ASYMPTOTIC STATISTICAL PROPERTIES OF AUTOREGRESSIVE MODEL FOR MIXED SPECTRUM ESTIMATION

Peter J. Sherman and Soon-Seng Lau

Iowa State University, Ames, IA 50010, U.S.A

ABSTRACT

This work addresses the influence of point spectrum on large sample statistics of the autoregressive spectral estimator. In particular, the asymptotic distributions of the AR coefficients, the innovations variance, and the spectral density estimator of a finite order AR(p) model for a mixed spectrum process are presented. Numerical simulations are performed to verify the analytical results.

1. INTRODUCTION

A mixed random process is one which includes both regular and a deterministic components. The most popular class of such processes is the class in which the deterministic component is a sum of sinusoids. This type of random process is commonplace in applications involving periodic phenomena. Examples include communication systems, biomedical signal processing, atmospheric sciences, and rotating machinery. In each example, one can easily identify particular cases of critical importance. For example, condition monitoring of a high speed turbomachinery is essential for ensuring safe operation of an aircraft.

In spite of the mixed nature of such processes, the vast majority of frequency domain analysis methods, such as FFT, AR, and ARMA methods rely on the erroneous assumption of a purely regular process. Autoregressive (AR) spectral analysis, in particular, has long been employed to model regular stationary random processes. For such processes, its statistical properties have been well studied. For example, Brockwell and Davis [2] used linear regression to establish various limiting results for AR model coefficients. Caines [3] applied convergence properties of martingale differences to derive a similar results. Percival and Walden [9], and Baggeroer [1] established confidence intervals for AR spectral estimator.

Several attempts had been made to generalize the properties of AR spectral analysis to a mixed spectrum process. Sakai [10] derived the asymptotic variance of the fixed p-th order AR spectral estimator by assuming that the estimation error is small. Mackisak and Poskitt [7] considered the case of a single sinusoid embedded in white noise. They showed that for fixed p, the least squares p-th order AR spectral estimator converges almost surely to the theoretical AR(p) spectrum associated with exact correlation interval.

However, one might well inquire as to the value of this result, since in the region of point spectrum the AR(p) theoretical spectrum is ill-conditioned, in the sense that it becomes unbounded as $p \to \infty$ [4]. One answer to this question is that if one has identified the point spectrum frequencies, then the AR(p) spectrum may still yield valuable information sufficiently far from these frequencies. In fact, as $p \to \infty$ the AR(p) spectrum converges almost everywhere to the continuous spectral density [5]. (As a side point, it should be noted that all traditional model order selection rules are inappropriate for AR models in this mixed spectrum setting.) Perhaps a more important reason for investigating the properties of AR spectral estimators in the mixed spectrum setting has to do with identification of point spectrum. Since the theoretical AR(p) spectrum has such distinctly different behavior at point spectrum frequencies, it may be possible to take advantage of this for detection of point spectrum. This could be attempted directly, or in conjunction with the associated MV(p) spectrum [12]. In any case, a necessary first step for use of the AR(p) spectral estimator in the mixed spectrum setting, either to estimate the power spectral density in a given region, or to detect point spectrum, is to arrive at its large sample distributional properties. As noted above, [7] has obtained such a result for white noise. Here we extend it to the colored noise situation.

In this work, a mixed spectrum process is defined as $y_n = x_n + \varepsilon_n$, where

$$x_n = \sum_{m=1}^{q} A_m \cos(\omega_m n + \phi_m), \qquad \varepsilon_n = \sum_{j=-\infty}^{\infty} \psi_j \zeta_{n-j}$$

where $\{A_m\}_{m=1}^{q}$ and $\{\omega_m\}_{m=1}^{q} \in (o, \pi)$ are unknown constants, $\{\zeta_n\}_{n=-\infty}^{\infty} \sim i.i.d.(0, \sigma_\zeta^2)$ with $E\{\zeta_n^4\} = \kappa\sigma_\zeta^4 < \infty$, $\sum |\psi_j| < \infty$, $\{\phi_m\}_{m=1}^{q} \sim i.i.d.\mathrm{U}(0, 2\pi]$ and independent of $\{\zeta_n\}$. The theoretical autocovariance of y_n is given by $r_\tau^y = r_\tau^x + r_\tau^\varepsilon$, where

