Journal of the Korean
Data & Information Science Society
1999, Vol. 10, No. 2, pp. 397 ~ 404

Simultaneous Confidence Regions for Spatial Autoregressive Spectral Densities [1]

Eunho Ha [2]

Abstract

For two-dimensional causal spatial autoregressive processes, we propose and illustrate a method for determining asymptotic simultaneous confidence regions using Yule-Walker, unbiased Yule-Walker and least squres estimators. The spectral density for first-order spatial autoregressive model are looked at in more detail. Finite sample properties based on simulation study are also presented.

Key Words and Phrases: Spatial Autoregressive Processes, Spectral Density, Yule-walker estimators, Least-Square Estimators

1. Introduction

The problem of searching for periodicities has played an important role in the development of one dimensional time series analysis(see Chapter 6 of Priestley(1981), Section 3.9 of Newton(1988), or Chapter 10 of Brockwell and Davis(1990), for recent surveys). Much of the effort has been devoted to finding and estimating the amplitude of pure sinusoids embedded in noise, and most methods have been based primarily on the periodogram. Most recently, spurred largely by developments in speech and communication theory, methods based on autoregressive spectral estimation have been developed(see Mackisack and Poskitt(1989), for example). Autoregressive spectral estimation for one dimensional time series has become an important method of spectral density estimation in recent years(Akaike(1969), Parzen(1974), Ulrych and Bishop(1975)) despite a lack of easily applied procedures for determining confidence bands on the function being estimated. In a series of papers, Newton and Pagano(1983, 1984) and Ensor and Newton(1988) provide point and interval

[1]This paper was supported by *NON DIRECTED RESEARCH FUND*, Korea Research Foundation, 96-04D0204
[2]Department of Statistics, Yonsei University, 234 Wonju-Maeji, Korea 220-710

estimators of such peak frequencies as well as finding simultaneous confidence bands for the entire spectral density.

The purpose of the present paper is to extend these results to spatial processes that are observed over a discrete grid in the plane. Such processes have been found to be useful in a variety of scientific areas including agricultural, meteorological, and geophysical(see Chapter 5 of Ripley(1981) for a survey). In Section 2, we consider two dimensional analogs of autoregressive processes and investigate the property of estimators for the coefficients in the process. In Section 3, we derive asymptotic simultaneous confidence bands for an causal spatial autoregressive processes. Section 4 gives the results of a simulation study for the simultaneous confidence regions for the entire spectral density.

2. Causal Spatial Autoregressive Processes and Estimators

Let $\{Y_{i,j} : (i,j) \in Z^2\}$ be a zero mean spatial processes with summable autocovariance function

$$R(s,t) = \mathrm{E}(Y_{i,j}Y_{i+s,j+t}),\ (s,t) \in Z^2 \tag{1}$$

and spectral density function

$$f(\omega_1,\omega_2) = \frac{1}{4\pi^2}\sum_{u=-\infty}^{\infty}\sum_{v=-\infty}^{\infty} R(u,v)e^{-iu\omega_1}e^{-iv\omega_2},\ (\omega_1,\omega_2) \in I_\omega \tag{2}$$

where Z is the set of integers, $Z^2 = Z \times Z$, and $I_\omega = [-\pi,\pi] \times [-\pi,\pi]$. We consider i and j to be the horizontal and vertical positions of $Y_{i,j}$ in the plane. A major difficulty in extending one dimensional autoregressive processes to the spatial cases is that the natural idea of expressing an observation as a linear combination of observations in the past(unilateral representations) is lost. However, there are two natural analogs of unilateral representations, called the causal and half-plane models, that have been found useful(see Tjostheim(1981) for a discussion of the utility of these models). The causal model expresses $Y_{i,j}$ as a function of Y's that are below or to the left of $Y_{i,j}$, that is, of $Y_{s,t}$'s for $s \le i$ and $t \le j$. The half-plane model expresses as a function of Y's that are in the same column or a column to the left of $Y_{i,j}$ except that it does not allow an Y directly above $Y_{i,j}$. In order to simplify the analysis described in this paper, we consider only the causal processes. As pointed out by Tjostheim(1981), under mild conditions on the function $f(\omega_1,\omega_2)$, one can arbitrarily closely approximate a spatial series by such a causal process for some (possibly large) order (p_1,p_2). Symbolically then, the causal autoregressive process of order (p_1,p_2) satisfies

