ON THE ASYMPTOTIC NORMALITY OF AUTOREGRESSIVE SPECTRAL DENSITY ESTIMATES FOR THE NOISE CORRUPTED CASE

D. F. Gingras

Naval Ocean Systems Center

San Diego, California

ABSTRACT

Asymptotic statistics for spectral density estimates of noise corrupted autoregressive (AR) series are evaluated. The "high-order" Yule-Walker equation estimates of the autoregressive parameters are used to form a spectral density estimate. The estimate is shown to be a consistent asymptotically normal (CAN) estimate. An expression for the variance of the limiting distribution in terms of the AR process parameters and the noise variance is provided.

INTRODUCTION

We consider the evaluation of asymptotic statistics associated with the estimation of the spectral density for an autoregressive process observed in additive white noise. The asymptotic statistics for the spectral density estimate without the presence of additive noise have been considered by Akaike [1], Kromer [2], and Berk [3]. For the noise corrupted case asymptotic statistics for the AR parameter estimates have been evaluated by Walker [4], Pagano [5], Gingras [6] and Lee [7]. The combined problem of evaluating the asymptotic statistics for the spectral density estimate in the presence of additive noise has not previously been considered.

In this paper we assume that the parameters of the AR process $\{\sigma^2, a_1, \cdots, a_p\}$ are estimated from observations of the noise corrupted series using the high-order Yule-Walker (Y-W) equations. The spectral estimate is formed by substitution of the parameter estimates into the AR spectral density function. We show that the resulting spectral density estimate is a consistent asymptotically normal (CAN) estimate. We calculate an exact expression for the variance of the limiting distribution. In its general form the variance expression is formidable, but its evaluation for specific cases should be straight forward.

This work was supported by the Office of Naval Research, Probability and Statistics Program.

AR PLUS NOISE MODEL

Assume that the observed process Y is a real process consisting of the sum of a stationary autoregressive (AR) process X and a noise process w, that is

$$Y = X + w. \tag{1}$$

The AR process X, assumed to be of known order p, is generated or adequately modeled by

$$X_n - a_1 X_{n-1} - \cdots - a_p X_{n-p} = \varepsilon_n \tag{2}$$

where the sequence $\{\varepsilon_n\}$ is assumed to be Gaussian i.i.d. with zero mean and variance σ_ε^2. The noise sequence $\{w_n\}$ is assumed to be wide sense stationary, Gaussian i.i.d. with zero mean and variance σ_w^2. The noise w and the AR process X are assumed to be uncorrelated.

Define the polynomial in z, z complex, by

$$A^p(z) = 1 - \sum_{j=1}^{p} a_j z^j . \tag{3}$$

The AR parameters $\{a_j\}_{j=1}^p$ are chosen such that the zeros of $A^p(z)$ lie outside of the unit circle on the z-plane. This guarantees that the AR process is stationary. The spectral density function for the stationary noise corrupted process Y is given by

$$\phi_Y(\lambda) = \frac{\sigma_w^2}{2\pi} + \frac{\sigma_\varepsilon^2}{2\pi\, A^p(e^{i\lambda})\, A^p(e^{-i\lambda})} . \tag{4}$$

Walker [4] and Pagano [5] showed that the AR plus noise process of (1) can be written as a special case of an autoregressive-moving average (ARMA) process. We can write (1) as

$$\begin{aligned} &Y_n - a_1 Y_{n-1} - \cdots - a_p Y_{n-p} \\ &\quad = \varepsilon_n + w_n - a_1 w_{n-1} - \cdots - a_p w_{n-p} . \end{aligned} \tag{5}$$

Define the covariance sequence for the Y process to be $\{r_k\}$, where $r_k = E[Y_n Y_{n-k}]$. Multiplying (5) through by Y_{n-k} and taking expectations term by term we obtain the following "Yule-Walker" (Y-W) equations:

$$r_0 - a_1 r_1 - \cdots - a_p r_p = \sigma_\varepsilon^2 + \sigma_w^2 \tag{6}$$

$$r_k - a_1 r_{k-1} - \cdots - a_p r_{k-p} = -a_k \sigma_w^2 \quad (1 \le k \le p) \tag{7}$$

$$r_k - a_1 r_{k-1} - \cdots - a_p r_{k-p} = 0 \quad (p+1 \le k \le 2p) . \tag{8}$$

The set of p equations (8) are referred to as the high-order Y-W equations and can be used in conjunction with estimates of the covariances r_k to provide unbiased estimates of the AR parameters of an ARMA process. Let $\underline{\Gamma}_p$ be a (pxp) covariance matrix with elements $\Gamma_{k,j} = r_{p+k-j}$, then

$$\underline{\Gamma}_p \triangleq \begin{bmatrix} r_p & r_{p-1} & \cdots & r_1 \\ r_{p+1} & r_p & \cdots & r_2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ r_{2p-1} & r_{2p-2} & \cdots & r_p \end{bmatrix} .$$