$$r_\tau^x = \sum_{m=1}^{q} \frac{A_m^2}{2} \cos(\omega_m \tau), \qquad r_\tau^\varepsilon = \sigma_\zeta^2 \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+\tau}$$

are the autocovariances of x_n and ε_n, respectively. The p-th order autoregressive predictor of y_n is

$$\hat{y}_n = -\sum_{k=1}^{p} a_k \cdot y_{n-k}$$

with the prediction error variance $E\{(\hat{y}_n - y_n)^2\} = \sigma_p^2$. The minimum variance parameters $\boldsymbol{a}_p' = [a_1, \; a_2, \; \cdots, \; a_p]$ which minimize σ_p^2 are obtained by solving the Yule-Walker equations: $\boldsymbol{a}_p = -\boldsymbol{R}_y^{-1}\boldsymbol{r}^y$, and $\sigma_p^2 = r_0^y + \boldsymbol{r}^{y\prime}\boldsymbol{a}_p$, where the $(p\times p)$ toeplitz matrix $\boldsymbol{R}_y = [r_{i-j}^y]_{i,j=1,\cdots,p}$ and the $(p\times 1)$ vector $\boldsymbol{r}^y = [r_1^y, \cdots, r_p^y]'$. Let $S_p(\omega)$ denote the minimum variance AR(p) spectrum and $\hat{S}_p(\omega)$ denote its least squares estimator. Then

$$S_p(\omega) = \frac{\sigma_p^2}{\left|\sum_{k=0}^{p} a_k e^{-ik\omega}\right|^2} \triangleq \frac{\sigma_p^2}{|\rho_p(\omega)|^2}$$

$$\hat{S}_p(\omega) = \frac{\hat{\sigma}_p^2}{\left|\sum_{k=0}^{p} \hat{a}_k e^{-ik\omega}\right|^2} \triangleq \frac{\hat{\sigma}_p^2}{|\hat{\rho}_p(\omega)|^2}$$

where $\hat{\boldsymbol{a}}_p$ and $\hat{\sigma}_p^2$ are the least squares estimators of $\boldsymbol{a}_p$ and σ_p^2, respectively, obtained by substituting r_τ^y by the sample autocovariance function defined by:

$$\hat{r}_\tau^y = \frac{1}{N}\sum_{n=1}^{N-\tau} y_n \cdot y_{n+\tau} \quad .$$

In the next section, the large sample distributions of $\hat{\boldsymbol{a}}_p$, $\hat{\sigma}_p^2$, and $\hat{S}_p(\omega)$ will be presented.

2. THEORETICAL RESULTS

The key to our analysis lies in the delta method for asymptotic analysis [2] and the limiting distribution of the sample autocovariance function [6]: Let $\boldsymbol{r}_p^y = [r_0^y, \cdots, r_p^y]'$, and $\hat{\boldsymbol{r}}_p^y$ be the corresponding finite sample estimator. Then $\sqrt{N}(\hat{\boldsymbol{r}}_p^y - \boldsymbol{r}_p^y) \xrightarrow{d} \mathrm{N}(0, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = [\sigma_{jk}]_{j,k=0,\dots,p}$ is:

$$\sigma_{jk} = \sum_{m=1}^{q} 2A_m^2 \cos(j\omega_m)\cos(k\omega_m) \sum_{\tau=-\infty}^{\infty} r_\tau^\varepsilon \cos(\omega_m \tau) + (\kappa-3) r_j^\varepsilon r_k^\varepsilon + \sum_{\tau=-\infty}^{\infty} \left\{ r_\tau^\varepsilon r_{\tau+j-k}^\varepsilon + r_{\tau+j}^\varepsilon r_{\tau-k}^\varepsilon \right\}$$

Our first result is an equivalent frequency domain expression for the variance-covariance matrix $\boldsymbol{\Sigma}$.