$$Y_{i,j} = \sum_{k=0}^{p_1}\sum_{l=0}^{p_2}\alpha_{kl}Y_{i-k,j-l} + \epsilon_{i,j} \tag{3}$$

where $\alpha_{00} = 0$, $\{\epsilon_{i,j}\}$ is a collection of independent random variables with zero mean and unit variance, and the complex valued polynomial

$$g(z_1, z_2) = 1 - \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} z_1^k z_2^l \tag{4}$$

is not zero for any z_1 and z_2 which simultaneously satisfy $|z_1| \leq 1$ and $|z_2| \leq 1$. Such a process is called a causal autoregressive process of order (p_1, p_2) and denoted by AR(p_1, p_2). The number of coefficients in the model (1), not counting $\alpha_{00} = 0$, is given by $d(p) = (p_1 + 1)(p_2 + 1) - 1$. We can write the spatial analogue of one-dimensional Yule-Walker equations as

$$R(s,t) = \sum_{k,l \in S_p} \alpha_{kl} R(s-k, t-l) \quad (s \geq 0, t \geq 0), \tag{5}$$

where $S_p = \{(i,j) : 0 \leq i \leq p_1, 0 \leq j \leq p_2\}$. If we let $T_p = S_p - \{(0,0)\}$, we can write (5) for all $(s,t) \in T_p$ as $R\alpha = r$, where R is a $d(p) \times d(p)$ matrix and α and r are vectors of length $d(p)$. Throughout this paper, we will use the same ordering of elements into the matrices and vectors as that used by Ha and Newton(1993). The α vector is thus $\alpha = (\alpha_{01}, \cdots, \alpha_{0,p_2}, \cdots, \alpha_{p1,0}, \cdots, \alpha_{p_1,p_2})$. For $s \geq 0$ and $t \geq 0$, define

$$\hat{R}(s,t) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=1}^{n-t} Y_{i,j} Y_{i+s,t+j}, \ \hat{R}(s,-t) = \frac{1}{mn} \sum_{i=1}^{m-s} \sum_{j=t+1}^{n} Y_{i,j} Y_{i+s,j-t}, \tag{6}$$

which are the sample autocovariances at lags (s,t) and $(s,-t)$, respectively. Note that $\hat{R}(s,t) = \hat{R}(-s,-t)$ and $\hat{R}(-s,t) = \hat{R}(s,-t)$. Tjostheim(1978) showed that the sample autocovariances are consistent estimator of the autocovariance function defined in (1). When we solve the Yule-Walker equations with the $\hat{R}$'s replacing the R's, we obtain the Yule-walker estimators, denoted by $\hat{\alpha}$. If instead of $\hat{R}$ we use

$$\tilde{R}(s,t) = \{mn/(m-s)(n-t)\}\hat{R}(s,t) \ (s,t \geq 0), \tag{7}$$

we obtain what we will call the 'unbiased Yule-Walker estimators', denoted by $\tilde{\alpha}$. The least squares estimator is defined as the estimator α which minimizes

$$\sum_{i=1}^{m+p_1} \sum_{j=1}^{n+p_2} \left(Y_{i,j} - \sum_{k=0}^{p_1} \sum_{l=0}^{p_2} \alpha_{kl} Y_{i-k,j-l} \right)^2 \tag{8}$$

where any unobserved Y is replaced by zero. We denote the least squares estimators by $\bar{\alpha}$. Ha and Newton(1993) showed that if $m/n \to c^2 (0 < c < \infty)$ as $m, n \to \infty$, the Yule-Walker estimator is consistent and has an asymptotic normal distribution with mean $b(\alpha) = R^{-1} R^* \alpha$ and covariance matrix R^{-1}. The R^* is the function of c and R and the exact formula of R^* is in Ha and Newton(1993). The unbiased Yule-Walker and least squares estimators have the same asymptotic properties as Yule-Walker estimator except that the asymptotic bias is zero(Ha and Newton(1993)).

