한국통계학회논문집 제2권 2호
1995년 12월 pp. 94-100

Comments on Functional Relations in the Parameters of Multivariate Autoregressive Process Observed with Noise[1)]

Jong Hyup Lee[2)] and Dong Wan Shin[3)]

Abstract

Vector autoregressive process disturbed by measurement error is a vector autoregressive process with nonlinear parametric restrictions on the parameter. A Newton-Raphson procedure for estimating the parameter which take advantage of the information contained in the restrictions is proposed.

1. Introduction

There are many situations in which a time series is subject to another measurement error. One example of such autoregressive(AR) process disturbed by noise can be founded in signal processing area. One send a signal process and then receiver may observe the signal with unknown noise. To handle this situation, several authors such as Pagano(1974), Sakai and Arase(1979) and Shin(1993) have considered a univariate AR signal observed with noise

$$y_t = x_t + u_t, \tag{1.1}$$

$$x_t + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} = \omega_t, \tag{1.2}$$

$t = 1,2,\cdots,n$. The series $\{y_t\}_{t=1}^{n}$ is a set of observations and the two independent error sequences $\{\omega_t\}$ and $\{u_t\}$ are independent, identically distributed(*i.i.d.*) with variances $\sigma_{\omega\omega}$ and σ_{uu}, respectively. They assumed stationarity for x_t by imposing that all the roots of the characteristic equation $1 + \phi_1 z + \cdots + \phi_p z^p = 0$ lie outside the unit circle.

In this note, we develop an estimating procedure for vector autoregressive process corrupted by white noise. There are several approaches for estimating unrestricted multivariate ARMA model. Spliid(1983) and Koreisha and Pukkila(1989) modified and extended the second stage of Hannan and Rissanen's(1982) procedure to the multivariate ARMA model. Reinsel et al.(1992) gave a Newton-Raphson iterative procedure to obtain maximum likelihood estimates, using the

---

1 The second author is partly supported by a grant from Korea Science and Engineering Foundation.
2 Department of Statistics, Sungshin Women's University, Seoul, 136-742, KOREA.
3 Department of Statistics, Ewha Womans University, Seoul, 120-750, KOREA.

initial estimates of Koreisha and Pukkila(1989) and the multivariate version of Hannan and Rissanen's(1982) estimates. On the other hand, Shin and Sarkar(1995) proposed a Newton-Raphson approximation to the restricted maximum likelihood estimator of the parameters which take advantage of the information contained in the restrictions. Also, they discussed the limiting distribution of the Newton-Raphson estimator and the numerical evaluation of the derivatives of likelihood function.

Since any order of vector autoregressive process is represented as first order vector autoregressive process(Anderson, 1959), it is sufficient to assume first order autoregressive process. In order to give motivation for our analysis, we first review the scalar first order autoregressive process.

The problem of estimating the parameters based on the data corrupted by unknown noise is equivalent to that of estimating an autoregressive moving average(ARMA) process parameters. Let L be the lag operator such that $L^j x_t = x_{t-j}$. Then, model (1.2) can be written $\phi(L)x_t = \omega_t$, where $\phi(L) = 1 + \phi_1 L$. Hence from (1.1)

$$\phi(L)y_t = \omega_t + \phi(L)u_t. \tag{1.3}$$

Since autocovariance function of the right side of (1.3) is zero when the lag is greater than one, we can find a first order moving average $e_t + \gamma_1 e_{t-1}$ whose autocovariance function is the same as that of the right side of (1.3). Thus we can write

$$\phi(L)y_t = \gamma(L)e_t, \tag{1.4}$$

where $\{e_t\}$ is a sequence of *i.i.d.* random variables with variance, say, σ_{ee} and $\gamma(L) = 1 + \gamma_1 L$. Note that since $\sigma_{\omega\omega} \neq 0$, $\phi(L)$ and $\gamma(L)$ are assumed to have no common roots. Also by Wilson's algorithm(Box and Jenkins(1976, p.203)), we find a unique $\gamma(L)$ which has characteristic roots outside the unit circle.

In the parameterization (1.4), we can take advantage of the ARMA estimation procedure established in the literature. However, when σ_{uu} is unknown, the number of parameters in $(\phi_1, \sigma_{uu}, \sigma_{\omega\omega})$ of model (1.3) is the same as that of parameters in $(\phi_1, \gamma_1, \sigma_{ee})$ of model (1.4). Therefore, there is no functional relations in the parameters of the transformed model (1.4).

