B(z)$ and $\sigma^{*2} \xrightarrow{\text{a.s.}} \sigma^2$.

Since $B^*(z) = 1 - \sum_{j=1}^p \beta_j^* z^j$ and $B(z) = 1 - \sum_{j=1}^p \beta_j z^j$, $\beta_j^* \xrightarrow{\text{a.s.}} \beta_j$ for

$j = 1, 2, \dots, p$ implies that $B^*(z) \xrightarrow{\text{a.s.}} B(z)$. Furthermore, it is already proved that σ^{*2} is a strongly consistent estimator of σ^2 by Corollary 2.4. Hence $f_X^*(\lambda)$ is a strongly consistent estimator of $f_X(\lambda)$.

If A3 removed, then it is possible to obtain the following theorems.

THEOREM 2.5. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. the random variables with mean zero and variance σ^2 . And let A1, A2, and A4 be hold. Then $R_X(l)$ is a consistent estimator of $\sigma_X(l)$, i.e.,*

$$R_X(l) \xrightarrow{P} \sigma_X(l).$$

Proof. It is only needed to show, in the proof of the Theorem 2.1, that the convergence of (2.1) is the convergence in probability. And hence it is only needed to show that the convergence of (2.2) is the convergence in probability. But, when $k = j + l$,

$$\begin{aligned} & \text{Var} \left(\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \right) \\ &= \text{Var} \left(\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \epsilon_{t-j} \epsilon_{t+l-k} \right) = \frac{1}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 \text{Var}(\epsilon_{t-j} \epsilon_{t+l-k}) \\ &= \frac{1}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 (E[(\epsilon_{t-j} \epsilon_{t+l-k})^2] - (E[\epsilon_{t-j} \epsilon_{t+l-k}])^2) \\ &= \frac{1}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 E[\epsilon_{t-j}^2 \epsilon_{t+l-k}^2] = \frac{\sigma^4}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 = o\left(\frac{1}{N}\right) \end{aligned}$$

Since ϵ_t are i.i.d. random variables. And if $k = j + l$,

$$\begin{aligned} & \text{Var} \left(\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \right) \\ &= \text{Var} \left(\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j}^2 - \sigma^2) \right) = \frac{1}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 \text{Var}(\epsilon_{t-j}^2) \\ &= \frac{1}{N^2} \sum_{t=1}^{N-l} m_t^2 m_{t+l}^2 (E[\epsilon_{t-j}^4] - (E[\epsilon_{t-j}^2])^2) = o\left(\frac{1}{N}\right) \end{aligned}$$

Since $E[\epsilon_t^4] < \infty$ and ϵ_t are i.i.d. random variables and

$$\begin{aligned} & E \left[\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \right] \\ &= \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (E[\epsilon_{t-j} \epsilon_{t+l-k}] - \sigma^2 \delta_{l,(k-j)}) = 0. \end{aligned}$$

Hence

$$\frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} (\epsilon_{t-j} \epsilon_{t+l-k} - \sigma^2 \delta_{l,(k-j)}) \xrightarrow{P} 0$$

as N tends to the infinity. Thus $R_Y(l) \xrightarrow{P} \sigma_Y(l)$. And by A2 and the Slutsky Theorem, $R_X(l) \xrightarrow{P} \sigma_X(l)$. \square

THEOREM 2.6. *Let X_t be the stationary autoregressive process of order p of the form (1.1) when ϵ_t are i.i.d. random variables with mean zero and variance σ^2 . And let A1, A2, and A4 be hold. Assume $B(e^{i\lambda})$ is nonzero, $-\pi < \lambda \leq \pi$. Then $f_X^*(\lambda)$ is a consistent estimator of $f_X(\lambda)$, i.e., $f_X^*(\lambda) \xrightarrow{P} f_X(\lambda)$.*

To prove this theorem, it is only need to show the following lemma. For, if the following Lemma 2.7 is proved, then the proof of this theorem is the same as the proof of Theorem 2.2 except the a.s. convergence being replaced by the convergence in probability.