Lemma 1. *If ζ_n is Gaussian white noise ($\kappa = 3$), the frequency domain expression for the variance-covariance matrix $\boldsymbol{\Sigma}$ of the limiting distribution of $\sqrt{N}(\hat{\boldsymbol{r}}_p^y - \boldsymbol{r}_p^y)$ is given by:*

$$\boldsymbol{\Sigma} = \frac{1}{\pi}\int_{-\pi}^{\pi} \Big\{ \big[2S_x(\omega) + S_\epsilon(\omega)\big] S_\epsilon(\omega) \boldsymbol{\Gamma}(\omega) \Big\} d\omega$$

where,

$$S_x(\omega) = 2\pi \sum_{m=1}^{q} \frac{A_m^2}{4}\delta(\omega \pm \omega_m)$$

$$S_\epsilon(\omega) = \sum_{\tau=-\infty}^{\infty} r_\tau^\epsilon e^{-i\tau\omega}$$

$$\boldsymbol{\Gamma}(\omega) = \boldsymbol{\gamma}(\omega)\boldsymbol{\gamma}(\omega)'$$

$$\boldsymbol{\gamma}(\omega)' = [1, \cos(\omega), \cdots, \cos(p\omega)].$$

The large sample distribution of $\hat{\boldsymbol{a}}_p$ is given in Theorem 2. Since it is done in a mixed spectrum setting, it presents an extension of the corresponding result for continuous spectrum processes (see [2, 3, 9]).

Theorem 2.

$$\sqrt{N}(\hat{\boldsymbol{a}}_p - \boldsymbol{a}_p) \xrightarrow{d} \mathrm{N}(\,\underline{0}\,,\ \boldsymbol{\Psi}_{\hat{\boldsymbol{a}}_p})$$

where, $\boldsymbol{\Psi}_{\hat{\boldsymbol{a}}_p} = \boldsymbol{R}_y^{-1}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'\boldsymbol{R}_y^{-1}$, and

$$\boldsymbol{A} \triangleq \begin{bmatrix} a_1 & a_2 & \cdots & a_p & 0 \\ a_2 & a_3 & \cdots & 0 & 0 \\ a_3 & a_4 & \cdots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{p-1} & a_p & \cdots & 0 & 0 \\ a_p & 0 & \cdots & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_0 & 0 & \cdots & 0 \\ 0 & a_1 & a_0 & \cdots & 0 \\ 0 & a_2 & a_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & a_{p-2} & a_{p-3} & \cdots & 0 \\ 0 & a_{p-1} & a_{p-2} & \cdots & a_0 \end{bmatrix}.$$

The limiting distribution of the error variance estimator $\hat{\sigma}_p^2$ is given below.

Theorem 3. *Defining $\boldsymbol{\nu} = [g_0^a, 2g_1^a, \cdots, 2g_p^a]$, where g_τ^a is the convolution of $[\cdots, 0, a_0, \cdots, a_p, 0, \cdots]$ with it's reverse at lag τ. Then,*

$$\sqrt{N}\,(\hat{\sigma}_p^2 - \sigma_p^2) = \boldsymbol{\nu}'\sqrt{N}(\hat{\boldsymbol{r}}_p^y - \boldsymbol{r}_p^y) + o_p(1) \qquad (1)$$

Hence,

$$\sqrt{N}\,(\hat{\sigma}_p^2 - \sigma_p^2) \xrightarrow{d} \mathrm{N}(\,0\,,\ \Psi_{\hat{\sigma}_p^2}\,)$$

where $\Psi_{\hat{\sigma}_p^2} = \boldsymbol{\nu}'\boldsymbol{\Sigma}\boldsymbol{\nu}$. If ζ_n is Gaussian white noise, the frequency domain expression for $\Psi_{\hat{\sigma}_p^2}$ can be expressed by using Lemma 1 as:

$$\Psi_{\hat{\sigma}_p^2} = \frac{\sigma_p^4}{\pi}\int_{-\pi}^{\pi} \Big\{ \big[2S_x(\omega) + S_\epsilon(\omega)\big] S_\epsilon(\omega) S_p^{-2}(\omega) \Big\} d\omega$$