On the other hand, in various fields such as engineering and economics, the vector ARMA process is one of the most applicable models for representing multivariate time series data. For a set of $r \times 1$ observation vectors $\{Y_t\}_{t=1}^{n}$, an r-variate AR model with error of the model (1.1)-(1.2) can be written as

$$Y_t = X_t + U_t, \tag{1.5}$$

$$X_t + A_1 X_{t-1} = \Omega_t, \tag{1.6}$$

where A_1 is $r \times r$ matrix of unknown parameters, $\{\Omega_t\}$ and $\{U_t\}$ are sequences of *i.i.d.* r-dimensional random vectors with mean zero and a nonsingular variance-covariance matrices $\Sigma_{\Omega\Omega}$ and Σ_{UU}, respectively. Let $a(L) = I + A_1 L$, where I is an identity matrix. Then, from (1.5)-(1.6)

$$a(L)Y_t = \Omega_t + a(L)U_t. \tag{1.7}$$

Similar to the univariate case of (1.4), we can represent model (1.7) as

$$a(L)Y_t = b(L)E_t, \tag{1.8}$$

where $\{E_t\}$ is a sequence of *i.i.d.* random vectors with a variance-covariance matrix, say, Σ_{EE} and $b(L) = I + B_1 L$, where B_1 is an $r \times r$ matrix.

When Σ_{UU} is unknown, the number of parameters in $(A_1, \Sigma_{UU}, \Sigma_{\Omega\Omega})$ of model (1.7) is $(r^2 + r(r+1))$ and the number of parameters in (A_1, B_1, Σ_{EE}) of model (1.8) is $(2r^2 + r(r+1)/2)$. Hence when $r \geq 2$, there are $r(r-1)/2 = (2r^2 + r(r+1)/2) - (r^2 + r(r+1))$ number of restrictions in the parameters (A_1, B_1, Σ_{EE}). It is necessary to find a method for estimating multivariate ARMA models with the parametric restrictions.

The aim of this paper is to investigate the parametric restrictions for $r(\geq 1)$-variate AR(1) signal process with noise and develop an estimating procedure which incoporates the restrictions. We derive the functional relations of the parameters for $r \geq 2$ and propose a restricted Newton-Raphson procedure. An explicit expressions of derivative of the restrictions are also given. Limiting distribution of the estimator is established.

2. Parametric Restrictions and a Newton-Raphson Procedure

Since $b(L)E_t$ and $\Omega_t + a(L)U_t$ should have the same autocovariance generating function, it follows that

$$b(L)\Sigma_{EE}[b(L^{-1})]' - \Sigma_{\Omega\Omega} - a(L)\Sigma_{UU}[a(L^{-1})]' = 0. \tag{2.1}$$

Let G_i , $i = 0,1$ be the coefficients of L^i in the expansion (2.1). Then the restriction (2.1) is $G_i = 0$, $i = 0,1$. Note that G_0 depends on $\Sigma_{\Omega\Omega}$ which can be obtained by letting $\Sigma_{\Omega\Omega} = \sum_{j=0}^{1} B_j \Sigma_{EE} B_j' - \sum_{j=0}^{1} A_j \Sigma_{UU} A_j'$ from $G_0 = 0$, where $A_0 = B_0 = I$. Therefore, $G_0 = 0$ is not a real parametric restriction on ξ. The parametric restrictions (2.1) are then given by

$$\begin{aligned} & G_1 = \Sigma_{EE} B_1' - \Sigma_{UU} A_1' = 0 \\ \leftrightarrow\ & A_1 \Sigma_{EE} B_1' = A_1 \Sigma_{UU} A_1' \\ \leftrightarrow\ & A_1 \Sigma_{EE} B_1' \text{ is symmetric.} \end{aligned} \tag{2.2}$$

Since there are $r(r-1)/2$ number of restrictions in the final expression of (2.2), (2.2) is the restrictions we are trying to establish. For a matrix M, we let vec(M) denote the vector obtained by stacking the columns of M, and vech(M) denote the vector obtained by stacking the columns of matrix M with all elements above the diagonal deleted. Let $\otimes$ denote the Kronecker product. Let $\alpha = \mathrm{vec}(A_1')$, $\beta = \mathrm{vec}(B_1')$, $\theta = (\alpha', \beta')'$, $\eta = \mathrm{vech}(\Sigma_{EE})$, and $\xi = (\alpha', \beta', \eta')'$. Letting $E_t(\theta) = 0$ for $t \le 0$, we can rewrite (1.8) to get