LEMMA 2.7. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. random variables with mean zero and variance σ^2 . And let A1, A2, and A4 be hold. Then $R_X^{-1} \xrightarrow{P} \Gamma_X^{-1}$.*

Proof. Since $R_X \xrightarrow{P} \Gamma_X$ as N goes to the infinity, $R_X \xrightarrow{\text{a.s.}} \Gamma_X$ for some subsequence N_k of sample size N with $\lim_{k \rightarrow \infty} N_k = \infty$. Hence by Lemma 2.3(b), $R_X^{-1} \xrightarrow{\text{a.s.}} \Gamma_X^{-1}$ for such subsequence N_k of sample size N . Hence $R_X^{-1} \xrightarrow{P} \Gamma_X^{-1}$ for such sequence. \square

Furthermore, the following special amplitude modulated process should be noted. Define the sequence $Y_t = m_t X_t$, $t = 0, \pm 1, \pm 2, \dots$ where

$$m_t = \begin{cases} 1 & \text{if } X_t \text{ is observed} \\ 0 & \text{if } X_t \text{ is missed.} \end{cases}$$

Then Theorem 2.1 and Theorem 2.2 hold under the assumptions A1-A4. Furthermore Theorem 2.5 and Theorem 2.6 hold under the assumptions A1, A2, and A4.

Some examples are given to show whether the estimate is consistent or not.

EXAMPLE 2.8. (Randomly missed observation) Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are

i.i.d. the random variables with mean zero and variance σ^2 . And assume that $X_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$, $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$. Consider the case when observations

are missed at random, i.e., $m_t = \begin{cases} 1 & \text{if } X_t \text{ is observed} \\ 0 & \text{if } X_t \text{ is missed} \end{cases}$ and $P[m_t = 1] =$

p , independent of t . Then $E[m_t] = p$ and $\sigma_m(l) = \begin{cases} p & \text{if } l = 0 \\ p^2 & \text{if } l \geq 1 \end{cases}$. Since

m_t are i.i.d. and $E[m_t] = p$, by the strong law of large numbers, as N

goes to infinity, $\frac{1}{N} \sum_{t=1}^N m_t \xrightarrow{\text{a.s.}} p$. Note that m_t^2 are i.i.d. and $m_t m_{t+l}$

are i.i.d. Hence by the strong law of large numbers, as N goes to infinity,

$$R_m(0) = \frac{1}{N} \sum_{t=1}^N m_t^2 \xrightarrow{\text{a.s.}} p \text{ and } R_m(l) = \frac{1}{N} \sum_{t=1}^{N-l} m_t m_{t+l} \xrightarrow{\text{a.s.}} p^2 \text{ if } l \geq 1.$$

Therefore, m_t is asymptotically stationary. And hence $\sigma_X^*(l) \xrightarrow{\text{a.s.}} \sigma_X(l)$ and $f_X^*(\lambda) \xrightarrow{\text{a.s.}} f_X(\lambda)$ if m_t and X_t are independent and $B(e^{i\lambda})$ is nonzero, $-\pi < \lambda \leq \pi$.

EXAMPLE 2.9. Consider the case when exactly every second observation is missing, i.e., $m_t = \begin{cases} 1 & \text{if } t \text{ is odd} \\ 0 & \text{if } t \text{ is even} \end{cases}$. Then it is possible to obtain

the result for $l = 1$, $\sum_{t=1}^{N-1} m_t m_{t+1} = 0$. Hence it cannot be considered

$$R_X(1) = \frac{R_Y(1)}{R_m(1)}. \text{ Furthermore, } \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t \text{ does not exist. Therefore it is impossible to apply the above theorem.}$$

EXAMPLE 2.10. Consider the case when N_0 is divided into blocks of length $2^j, j \in N_0$ where a block of length 2^{2j} in which every second observation is missing is followed by a block of length 2^{2j+1} with completely available observations, i.e.

$$m_t = \begin{cases} 0 & \text{if } t = 0 \\ 0 & \text{if } 2^{2j} \leq t \leq 2^{2j+1} - 1 \text{ and } t \text{ odd} \\ 1 & \text{if } 2^{2j} \leq t \leq 2^{2j+1} - 1 \text{ and } t \text{ even} \\ 1 & \text{if } 2^{2j+1} \leq t \leq 2^{2j+2} - 1 \end{cases}$$