Theorems 2 and 3 are generalizations of the results in [2, 3, 9]. For the case where $x_n = 0$ and ε_n is a causal AR(q) process of order $q \le p$, Theorem 2 reduces to $\sqrt{N}(\hat{\boldsymbol{a}}_p - \boldsymbol{a}_p) \xrightarrow{d} \mathrm{N}(\underline{0}, \sigma_\zeta^2 \boldsymbol{R}_y^{-1})$. If, in addition, ε_n is causal Gaussian AR(q), Theorem 3 reduces to $\sqrt{N}(\hat{\sigma}_p^2 - \sigma_\zeta^2) \xrightarrow{d} \mathrm{N}(0, 2\sigma_\zeta^4)$.

The previous results set the stage for our ultimate goal of this work, which is the asymptotic distribution of the AR spectral estimator $\hat{S}_p(\omega)$. To this end, we arrived at the following lemma.

Lemma 4.

$$\sqrt{N}\left[|\hat{\rho}_p(\omega)|^2 - |\rho_p(\omega)|^2\right] = 2\left[\boldsymbol{A}\boldsymbol{\gamma}(\omega)\right]' \sqrt{N}(\hat{\boldsymbol{a}}_p - \boldsymbol{a}_p) + o_p(1) \qquad (2)$$

Thus, $\sqrt{N}\left[|\hat{\rho}_p(\omega)|^2 - |\rho_p(\omega)|^2\right] \xrightarrow{d} \mathrm{N}(0, \Psi_{|\hat{\rho}_p|^2})$, *where,*

$$\Psi_{|\hat{\rho}_p|^2} = 4\left[\boldsymbol{A}\boldsymbol{\gamma}(\omega)\right]' \boldsymbol{R}_y^{-1} \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}' \boldsymbol{R}_y^{-1} \left[\boldsymbol{A}\boldsymbol{\gamma}(\omega)\right]$$

Using (1) and (2), we arrived at the following result.

Theorem 5.

$$\sqrt{N}\left[\hat{S}_p(\omega) - S_p(\omega)\right] \xrightarrow{d} \mathrm{N}\left(\, 0 \,,\ \Psi_{\hat{S}_p} \,\right)$$

where,

$$\Psi_{\hat{S}_p} = \frac{S_p^2(\omega)}{\sigma_p^4}\left\{\Psi_{\hat{\sigma}_p^2} + S_p^2(\omega)\Psi_{|\hat{\rho}_p|^2} - 2S_p(\omega)\Psi_{\hat{\sigma}_p^2, |\hat{\rho}_p|^2}\right\}$$
$$\Psi_{\hat{\sigma_p}^2, |\hat{\rho}_p|^2} = -2\boldsymbol{\nu}'\boldsymbol{\Sigma}\boldsymbol{A}'\boldsymbol{R}_y^{-1}\boldsymbol{A}\boldsymbol{\gamma}(\omega)$$

Theorem 5 is the first result we know of that allows one to study the theoretical statistical behavior of the AR(p) spectral estimator at signal frequencies. Unfortunately, it does not offer immediate insight into this behavior. Therefore, at this stage we will examine Theorem 5 via various simulations.

3. NUMERICAL SIMULATION

In this section, we verify our results via extensive numerical simulations. Consider the mixed spectrum process $y_n = x_n + \varepsilon_n$, where

$$\begin{aligned} x_n &= \cos\left(\frac{n\pi}{2}\right) + \cos\left(\frac{3n\pi}{4}\right) \\ \varepsilon_n &= 0.3\varepsilon_{n-1} - 0.9\varepsilon_{n-2} + \zeta_n \\ \zeta_n &\sim i.i.d.\mathrm{N}(0,1) \end{aligned}$$

The signal-to-noise ratio equals to -7.3221 dB in this case. Consider fitting AR(p) models of order $p = 5$ and $p = 50$ to 1000 realizations of the process, each with $N = 1000$ samples. The simulation and theoretical means and variances of $\hat{\sigma}_p^2$ and of the first three elements of $\hat{\boldsymbol{a}}_p^2$ are tabulated in Table 1 and Table 2, respectively. In either case, the theoretical values are enclosed in parenthesis. These results show small deviations between the simulation and theoretical values and verify the theoretical results.