3. Simultaneous Confidence Regions for Spectral Densities

Suppose Y is a stable causal spatial AR(p_1, p_2) process. We can write the spectral density of the process as

$$f(\omega_1, \omega_2) = \frac{1}{4\pi^2} \frac{1}{|g(e^{i\omega_1}, e^{i\omega_2})|}, \quad (\omega_1, \omega_2) \in I_\omega \tag{9}$$

The reciprocal of $f(\omega_1, \omega_2)$ can be written by

$$\begin{aligned} h(\omega_1, \omega_2) &= \frac{1}{f(\omega_1, \omega_2)} \\ &= 4\pi^2 - 8\pi^2 \sum_{j,k \in T_p} \alpha_j k \cos(j\omega_1 + k\omega_2) \\ &\quad +4\pi^2 \sum_{j,k \in T_p} \sum_{l,m \in T_p} \alpha_{jk}\alpha_{lm} \cos[(j-l)\omega_1 + (k-m)\omega_2] \\ &= 4\pi^2[1 + X^T(\alpha, \omega_1, \omega_2)\alpha], \end{aligned} \tag{10}$$

where $X(\alpha, \omega_1, \omega_2)$ and α are vectors of length $d(p) = (p_1+1)(p_2+1) - 1$. The element of $X(\alpha, \omega_1, \omega_2)$ corresponding to the α_{jk}, is

$$X_{jk}(\alpha, \omega_1, \omega_2) = 2\cos(j\omega_1 + k\omega_2) + \sum_{l,m} \alpha_{lm} \cos[(j-l)\omega_1 + (k-m)\omega_2]. \tag{11}$$

In a theorem below we use three estimators $\hat{\alpha}, \tilde{\alpha}, \bar{\alpha}$ of the coefficients α for the causal spatial AR(p_1, p_2). Then the causal spatial autoregressive spectral estimator $\hat{f}(\omega_1, \omega_2)$ of $f(\omega_1, \omega_2)$ consists of evaluating (9) with $\hat{\alpha}, \tilde{\alpha}, \bar{\alpha}$ replacing α. Then asymptotic simultaneous confidence bands for the spectral density function using Scheffe projections are given by the following theorem.

Theorem Let $\{Y_{i,j} : i = 1, \cdots, m, j = 1, \cdots, n\}$ be data from a causal autoregressive process of order (p_1, p_2) and let $\hat{\alpha}$ be the Yule-Walker estimators defined in Section 2. Let

$$\hat{h}(\omega_1, \omega_2) = \frac{1}{\hat{f}(\omega_1, \omega_2)} = 4\pi^2[1 + X^T(\hat{\alpha}, \omega_1, \omega_2)\hat{\alpha}]. \tag{12}$$

Then if $m/n \to c^2 (0 < c < \infty)$ as $m, n \to \infty$, we have the following
(a) The probability is at least $1-\beta$ that simultaneously for all (ω_1, ω_2) in the interval $(\omega_1, \omega_2) \in I_\omega = [-\pi, \pi] \times [-\pi, \pi]$,

$$\frac{1}{\hat{h}(\omega_1, \omega_2) + s(\omega_1, \omega_2)} \le f(\omega_1, \omega_2) \le \frac{1}{\hat{h}(\omega_1, \omega_2) - s(\omega_1, \omega_2)}, \tag{13}$$

where if $\hat{h}(\omega_1, \omega_2) - s(\omega_1, \omega_2) \leq 0$ we use infinity as the upper limit, and

$$s^2(\omega_1, \omega_2) = \frac{\chi^2_{\beta,d(p)}(\hat{\lambda})}{mn} X^T(\hat{\alpha}, \omega_1, \omega_2)\hat{R}^{-1}X(\hat{\alpha}, \omega_1, \omega_2) \tag{14}$$

with $\chi^2_{\beta,d(p)}(\hat{\lambda})$ being the upper β critical value of a noncentral chi-squared distribution having $d(p)$ degree of freedom and noncentrality parameter $\hat{\lambda} = b^T(\hat{\alpha})\hat{R}b(\hat{\alpha})$. The $b(\alpha)$ and R is defined in Section 2.

