The Korean Communications
in Statistics Vol. 4, No. 2, 1997
pp. 345 - 351

Fixed Accuracy Confidence Set for the Autocorrelations of Linear Processes [1)]

Sangyeol Lee [2)]

Abstract

This paper considers the problem of sequential fixed accuracy confidence set procedure of the autocorrelations of stationary linear processes. The proposed procedure for fixed-width confidence set is shown to be both asymptotically consistent and asymptotically efficient as the size of the width approaches zero.

1. Introduction

Let $\{X_t ; t \in Z\}$, Z={0, ± 1, ± 2, ...}, be a stationary linear process defined on a probability space (Ω, F, P) of the form:

$$X_t = \sum_{i=0}^{\infty} a_i \varepsilon_{t-i}, \quad t \in Z, \tag{1.1}$$

where the real sequence $\{a_i\}$ satisfies the absolute summablity condition $\sum_{i=0}^{\infty} |a_i| < \infty$, and $\{\varepsilon_t : t = 0, \pm 1, \ldots\}$ are unobservable iid random variables with $E\varepsilon_1 = 0$ and $E\varepsilon_1^{4\alpha} < \infty$ for some $\alpha > 1$. The linear processes include a general class of stationary processes covering ARMA (autoregressive and moving average) and infinite order autoregressive models. Applications to economics, engineering and physical sciences are broad. Moreover, most of standard texts like Fuller (1976) and Brockwell and Davis (1990) put the linear process in the central position for asymptotic studies.

In this paper we are concerned with the problem of constructing cofidence set for the autocorrelations of linear processes. The topic has long been of primary interest among

1) This Research was supported by the Sookmyung Women's University Research Grants, 1996
2) Assistant Professor, Department of Statistics, Sookmyung Women's University, Seoul, 140-742, Korea

researchers for its importance. For example, accurate estimation of autocorrelations is very important in selecting an appropriate ARMA model (cf. Box and Jenkins, 1976). Here, we will particularly consider the sequential method initiated by Robbins (1959).

Compare to iid cases, the literature on sequential estimation in time series emerged recently. See Fakhre-Zakeri and Lee (1992, 1993), Lee (1994) and the papers cited therein. Lee (1996) considered the sequetial point estimation problem on the autocorrelations of the linear process in (1.1). Some preliminary results in Lee (1996) are also useful in our setting (cf. Lemmas 3.1 below). Particularly the random central limit theorem for the autocorrelation vector, which is established in the Appendix of Lee (1996), plays an important role.

In Section 2, we propose a sequential procedure to deal with fixed accuracy confidence set with prescribed coverage probability. The proofs are given in Section 3.

2. Main results

Let $X_1, \ldots, X_n$ be n consecutive observations following the model (1.1) and denote by $\gamma(k)$ and $\rho(k)$ the autocovariance and autocorrelation at lag k, respectively. As estimates of $\gamma(k)$ and $\rho(k)$, we use the sample autocovariances and autocorrelations

$$\hat{\gamma}_n(k) = n^{-1} \sum_{t=1}^{n-k} X_t X_{t+k}, \quad 0 \le k \le n-1, \tag{2.1}$$

and

$$\hat{\rho}_n(k) = \hat{\gamma}_n(k) / \hat{\gamma}_n(0), \tag{2.2}$$

respectively. It is well known that those random sequences are strongly consistent estimates of true parameters when the indices k are fixed. Moreover, it is known that if $\hat{\underline{\rho}}_n(r) = (\hat{\rho}_1, \ldots, \hat{\rho}_n(r))'$ and $\underline{\rho} = (\rho(1), \ldots, \rho(r))'$, $r = 1, 2, \ldots$, then under the moment condition $E\varepsilon_1^4 < \infty$, as $n \to \infty$,

$$n^{1/2}(\hat{\underline{\rho}}_n(r) - \underline{\rho}(r)) \to (Y_1, \ldots, Y_r)' \tag{2.3}$$

where

$$Y_k = \sum_{i=1}^{\infty} \{ \rho(k+i) + \rho(k-i) - 2\rho(k)\rho(i) \} Z_i \tag{2.4}$$

with Z_i being iid $N(0,1)$ random variables (see Brockwell and Davis, 1990, Theorem 7.2.1, P. 221). The following procedures will heavily rely on the fact (2.3).