The estimated means of $\hat{S}(\omega)$ (dashed line) for $p = 5$ and $p = 50$ are plotted against the theoretical means (solid line) in Figure 1 and Figure 2, respectively. That for the estimated variances are plotted in Figure 3 and Figure 4. Again, there is a close match between the numerical and theoretical spectral means and variance.

Table 1: Estimated and theoretical means and variances of innovations variance.

AR order	Innovations variance, $\hat{\sigma}^2$	
p	$E\{\hat{\sigma}^2\}$	$N \cdot Var\{\hat{\sigma}^2\}$
5	1.638	6.085
	(1.641)	(5.420)
50	1.007	2.464
	(1.062)	(2.304)

Table 2: Estimated and theoretical means and variances of AR coefficients.

AR order	AR coefficients, $\hat{\boldsymbol{a}}$			
p	$E\{\hat{\boldsymbol{a}}\}$	$N \cdot Var\{\hat{\boldsymbol{a}}\}$		
5	0.092	0.535	0.160	0.402
	(0.092)	(0.497)	(0.143)	(0.370)
	0.814	0.160	0.541	0.386
	(0.815)	(0.143)	(0.525)	(0.371)
	-0.015	0.402	0.386	0.778
	(-0.016)	(0.370)	(0.371)	(0.729)
50	-0.273	1.229	-0.347	1.151
	(-0.272)	(1.016)	(-0.272)	(0.907)
	0.916	-0.347	1.233	-0.639
	(0.913)	(-0.272)	(1.060)	(-0.505)
	-0.010	1.151	-0.639	2.286
	(-0.009)	(0.907)	(-0.505)	(1.890)

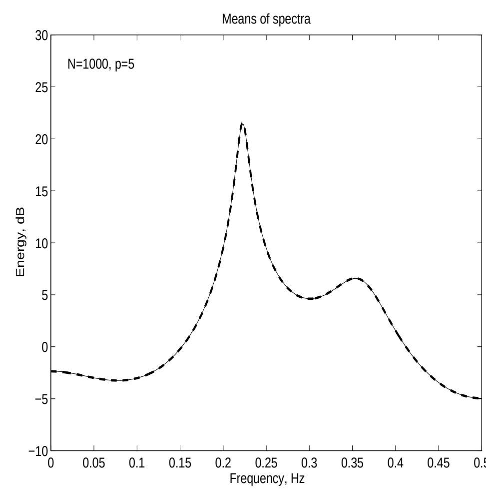

Figure 1: Simulation and theoretical spectral means for $p = 5$.

4. CONCLUSION

The large sample properties of an AR estimator for a regular process with an unspecified continuous spectrum were generalized to those for a stochastic process containing a mixed spectrum. In the process, asymptotic normality of

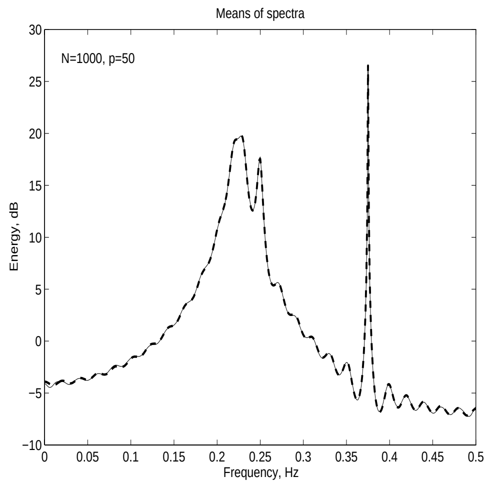

Figure 2: Simulation and theoretical spectral means for $p = 50$.