(b) If we use the unbiased Yule-walker estimators and least squares estimators, we have the same asymptotic properties as in (a) except that the noncetrality parameter is zero.

Proof: (a) Since $\hat{\alpha}$ is consistent estimator of α , using Slutsky's theorem(Rac, 1973, p. 122)

$$\sqrt{mn}(\hat{h}(\hat{\alpha}, \omega_1, \omega_2) - h(\alpha, \omega_1, \omega_2)) = 4\pi^2\sqrt{mn}(X^T(\hat{\alpha}, \omega_1, \omega_2)\hat{\alpha} - X^T(\alpha, \omega_1, \omega_2)\alpha), \tag{15}$$

has the same asymptotic distribution as

$$4\pi^2\sqrt{mn}(X^T(\alpha, \omega_1, \omega_2)\hat{\alpha} - X^T(\alpha, \omega_1, \omega_2)\alpha). \tag{16}$$

Since $\sqrt{mn}(\hat{\alpha}-\alpha)$ has an asymptotic normal distribution with mean $b(\alpha) = R^{-1}R^*c$ and covariance matrix R^{-1}, $mn(\hat{\alpha} - \alpha)^T R(\hat{\alpha} - \alpha)$ is asymptotically noncentral chi-square variable with degree of freedom $d(p)$ and noncentrality parameter $\lambda = b^T(\alpha)Rb(\alpha)$. Since $\hat{R}$ is the consistent estimator of R, $mn(\hat{\alpha} - \alpha)^T\hat{R}(\hat{\alpha} - \alpha)$ also has the same asymptotic distribution as $mn(\hat{\alpha}-\alpha)^T R(\hat{\alpha}-\alpha)$. Thus the probability is asymptotically $1 - \beta$ that the true parameter α lies inside the ellipsoid defined as the set of vectors $(\hat{\alpha} - l)D(\lambda)(\hat{\alpha} - l) \leq 1$, where $D(\lambda) = mn\hat{R}/\chi^2_{d(p)}(\lambda)$. Since $b(\hat{\alpha})$ is consistent estimator of $b(\alpha)$, we also have $(\hat{\alpha} - l)D(\hat{\lambda})(\hat{\alpha} - l) \leq 1$, where $D(\hat{\lambda}) = mn\hat{R}/\chi^2_{d(p)}(\hat{\lambda})$. But (Scheffe, 1957, p. 407), α is in this ellipsoid if and only if $|Z^T(\hat{\alpha} - \alpha)| \leq (Z^T D^{-1}(\hat{\lambda})Z)^{1/2}$ for all $d(p)$ dimensional vectors Z. Thus, in particular only if $|X^T(\hat{\alpha}, \omega_1, \omega_2)(\hat{\alpha} - \alpha)| \leq (X^T(\alpha, \omega_1, \omega_2)D^{-1}(\hat{\lambda})X(\alpha, \omega_1, \omega_2))^{1/2}$ for all $d(p)$ dimensional vectors Z. Since $\hat{\alpha}$ is consistent estimator of α, we thus have

$$|X^T(\hat{\alpha}, \omega_1, \omega_2)\hat{\alpha} - X^T(\alpha, \omega_1, \omega_2)\alpha| \leq (X^T(\hat{\alpha}, \omega_1, \omega_2)D^{-1}(\hat{\lambda})X(\hat{\alpha}, \omega_1, \omega_2))^{1/2}. \tag{17}$$

This gives a $1 - \beta$ simultaneous confidence bands for

$$h(\omega_1, \omega_2) = 1/f(\omega_1, \omega_2) = 4\pi^2 + 4\pi^2 X^T(\alpha, \omega_1, \omega_2)\alpha, \tag{18}$$

because

$$\hat{h}(\omega_1, \omega_2) - s(\omega_1, \omega_2) \leq h(\omega_1, \omega_2) \leq \hat{h}(\omega_1, \omega_2) + s(\omega_1, \omega_2), \tag{19}$$

the reciprocal of which is the simultaneous confidence band for $f(\omega_1, \omega_2)$. It could happen that $\hat{h}(\omega_1, \omega_2) - s(\omega_1, \omega_2) < 0$ for some (ω_1, ω_2) in which case one can use

infinity for the upper limit which does not diminish the probability content of the bands.
(b) Since the unbiased Yule-Walker estimators $\tilde{\alpha}$ and the least squares estimators $\bar{\alpha}$ have the same asymptotic properties as the Yule-Walker estimators $\hat{\alpha}$ except that the asymptotic bias $b(\alpha)$ is zero, the part (b) in the theorem is satisfied.