Suppose that one wishes to construct a confidence set for $\rho(r)$ in r-dimensional Euclidean space R^r with the maximum diameter $2d$ $(d>0)$, based on n observations, of which the probability of coverage is at least equal to $1-\alpha$ $(0<\alpha<1)$ as d goes to zero. In view of (2.3) we propose an ellipsoidal confidence region for the unknown $\rho(r)$ at sample size n:

$$R_n = \{ \underline{\theta} \in R^r; (\underline{\theta} - \hat{\underline{\rho}}_n(r))' H (\underline{\theta} - \hat{\underline{\rho}}_n(r)) \le d^2 \lambda_{\min}(H) \} ,$$

where H is a known $r \times r$ positive definite matrix and $\lambda_{\min}(A)$ $(\lambda_{\max}(A))$ denotes the smallest (largest) eigenvalue of A. Then R_n defines an ellipsoid with maximum axis $2d$ and choice of d and H determines the size and shape of the confidence set, respectively.

Let Γ denote the $r \times r$ matrix with the (k, l)-th entry being

$$\omega(k, l) = \sum_{i=1}^{\infty} \{ \rho(k+i) + \rho(k-i) - 2\rho(k)\rho(i) \} \{ \rho(l+i) + \rho(l-i) - 2\rho(l)\rho(i) \}, \tag{2.5}$$

and denote by $\chi_r^2(1-\alpha)$ the upper $1-\alpha$ point of a chi-square distribution with r degree of freedom. Set $L = \chi_r^2(1-\alpha)\lambda_{\min}^{-1}(H)\lambda_{\max}(H)$. Following the arguments as in Srivastava (1967), page 136, we have that

$$P(\underline{\rho}(r) \in R_n) \ge P\{ n(\hat{\underline{\rho}}_n - \underline{\rho}(r))' \Gamma^{-1} (\hat{\underline{\rho}}_n(r) - \underline{\rho}(r)) \le n d^2 \lambda_{\max}^{-1}(\Gamma) \lambda_{\min}(H) \lambda_{\max}^{-1}(H) \}.$$

By (2.3) the right hand side of the above inequality goes to $1-\alpha$ provided that

$$n \simeq d^{-2} L \lambda_{\max}(\Gamma) \tag{2.6}$$

since we have

$$n^{1/2}(\hat{\underline{\rho}}_n(r) - \underline{\rho}(r)) \rightarrow N(\underline{0}, \Gamma) \quad \text{as} \quad n \rightarrow \infty. \tag{2.7}$$

Now, letting $\hat{\Gamma}_n$ be the random matrices with the (k, l)-th entry being

$$\begin{aligned} \hat{\omega}_n(k, l) = \sum_{i=1}^{h_n} \{ & \hat{\rho}_n(k+i) + \hat{\rho}_n(k-i) - 2\hat{\rho}_n(k)\hat{\rho}_n(i) \} \\ & \times \{ \hat{\rho}_n(l+i) + \hat{\rho}_n(l-i) - 2\hat{\rho}_n(l)\hat{\rho}_n(i) \}, \end{aligned} \tag{2.8}$$

where h_n is a sequence of positive integers such that

$$h_n \rightarrow \infty, \; h_n = O(n^{\beta}), \; \beta \in (0, (\alpha-1)/2\alpha). \tag{2.9}$$

Define the stopping rule, in analogy of (2.6), as follows :

$$T_d = \inf\{ n \geq m_0 ; n \geq Ld^{-2}(\lambda_{\max}(\hat{\Gamma}_n) + n^{-\lambda})\} , \ \lambda > 0 . \tag{2.10}$$

Here m_0 is initial size and $n^{-\lambda}$ is the delay factor as seen in Chow and Yu (1981). Later, we will show that $P(\underline{\rho}(r) \in R_{T_d}) \geq 1-\alpha$ *as* $d \to 0$.