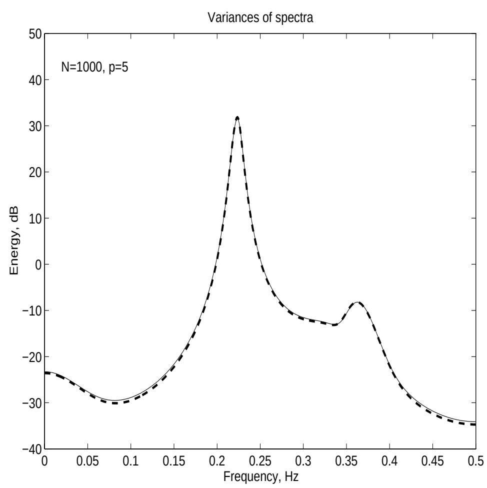

Figure 3: Simulation and theoretical spectral variances for $p = 5$.

the AR coefficients, innovation variances, and AR spectral estimate of a finite order autoregression were derived. Numerical simulations verified the reliability of the analytical results.

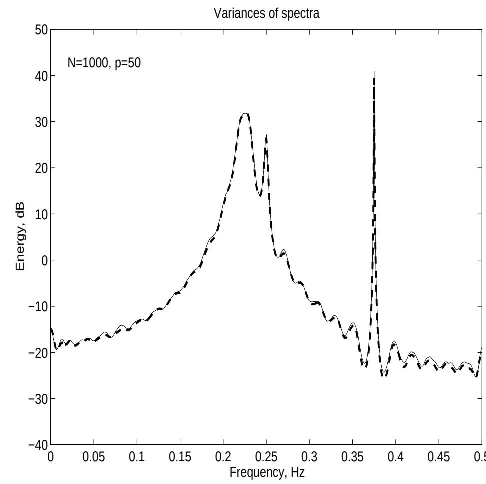

Figure 4: Simulation and theoretical spectral variances for $p = 50$.

5. REFERENCES

[1] Baggeroer, A.B. “Confident interval for regression (MEM) spectral estimates.” *IEEE Trans. IT*, 22, 534–545, 1976.

[2] Brockwell, P.J. and Davis, R.A. *Time Series: Theory and Methods.* 2nd ed., Springer-Verlag, New York, 1991.

[3] Caines, P.E. *Linear Stochastics Systems.* John Wiley & Sons, Inc. 1988.

[4] Foias, C., Frazho, A., and Sherman, P. “A geometric approach to the maximum likelihood spectral estimator for sinusoids in noise.” *IEEE Trans. IT*, 34(5), 1066–1070, Sept. 1988.

[5] Grenander, V. and Szego, G. *Toeplitz Forms and Their Applications.* Univ. of California Press, Berkeley, 1958.

[6] Li, T.H., Kedem, B. and Yakowitz, S. “Asymptotic normality of sample autocovariances with an application in frequency estimation.” *Stoc. Process, Appl.*, 52, 329–349, 1994.

[7] Mackisack, M. and Poskitt, D.S. “Some properties of autoregressive estimates for processes with mixed spectra.” *J. Time Series Anal.*, 11(4), 325–337, 1990.

[8] Marple, L.S. *Digital Spectral Analysis with Applications.* Prentice Hall, 1987.

[9] Percival, D.B. and Walden, A.T. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques.* Cambridge University Press, Great Britain, 1993.

[10] Sakai, H. “Statistical properties of AR spectral analysis.” *IEEE Trans. ASSP*, 27(4), 402–409, Aug. 1979.

[11] Sherman, P., White, L., W., Spanjaard, J., and Bitmead R. “Asymptotic statistics of AR spectral estimators for processes containing mixed spectrum.” *8th IEEE Sig. Process. Workshop on Stat. Array Process, Corfu, Greece*, 48–51, 1996.

[12] Sherman, P. and Lou, K.L. “On the family of ML spectral estimates for mixed spectrum identification.” *IEEE Trans. ASSP*, 39(3), 644–655, March 1991.