4. Numerical Examples for Spatial AR(1,1) Process

Let Y be an causal spatial autoregressive process of order (1,1) with coefficients $\alpha_{01}, \alpha_{10}, \alpha_{11}$, that is

$$Y_{i,j} = \alpha_{01} Y_{i,j-1} + \alpha_{10} Y_{i-1,j} + \alpha_{11} Y_{i-1,j-1} + \epsilon_{ij} \tag{20}$$

where ϵ is a white noise series of uncorrelated zero mean random variables having unit variance. The reciprocal of spectral density function is

$$\begin{aligned} h(\omega_1, \omega_2) &= \frac{4\pi^2}{f(\omega_1, \omega_2)} \\ &= 4\pi^2[1 + X^T(\alpha, \omega_1, \omega_2)\alpha], \end{aligned} \tag{21}$$

where the α vector is $\alpha^T = (\alpha_{01}, \alpha_{10}, \alpha_{11})$ and X vector is

$$X(\alpha, \omega_1, \omega_2) = \begin{bmatrix} -2\cos(\omega_2) + 2\alpha_{10}\cos(\omega_1 - \omega_2) + \alpha_{01} \\ -2\alpha\cos(\omega_1) + 2\alpha_{11}\cos(\omega_2) + \alpha_{10} \\ -2\cos(\omega_1 + \omega_2) + 2\alpha_{01} + \alpha_{11} \end{bmatrix}. \tag{22}$$

We thus have the simultaneous confidence bands of spectral density of level 0.95 for all (ω_1, ω_2) in the interval $(\omega_1, \omega_2) \in I_\omega = [-\pi, \pi] \times [-\pi, \pi]$ with Yule-Walker estimators $\hat{\alpha}$, where

$$s^2(\omega_1, \omega_2) = \frac{\chi^2_{0.05,3}(\hat{\lambda})}{mn} X^T(\hat{\alpha}, \omega_1, \omega_2)\hat{R}^- 1 X(\hat{\alpha}, \omega_1, \omega_2) \tag{23}$$

with $\chi^2_{0.05,3}(\hat{\lambda})$ being the upper 0.05 critical value of a noncentral chi-squared distribution having 3 degree of freedom and noncentrality parameter $\hat{\lambda} = b^T(\hat{\alpha})\hat{R}b(\hat{\alpha})$. To complete the simultaneous confidence bands for spectral density, we need to compute the estimated noncetrality parameter $\hat{\lambda} = b^T(\hat{\alpha})\hat{R}b(\hat{\alpha})$, which is the estimator of true noncentrality parameter $\lambda = b^T(\alpha)Rb(\alpha)$. Ha and Newton(1993) showed the explicit formula of $b(\alpha)$, so we can obtain the estimator of noncentrality parameter by plugging $\hat{\alpha}$ into α and using $\hat{R}$ instead of R. We also can similarly construct the simultaneous confidence bands of spectral density using the unbiased Yule-Walker estimators and least squares estimators. To investigate the finite sample behaviour of the proposed confidences bands we first considered their performance on 100 sample realizations for each of three estimators $\hat{\alpha}, \tilde{\alpha}, \bar{\alpha}$ using the $\alpha_{01} = 0.3, \alpha_{10} = 0.5, \alpha_{11} = 0.1$. This model is used because the magnitude of the biases of this model are larger and changes(with changing c) are also larger. Given the parameter estimators we evaluated the true spectra, the estimated spectra, and the estimated upper and lower 0.95 confidence bands at the 31 equally spaced frequencies between $\omega_1 \in [-\pi, \pi]$ and the 31 equally spaced frequencies between $\omega_2 \in [-\pi, \pi]$. We then counted the number of times the true spectra was contained in these estimated bands at all frequencies. For the large sample sizes, the simulation results was consistent with the stated confidence level. Next we look at some simulation results for moderate grid sizes to see the effect of bias of the Yule-Walker estimators. The simulations are performed for the grid sizes $(m, n) = (20, 20), (16, 20), (15, 25), (12, 30), (15, 60)$ to represent a range of $0.5 \le c \le 1.0$. Table 1 gives the values of the asymptotic biases) $b(\alpha) = (b_{01}, b_{10}, b_{11})$, $c^2 = \lim_{m,n \to \infty} m/n$, the asymptotic noncentrality parameter λ, and the frequency of 0.95 confidence band coverage for the Yule-Walker estimators using 100 sample realizations. We can see from Table 1 that the biases are larger when has a smaller value. The noncetrality parameter also large when the bias is