Then if $N = 2^{2n}$,

$$\begin{aligned} \frac{1}{N} \sum_{t=1}^N m_t &= \frac{(2^1 + 2^3 + \dots + 2^{2n-1}) + \frac{1}{2}(2^2 + 2^4 + \dots + 2^{2n-2})}{2^{2n}} \\ &= \frac{\frac{2}{3}(2^{2n} + 2^{2n-2} - 2)}{2^{2n}} \rightarrow \frac{5}{6}. \end{aligned}$$

If $N = 2^{2n+1}$,

$$\begin{aligned} \frac{1}{N} \sum_{t=1}^N m_t &= \frac{(2^1 + 2^3 + \dots + 2^{2n-1}) + \frac{1}{2}(2^2 + 2^4 + \dots + 2^{2n})}{2^{2n+1}} \\ &= \frac{\frac{2}{3}(2^{2n} - 1 + 2^{2n} - 1)}{2^{2n+1}} \rightarrow \frac{2}{3}. \end{aligned}$$

Hence there does not exist $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N m_t$. This implies that m_t is not asymptotically stationary. Therefore it cannot be guaranteed the strong consistency.

These examples show that if the observations are periodically missing, it is not able to guarantee the strong consistency.

EXAMPLE 2.11. Consider the case when the first $N^\alpha (\alpha < 1)$ data of N are available and the rest are missing, i.e., $m_t = \begin{cases} 1 & \text{if } t \leq N^\alpha \\ 0 & \text{if } N^\alpha < t \leq N \end{cases}$. Then in this case, if we take the sample size N^α instead of N , $R_X(l) = \tilde{C}_X(l)$, $f_X^*(\lambda) = \tilde{f}_X(\lambda)$ are strongly consistent estimators of $\sigma_X(l)$ and $f_X(\lambda)$, respectively.

Appendix

LEMMA 2.3. Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. random variables with mean zero and variance σ^2 . And let A1 – A4 be hold. Then

$$(a) \quad R_X \xrightarrow{\text{a.s.}} \Gamma_X \quad (b) \quad R_X^{-1} \xrightarrow{\text{a.s.}} \Gamma_X^{-1}.$$

Proof. (a) By Theorem 2.1, the following relation can be obtained $r_X \xrightarrow{\text{a.s.}} \gamma_X$ i.e., r_X is a strongly consistent estimator of γ_X . And $R_X(0) \xrightarrow{\text{a.s.}} \sigma_X(0)$. Hence $R_X \xrightarrow{\text{a.s.}} \Gamma_X$ (b) By (a), $R_X \xrightarrow{\text{a.s.}} \Gamma_X$. Then $R_X \rightarrow \Gamma_X$ for all sample point x outside a set of probability zero. By the definition of convergence of a sequence of matrices, any entry of R_X converges to the corresponding entry of Γ_X for all x outside a set of probability zero. Since the number of entries of R_X is p^2 , $\det(R_X)$ converges to $\det(\Gamma_X)$ for all such x . Since Γ_X^{-1} exists according to the positive definiteness of Γ_X , this implies that $\det(R_X)$ is nonzero for all such x and for all sufficiently large sample size N . Hence R_X^{-1} exists for all large N . Now it will be proved that $R_X^{-1} \rightarrow \Gamma_X^{-1}$ for all x . For notational convenience, we set $R_X = [a_{ij}^{(N)}]_{p \times p}$ and $\Gamma_X = [a_{ij}]_{p \times p}$. Suppose $\lim_{N \rightarrow \infty} \sum_{i,j} (a_{ij}^{(N)} - a_{ij})^2 = 0$. Then all the entries of the matrix R_X^{-1} have a

bound independence of N . Let $C_{ij}^{(N)}$ be the cofactor of $a_{ij}^{(N)}$ in $\det(R_X)$ and C_{ij} the cofactor of a_{ij} in $\det(\Gamma_X)$. Then we have

$$\begin{aligned} R_X^{-1} - \Gamma_X^{-1} &= R_X^{-1}(\Gamma_X - R_X)\Gamma_X^{-1} \\ &= \frac{1}{\det(R_X)\det(\Gamma_X)} [C_{ij}^{(N)}]^t (\Gamma_X - R_X) [C_{ij}]^t \\ &= (\det(R_X))^{-1} (\det(\Gamma_X))^{-1} [C_{ij}^{(N)}]^t (\Gamma_X - R_X) [C_{ij}]^t \end{aligned}$$