Define the (1xp) vectors

$$\underline{a}^T \triangleq [a_1, a_2, \cdots, a_p]$$

$$\underline{R}_{p+1}^T \triangleq [r_{p+1}, r_{p+2}, \cdots, r_{2p}]$$

then the p equations of (8) can be written as

$$\underline{\Gamma}_p \underline{a} = \underline{R}_{p+1} . \tag{9}$$

Gersch [8] proved that the nonsymmetric Toeplitz matrix $\underline{\Gamma}_p$, p finite, is nonsingular, thus a solution for (9) always exists.

Given a finite set of observations of the noise corrupted process Y, that is $\{Y_n\}_{n=1}^N$ N > 2p we estimate the covariance sequence $\{r_k\}$ by

$$\hat{r}_k = \begin{cases} \dfrac{1}{(N-|k|)} \displaystyle\sum_{n=1}^{N-|k|} Y_n Y_{n+|k|} & |k| \le N-1 \\ 0 & |k| > N-1 . \end{cases} \tag{10}$$

When the covariances r_k in the matrix $\underline{\Gamma}_p$ and the vector $\underline{R}_{p+1}$ are replaced by their corresponding estimates using (10) the estimated matrix and vector will be denoted by $\hat{\underline{\Gamma}}_p$ and $\hat{\underline{R}}_{p+1}$, respectively. The high-order Y-W equations (9) can be expressed in terms of the estimated covariances as

$$\hat{\underline{\Gamma}}_p \hat{\underline{a}} = \hat{\underline{R}}_{p+1} . \tag{11}$$

To estimate the spectral density we require estimates of the AR parameters from observations of the noise corrupted process Y, such as by (11) and an estimate of σ_ε^2. Because of the presence of the noise w the usual estimate of the variance of the AR process from (6) is not adequate. For the noise corrupted case (6) will provide an estimate of $\sigma_\varepsilon^2 + \sigma_w^2$, thus one of the equations of (7) must be used to estimate σ_w^2. Using this approach, with the covariance estimates of (10), and the estimates of the AR parameters of (11) we have

$$\hat{\sigma}_\varepsilon^2 = \sum_{j=0}^{p} \hat{a}_j \hat{r}_j + (1/\hat{a}_p) \sum_{j=0}^{p} \hat{a}_j \hat{r}_{p-j} \tag{12}$$

where $a_0 = -1$ and $a_p \neq 0$.

SPECTRAL ESTIMATE STATISTICS

Define the parameter vector $\underline{\theta}^T$ by

$$\underline{\theta}^T = [\sigma_\varepsilon^2, a_1, \cdots, a_p].$$

By (11) and (12) we form estimates for $\underline{\theta}^T$ and use these to form the spectral density estimate, i.e.,

$$\hat{\phi}_X(\lambda, \hat{\underline{\theta}}) = \frac{\hat{\sigma}_\varepsilon^2}{2\pi\, \hat{A}^p(e^{i\lambda})\, \hat{A}^p(e^{-i\lambda})} \tag{13}$$

where $\hat{A}^p(e^{i\lambda})$ is formed by substituting the AR parameter estimates into (3) and evaluating at $z = e^{i\lambda}$. In order to evaluate asymptotic statistics for $\hat{\phi}_X(\lambda, \hat{\underline{\theta}})$ we first establish asymptotic statistics for $\hat{\underline{\theta}}$.

Define the vectors and matrix:

$$\underline{0} \triangleq [0, 0, \cdots, 0] \text{ (1xp)}$$

$$\tilde{\underline{a}}^T \triangleq [a_p, a_{p-1}, \cdots, a_1]$$

$$\underline{R}^T \triangleq [r_1, \cdots, r_{2p}]$$

$$\underline{R}_0^T \triangleq [r_0, r_1, \cdots, r_p]$$

$$\underline{R}_p^T \triangleq [r_p, r_{p+1}, \cdots, r_{2p-1}]$$

$$\underline{D} \triangleq \begin{bmatrix} -a_p & -a_{p-1} & \cdots & -a_1 & 1 & 0 & \cdots & 0 \\ 0 & -a_p & & & -a_1 & 1 & 0 \cdots & 0 \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ \cdot & & & & & & & \cdot \\ 0 & & \cdots & -a_p & & \cdots & -a_1 & 1 \end{bmatrix} .$$