Table 1: Frequency of 0.95 confidence band coverage

(m, n)	c	b_{01}	b_{10}	b_{11}	λ	Frequency
(20,20)	1.000	0.590	-0.030	-1.049	1.455	98
(16,20)	0.894	0.726	-0.139	-1.092	1.747	95
(15,25)	0.775	0.916	-0.282	-1.169	2.323	97
(12,30)	0.632	1.217	-0.494	-1.319	3.606	95
(15,60)	0.500	1.629	-0.765	-1.561	6.055	90

large. When the bias is not large, the simulation results for moderate grid sizes is reasonably good although not excellent. For the grid size (15, 60), the bias is large compared to the other grid sizes and the frequency of 0.95 confidence band coverage is 90 which does not attain the stated confidence level for simultaneous confidence bands of spectral density. In calculating the confidence bands, for several of the simulated series, $\hat{h}(\omega_1, \omega_2) - s(\omega_1, \omega_2)$ was negative for several frequencies. This behaviour may be indicating the existence of deterministic components in which case the spectral density does not exist at these frequencies.

References

1. Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, Tokyo, 21, 243-7.

2. Brockwell, P. J. and Davis, R. A. (1990). *Time Series: Theory and Methods, 2nd edition,* New York, Springer-Verlag

3. Ensor, K. B. and Newton, H. J. (1988). The effect of order estimation on estimating the peak frequency of an autoregressive spectral density. *Biometrika,* 75, 587-589.

4. Ha, E. and Newton, H. J. (1993). The bias of estimators of causal spatial autoregressive processes. *Biometrika,* 80, 242-5.

5. Mackisack, M. S. and Poskitt, D. S. (1989). Autoregressive frequency estimation. *Biometrika,* 76, 565-576.

6. Newton, N. H. (1988). *TIMESLAB: A Time Series Analysis Laboratory,* Pacific Grove, California, Wadsworth and Brooks/Cole.

7. Newton, H. J. and Pagano, M. (1983). A method for determining periods in time series. *Journal of the American Statistical Association,* 78, 152-157.

8. Newton, H. J. and Pagano, M. (1984). Simultaneous confidence bands for autoregressive spectra. *Biometrika,* 70, 197-202.

9. Parzen, E. (1974). Recent advances in time series modelling. *I.E.E.E. Trans. Auto. Control,* 19, 723-30.

10. Priestley, M. B. (1981). *Spectral Analysis of Time Series,* New York, Academic Press

11. Rao, C. R. (1973). *Linear Statistical Inference and its Applications, 2nd edition.* New York, Wiley.

12. Ripley, B. D. (1981). *Spatial Statistics,* New York, Wiley

13. Scheffe, H. (1959). *The Analysis of Variance,* New York, Wiley.

14. Tjostheim, D. (1981). Autoregressive modelling and spectral analysis of array data in the plane. *IEEE Transactions in Geoscience Remote Sensing,* GE-19, 15-24.

15. Ulrych, T. L. and Bishop, T. N. (1975). Maximum entropy spectral analysis and autoregressive decomposition. *Review of Geophysics and Space Physics,* 13, 183-200.