Since $\lim_{N \rightarrow \infty} (\det(R_X)) = \det(\Gamma_X)$ and $\lim_{N \rightarrow \infty} \sum_{i,j} (a_{ij}^{(N)} - a_{ij})^2 = 0$, the matrix norms of the matrices $(\det R_X)^{-1} [C_{ij}^{(N)}]^t$, $(\det \Gamma_X)^{-1} [C_{ij}]^t$ are bounded by a constant independence of N . Since $R_X \xrightarrow{\text{a.s.}} \Gamma_X$, this implies $R_X^{-1} \xrightarrow{\text{a.s.}} \Gamma_X^{-1}$. □

COROLLARY 2.4. *Let X_t be the stationary autoregressive process of order p of the form (1.1) where ϵ_t are i.i.d. the random variables with mean zero and variance σ^2 . And let A1 – A4 be hold. Then*

$$(a) \quad \beta_j^* \xrightarrow{\text{a.s.}} \beta_j, \quad j = 1, 2, \dots, p \quad (b) \quad \sigma^{*2} \xrightarrow{\text{a.s.}} \sigma^2.$$

Proof. (a) Note that $R_X \beta^* = r_X$ and $\Gamma_X \beta = \gamma_X$ by the Yule-Walker equation. Hence this fact and Lemma 2.3 imply that $\beta_j^* \xrightarrow{\text{a.s.}} \beta_j, \quad j = 1, 2, \dots, p$.

(b) Note that $\sigma^{*2} = R_X(0) - \beta^{*t} r_X$. Hence the result immediately follows from Lemma 2.3 and (a). □

References

- [1] K. N. Berk, *Consistent autoregressive spectral estimates.*, Ann. Statist. **2** (1974), no. 3, 489–502.
- [2] P. Bloomfield, *Spectral analysis with randomly missed observations*, J. R. Stat. Soc. Ser. B Stat. Methodol. **32** (1970), 369–380.
- [3] R. Dahlhaus, *Nonparametric spectral analysis with missing observations*, Sankhyā Ser. A. **49** (1987), 347–367.
- [4] W. Dunsmuir and P. M. Robinson, *Asymptotic theory for time series containing missing and amplitude modulated observations.*, Sankhyā Ser. A **43** (1981a), 260–281.
- [5] ———, *Parametric estimators for stationary time series with missing observations.*, Adv. Appl. Prob. **13** (1981b), 126–146.
- [6] W. A. Fuller, *Introduction to statistical time series*, John Wiley, New York (1976).
- [7] H. Hall and C. C. Heyde, *Martingale limit theory and its application.*, Academic Press (1980).
- [8] S. Haykin, *Nonlinear methods of spectral analysis.*, Springer-Verlag (1979).
- [9] R. H. Jones, *Spectral analysis with regularly missed observations*, Ann. Math. Statist. **32** (1962), 455–461.
- [10] R. E. Kromer, *Asymptotic properties of the autoregressive spectral estimator.*, Ph. D. Dissertation, Stanford Univ. (1970).
- [11] R. J. Marshall, *Autocorrelation estimation of time series with randomly missing observations.*, Biometrika **67** (1980), no. 3, 567–570.
- [12] E. Parzen, *On spectral analysis with missing observations and amplitude modulation.*, Tech. Rep. **46** (1962).
- [13] M. B. Priestley, *Spectral analysis and time series*, Academic Press (1981), 321–330.
- [14] P. Revesz, *The laws of large numbers*, Academic Press (1968).
- [15] P. A. Scheinok, *Spectral analysis with randomly missed observations: the binomial case.*, Ann. Math. Statist. **36** (1965), 971–977.

Department of Mathematics
Yonsei University
Seoul 120-749, Korea
E-mail: jhmslee@yonsei.ac.kr