The following is the main theorem asserted by the performance of the sequential methods described above.

Theorem 1. (Fixed accuracy confidence set with prescribed coverage probability) As $d \to 0$,

$$T_d / k_0 \to 0 \quad a.s. \tag{2.11}$$

$$E | T_d / k_0 - 1 | \to 0 \quad a.s. \tag{2.12}$$

$$T_d^{1/2}(\hat{\underline{\rho}}_{T_d} - \underline{\rho}(r)) \to N(0, \Gamma) \tag{2.13}$$

$$P(\underline{\rho}(r) \in R_{T_d}) \geq 1 - \alpha \ , \tag{2.14}$$

where $k_0 = [Ld^{-2}\lambda_{\max}(\Gamma)] + 1$.

3. Proofs

We start this section with two lemmas which can be found in Lee (1996), Lemma 4.5 and Theorem A.1 in the Appendix.

Lemma 3.1. The following holds under the condition of Theorem 1:

as $n \to \infty$,

$$\max_{1 \leq k \leq 2h_n} \| \hat{\rho}_n(k) - \rho(k) \|_{2\alpha} = O(n^{-1/2}),$$

where for any random variable Z, $\|Z\|_{2\alpha} = (EZ^{2\alpha})^{1/2\alpha}$.

Lemma 3.2. Let $\{N_n\}$ be a sequence of positive random variables such that $N_n / n \to N$ *as* $n \to \infty$ in probability, where N is a positive random variable with $P(N < \infty) = 1$. Then, as $n \to \infty$,

$$N_n^{1/2}(\hat{\underline{\rho}}_n(r) - \underline{\rho}(r)) \to N(\underline{0}, \Gamma).$$

Lemma 3.3. Under the same conditions of Theorem 1,

$$P(|\lambda_{\max}(\hat{\Gamma}_n) - \lambda_{\max}(\Gamma)| > \delta) = O(n^{-\alpha} h_n^{2\alpha}). \tag{3.1}$$

Proof. Let

$$\phi(k,i) = \rho(k+i)+\rho(k-i)-2\rho(k)\rho(i),$$

$$\hat{\phi}(k,i) = \hat{\rho}_n(k+i)-\hat{\rho}_n(k-i)-2\hat{\rho}_n(k)\hat{\rho}_n(i).$$

Note that

$$\begin{aligned} &| \hat{w}_n(k,l)-w(k,l)| \\ \le\ & \sum_{i=1}^{h_n}\{|\hat{\phi}_n(k,i)-\phi(k,i)|\times|\hat{\phi}(l,i)|+|\phi(k,i)|\times|\hat{\phi}_n(l,i)-\phi(l,i)|\} \\ +\ & \sum_{i=h_n+1}^{\infty}|\phi(k,i)\,\phi(l,i)| \\ =\ & I_1+I_2 \ \ (\text{say}). \end{aligned}$$

I_2 goes to 0 since $\sum_{i=0}^{\infty}|\rho(i)| < \infty$ due to the abosolute summability condition on $\{a_j\}$. On the other hand, for all sufficiently large n,

$$\|I_1\|_{2\alpha} \le 8h_n \max_{1\le i\le 2h_n}\|\hat{\rho}_n(j)-\rho(j)\|_{2\alpha} = O(h_n n^{-1/2})$$

by Lemma 3.1. Thus, for any $\delta > 0$,

$$\max_{1\le k,l\le r} P(|\hat{w}_n(k,l)-w(k,l)| > \delta) = O(h_n^{2\alpha}n^{-\alpha}),$$

which is $o(1)$ by (2.9).