$$\begin{aligned} E_t(\theta) &= Y_t + A_1 Y_{t-1} - B_1 E_{t-1}(\theta) \\ &= Y_t + (I_r \otimes Y_{t-1}')\alpha - (I_r \otimes E_{t-1}'(\theta))\beta , \end{aligned} \tag{2.3}$$

where I_r is the r-dimensional identity matrix. To define the conditional Gaussian likelihood of $\{Y_t\}_{t=1}^{n}$, we have assumed, as in Reinsel et al.(1992), that the initial observation Y_0 is available and is considered fixed, and the intial disturbance $E_0 = 0$. Then, the negative logarithm of the conditional Gaussian likelihood function of $\{Y_t\}_{t=1}^{n}$ is approximated by

$$L_n = L_n(\xi) = 2^{-1} \sum_{t=1}^{n} E_t'(\theta) \Sigma_{EE}^{-1} E_t(\theta) + 2^{-1} n \log \det[\Sigma_{EE}]. \tag{2.4}$$

From Reinsel, Basu, and Yap(1992), an unrestricted Newton-Raphson procedure for estimating the model (1.8) is

$$\xi^{+} = \tilde{\xi} - \hat{H}_{\xi\xi}^{-1} \hat{H}_{\xi} . \tag{2.5}$$

Here $\tilde{\xi}$ is the consistent initial estimator of ξ such as Hannan and Rissanen's(1982) estimator. $\hat{H}_{\xi\xi} = \partial^2 L_n(\tilde{\xi}) / \partial \xi \partial \xi'$, and $\hat{H}_{\xi} = \partial L_n(\tilde{\xi}) / \partial \xi$.

Now we modify the Newton-Raphson estimator ξ^{+} in (2.5) to accommodate the restriction (2.1) which reduces to symmetry of $A_1 \Sigma_{EE} B_1'$ as in (2.2). Let g_1 be the vector of restrictions in (2.2) which will be stated explicitly later. The restricted estimator is obtained by minimizing the Lagrangian function $L_n + g_1' \lambda$, $\lambda \in R^k$, $k = r^2$ with respect to ξ and λ. Setting the first partial derivatives of $(L_n + g_1' \lambda)$ with respect to ξ and λ to zero, we get

$$L_{\xi} + F' \lambda = 0 \quad \text{and} \quad g_1 = 0 , \tag{2.6}$$

where $L_\xi = \partial L_n/\partial \xi$ and $F = \partial g_1/\partial \xi'$. By Shin and Sarkar(1995), the Newton-Raphson estimator is obtained as

$$\begin{bmatrix} \hat{\xi} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \tilde{\xi} \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{H}_{\xi\xi} & \tilde{F}' \\ \tilde{F} & 0 \end{bmatrix}^{-1} \begin{bmatrix} \tilde{H}_{\xi} \\ \tilde{g}_1 \end{bmatrix}. \qquad (2.7)$$

We give an explicit expression for F in (2.7). Note that

$$F = \partial g_1/\partial \xi' = [\,\partial g_1/\partial \alpha', \partial g_1/\partial \beta', \partial g_1/\partial \eta'\,].$$

Now the real $r(r-1)/2$ restrictions in (2.2) is vech $(A_1 \Sigma_{EE} B_1')$ = vech $(B_1 \Sigma_{EE} A_1')$ except the identities corresponding to diagonal elements of $A_1 \Sigma_{EE} B_1' = B_1 \Sigma_{EE} A_1'$. Let

$$D = \begin{bmatrix} 0_{r-1,1} & I_{r-1} & 0_{r-1,1} & 0_{r-1,r-2} & \cdots & 0_{r-1,1} & 0_{r-1,1} \\ 0_{r-2,1} & 0_{r-2,r-1} & 0_{r-2,1} & I_{r-2} & \cdots & 0_{r-2,1} & 0_{r-2,1} \\ \vdots & & \vdots & & & \vdots & \\ 0_{1,1} & 0_{1,r-1} & 0_{1,1} & 0_{1,r-2} & \cdots & I_1 & 0_{1,1} \end{bmatrix}$$

be a $[(r-1)r/2] \times [r(r+1)/2]$ matrix, where $0_{i,j}$ is the $i \times j$ zero matrix. Then D vech(M) is the vector obtained by stacking columns of a $r \times r$ matrix M with all diagonal elements and above-diagonal elements removed. Hence the restriction we are trying to clarify is