From the development in [6] we have

$$N^{\frac{1}{2}}(\hat{\underline{a}} - \underline{a}) \sim N^{\frac{1}{2}} \underline{\Gamma}_p^{-1} \underline{D} (\hat{\underline{R}} - \underline{R}) \tag{14}$$

where $\sim$ indicates that the limit distribution, as $N \to \infty$, is identical for both random vectors. By (12) and the fact that the high-order Y-W equation estimates of the AR parameters converge in probability to the true parameters we have

$$N^{\frac{1}{2}}(\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2) \sim N^{\frac{1}{2}} \{[-1, \underline{a}^T][\hat{\underline{R}}_0 - \underline{R}_0] - (1/a_p)[\tilde{\underline{a}}^T, -1][\hat{\underline{R}}_0 - \underline{R}_0]\} \quad . \tag{15}$$

Under our assumption that the Y process is a linear Gaussian process, by Hannan [9] and Walker [10], we have that the vector $N^{\frac{1}{2}}(\hat{\underline{R}} - \underline{R})$ is asymptotically jointly normal and

$$\lim_{N\to\infty} E[N^{\frac{1}{2}}(\hat{\underline{R}} - \underline{R}), N^{\frac{1}{2}}(\hat{\underline{R}} - \underline{R})^T] = 2\pi \int_{-\pi}^{\pi} \underline{U}^2 \phi_Y(\lambda) d\lambda \tag{16}$$

where the matrix $\underline{U}$ is defined by

$$\underline{U} \triangleq \quad e^{i(k+j)\lambda} + e^{i(k-j)\lambda} \qquad \begin{matrix} k=1, \cdots, 2p \;; \\ j=1, \cdots, 2p \;. \end{matrix}$$

By (14) and (15) and the above result it is straight forward to conclude that

$$N^{\frac{1}{2}}(\hat{\underline{\theta}} - \underline{\theta}) \xrightarrow[N\to\infty]{\mathscr{L}} N_{p+1}(\underline{0}, \underline{\Sigma}) \tag{17}$$

that is, the parameter error vector converges in distribution to a zero mean p+1 variate normal random vector with limiting covariance $\underline{\Sigma}$. We now evaluate the form of the covariance matrix $\underline{\Sigma}$.

Define ζ and $\underline{Z}$ by

$$\zeta \triangleq [-1, \underline{a}^T][\hat{\underline{R}}_0 - \underline{R}_0] - (1/a_p)[\tilde{\underline{a}}^T, -1][\hat{\underline{R}}_0 - \underline{R}_0]$$

$$\underline{Z} \triangleq \underline{\Gamma}_p^{-1} \underline{D} (\hat{\underline{R}} - \underline{R})$$

then

$$\underline{\Sigma} = \lim_{N\to\infty} N\, E \left[\begin{bmatrix} \zeta \\ \underline{Z} \end{bmatrix} [\zeta, \underline{Z}^T]\right] \triangleq \left[\begin{array}{c|c} v^2 & \underline{C}^T \\ \hline \underline{C} & \underline{\Phi} \end{array}\right].$$

Now, using (15) and (16) we can write

$$v^2 = \lim_{N\to\infty} N\, E[\zeta^2]$$

$$= \sigma_\varepsilon^4 + \sigma_w^2 + 2\sigma_w^2\sigma_\varepsilon^2 - 2\sigma_w^4 - (2/a_p)\sigma_\varepsilon^2 r_p$$

$$+ [1 + (1/a_p)^2][\sigma_w^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 r_0 + \sigma_w^4 \sum_{j=0}^{p} a_j^2]. \tag{18}$$

$$\underline{C}^T = \lim_{N\to\infty} N\, E[\zeta\, \underline{Z}^T]$$

$$= \sigma_\varepsilon^2 \underline{R}_p^T (\underline{\Gamma}_p^{-1})^T - (1/a_p)\underline{P}^T(\underline{\Gamma}_p^{-1})^T \tag{19}$$

where

$$\{\underline{P}^T\}_j \triangleq [(a_p)^2\sigma_w^4\delta(j) + \sigma_\varepsilon^2 r_j + \sigma_w^4 \sum_{k=0}^{p-|j|} a_k a_{k+j}] \qquad j=0, 1, \cdots, p-1$$

and $\delta(j) = \begin{cases} 1; & j=0 \\ 0; & j\neq 0 \end{cases}$

Using (14) and (16)

$$\underline{\Phi} = \lim_{N\to\infty} N\, E[\underline{Z}\, \underline{Z}^T]$$

$$= \sigma_\varepsilon^2 \underline{\Gamma}_p^{-1} \underline{\Gamma}_0 (\underline{\Gamma}_p^{-1})^T + \sigma_w^2 \underline{\Gamma}_p^{-1}[\sigma_\varepsilon^2 \underline{I} + \sigma_w^2 \underline{Q}](\underline{\Gamma}_p^{-1})^T \tag{20}$$

where

$$\underline{Q} \triangleq \begin{bmatrix} \sum_{m=0}^{p} a_m^2 & \cdots & \sum_{m=0}^{1} a_m a_{m+(p-1)} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \sum_{m=0}^{1} a_m a_{m+(p-1)} & \cdots & \sum_{m=0}^{p} a_m^2 \end{bmatrix}$$