Consequently, for all $\varepsilon > 0$, $\theta > 0$, by (3.2),

$$\begin{aligned} & P(|\lambda_{\max}(\hat{\Gamma}_n)-\lambda_{\max}(\Gamma)| \ge \delta) \\ \le\ & P(|\lambda_{\max}(\Gamma_n)-\lambda_{\max}(\Gamma)| \ge \delta),\ |\hat{w}_n(k,l)-w(k,l) \le \theta \ \text{ for all } k,l) \\ +\ & \sum_{k,l} P(|\hat{w}_n(k,l)-w(k,l)| > \theta) \\ =\ & p(\theta,\delta)+o(1), \end{aligned}$$

where $p(\theta,\delta)$ are the positive real numbers such that $\lim_{\theta\to 0} p(\theta,\delta)=0$. This asserts (3.1).

Lemma 3.4. The family $\{T_d/k_0\}$ is uniformly integrable.

Proof. Let $\delta > 0$ and $l_d = Ld^{-2}(\lambda_{\max}(\Gamma)+2\delta)+1$. Then for $n > l_d$ and sufficiently large n such that $n^{-\lambda} < \delta$,

$$
\begin{aligned}
P(K_d > n) &\le P[Ld^{-2}(\lambda_{\max}(\Gamma_n) + n^{-\lambda}) \ge n] \\
&\le P[Ld^{-2}(\lambda_{\max}(\Gamma_n) + n^{-\lambda}) \ge l_d] \\
&\le P[|\lambda_{\max}(\hat{\Gamma}_n) - \lambda_{\max}(\Gamma)| > \delta] \\
&= \gamma_n \quad (\text{say}),
\end{aligned}
$$

Then γ_n , $n \ge 1$, are independent of $d > 0$ and summable by Lemma 3.3. Now the assertion follows from Lemma 4.4 of Woodroofe (1982).

Proof of Theorem 1. By Lemma 3.3 and (2.9), for any $\delta > 0$, we have

$$
\sum_{n=1}^{\infty} P(|\lambda_{\max}(\Gamma_n) - \lambda_{\max}(\Gamma)| > \delta) = \sum_{n=1}^{\infty} O(n^{-\alpha+2\alpha\beta}) < \infty ,
$$

so that $\lambda_{\max}(\Gamma_n) \to \lambda_{\max}(\Gamma)$ $a.s.$ as $n \to \infty$. Then it follows from the definition of T_d that $T_d / k_0 \to 1$ $a.s.$, which in turn implies (2.12) and (2.13) using Lemmas 3.4 and 3.2, respectively. (2.14) is an immediate result of (2.11) and (2.6).

References

[1] Brockwell, P. J. and Davis, R. A. (1990). *Time Series : Theory and Methods*. 2ed. Springer, New York.

[2] Box, G. E. and Jenkins, G. M. (1976). *Time Series Analysis ; Forecastings and Control*, Holden-Day, San Francisco.

[3] Chow, Y. S. and Yu, K. F. (1981). The performance of a sequential procedure for the estimation of the mean. *Annals of Statistics*, **9**, 184-189.

[4] Fakhre-Zakeri, I. and Lee, S. (1992). Sequential estimation of the mean of a linear process. *Sequential Analysis*, **11**, 181-197.

[5] Fakhre-Zakeri, I. and Lee, S. (1993). Sequential estimation of the mean vector of a multivariate linear process. *Journal of Mltivariate Analysis* **47**, 196-209.

[6] Fuller, W. (1976) *Introduction to Statistical Time Series*. Wiley, New York.

[7] Lee, S. (1994). Sequential estimation for parameters of stationary autoregressive models. *Sequential Analysis*. **13**, 301-317.

[8] Lee, S. (1996). Sequential estimation for the autocorrelations of linear processes. *Annals of Statistics*, **24**, 2233-2249.

[9] Robbins, H. (1959). Sequential estimation of the mean of a normal population. In *Probability* and *Statistics* (U. Grenander, ed) 235-245 Wiley, New York.

[10] Srivastava, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *Journal of Royal Statististical Society Series*, *B*, **29**, 132-140.

[11] Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis.* CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philladalphia.