$$g_1 = D\Psi \quad \text{vec}\,(A_1 \Sigma_{EE} B_1') - D\Psi \quad \text{vec}\,(B_1 \Sigma_{EE} A_1') = 0, \qquad (2.8)$$

where Ψ is a matrix such that vech $(M) = \Psi$ vec (M) given in Fuller(1987, p383). Differentiating (2.8) gives

$$\partial g_1/\partial \alpha' = D\Psi[(B_1 \Sigma_{EE}) \otimes I]K - D\Psi[I \otimes (B_1 \Sigma_{EE})],$$

$$\partial g_1/\partial \beta' = D\Psi[I \otimes (A_1 \Sigma_{EE})] - D\Psi[(A_1 \Sigma_{EE}) \otimes I]K,$$

$$\partial g_1/\partial \eta' = D(A_1 \otimes B_1')\Phi - D(B_1 \otimes A_1')\Phi,$$

where K is a matrix such that vec $(M) = K$vec (M') and Φ a matrix such that vec (Σ_{EE}) = Φ vech (Σ_{EE}) which can be found in Fuller(1987, p383). Now we can compute an estimate of A_1 for model (1.7) by applying (2.7) with F and g_1 described above.

In Theorem 1 below, we finally give the limiting distribution of our estimator.

Theorem 1. Assume that model (1.6) satisfies the following conditions: (i) All roots of det[$a(L)$] = 0 lie outside the unit circle. (ii) The matrix A_1 is of full rank. Then

$$\sqrt{n}(\hat{\xi} - \xi) \xrightarrow{d} N(0, \Gamma),$$

where Γ is the $(2r^2+r(r+1)/2)\times(2r^2+r(r+1)/2)$ upper left block of matrix $\begin{bmatrix} V & F' \\ F & 0 \end{bmatrix}^{-1}$,

$V=$ diag $(I(\theta),\Sigma_{EE}^{-1})$, and $I(\theta)$ is the information matrix of θ in the unrestricted model (1.8).

Proof. As indicated in Hallin(1984, p.190), $b(L)$ can be uniquely chosen such that all the roots of det$[b(L)]$ have absolute value greater than one. Since det$[\Sigma_{\omega\omega}]\neq 0$, $b(L)$ and $a(L)$ can not have a left common factor which has no constant determinant. Therefore, the conditions C1 - C3 of Shin and Sarkar(1995) are satisfied and Theorem 1 of Shin and Sarkar is applicable to give the limiting distribution.

References

[1] Anderson, T. W. (1959). On asymptotic distributions of estimates of parameters of stochastic difference equations, *Annals of Mathematical Statistics,* Vol. 21, 243-247.

[2] Box, G.E.P. and Jenkins, G.M. (1976). *Time series analysis:Forecasting and Control,* Holden-Day, San Francisco.

[3] Dunsmuir, W. and Hannan, E.J. (1976). Vector linear time series models, *Advances in Applied Probability,* Vol. 8, 339-364.

[4] Fuller, W.A. (1987). *Measurement error models,* New York: John Wiley.

[5] Hallin, M. (1984). Spectral fatorization of nonstationary moving average process, *Annals of Statistics,* Vol. 12, 172-192.

[6] Hannan, E.J. and Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order, *Biometrika,* Vol. 69, 81-94.

[7] Koreisha, S. and Pukkila, T. (1989). Fast linear estimation methods for vector autoregressive moving-average models, *Journal of Time Series Analysis,* Vol. 10, 325-339.

[8] Pagano, M. (1974). Estimation of models of autoregressive signal plus white noise, *Annals of Statistics,* Vol. 2, 99-108.

[9] Reinsel, G.C., Basu, S., and Yap, S.F. (1992). Maximum likelihood estimators in the multivariate autoregressive moving-average model from a generalized least squares view point, *Journal of Time Series Analysis,* Vol. 13, 133-145.

[10] Sakai, H. and Arase, M. (1979). Recursive parameter estimation of an autoregressive process disturbed by white noise, *International Journal of Control,* Vol. 30, 949-966.

[11] Shin, D.W. (1993). Maximum likelhood estimation for autoregressive process disturbed by a moving average, *Journal of Time Series Analysis,* Vol. 14, 629-643.

[12] Shin, D.W. and Sarkar, S. (1995). Estimation of the multivariate autoregressive moving average having parameter restrictions and an application to rotational sampling, *Journal of Time Series Analysis,* Vol. 16, 431-444.

[13] Spliid, H. (1983). A fast estimation method for the vector autoregressive moving average model with exogenous variables, *Journal of American Statistical Association,* Vol. 78, 843-849.