We can now proceed to show that our estimate of the spectral density $\phi_X(\lambda,\underline{\theta})$ is a consistent asymptotically normal estimate and evaluate the variance of the limiting distribution. Since the function $\phi_X(\lambda,\underline{\theta})$ is totally differentiable then using (17) and a convergence theorem in Rao [11] we have that

$$N^{\frac{1}{2}}[\hat{\phi}_X(\lambda,\hat{\underline{\theta}}) - \phi_X(\lambda,\underline{\theta})] \xrightarrow[N\to\infty]{\mathscr{L}} N\,(0,\ \underline{\rho}^T\ \underline{\Sigma}\ \underline{\rho})$$

where $\underline{\rho}$ is a gradient vector given by

$$\underline{\rho}^T = \left[\frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial\sigma_\varepsilon^2},\ \frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial a_1},\ \cdots,\ \frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial a_p}\right].$$

Let

$$b \triangleq \frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial\sigma_\varepsilon^2} = \frac{1}{2\pi\ A^p(e^{i\lambda})A^p(e^{-i\lambda})}$$

$$\underline{B}^T(\lambda) \triangleq \left[\frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial a_1},\ \cdots,\ \frac{\partial\phi_X(\lambda,\underline{\theta})}{\partial a_p}\right]$$

$$= 2\phi_X(\lambda,\underline{\theta})\left[\mathrm{Re}\left\{\frac{e^{i\lambda}}{A(e^{i\lambda})}\right\},\ \cdots,\ \mathrm{Re}\left\{\frac{e^{ip\lambda}}{A(e^{i\lambda})}\right\}\right].$$

Thus, the variance of the limiting distribution for the spectral density estimate is given by

$$\underline{\rho}^T\ \underline{\Sigma}\ \underline{\rho} = b^2v^2 + b\underline{B}^T(\lambda)\underline{C} + b\underline{C}^T\underline{B}(\lambda) + \underline{B}^T(\lambda)\ \underline{\Phi}\ \underline{B}(\lambda)$$

CONCLUSIONS

We have shown that given observations of a noise corrupted AR process the high-order Y-W equation estimates of the AR parameters produce an estimate of the spectral density that is a consistent asymptotically normal (CAN) estimate. We have also developed an exact expression for the variance of the limiting normal distribution. This general variance expression is formidable, but specific low order cases can be evaluated relatively easily.

For the noise corrupted case we now have a method for comparing the estimator stability for finite sample size experiments with a theoretical limiting value specified in terms of the AR process parameters and the noise variance.

REFERENCES

[1] H. Akaike, "Power spectrum estimation through autoregressive model fitting." Ann.Inst. Statist.Math,vol.21,407-419, 1969.

[2] R. Kromer, "Asymptotic properties of the autoregressive spectral estimator." Ph.D. thesis, Dept. of Statistics, Stanford Univ.,Stanford, CA,1969.

[3] K.N. Berk, "Consistent autoregressive spectral estimates." Ann.Statist.,vol.2,no.3,489-502, 1974.

[4] A.M. Walker, "Some consequences of superimposed error in time series analysis." Biometrika,vol.47,33-43,1960.

[5] M. Pagano, "Estimation of models of autoregressive signal plus white noise." Ann. Statist.,vol.2,99-108,1974.

[6] D.F. Gingras, "Estimation of the autoregressive parameters from observations of a noise corrupted autoregressive time series." Proc.IEEE Int.Conf.Acoust.,Speech,Signal Processing,Paris,France, 228-231,1982.

[7] T.S. Lee, "Large sample identification and spectral estimation of noisy multivariate autoregressive processes." IEEE Trans.Acoust., Speech,Signal Processing,vol.ASSP-31,76-82, 1983.

[8] W. Gersch, "Estimation of the autoregressive parameters of a mixed autoregressive moving average time series." IEEE Trans.Automat. Contr.,vol.AC-15, 583-588,1970.

[9] E.J. Hannan, Time Series Analysis, London: Methuen, 1960.

[10] A.M. Walker, "The asymptotic distribution of serial correlation coefficients for autoregressive processes with dependent residuals." Proc.Camb.Phil.Soc.,vol.50,60-64, 1954.

[11] C.R. Rao, Linear Statistical Inference and Its Application. New York: John Wiley and Sons, Inc., 1965.